# The Problem of Measure Sensitivity Redux

Peter Brössel*†

Fitelson demonstrates that the validity of various central arguments within Bayesian confirmation theory depends on which confirmation measure is adopted. This article adds to the results set out by Fitelson, expanding on them in two principal respects. First, it considers more confirmation measures. Second, it shows that there are important arguments within Bayesian confirmation theory and that there is no confirmation measure that renders them all valid. Finally, the article reviews the ramifications that this "strengthened problem of measure sensitivity" has for Bayesian confirmation theory and discusses whether it points at pluralism about notions of confirmation.

**1. Introduction.** A plurality of Bayesian confirmation measures have been proposed as a means of providing a theory of confirmation. However, this plurality of measures of (incremental) confirmation is now seen as bringing about one of the biggest challenges with which Bayesian confirmation theory is faced: the problem of measure sensitivity. Eells and Fitelson, for example, write that "a plethora of non-equivalent measures of evidential support (or confirmation) have been proposed and defended in the philosophical literature. This plurality of measures of support is problematic, since it affects a great many arguments surrounding Bayesian confirmation theory (and inductive logic, generally)" (2002, 129).

Received July 2012; revised January 2013.

*To contact the author, please write to: Institut für Philosophie II, Universität Bochum; e-mail: peter.broessel@rub.de.

†I am very grateful to four exceedingly helpful referees of this journal, especially for spotting a mistake in one of my proofs and encouraging me to discuss, among other issues, the ramifications of the problem of measure sensitivity in more detail in section 4. In addition, I would like to thank Ralf Busse, Vincenzo Crupi, Anna-Maria A. Eder, Branden Fitelson, Franz Huber, James M. Joyce, and Hannes Leitgeb for their invaluable feedback on various versions of this article. I also want to emphasize my gratitude to James M. Joyce for the permission to cite and quote his unpublished manuscript "On the Plurality of Probabilist Measures of Evidential Relevance." Finally, I also would like to thank Robert Lehnert and Ben Young for proofreading the manuscript.

Philosophy of Science, 80 (July 2013) pp. 378–397. 0031-8248/2013/8003-0007$10.00
Copyright 2013 by the Philosophy of Science Association. All rights reserved.



https://doi.org/10.1086/671172 Published online by Cambridge University Press

To date, the most forceful display of the problem of measure sensitivity is one provided by Fitelson (1999). In his paper, Fitelson shows that the validity of various central arguments within Bayesian confirmation theory depends on which Bayesian measure of confirmation is adopted. The current article adds to the results set out in Fitelson (1999), expanding on them in two principal respects. First, this article considers more confirmation measures; in Fitelson (1999) only a proper subclass of those confirmation measures considered here is taken into account. Second, this article shows that the problem of measure sensitivity actually runs deeper and is far more severe than is noted in Fitelson (1999).

The following section introduces the plurality of confirmation measures of Bayesian confirmation theory. Section 3 introduces the problem of measure sensitivity. In particular, this section considers four prominent arguments in Bayesian confirmation theory and shows that there is no confirmation measure that validates all of them. Subsections 3.1 and 3.2 introduce arguments already considered in Fitelson (1999), while subsections 3.3 and 3.4 consider two new arguments from Fitelson (2006) and Huber (2005), respectively. Section 4 discusses the ramifications that the strengthened problem of measure sensitivity has for Bayesian confirmation theory.

**2. The Plurality of Bayesian Measures of Confirmation.** Providing a theory of confirmation is often taken to be one of the pivotal tasks of Bayesian philosophy of science and epistemology since providing such a theory is considered to be central for understanding various aspects of scientific reasoning. Intuitively, how worthy of belief certain theories or hypotheses are depends on how well confirmed they are, and which experiments scientists should conduct depends on whether the possible experimental outcomes would confirm the theory or hypothesis in question. Others take Bayesian confirmation theory to govern "the steps in the *arguments* by which our current opinions are to be justified by the evidence that we have already assembled" (Lange 1999, 295). However, whether a Bayesian theory of confirmation can meet the expectations and hopes placed on it by Bayesian philosophers of science and epistemologists cannot be investigated until a complete theory of confirmation has been provided.

Bayesian confirmation theory holds that confirmation is best understood as a relation between evidence $E$ and hypothesis $H$, relative to some probability measure Pr and background assumptions $B$. A complete Bayesian theory of confirmation requires that one provide an interpretation of the probability measure Pr and say when and to what degree a hypothesis $H$ is confirmed by the evidence $E$ relative to the background assumptions $B$ (henceforth, mention of the background assumptions is suppressed). The current article will not address the former task, that is, answering the question of how one should interpret the probabilities (evidential proba-

https://doi.org/10.1086/671172 Published online by Cambridge University Press

bilities, subjective probabilities, logical probabilities, etc.) relevant in the context of Bayesian confirmation theory. Instead, it focuses on the latter task of saying when and to what degree a hypothesis is confirmed by the evidence relative to some probability function. To this day, no confirmation measure is widely enough accepted that it can be considered to be part of the canon of Bayesian confirmation theory. However, two points are now commonly agreed on. First, almost all Bayesian confirmation theorists argue that the only adequate (or at least interesting) notion of confirmation is incremental confirmation. Evidence $E$ is said to incrementally confirm hypothesis $H$ relative to some background assumptions $B$ and probability function Pr, if and only if $\Pr(H|E \cap B) > \Pr(H|B)$. Second, most Bayesian epistemologists lean explicitly or implicitly toward a monistic position regarding confirmation measures (Milne 1996; Christensen 1999; Fitelson 2001; Eells and Fitelson 2002). Monism of confirmation measures is the position that there is only one adequate confirmation measure (up to ordinal equivalence). For most Bayesian confirmation theorists, the only question that remains to be answered is which incremental confirmation measure to accept.

While Bayesian confirmation theorists are in agreement about when a hypothesis is incrementally confirmed by the evidence, they disagree about how the quantitative notion of degree of incremental confirmation provided by some evidence is to be explicated. To this end, various probabilistic measures of incremental confirmation have been proposed. Earman (1992, chap. 5), Fitelson (1999, 2001), and Crupi, Tentori, and Gonzalez (2007) provide excellent overviews of the proposed measures of incremental confirmation and the suggested conditions of adequacy for an explication of "degree of incremental confirmation." As already said, none of these confirmation measures can be considered to be part of the canon of Bayesian confirmation theory. Nevertheless, there are two requirements that any Bayesian confirmation theorist agrees with:

> **Requirement 2.1** (Demarcation). A function $\mathfrak{c}(\cdot, \cdot)$ is a measure of (incremental) confirmation relative to probability function Pr, only if there is some (marker or neutral value) $r \in \mathbb{R}$ so that, for all hypotheses $H$ and evidence $E$, if $1 > \Pr(H|E) > 0$ and $\Pr(E) > 0$, then
>
> $$\mathfrak{c}(H, E) \begin{cases} > r, & \Pr(H|E) > \Pr(H) \\ = r, & \Pr(H|E) = \Pr(H) \\ < r, & \Pr(H|E) < \Pr(H) \end{cases}.$$

This first requirement on confirmation measures is taken from Fitelson (2001) and requires that they should be measures of incremental confir-

https://doi.org/10.1086/671172 Published online by Cambridge University Press

mation.[1] In particular, it requires that if the evidence does not guarantee that the hypothesis in question is true or false, then there is a marker or neutral value $r$, such that the degree of confirmation of hypothesis $H$ by evidence $E$ is higher than the marker $r$, if and only if $H$ and $E$ are positively correlated with respect to probability function Pr. The degree of confirmation of hypothesis $H$ by evidence $E$ is lower than the marker $r$, if and only if $H$ and $E$ are negatively correlated with respect to probability function Pr. The degree of confirmation of hypothesis $H$ by evidence $E$ equals the marker $r$, if and only if $H$ and $E$ are probabilistically independent with respect to probability function Pr.

In addition, confirmation measures should satisfy what Joyce (2003a) calls the weak law of likelihood requirement, which "encapsulates a minimal form of Bayesianism to which all parties can agree" (sec. 3).

**Requirement 2.2** (Weak Law of Likelihood). If $\Pr(E|H_1) > \Pr(E|H_2)$ and $\Pr(E|\neg H_1) < \Pr(E|\neg H_2)$, then $\mathfrak{c}(H_1, E) > \mathfrak{c}(H_2, E)$.

Joyce (2003a, sec. 3) even contends that "the weak likelihood principle must be an integral part of any account of evidential relevance that deserves the title 'Bayesian.' To deny it is to misunderstand the central message of Bayes' Theorem for questions of evidence: namely, that hypotheses are confirmed by data they predict" (see the appendix). This second requirement holds that when the evidence $E$ is more probable in the light of hypothesis $H_1$ than in the light of $H_2$ and if the evidence $E$ is also less probable under the assumption that $H_1$ is false than it is under the assumption that $H_2$ is false with respect to probability function Pr, then $E$ confirms $H_1$ more than it confirms $H_2$.

What follows are among the most prominent Bayesian confirmation measures that satisfy both requirements.

(Carnap 1962):

$$d(H, E) = \Pr(H|E) - \Pr(H),$$

if $\Pr(E) > 0$.

(Carnap 1962):

$$\mathfrak{r}(H, E) = \Pr(H \cap E) - \Pr(H) \times \Pr(E).$$

1. "Demarcation" has been adopted following Crupi et al. (2007).

https://doi.org/10.1086/671172 Published online by Cambridge University Press

(Christensen 1999; Joyce 1999):

$$S(H, E) = \Pr(H|E) - \Pr(H|\neg E),$$

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

(Crupi et al. 2007):

$$Z(H, E) = \begin{cases} \dfrac{d(H, E)}{1 - \Pr(H)} & \text{if } \Pr(H|E) \geq \Pr(H) > 0 \\ \dfrac{d(H, E)}{1 - \Pr(\neg H)} & \text{if } \Pr(H|E) < \Pr(H) \\ 1 & \text{if } \Pr(H) = 0 \end{cases}.$$

(Good 1960; Fitelson 2001):

$$l(H, E) = \begin{cases} \log\left[\dfrac{\Pr(E|H)}{\Pr(E|\neg H)}\right] & \text{if } \Pr(E|H) > 0 \text{ and } \Pr(E|\neg H) > 0 \\ \infty & \text{if } \Pr(E) > 0 \text{ and } \Pr(E|\neg H) = 0 \\ -\infty & \text{if } \Pr(E) = 0 \text{ or } \Pr(E|H) = 0 \end{cases}.$$

(Kemeny and Oppenheim 1952):

$$F(H, E) = \begin{cases} \dfrac{\Pr(E|H) - \Pr(E|\neg H)}{\Pr(E|H) + \Pr(E|\neg H)} & \text{if } \Pr(E|H) > 0 \text{ and } \Pr(E|\neg H) > 0 \\ 1 & \text{if } \Pr(E) > 0 \text{ and } \Pr(\neg H) = 0 \\ -1 & \text{if } \Pr(E) = 0 \text{ or } \Pr(H) = 0 \end{cases}.$$

(Milne 1996):

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

(Mortimer 1988):

$$M(H, E) = \Pr(E|H) - \Pr(E),$$

if $\Pr(H) > 0$.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

(Nozick 1981):

$$N(H, E) = \Pr(E|H) - \Pr(E|\neg H),$$

if $1 > \Pr(H) > 0$.

As already noted, this plurality of Bayesian measures of (incremental) confirmation is generally taken to bring about one of the biggest challenges for Bayesian confirmation theory: the problem of measure sensitivity. This problem is the subject of the next section.

**3. The Problem of Measure Sensitivity.** In Fitelson's own words, the problem of measure sensitivity is as follows: "I will show that seven well-known arguments surrounding contemporary Bayesian confirmation theory are sensitive to choice of measure. I will argue that this exposes a weakness in the theoretical foundation of Bayesian confirmation theory which must be shored-up. I call this problem *the problem of measure sensitivity*" (1999, S364). This section of the article demonstrates how serious the problem is, mainly by following Fitelson's lead. The aim is to show that the validity of different but equally important arguments within confirmation theory depends on which measure of confirmation is adopted. However, the results displayed here are in two respects more powerful than those achieved in Fitelson (1999). First, this article considers more confirmation measures. In Fitelson's considerations, only a proper subclass of the confirmation measures introduced in section 2 is taken into account. In particular, Fitelson considers only the confirmation measures $d$, $\mathfrak{r}$, $r$, and $l$. Second, this article shows that the problem of measure sensitivity actually runs deeper and is far more severe than is noted in Fitelson (1999). Fitelson draws only the following, relatively tentative, conclusion: "Many well-known arguments in quantitative Bayesian confirmation theory are valid only if the difference measure $d$ is to be preferred over other relevance measures (at least, in the confirmational contexts in question). I have also shown that there are compelling reasons to prefer $d$ over the log-ratio measure $r$. Unfortunately, like Rosenkrantz (1981), I have found *no* compelling reasons offered in the literature to prefer $d$ over the log-*likelihood* ratio measure $l$" (S371).

According to this, it is possible to adopt the difference measures of confirmation, and no argument of confirmation theory considered in Fitelson (1999) would be affected. The only problem Fitelson has with adopting $d$ is that there is no compelling argument against the $l$ measure of confirmation. However, the arguments Fitelson reviews are central to Bayesian confirmation theory insofar that almost all Bayesians are at least tempted to argue that validating these arguments is a desideratum that any adequate measure of confirmation should satisfy. If only one confirmation measure

https://doi.org/10.1086/671172 Published online by Cambridge University Press

validates all these arguments, this is a reason to believe that it is the correct measure of confirmation.

However, the following subsections consider four prominent arguments in Bayesian confirmation theory and show that there is no confirmation measure that validates all of them. This strengthens the results achieved in Fitelson (1999) since it shows that there cannot be a single confirmation measure that validates all of those prominent arguments. Hence, it is impossible to argue that validating these four arguments is a desideratum that any adequate measure of confirmation should satisfy. Subsections 3.1 and 3.2 introduce arguments already considered in Fitelson (1999), while subsections 3.3 and 3.4 consider two new arguments from Fitelson (2006) and Huber (2005), respectively.

*3.1. Earmann and the Problem of Irrelevant Conjunction.* First, let us look at what is called the problem of irrelevant conjunction.[2] There are two (closely related) formulations of this problem. The first is the traditional formulation, and the second is Fitelson's (2002) version. The traditional formulation of the problem is as follows:

> **The Problem of Irrelevant Conjunction 1.** If $H$ implies $E$, then $E$ confirms $H \cap I$ for arbitrary $I$, provided $\Pr(H \cap I) > 0$.

Fitelson (2002) introduces a slightly different formulation of the problem. According to him, the problem of irrelevant conjunction is as follows:

> **The Problem of Irrelevant Conjunction 2.** If $E$ confirms $H$, and $I$ is confirmationally irrelevant to $H$, $E$, and $H \cap E$, then $E$ also confirms $H \cap I$.

Both formulations are statements of the problem of irrelevant conjunction because one can tack a confirmationally irrelevant conjunct $I$ (the irrelevant conjunct) to the hypothesis $H$ (the relevant conjunct) and still get the result that $H \cap I$ (the irrelevant conjunction) is confirmed.

This first formulation of the problem was first discussed in Hempel (1945) in connection with the hypothetico-deductive method of confirmation. If one follows Hempel in adopting the special consequence condition, which requires that "if an observation report confirms a hypothesis $H$, then it also confirms every consequence of $H$" (31), one must reject the first formulation of the problem. The reason is that taken together they imply that the evidence $E$ confirms even the arbitrarily chosen conjunct $I$.

2. I am grateful to Vincenzo Crupi for a helpful discussion about the problem of irrelevant conjunction in general and requirements 3.1–3.3 in particular.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

This would trivialize the notion of confirmation: every piece of evidence $E$ would confirm every hypothesis $I$.

The standard Bayesian conception of qualitative confirmation validates the two formulations of the problem of irrelevant conjunction, but it does not validate Hempel's special consequence condition. Accordingly, the standard Bayesian conception of qualitative confirmation does not threaten to trivialize the notion of confirmation. Nevertheless, Bayesian confirmation theorists typically consider the result that their qualitative notion of confirmation validates both formulations of the problem to be counterintuitive. In particular, in the literature one can find examples that seem to strongly support the claim that this result is counterintuitive. For example, Glymour (1980) writes, "Nor are we greatly impressed by a theory of evidence which claims, for instance, that the Stark effect confirms the conjunction of quantum mechanics and the principle that the president of the Church of Jesus Christ of Latter Day Saints is infallible, all the while reassuring us that the Stark effect does not confirm the latter principle alone" (31).

However, Earman (1992) argues that examples like this are not too bad, as long as according to any reasonable confirmation measure the irrelevant conjunction $H \cap I$ is less confirmed than the relevant conjunct $H$. This is supposed to "soften the impact" (as Fitelson [2002] puts it) of this counterintuitive result regarding the qualitative conception of incremental confirmation. Thus, aiming at cushioning the impact of the problem in its traditional formulation, Earman (1992) requires that any reasonable confirmation measure satisfy the following condition:

> **Requirement 3.1** (Irrelevant Conjunction I). If $\mathfrak{c}$ is a reasonable confirmation measure, $H \vDash E$, and $I$ is logically independent of $H$, $E$, and $H \cap E$, then:
>
> $$\mathfrak{c}(H, E) > \mathfrak{c}(H \cap I, E).$$

Hawthorne and Fitelson (2004) suggest an alternative and, in their view, better-suited requirement to "cushion" the impact of the problem of irrelevant conjunction in its formulation by Fitelson (2002).

> **Requirement 3.2** (Irrelevant Conjunction II). If $\mathfrak{c}$ is a reasonable confirmation measure, $\Pr(H|E) > \Pr(H)$, $\Pr(E|H \cap I) = \Pr(E|H)$, and $\Pr(I|H) \neq 1$, then:
>
> $$\mathfrak{c}(H, E) > \mathfrak{c}(H \cap I, E).$$

It is an interesting question which confirmation measures of section 2 satisfy these requirements. The following observation provides a list of which confirmation measures satisfy requirements 3.1 and 3.2 and which do not.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

**Observation 1.** Confirmation measures $d$, $\mathfrak{r}$, $S$, $Z$, $l$, $F$, and $N$ satisfy requirements 3.1 and 3.2; $r$ and $M$ do not satisfy requirements 3.1 and 3.2.[3]

One might be tempted to argue that observation 1 shows that $(a)$ the ratio measure $r$ and the Mortimer measure $M$ are not adequate measures of confirmation, and $(b)$ all other Bayesian measures of confirmation are candidates for being the one true measure of confirmation. This is because they "soften the impact" of the counterintuitive result insofar as they can be used to explain why philosophers have the preexplicatory intuition that irrelevant conjunctions are or should not be confirmed. The explanation is that they are confirmed to such a minimal degree that, preexplicatory, many philosophers and scientists do not have the intuition that they are confirmed. Thus, Bayesian confirmation theory can explain (away) these intuitions, given that they reject the ratio measure of confirmation.

However, inferring from observation 1 that all other Bayesian measures of confirmation except $r$ and $M$ are candidates for being the one true measure of confirmation would be hasty. The above requirements 3.1 and 3.2 apply only in the case of confirmation. Accordingly, it is an interesting question whether a similar requirement holds for the case of disconfirmation. If one accepts Hawthorne's and Fitelson's requirement 3.2 for the case of confirmation, one might be tempted to accept an analogous requirement for the case of disconfirmation. More specifically, one might be tempted to argue that it should hold in general that if $E$ confirms or disconfirms hypothesis $H$, then $E$ confirms or disconfirms $H$ to a higher degree than it confirms the conjunction $H \cap I$ with the confirmationally irrelevant conjunction $I$. The following requirement is a slightly weaker version of a requirement discussed (but ultimately not endorsed) in Crupi and Tentori (2010).

**Requirement 3.3** (Irrelevant Conjunction III).[4] If $\mathfrak{c}$ is a reasonable confirmation measure with marker or neutral value $r$, $\Pr(H|E) < \Pr(H)$, and $\Pr(E|H \cap I) = \Pr(E|H) > 0$ and $\Pr(I|H) \neq 1$, then:

$$|\mathfrak{c}(H, E) - r| > |\mathfrak{c}(H \cap I, E) - r|.$$

According to this requirement, "if $E$ has a positive [or negative] confirmatory impact on $H$ and $I$ is a confirmationally irrelevant conjunct to

3. Proofs for all observations can be found on the author's website http://rub.academia.edu/PeterBrössel.

4. In particular, this requirement is logically weaker than the one considered in Crupi and Tentori (2010) since it is restricted to cases of (dis)confirmation where the evidence does not falsify $H$ and $H \cap I$, i.e., if $\Pr(E|H \cap I) = \Pr(E|H) > 0$ (without this caveat the requirement would be satisfied only by the measures $d$, $\mathfrak{r}$, $S$, and $N$). It is also interesting to note that this requirement is implied by the strengthened weak law of likelihood (requirement 2.3) discussed in the appendix.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

$H$ with regard to $E$, then $E$ will have the same kind of impact (namely, positive [or negative]) on $H \cap I$ but to a minor extent" (Crupi and Tentori 2010, 4).

> **Observation 2.** Confirmation measures $d$, $\mathfrak{r}$, $S$, $l$, $F$, and $N$ satisfy requirement 3.3; $Z$, $r$, and $M$ do not satisfy requirement 3.3.

It is interesting to note that the $Z$ confirmation measure is the only confirmation measure that satisfies requirement 3.2 but that does not satisfy the analogous requirement for the case of disconfirmation. The important question is whether observation 2 can be taken to show that (*a*) the $Z$ measure, the ratio measure $r$, and the Mortimer measure $M$ are not adequate measures of confirmation, and (*b*) all other Bayesian measures of confirmation are candidates for being the one true measure of confirmation. I do not think so. The problem with this interpretation of the observations is that equally reasonable considerations exclude other confirmation measures.

*3.2. Horwich and the Paradox of the Ravens.* Another requirement on confirmation measures arises from arguments surrounding the well-known paradox of the ravens (Hempel 1945). This paradox is fueled by the following inconsistent intuitions:

> **Ravens Paradox 1.** (i) A nonblack nonraven (e.g., a white shoe) does not confirm the hypothesis "All ravens are black." (ii) A nonblack nonraven confirms the hypothesis "All nonblack things are nonravens." (iii) Logically equivalent hypotheses are equally confirmed by a given piece of evidence (Equivalence Condition).

Now let $H_1$ be the hypothesis "All ravens are black" and let $H_2$ be the hypothesis "All nonblack things are nonravens." Obviously, $H_1$ and $H_2$ are logically equivalent. According to i, nonblack nonravens do not confirm $H_1$. Hence, by the equivalence condition iii, nonblack nonravens do not confirm $H_2$. This contradicts intuition ii, according to which nonblack nonravens should confirm $H_2$ and black ravens should not confirm it.

This shows that these intuitions are logically inconsistent. The standard Bayesian solution of the ravens paradox is similar to the solutions suggested for the problem of irrelevant conjunction. In particular, the idea is that the observation of a nonblack nonraven can confirm the hypothesis "All ravens are black" but does so less strongly than the observation of black ravens. This "softens the impact" of the counterintuitive result that the observation of white shoes confirms "All ravens are black" at all. Horwich (1982) relies on the following requirement to show that Bayesian confirmation theory can soften the impact of the ravens paradox.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

**Requirement 3.4** (The Ravens Paradox). If $\mathfrak{c}$ is a reasonable confirmation measure and $\Pr(H|E_1) > \Pr(H|E_2)$ then:

$$\mathfrak{c}(H|E_1) > \mathfrak{c}(H|E_2).$$

**Observation 3.** Confirmation measures $d$, $Z$, $l$, $F$, and $r$ satisfy requirement 3.4; $\mathfrak{r}$, $S$, $M$, and $N$ do not satisfy requirement 3.4.

All confirmation measures that are in agreement with requirement 3.4 can explain why philosophers have the preexplicatory intuition that a nonblack nonraven does not confirm "All ravens are black." The explanation is that a nonblack nonraven confirms this hypothesis to such a minimal degree that philosophers and scientists do not have the preexplicatory intuition that it confirms it at all. However, the problem of irrelevant conjunction and the paradox of the ravens together do not show that Bayesians can exclude $\mathfrak{r}$, $S$, $r$, $M$, and $N$ and that all other measures are reasonable confirmation measures. Again, other equally reasonable considerations exclude other measures.

*3.3. Fitelson, Confirmation, and Inductive Logic.* Other arguments in Bayesian confirmation theory deal with the question of what is the purpose of Bayesian confirmation theory. For example, Fitelson (2006) stresses that, since Carnap (1962), the concept of confirmation has been closely linked to that of inductive logic. Carnap's project was to provide a theory of confirmation or inductive logic that can be considered as an extension of deductive logic (Fitelson 2006; Crupi et al. 2007). Fitelson (2006, 501) argues that this requires that confirmation (i.e., partial entailment, in Carnap's view) must be analogous to deductive entailment regarding certain aspects. In particular, the evaluation of inductive arguments should retain the idea that there is no better inductive argument for some proposition or hypothesis than a deductive argument. Similarly, there is no worse argument than an argument with premises that imply the falsity of the conclusion. In accordance with this consideration, Fitelson suggests that any theory of confirmation that is supposed to be an account of inductive logic has to satisfy the requirement of logicality.[5]

5. Crupi et al. (2007) discuss another interesting requirement whose satisfaction they take to be necessary for considering confirmation and inductive logic as an extension of deductive logic. However, for present purposes it is sufficient to consider the relatively weak logicality requirement. In addition, Crupi et al. (2007) agree in principle with the latter requirement. Accordingly, it is safe to assume that the logicality requirement is a minimal requirement that any confirmation measure must satisfy for considering confirmation theory an extension of deductive logic.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

**Requirement 3.5** (Logicality). If $\mathfrak{c}$ is an adequate confirmation measure, then (i) if $E$ implies $H$, then $\mathfrak{c}(H, E) = \max$, and (ii) if $E$ implies $\neg H$, then $\mathfrak{c}(H, E) = \min$.

**Observation 4.** Confirmation measures $Z$, $l$, and $F$ satisfy requirement 3.5; $d$, $\mathfrak{r}$, $S$, $r$, $M$, and $N$ do not satisfy requirement 3.5.

The confirmation measures $d$, $\mathfrak{r}$, $S$, $r$, $M$, and $N$ fail to assign the maximum value if $H$ is implied by $E$. Measures $\mathfrak{r}$, $d$, $S$, $M$, and $N$ also fail to assign the minimum value if hypothesis $H$ is refuted by the evidence $E$.

Again, this observation together with the preceding one can be taken to show that $Z$, $l$, and $F$ are the only candidates for being an adequate confirmation measure. However, one can find another important argument within Bayesian confirmation theory that is not valid according to these confirmation measures.

*3.4. Huber and the Assessment of Theories.* Huber (2007) focuses on another, more venerable question that arises in the context of Bayesian confirmation theory and inductive logic: Hume's problem of induction. Huber writes: "Let us grant that Bayesian confirmation theory adequately explicates the concept of confirmation. . . . The question remains what happened to Hume's problem of the justification of induction. We know—by definition—that the conclusion of an inductively strong argument is well-confirmed by its premises. But does that also justify our acceptance of that conclusion? Don't we first have to justify our definition of confirmation before we can use it to justify our inductive inferences?" (sec. 7). In an earlier article, Huber suggests an answer to this question, and the answer Huber gives should be good news for Bayesian confirmation theory: "An examination of two popular measures of incremental confirmation suggests the view that one should stick to incrementally well confirmed theories, because incremental confirmation takes one to (the most) informative (among all) true theories. However, incremental confirmation does not further this goal in general" (2005, 1146).

That some confirmation measures lead one to accept the most informative among all true theories is truly good news. However, the last sentence adds to the problem at hand: the plurality of confirmation measures and the problem of measure sensitivity. Again, one interesting and valuable argument in favor of Bayesian confirmation theory depends on the choice of the confirmation measure. Skipping the details, Huber (2005) proposes a requirement dubbed "continuity" on confirmation measures so that they take one to "(the most) informative (among all) true theories." Continuity requires that "any surplus in informativeness succeeds, if the difference in plausibility is small enough" (1156). Following philosophers such as Carnap and Bar-

https://doi.org/10.1086/671172 Published online by Cambridge University Press

Hillel (1952), Popper (1954), and Hempel (1960), Huber (2005, 2008) suggests measuring the amount of information provided by a hypothesis with the help of the measure $i(H) = \Pr(\neg H)$. The idea of this measure of information is that a hypothesis is the more informative, the more logical possibilities it excludes, and it is the more probable, the fewer possibilities it excludes. The informativity and the probabilities of some hypotheses are therefore "conflicting in the sense that the former is an increasing and the latter a decreasing function of the logical strength of the theory to be assessed" (Huber 2008, 91). With the help of this measure of the informativity of confirmation measures at hand, Huber (2005, 2008) suggests the following continuity requirement (here a minimally weaker version of Huber's requirement is considered).

> **Requirement 3.6** (Continuity). For all probability functions Pr, if $\mathfrak{c}$ is an adequate confirmation measure defined in terms of Pr, then $\forall \varepsilon > 0 \exists \delta_\varepsilon > 0$, so that if $i(H_1) \geq i(H_2) + \varepsilon$, $\Pr(H_1|E) \geq \Pr(H_2|E) - \delta_\varepsilon$, and $\Pr(H_1|E) > 0$, then $\mathfrak{c}(H_1, E) > \mathfrak{c}(H_2, E)$.

With the help of the continuity requirement, Huber (2005) is able to provide a strong argument for Bayesian confirmation theory. Again skipping the details, if measures of confirmation satisfy, among other requirements, the continuity requirement, they lead a Bayesian agent to "(the most) informative (among all) true theories." However, as already noted in the quote from Huber, not all incremental confirmation measures satisfy the continuity requirement.

> **Observation 5.** Confirmation measures $d$, $\mathfrak{r}$, $S$, $r$, $M$, and $N$ satisfy requirement 3.6; $Z$, $l$, and $F$ do not satisfy requirement 3.6.

That the confirmation measures $Z$, $l$, and $F$ do not satisfy requirement 3.5 and that these confirmation measures do not lead Bayesian agents to "(the most) informative (among all) true theories," as Huber puts it, can be considered to be a powerful argument against these measures of confirmation. If Bayesians subscribe to the point of view that "inductive procedures . . . lead to the *acceptance* of certain empirical hypotheses on the basis of evidence that gives them more or less strong . . . support" (Hempel 1960, 462), it seems Bayesians should reject the confirmation measures $Z$, $l$, and $F$. Instead, Bayesians should choose between those measures that satisfy requirement 3.6 and that can lead one to "(the most) informative (among all) true theories."

*3.5. The Strengthened Problem of Measure Sensitivity.* The previous subsections successfully demonstrate that the plurality of confirmation mea-

https://doi.org/10.1086/671172 Published online by Cambridge University Press

sures and the problem of the measure sensitivity of various arguments surrounding confirmation theory present an enormous challenge for the standard Bayesian conception of confirmation. No single confirmation measure satisfies all the requirements; hence, no confirmation measure renders all the above-mentioned arguments in Bayesian confirmation theory valid. Given the usual assumption that monism of confirmation measures holds true, this strengthened problem of measure sensitivity is especially problematic. As already noted, monism of confirmation measures says that there is only one adequate confirmation measure (up to ordinal equivalence). Hence, monism requires providing plausible desiderata that single out the "one true measure of confirmation," as Milne (1996) puts it.

Table 1 displays whether the different confirmation measures introduced in section 2 satisfy the requirements presupposed in the context of various arguments of Bayesian confirmation theory. However, one might object that these considerations do not show that there is no confirmation measure that validates all of them. The considerations could just demonstrate that we have not found the relevant confirmation measure yet. More specifically, there might be different confirmation measures as yet unproposed (and therefore not mentioned in sec. 2) that validate all four arguments by satisfying the six requirements 3.1–3.6. However, one can show that this is not possible. Closer inspection of table 1 reveals that there are confirmation measures that satisfy the requirements 3.1–3.5, that is, the confirmation measures $l$ and $F$, and there is one confirmation measure that satisfies requirements 3.1–3.4 and requirement 3.6, that is, the confirmation measure $d$. However, there is no confirmation measure that satisfies requirements 3.5 and 3.6. The reason for this is that both requirements are inconsistent.

> **Observation 6.** It is logically impossible to construct a confirmation measure $\mathfrak{c}$ that satisfies requirement 3.5 and requirement 3.6.

According to requirement 3.5 (Fitelson's logicality requirement), two hypotheses $H_1$ and $H_2$ should be assigned the maximum degree of confir-

TABLE 1. Confirmation Measures' Requirement Satisfaction

| Requirement | $d$ | $\mathfrak{r}$ | $S$ | $Z$ | $l$ | $F$ | $r$ | $M$ | $N$ |
|---|---|---|---|---|---|---|---|---|---|
| 3.1 | Yes | Yes | Yes | Yes | Yes | Yes | No | No | Yes |
| 3.2 | Yes | Yes | Yes | Yes | Yes | Yes | No | No | Yes |
| 3.3 | Yes | Yes | Yes | No | Yes | Yes | No | No | Yes |
| 3.4 | Yes | No | No | Yes | Yes | Yes | Yes | No | No |
| 3.5 | No | No | No | Yes | Yes | Yes | No | No | No |
| 3.6 | Yes | Yes | Yes | No | No | No | Yes | Yes | Yes |
| 3.1–3.6 | No | No | No | No | No | No | No | No | No |

Note.—"Yes" or "no" denotes whether a confirmation measure satisfies a requirement.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

mation if they are implied by the evidence, independently of any surplus in informativeness between $i(H_1)$ and $i(H_2)$ (for the sake of the argument, let us assume there is such a difference in informativeness). However, according to requirement 3.6 (Huber's continuity requirement), any surplus in informativeness between $i(H_1)$ and $i(H_2)$ (i.e., $\Pr(\neg H_1)$ and $\Pr(\neg H_2)$) succeeds if the difference in plausibility is small enough. Since by assumption both hypotheses are implied by the evidence, there is no difference in plausibility since $\Pr(H_1|E) = \Pr(H_1|E) = 1$. Hence, the surplus in informativeness should succeed, but this contradicts what is required by logicality. Accordingly, one can give up searching for confirmation measures that satisfy all six requirements.

In conclusion, the strengthened problem of measure sensitivity teaches two important lessons. First, one cannot use the above requirements as desiderata that an adequate measure of confirmation must satisfy to single out the one true confirmation measure. According to the arguments considered in Fitelson (1999), it would be possible to adopt the confirmation measure $d$ and have all arguments be valid. Indeed, many Bayesian confirmation theorists would at least be tempted to argue that validating these arguments is a desideratum that any adequate measure of confirmation should satisfy. This is not possible with the arguments and requirements considered here. None of the confirmation measures introduced in section 2 satisfies them all, and it is impossible to construct such a confirmation measure. This makes the plurality of confirmation measures in conjunction with the strengthened problem of measure sensitivity a real chestnut of confirmation theory.

Second, given monism of confirmation measures, one of the four arguments considered above is not valid. Either Bayesian confirmation theory cannot be used to soften the impact of the problem of irrelevant conjunction and the paradox of the ravens, or one has to give up considering confirmation theory as an inductive logic (i.e., as an extension of deductive logic in the sense of Fitelson [2006]), or one has no argument that it leads one to accept the most informative of all true theories. This is true because none of the confirmation measures introduced in section 2 satisfy all six requirements. Of course, there might be different arguments for similar conclusions with weaker or different requirements. However, according to the state of the art in Bayesian confirmation theory, monism requires that at least one of the above requirements is given up and that the arguments that presuppose these requirements are therefore not sound. Assuming that the Bayesian monist aims at maximizing the number of arguments validated, two options present themselves: option 1 is to choose measure $d$, which softens the impact of the problem of irrelevant conjunction and the paradox of the ravens and which leads us to accept the most informative of all true theories. Option 2 is to choose measure $l$ (or $F$), which softens the impact of the problem of irrelevant conjunction and the paradox of the ravens and which

https://doi.org/10.1086/671172 Published online by Cambridge University Press

we can consider an extension of deductive logic in the sense of Fitelson (2006).

**4. The Problem of Measure Sensitivity as an Argument for Pluralism?** However, there is yet another plausible alternative. This alternative is to give up monism and to embrace the plurality of Bayesian confirmation measures (Joyce 1999, 2003a; Steel 2007; Huber 2008). Then one could argue that indeed all four arguments considered in section 3 are valid, even though they are not valid with respect to one notion of confirmation. Instead, one could argue that there are at least two (and possibly more) diverging notions of confirmation, and for each of the arguments considered in section 3 there is at least one notion of confirmation that validates it. For example, one could argue that there is one notion of confirmation that satisfies requirements 3.1–3.5 and one notion of confirmation that satisfies requirements 3.1–3.4 and requirement 3.6. Joyce (2003b) labels this position pluralism and describes it as follows: "The various measures of confirmation capture different, but entirely legitimate, concepts of evidence, concepts that must play a central role in any adequate epistemology. . . . We need a theory of confirmation that is able to characterize a variety of evidential relationships that hold across a wide range of states of background knowledge and opinion. Only a probabilistic theory that provides many measures of incremental confirmation is up to the task" (1).

One might be tempted to argue that pluralism must be rejected since it is not possible to adopt more than one explication for the one term "evidence $E$ incrementally confirms hypothesis $H$ (with respect to background knowledge $B$) to degree $r$." However, on this question Joyce has Carnap on his side. Carnap writes: "It has repeatedly occurred in the history of science that a vehement but futile controversy arose between the proponents of two or more *explicata* who shared the erroneous belief that they had the same explicandum; when finally it became clear that they meant different explicanda, unfortunately designated by the same term, and that the different explicate were hence compatible and moreover were found to be equally fruitful scientific concepts, the controversy evaporated into nothing" (1962, 26).[6] Accordingly, it might be the case that the various confirmation theorists have different explicanda in mind but use the same term "incremental confirmation" to refer to them. This assumption is the most plausible explanation for

6. I am grateful to an anonymous referee for making me aware of this passage. In addition, the referee correctly pointed out that Carnap's distinction between absolute confirmation and incremental confirmation is a historical instance of such a controversy. In particular, Carnap was able to show that the triviality result achieved in Hempel (1945) is based on the erroneous belief that there is just one explicandum designated by the term "confirmation," instead of the two explicanda nowadays designated by the terms absolute confirmation and incremental confirmation.

https://doi.org/10.1086/671172 Published online by Cambridge University Press

the findings of section 3. However, proponents of pluralism cannot simply assume that all the different measures introduced in section 2 "play a central [and irreplaceable] role in any adequate epistemology," as Joyce puts it, or, in Carnap's words, that they are "equally fruitful scientific concepts." In particular, they have to show that these (ordinally) different confirmation measures reflect equally important but nevertheless different aspects of scientific reasoning. Hence, proponents of pluralism have to show that an adequate confirmation theory requires at least two diverging notions of confirmation and therefore also two confirmation measures.[7] In addition, they have to provide us with a set of criteria that allow us to distinguish confirmation measures that capture important evidential relationships from probabilistic measures that do not capture important evidential relationships. Hence, proponents of pluralism have to provide criteria that exclude measures that are not required by an adequate confirmation theory.

**5. Summary.** The current article strengthens the problem of measure sensitivity by showing that there are four valuable and important arguments within Bayesian confirmation theory that are not validated by any single measure of confirmation. These considerations show that providing some form of solution to Fitelson's (1999) problem of measure sensitivity is central to Bayesian confirmation theory.

It is important to note that I do not take the strengthened version of the problem of measure sensitivity to provide any reason to abandon Bayesian confirmation theory. What these considerations show, however, is that any solution to the problem of measure sensitivity will require enormous modifications of the standard Bayesian conception of confirmation, beyond anything anticipated in Fitelson (1999): either one has to give up monism of confirmation measures or one has to give up the validity of one of the four important arguments considered in section 3 of the article.

More specifically, the first option is giving up monism of confirmation measures—which I take to be part of the standard Bayesian conception of confirmation—and embracing pluralism as suggested by Joyce and others. This I consider to be a major (but nevertheless very attractive) deviation from the standard Bayesian conception of confirmation. The challenge connected with this option is to show that a complete Bayesian confirmation theory requires at least two notions of confirmation to serve all the purposes of confirmation theory and to present criteria that distinguish between confirmation measures that capture important evidential relationships and probabilistic measures that do not capture important evidential relationships.

7. An argument to the effect that an adequate theory of confirmation needs at least two measures of confirmation is presented in Brössel (2012).

https://doi.org/10.1086/671172 Published online by Cambridge University Press

The second option is to uphold monism of confirmation measures and to give up the validity of one of the four arguments considered in section 3. The incompatibility of requirement 3.5 and requirement 3.6 suggests that one has to give up considering confirmation theory as an inductive logic, or one is left with no argument that it leads one to accept the most informative of all true theories. Although there is the possibility that there are different arguments for similar conclusions with weaker or different requirements, according to the state of the art in Bayesian confirmation theory, monism requires that at least one of the two requirements is given up and that the arguments that presuppose these requirements are therefore not sound. The challenge connected with this option is to say which argument should be given up and to provide a justification for why a confirmation measure that does not validate this argument is nevertheless adequate. This includes, in particular, showing that the "one true" confirmation measure can serve all the purposes associated with Bayesian confirmation theory, even if it does not validate all the above arguments. Hence, proponents of monism have to show that an adequate Bayesian theory of confirmation does not require two measures of confirmation. This presents an enormous challenge to those Bayesian confirmation theorists who want to uphold monism.

## Appendix

In the light of Joyce's strong claims about the weak law of likelihood, it is interesting to note that Crupi, Festa, and Buttasi (2010) are able to show that the weak law of likelihood is logically independent of the demarcation requirement (requirement 2.1). It is also interesting to note that minimally strengthening this "integral part of any account of evidential relevance" leads to a requirement that not all of the measures of confirmation discussed in the literature are able to satisfy. The strengthened weak law of likelihood I have in mind is this:

> **Requirement 2.3** (Strengthened Weak Law of Likelihood). If $0 < \Pr(E|H_1) \geq \Pr(E|H_2)$ and $\Pr(E|\neg H_1) < \Pr(E|\neg H_2)$, or if $\Pr(E|H_1) > \Pr(E|H_2)$ and $0 < \Pr(E|\neg H_1) \leq \Pr(E|\neg H_2)$, then $\mathfrak{c}(H_1, E) > \mathfrak{c}(H_2, E)$.

Joyce's Weak Law of Likelihood requires that if the evidence $E$ is strictly more probable in the light of hypothesis $H_1$ than in the light of $H_2$ and if it is strictly less probable in the light of $\neg H_1$ than in the light of $\neg H_2$, then $E$ confirms $H_1$ more than $H_2$. The strengthened weak law of likelihood slightly weakens the antecedent of Joyce's requirement and requires that $E$ confirms $H_1$ more than $H_2$, if the evidence $E$ is strictly more probable in the light of

https://doi.org/10.1086/671172 Published online by Cambridge University Press

hypothesis $H_1$ than in the light of $H_2$ and if $E$ is strictly less or equally probable in the light of $\neg H_1$ than $\neg H_2$ or if the evidence $E$ is strictly more or equally probable in the light of hypothesis $H_1$ than in the light of $H_2$ and if $E$ is strictly less probable in the light of $\neg H_1$ than $\neg H_2$. This slightly weakened antecedent seems to suffice to underscore the claim that $H_1$ predicts $E$ better than $H_2$ and thus should be confirmed to a higher degree. Confirmation measures discussed in the literature that do not satisfy the strengthened weak law of likelihood are the $Z$ measure by Crupi et al. (2007), the $r$ measure of confirmation defended in Milne (1996), and the $M$ measure of confirmation by Mortimer (1988). These confirmation measures are introduced in detail section 2.

## REFERENCES


Brössel, Peter. 2012. "Rethinking Bayesian Confirmation Theory." PhD diss., University of Konstanz.

Carnap, Rudolf. 1962. *The Logical Foundations of Probability*. 2nd ed. Chicago: University of Chicago Press.

Carnap, Rudolf, and Yoshua Bar-Hillel. 1952. "An Outline of a Theory of Semantic Information." Technical Report no. 247, Research Laboratory of Electronics, MIT.

Christensen, David. 1999. "Measuring Confirmation." *Journal of Philosophy* 96:437–61.

Crupi, Vincenzo, Roberto Festa, and Carlo Buttasi. 2010. "Towards a Grammar of Bayesian Confirmation." In *Epistemology and Methodology of Science*, ed. Mauricio Suárez, Mauro Dorato, and Miklós Rédei, 73–93. Dordrecht: Springer.

Crupi, Vincenzo, and Katya Tentori. 2010. "Irrelevant Conjunction: Statement and Solution of a New Paradox." *Philosophy of Science* 77:1–13.

Crupi, Vincenzo, Katya Tentori, and Michel Gonzalez. 2007. "On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues." *Philosophy of Science* 74:229–52.

Earman, John. 1992. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.

Eells, Ellery, and Branden Fitelson. 2002. "Symmetries and Asymmetries in Evidential Support." *Philosophical Studies* 107:129–42.

Fitelson, Branden. 1999. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (Proceedings): S362–S378.

———. 2001. "Studies in Bayesian Confirmation Theory." PhD diss., University of Wisconsin–Madison.

———. 2002. "Putting the Irrelevance Back into the Problem of Irrelevant Conjunction." *Philosophy of Science* 69:611–22.

———. 2006. "Logical Foundations of Evidential Support." *Philosophy of Science* 73:500–512.

Glymour, Clark. 1980. *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Good, Irving John. 1960. "Weight of Evidence, Corroboration, Explanatory Power, Information and the Utility of Experiments." *Journal of the Royal Statistical Society* B 22:319–31.

Hawthorne, James, and Branden Fitelson. 2004. "Re-solving Irrelevant Conjunction with Probabilistic Independence." *Philosophy of Science* 71:505–14.

Hempel, Carl Gustav. 1945. "Studies in the Logic of Confirmation." Pt. 1. *Mind* 54:1–26.

———. 1960. "Inductive Inconsistencies." *Synthese* 12:439–69.

Horwich, Paul. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.

Huber, Franz. 2005. "What Is the Point of Confirmation?" *Philosophy of Science* 72:1146–59.

———. 2007. "Confirmation and Induction." In *Internet Encyclopedia of Philosophy*, ed. James Fieser and Bradley Dowden. http://www.iep.utm.edu/conf-ind/.

———. 2008. "Milne's Argument for the Log-Ratio Measure." *Philosophy of Science* 75:413–20.


https://doi.org/10.1086/671172 Published online by Cambridge University Press

Joyce, James. 1999. *Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
———. 2003a. "Bayes' Theorem." In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. http://plato.stanford.edu/entries/bayes-theorem/.
———. 2003b. "On the Plurality of Probabilist Measures of Evidential Relevance." Paper presented at the 26th Annual International Wittgenstein Symposium of the Austrian Ludwig Wittgenstein Society, Kirchberg.
Kemeny, John, and Paul Oppenheim. 1952. "Degree of Factual Support." *Philosophy of Science* 19:307–24.
Lange, Mark. 1999. "Calibration and the Epistemological Role of Bayesian Conditionalization." *Journal of Philosophy* 96:294–324.
Milne, Peter. 1996. "log$[p(h/eb)/p(h/b)]$ Is the One True Measure of Confirmation." *Philosophy of Science* 63:21–26.
Mortimer, Halina. 1988. *The Logic of Induction*. Paramus, NJ: Prentice Hall.
Nozick, Robert. 1981. *Philosophical Explanations*. Oxford: Clarendon.
Popper, Karl. 1954. "Degree of Confirmation." *British Journal of the Philosophy of Science* 5: 143–49.
Rosenkrantz, Roger. 1981. *Foundations and Applications of Inductive Probability*. Atascadero, CA: Ridgeview.
Steel, Daniel. 2007. "Bayesian Confirmation Theory and the Likelihood Principle." *Synthese* 156:55–77.

https://doi.org/10.1086/671172 Published online by Cambridge University Press