Brit. J. Phil. Sci. **64** (2013), 189–204

New Axioms for Probability and Likelihood Ratio Measures

Vincenzo Crupi, Nick Chater, and Katya Tentori

ABSTRACT

Probability ratio and likelihood ratio measures of inductive support and related notions have appeared as theoretical tools for probabilistic approaches in the philosophy of science, the psychology of reasoning, and artificial intelligence. In an effort of conceptual clarification, several authors have pursued axiomatic foundations for these two families of measures. Such results have been criticized, however, as relying on unduly demanding or poorly motivated mathematical assumptions. We provide two novel theorems showing that probability ratio and likelihood ratio measures can be axiomatized in a way that overcomes these difficulties.

1 Introduction

In what follows, we will denote the quantity below as the probability ratio:

$$\frac{P(h|e)}{P(h)}$$

To ensure mathematical definiteness, we will assume throughout that *h* and *e* are contingent statements and *P* is a regular probability function (so that $0 < P(h), P(e) < 1$).

From Keynes ([1921], pp. 165ff) to Kuipers ([2000], pp. 49ff), the probability ratio, or strictly increasing functions of it, have often been said to measure the degree of inductive support or confirmation that evidence *e* provides to hypothesis *h*. Similarly, the probability ratio has been employed to characterize the strength of inductive arguments in the psychological study of human

doi:10.1093/bjps/axs018

Advance Access published on October 15, 2012

reasoning (see Viale and Osherson [2000]; Lo et al. [2002], p. 186), with e and h denoting premise and conclusion, respectively. Further, probability ratio measures have been invoked to quantify the explanatory power of hypothesis h with regards to datum e (Popper [1954], p. 147; McGrew [2003]), the severity of e as a (positive) test of hypothesis h (Popper [1963], p. 526), the amount of information transmitted by statement e relative to statement h (Hintikka [1968]), and the degree of coherence between statements e and h (Shogenji [1999]; Schupbach [2011]). Moreover, the notion of relative change in risk as employed in epidemiology also amounts to a probability ratio measure.[1]

On the other hand, we will denote the following quantity as the likelihood ratio:

$$\frac{P(e|h)}{P(e|\neg h)}$$

The likelihood ratio, and strictly increasing functions of it, have no less a remarkable historical record than probability ratio measures. As reported by Irving John Good, the likelihood ratio was crucially employed by Alan Turing in a 'vital cryptanalytic application' during the Second World War and said to quantify 'the factor in favour of h provided by e' (Good [1985], p. 252; also see Good [1950], pp. 62–3). Likelihood ratio measures have since been used to explicate or represent a number of related notions, namely, weight of evidence (Good [1950]; Minsky and Selfridge [1961]), belief change (Heckerman [1988]), corroboration (Good [1960], [1968]), and once again support or confirmation (for example, Kemeny and Oppenheim [1952]; Watanabe [1969], p. 374; Fitelson [2001]), as well as inductive argument strength (Oaksford and Hahn [2007], pp. 286 ff; also see Tenenbaum and Griffiths [2001]). Likelihood ratio measures have also played a central role in the definition and interpretation of some classical expert systems (Duda et al. [1976]; Heckerman [1986]; Heckerman and Shortliffe [1992]) as well as in other areas of artificial intelligence (for example, Dembczynski et al. [2007]). According to Fitelson and Hitchcock ([2011]), the likelihood ratio reflects a measure of causal strength in the spirit of Lewis ([1986]) (with h and e here denoting an antecedent causal factor and an associated subsequent event, respectively). Finally, the likelihood ratio occurs as a summary measure of the diagnostic value of a test result in the medical literature (Deeks and Altman [2004]) and as a measure of the probative value of a datum in legal reasoning and forensic science (Kaye and Koehler [2003]; Taroni et al. [2006]).

[1] To see why, let $P(h)$ be the so-called control event rate and $P(h|e)$ the event rate on e (the latter being a relevant experimental intervention or environmental exposure). The standard definition of relative change in risk (increase/reduction) becomes $(P(h|e) - P(h))/P(h)$ (see, for example, Barratt et al. [2004]), which is a simple increasing function of the probability ratio, namely, $P(h|e)/P(h) - 1$.

Seeking theoretical clarification, a natural goal is to axiomatize these two families of measures, i.e. to identify conditions that are necessary and sufficient to single out each of them as capturing a target notion. Indeed, various results of this kind have been presented. For our purposes, Heckerman ([1988]) provides a useful starting point. Heckerman was following Cox ([1946]), who derived probability as an essentially unique quantitative representation that fulfilled a few axioms held to characterize belief. In the attempt to provide a parallel result concerning probabilistic change in belief (or belief update), here labelled $C(h, e)$, Heckerman offered an axiomatic foundation of likelihood ratio measures involving the following principle:

(H) $P(h|e)$ is a continuous function of $C(h, e)$ and $P(h)$ only, and is increasing in each argument when the other is held constant.

If P is required to range over a continuous interval, however, its domain must also be non-denumerable. As a consequence, Heckerman's ([1988]) theorem does not cover as simple a case as, say, a well-specified urn setting. This point was forcefully noted by Halpern ([1999]), who traced the problem back to Cox ([1946]) himself. According to Fitelson ([2006]), the same difficulty afflicts two further contributions that are rather well-known in the philosophy of science, i.e. Good's ([1960]) earlier axiomatic derivation of likelihood ratio measures for the weight of evidence (also see Good [1968], [1984]) and Milne's ([1996]) derivation of the log probability ratio measure for confirmation. In fact, a common trait of all these contributions is the reliance on certain properties of functional equations (see Aczel [1966]) which presuppose the functions involved to be continuous. A possible resolution had been envisaged by Halpern ([1999]) himself: in order to defend the technical assumptions needed, one might emphasize that the function representing belief (or change in belief, for that matter) has to 'apply uniformly to all domains' (p. 80). In Huber's ([2008]) view, for instance, this line of argument would make continuity assumptions such as those in Milne ([1996]) 'perfectly reasonable' (p. 419). Be that as it may, they still would 'have no intuitive connection to material desiderata for inductive logic', as Fitelson ([2006]) pointed out (p. 506, fn 12).

Here we argue that the above difficulty can be overcome altogether. As we will show, probability ratio and likelihood ratio measures can be axiomatized just by means of principles that are philosophically significant while mathematically undemanding, with no need for the domain of the probability function to be non-denumerable. Most of the points summarized above arose in discussions of probabilistic measures of confirmation, and we will also adopt the same framework for our presentation. However, this does not prevent our results and their consequences from being exploited in other contexts (for example, probabilistic measures of explanatory power, coherence, causal strength, and so on).

2 Axioms for Probability Ratio Measures

Let L be a propositional language and L_c the set of the contingent formulae in L, i.e. those expressing neither logical truths nor logical falsehoods. Further, let $\boldsymbol{P}$ be the set of all regular probability functions that can be defined over L. Each element $P \in \boldsymbol{P}$ can thus be seen as representing a possible (non-dogmatic, see Howson [2000], p. 70) state of belief concerning a domain described in L. In order to run our argument, we will need to explicitly represent the dependence of confirmation on a given probability distribution, thus positing C: $\{L_c \times L_c \times \boldsymbol{P}\} \rightarrow \Re$ and adopting the notation $C_P(h, e)$, with $h, e \in L_c$.[2] Our first axiom is as follows:

A0. *Formality*.

There exists a function g such that, for any $h, e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = g[P(h \wedge e), P(h), P(e)]$.

Note that the probability distribution over the algebra generated by h and e is entirely determined by $P(h \wedge e)$, $P(h)$ and $P(e)$. So A0 simply states that $C_P(h, e)$ depends on that distribution, and nothing else. This is a widespread (albeit often tacit) assumption in discussions of confirmation in a probabilistic framework. Under slightly different renditions, it is also explicitly subscribed to by both Good ([1960], p. 322, [1968], pp. 127–8) and Milne ([1996], p. 21). The label formality is drawn from Tentori et al. ([2007], [2010]).

Now consider the following:

A1. *Final probability incrementality*.

For any h, e_1, $e_2 \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e_1) \gtreqless C_P(h, e_2)$ iff $P(h|e_1) \gtreqless P(h|e_2)$.

A2. *Law of likelihood*.

For any h_1, h_2, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h_1, e) \gtreqless C_P(h_2, e)$ iff $P(e|h_1) \gtreqless P(e|h_2)$.

A1 is closely related to the homonymous condition in Crupi et al. ([2010], pp. 77–9), there identified as describing a very basic property of the Bayesian notion of confirmation. The principle appears under the label 'law of conditional probability' in Hájek and Joyce ([2008], p. 122). The right-to-left part

[2] As usual, a further term, B, should be included, representing relevant background knowledge and assumptions, thus having $C_P(h, e|B)$. Such a term will be omitted from our notation for simple reasons of convenience, as it is inconsequential for our discussion.

of the biconditional also occurs in Steel ([2003], pp. 219–21) (who provides a few further references) and in Eells and Fitelson ([2000], p. 670), who remark: 'it is not an exaggeration to say that most Bayesian confirmation theorists' would accept it 'as a desideratum for Bayesian measures of confirmation' (see Fitelson [2006], p. 506, for yet another occurrence and a similar comment). As Eells and Fitelson ([2000]) further point out, the principle is crucially involved in classical Bayesian analyses such as the solution of the ravens paradox offered by Horwich ([1982], pp. 54–63).

A2 is endorsed by both Edwards ([1972], pp. 30–1) and Milne ([1996]). The label 'law of likelihood' goes back to Hacking ([1965]) and also occurs in Hájek and Joyce ([2008], p. 122) as well as in Crupi et al. ([2010], pp. 82–3).[3]

The following can be proved (essentially the same theorem was anticipated in Chater and Oaksford [2008, unpublished]):

Theorem 1 A0–A2 iff there exists a strictly increasing function, f, such that $C_P(h,e) = f\left(\frac{P(h|e)}{P(h)}\right)$.

A proof of Theorem 1 is provided in the Appendix.

3 Axioms for Likelihood Ratio Measures

In the set of axioms below, A2 has been replaced by a different statement A2*:

A0. *Formality*.

There exists a function g such that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h,e) = g(P(h \wedge e), P(h), P(e))$.

A1. *Final probability incrementality*.

For any h, e_1, $e_2 \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e_1) \gtreqless C_P(h, e_2)$ iff $P(h|e_1) \gtreqless P(h|e_2)$.

A2*. *Modularity (for conditionally independent data)*.

For any h, e_1, $e_2 \in L_c$ and any $P \in \boldsymbol{P}$, if $P(e_1|\pm h \wedge e_2) = P(e_1|\pm h)$, then $C_P(h, e_1|e_2) = C_P(h, e_1)$.

The notion of conditional confirmation in A2* is meant as usual, i.e. with all the relevant values from P being conditionalized on e_2. The label for A2* is freely adapted after Heckerman ([1988], pp. 18–19). The same principle also appears in Good ([1968], p. 134) and Fitelson ([2001], p. S130).

[3] Sober ([2008], pp. 32ff) deserves separate mention here, for he deliberately handles the 'law of likelihood' as allowing for a non-committal attitude towards Bayesian priors, so that his rendition does not imply our condition A2. Also notice that A2 is not restricted to pairs of mutually exclusive hypotheses, h_1 and h_2 (see the discussion in Steel [2007]).

The following can be shown:

Theorem 2 A0–A2* iff there exists a strictly increasing function, f, such that $C_P(h, e) = f\left(\frac{P(e|h)}{P(e|\neg h)}\right)$.

A proof of Theorem 2 is also provided in the Appendix.

Heckerman's ([1988]) main result closely resembles Theorem 2, but with his demanding principle (H) above instead of A1. The connection between A1 and (H) is thus of particular interest here. On the one hand, as we already know, A1 does not require any continuous probability function and thus no commitment to a non-denumerable domain. But A1 does convey one core philosophical tenet of (H), i.e. that for any hypothesis, h, posterior and confirmation always move in the same direction in the light of data, e. As pointed out above, this seems intuitively compelling, and certainly so within a Bayesian perspective.

4 Discussion

While we confined full-fledged proofs to the Appendix, an informal overview of the proof strategy involved might be of interest, in that it indirectly elucidates how functional equations and continuity assumptions can be dispensed with.[4] The left-to-right implication for the probability ratio case (i.e. Theorem 1 above) will serve as an example. (The right-to-left side of both theorems is indeed rather straightforward.)

As a first step, one relies on axiom A0 to point out that $C_P(h, e)$ demonstrably is a function of $P(h)$ and $P(e)$ along with the target quantity relating h and e, in this case the probability ratio itself $P(h|e)/P(h)$, so that $C_P(h, e) = j(P(h|e)/P(h), P(h), P(e))$. One proceeds to show, however, that were $C_P(h, e)$ to actually depend on $P(e)$, it would then give different values for some distinct e_1 and e_2 with $P(e_1) \neq P(e_2)$ even if $P(h|e_1) = P(h|e_2)$, thus violating axiom A1. So A1 rules out $P(e)$ as a separate relevant variable for $C_P(h, e)$ and in fact dictates that $C_P(h, e) = k(P(h|e)/P(h), P(h))$. By a similar move, one can now show that, were $C_P(h, e)$ to actually depend on $P(h)$, it would then give different values for some distinct h_1 and h_2 with $P(h_1) \neq P(h_2)$ even if $P(h_1|e)/P(h_1) = P(h_2|e)/P(h_2)$, namely—by Bayes's theorem—$P(e|h_1) = P(e|h_2)$, thus violating axiom A2. So A2 further rules out $P(h)$ as a separate relevant variable for $C_P(h, e)$ and in fact dictates that $C_P(h,e) = f(P(h|e)/P(h))$. Finally, the strictly increasing behaviour of f easily follows too, once again from A1. Essentially the same pattern of derivation applies in the likelihood ratio proof (Theorem 2 above) and can be exploited in still other cases (see Crupi and Tentori [forthcoming (a), (b)] for other applications).

[4] We thank an anonymous reviewer for suggesting this illustration.

As concerns further implications of our current results, let us first take the connection between probability ratio measures and the law of likelihood. This had not escaped notice in the literature (see Fitelson [2007]). The connection is only partial, however, as shown by consideration of Mortimer's ([1988], Section 11.1) measure of confirmation (formally equivalent to Suppes's [1970] measure of causal strength as construed by Fitelson and Hitchcock [2011]):

$$P(e|h) - P(e).$$

This quantity does satisfy our axiom A2, i.e. the law of likelihood, while being ordinally divergent from probability ratio measures (see Crupi et al. [2007], p. 231, for a proof of this divergence). Our Theorem 1 above neatly identifies what more, beyond A2, is necessary and sufficient to isolate probability ratio measures, i.e. axioms A0 and A1 (indeed, Mortimer's measure demonstrably breaks with axiom A1.)

Similarly, consider A2*, i.e. modularity for conditionally independent data. Its connection with likelihood ratio measures has itself been repeatedly noticed (see Good [1968]; Fitelson [2001]). Yet that principle is also satisfied by a quantity that is ordinally distinct (see again Crupi et al. [2007], p. 231), i.e. Nozick's [1981] measure of confirmation (formally equivalent to Eells's [1991] measure of causal strength, as construed by Fitelson and Hitchcock [2011]), namely, the following:

$$P(e|h) - P(e|\neg h).$$

Theorem 2 above tells us what more, beyond A2*, is necessary and sufficient to single out likelihood ratio measures, i.e. again Axioms A0 and A1 (Nozick's measure also breaks with Axiom A1; see Crupi et al. [2010], p. 81.)

More broadly, analyses in the literature can be roughly distinguished as favouring 'monism' (for example, Good [1984]; Heckerman [1988]; Milne [1996]; Fitelson [2001]; Crupi and Tentori [2010]) versus more or less extreme forms of pluralism concerning alternative measures of confirmation (for example, Festa [1999]; Howson [2000], pp. 184–5; Steel [2007]; Hájek and Joyce [2008]; Huber [2008]). A similar situation arises in related domains, such as the probabilistic formalization of explanatory power (see, for example, Schupbach and Sprenger [2011]; Crupi and Tentori [2012]). In this contribution, we do not enter these debates. We do believe, however, that axiomatizing alternative candidate explications of epistemological concepts fosters insight in their properties and discussion of their implications. So far, it has appeared that such insights come at the cost of accepting mathematical assumptions of continuity that were debatable and yet required for technical reasons. Not so.

Acknowledgements

We are grateful to Branden Fitelson, David Heckerman, Hannes Leitgeb, Tommaso Mastropasqua, and Dan Osherson for very useful exchanges on the results presented here. This work is supported by Italian Ministry of Scientific Research (PRIN grant 20083NAH2L_001); Spanish Department of Science and Innovation (FFI2008-01169/FISO); Alexander von Humboldt Foundation.

Vincenzo Crupi
Department of Philosophy and Education
University of Turin, via S. Ottavio 20
10124 Torino, Italy
and
Munich Center for Mathematical Philosophy
Ludwig Maximilian University
Ludwigstrasse 31, 80539 Munich, Germany
vincenzo.crupi@unito.it

Nick Chater
Warwick Business School
University of Warwick
Coventry CV4 7AL, UK
nick.chater@wbs.ac.uk

Katya Tentori
Center for Mind/Brain Science
and
Department of Cognitive Science and Education
University of Trento, corso Bettini 31
38068 Rovereto, Italy
katya.tentori@unitn.it

Appendix

Proof of Theorem 1: A0–A2 iff there exists a strictly increasing function, f, such that $C_P(h, e) = f\left(\frac{P(h|e)}{P(h)}\right)$.

The proof provided concerns the left-to-right implication in the theorem (verification of the right-to-left implication is straightforward).

Notice that $P(h \wedge e) = (P(h|e)/P(h))P(h)P(e)$. As a consequence, by A0, there exists a function, j, such that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = j(P(h|e)/P(h), P(h), P(e))$. With no loss of generality, we will convey probabilistic coherence and regularity by constraining the domain of j to include triplets of values (x, y, z) such that the following conditions are jointly satisfied:

(i) $0 < y, z < 1$;

(ii) $x \geq 0$, by which $x = P(h|e)/P(h) \geq 0$, so that $P(h|e) \geq 0$, and thus $P(h \wedge e) \geq 0$;

(iii) $x \leq 1/y$, by which $xy = P(h|e) \leq 1$, so that $P(h \wedge e) \leq P(e)$, and thus $P(\neg h \wedge e) \geq 0$;

(iv) $x \leq 1/z$, by which $xz = P(e|h) \leq 1$, so that $P(h \wedge e) \leq P(h)$, and thus $P(h \wedge \neg e) \geq 0$;

(v) $x \geq (y + z - 1)/yz$, by which $xyz = P(h \wedge e) \geq P(h) + P(e) - 1 = y + z - 1$, and thus $P(h \wedge e) + P(\neg h \wedge e) + P(h \wedge \neg e) \leq 1$.

We thus posit j: $\{(x, y, z) \in \{\Re^+ \cup \{0\}\} \times (0, 1)^2 \,|\, (y + z - 1)/yz \leq x \leq 1/y, 1/z\} \to \Re$ and denote the domain of j as D_j.

Lemma 1.1: For any x, y, z_1, z_2 such that $x \in \Re^+ \cup \{0\}$, $y, z_1, z_2 \in (0,1)$, and $(y + z_1 - 1)/yz_1, (y + z_2 - 1)/yz_2 \leq x \leq 1/y, 1/z_1, 1/z_2$, there exist $h, e_1, e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $P'(h|e_1)/P'(h) = P'(h|e_2)/P'(h) = x$, $P'(h) = y$, $P'(e_1) = z_1$ and $P'(e_2) = z_2$.

Proof: The equalities in Lemma 1.1 arise from the following scheme of probability assignments:

$P'(h \wedge e_1 \wedge e_2) = (xz_1)(xz_2)y$; $\quad P'(\neg h \wedge e_1 \wedge e_2) = \frac{(1-xy)^2 z_1 z_2}{(1-y)}$;

$P'(h \wedge e_1 \wedge \neg e_2) = (xz_1)(1 - xz_2)y$; $\quad P'(\neg h \wedge e_1 \wedge \neg e_2) = (1 - xy)z_1\left(1 - \frac{(1-xy)z_2}{(1-y)}\right)$;

$P'(h \wedge \neg e_1 \wedge e_2) = (1 - xz_1)(xz_2)y$; $\quad P'(\neg h \wedge \neg e_1 \wedge e_2) = \left(1 - \frac{(1-xy)z_1}{(1-y)}\right)(1 - xy)z_2$;

$P'(h \wedge \neg e_1 \wedge \neg e_2) = (1 - xz_1)(1 - xz_2)y$; $\quad P'(\neg h \wedge \neg e_1 \wedge \neg e_2) = \left(1 - \frac{(1-xy)z_1}{(1-y)}\right)\left(1 - \frac{(1-xy)z_2}{(1-y)}\right)(1 - y)$.

Suppose there exist $(x, y, z_1), (x, y, z_2) \in D_j$ such that $j(x, y, z_1) \neq j(x, y, z_2)$. Then, by Lemma 1.1 and the definition of D_j, there exist $h, e_1, e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $P'(h|e_1)/P'(h) = P'(h|e_2)/P'(h) = x$, $P'(h) = y$, $P'(e_1) = z_1$, and $P'(e_2) = z_2$. Clearly, if the latter equalities hold, then $P'(h|e_1) = P'(h|e_2)$. Thus, there exist $h, e_1, e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $C_{P'}(h, e_1) = j(x, y, z_1) \neq j(x, y, z_2) = C_{P'}(h, e_2)$ even if $P'(h|e_1) = P'(h|e_2)$, contradicting A1. Conversely, A1 implies that, for any $(x, y, z_1), (x, y, z_2) \in D_j$, $j(x, y, z_1) = j(x, y, z_2)$. So, for A1 to hold, there must exist k such that, for any $h, e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = k(P(h|e)/P(h), P(h))$ and $k(x, y) = j(x, y, z)$. We thus posit k: $\{(x, y) \in \{\Re^+ \cup \{0\}\} \times (0, 1) \,|\, x \leq 1/y\} \to \Re$ and denote the domain of k as D_k.

Lemma 1.2: For any x, y_1, y_2 such that $x \in \Re^+ \cup \{0\}$, $y_1, y_2 \in (0, 1)$ and $x \leq 1/y_1$, $1/y_2$, there exist h_1, h_2, $e \in L_c$ and $P'' \in \boldsymbol{P}$ such that $P''(h_1|e)/P''(h_1) = P''(h_2|e)/P''(h_2) = x$, $P''(h_1) = y_1$, and $P''(h_2) = y_2$.

Proof: Let $z \in (0, 1)$ be given so that $z \leq 1/x$, $(1-y_1)/(1-xy_1)$, $(1-y_2)/(1-xy_2)$ (as the latter quantities must all be positive, z exists). The equalities in Lemma 1.2 arise from the following scheme of probability assignments:

$$P''(h_1 \wedge h_2 \wedge e) = (xy_1)(xy_2)z; \qquad P''(\neg h_1 \wedge h_2 \wedge e) = (1-xy_1)(xy_2)z;$$

$$P''(h_1 \wedge h_2 \wedge \neg e) = \frac{(1-xz)^2 y_1 y_2}{(1-z)}; \qquad P''(\neg h_1 \wedge h_2 \wedge \neg e) = \left(1 - \frac{(1-xz)y_1}{(1-z)}\right)(1-xz)y_2;$$

$$P''(h_1 \wedge \neg h_2 \wedge e) = (xy_1)(1-xy_2)z; \qquad P''(\neg h_1 \wedge \neg h_2 \wedge e) = (1-xy_1)(1-xy_2)z;$$

$$P''(h_1 \wedge \neg h_2 \wedge \neg e) = (1-xz)y_1\left(1 - \frac{(1-xz)y_2}{(1-z)}\right); \qquad P''(\neg h_1 \wedge \neg h_2 \wedge \neg e) = \left(1 - \frac{(1-xz)y_1}{(1-z)}\right)\left(1 - \frac{(1-xz)y_2}{(1-z)}\right)(1-z).$$

Suppose there exist $(x, y_1), (x, y_2) \in D_k$ such that $k(x, y_1) \neq k(x, y_2)$. Then, by Lemma 1.2 and the definition of D_k, there exist h_1, h_2, $e \in L_c$ and $P'' \in \boldsymbol{P}$ such that $P''(h_1|e)/P''(h_1) = P''(h_2|e)/P''(h_2) = x$, $P''(h_1) = y_1$, and $P''(h_2) = y_2$. By the probability calculus, if the latter equalities hold, then $P''(e|h_1) = P''(e|h_2)$. Thus, there exist h_1, h_2, $e \in L_c$ and $P'' \in \boldsymbol{P}$ such that $C_{P''}(h_1, e) = k(x, y_1) \neq k(x, y_2) = C_{P''}(h_2, e)$ even if $P''(e|h_1) = P''(e|h_2)$, contradicting A2. Conversely, A2 implies that, for any $(x, y_1), (x, y_2) \in D_k$, $k(x, y_1) = k(x, y_2)$. So, for A2 to hold, there must exist f such that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = f(P(h|e)/P(h))$ and $f(x) = k(x, y)$. We thus posit f: $\{\Re^+ \cup \{0\}\} \rightarrow \Re$ and denote the domain of f as D_f.

Lemma 1.3: For any $x_1, x_2 \in \Re^+ \cup \{0\}$, there exist h, e_1, $e_2 \in L_c$ and $P''' \in \boldsymbol{P}$ such that $P'''(h|e_1)/P'''(h) = x_1$ and $P'''(h|e_2)/P'''(h) = x_2$.

Proof: Let y, z_1, $z_2 \in (0, 1)$ be given so that $y \leq 1/x_1, 1/x_2$ (as the latter quantities must all be positive, y exists), $z_1 \leq 1/x_1$, $(1-y)/(1-x_1y)$ (as the latter quantities must all be positive, z_1 exists), and $z_2 \leq 1/x_2$, $(1-y)/(1-x_2y)$ (as the latter quantities must all be positive, z_2 exists). The equalities in Lemma 1.3 arise from the following scheme of probability assignments:

$$P'''(h \wedge e_1 \wedge e_2) = (x_1z_1)(x_2z_2)y; \qquad P'''(\neg h \wedge e_1 \wedge e_2) = \frac{(1-x_1y)(1-x_2y)z_1z_2}{(1-y)};$$

$$P'''(h \wedge e_1 \wedge \neg e_2) = (x_1z_1)(1-x_2z_2)y; \qquad P'''(\neg h \wedge e_1 \wedge \neg e_2) = (1-x_1y)z_1\left(1 - \frac{(1-x_2y)z_2}{(1-y)}\right);$$

$$P'''(h \wedge \neg e_1 \wedge e_2) = (1-x_1z_1)(x_2z_2)y; \qquad P'''(\neg h \wedge \neg e_1 \wedge e_2) = \left(1 - \frac{(1-x_1y)z_1}{(1-y)}\right)(1-x_2y)z_2;$$

$$P'''(h \wedge \neg e_1 \wedge \neg e_2) = (1-x_1z_1)(1-x_2z_2)y; \qquad P'''(\neg h \wedge \neg e_1 \wedge \neg e_2) = \left(1 - \frac{(1-x_1y)z_1}{(1-y)}\right)\left(1 - \frac{(1-x_2y)z_2}{(1-y)}\right)(1-y).$$

Suppose there exist x_1, $x_2 \in D_f$ such that $x_1 > x_2$ and $f(x_1) \leq f(x_2)$. Then, by Lemma 1.3 and the definition of D_f, there exist h, e_1, $e_2 \in L_c$

and $P''' \in \boldsymbol{P}$ such that $P'''(h|e_1)/P'''(h) = x_1$ and $P'''(h|e_2)/P'''(h) = x_2$. Clearly, if the latter equalities hold, then $P'''(h|e_1) > P'''(h|e_2)$. Thus, there exist h, e_1, $e_2 \in L_c$ and $P''' \in \boldsymbol{P}$ such that $C_{P'''}(h, e_1) = f(x_1) \leq f(x_2) = C_{P'''}(h, e_2)$ even if $P'''(h|e_1) > P'''(h|e_2)$, contradicting A1. Conversely, A1 implies that, for any x_1, $x_2 \in \Re^+ \cup \{0\}$, if $x_1 > x_2$ then $f(x_1) > f(x_2)$. By a similar argument, A1 also implies that, for any x_1, $x_2 \in \Re^+ \cup \{0\}$, if $x_1 = x_2$ then $f(x_1) = f(x_2)$. So, for A1 to hold, it must be that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = f(P(h|e)/P(h))$ and f is a strictly increasing function. □

Proof of Theorem 2: A0–A2* iff there exists a strictly increasing function, f, such that $C_P(h, e) = f\left(\frac{P(e|h)}{P(e|\neg h)}\right)$.

The proof provided concerns the left-to-right implication in the theorem (verification of the right-to-left implication is straightforward).

Notice that

$$P(h \wedge e) = \frac{P(h)\frac{P(e|h)}{P(e|\neg h)}}{P(h)\frac{P(e|h)}{P(e|\neg h)} + P(\neg h)} P(e)$$

As a consequence, by A0, there exist a function j such that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = j(P(e|h)/P(e|\neg h), P(h), P(e))$. With no loss of generality, we will convey probabilistic coherence and regularity by constraining the domain of j to include triplets of values (x, y, z) such that the following conditions are jointly satisfied:

(i) $0 < z < 1$;

(ii) $x \geq 0$, by which $xy/(xy + (1 - y)) = P(h|e) \geq 0$, and thus $P(h \wedge e) \geq 0$;

(iii) $0 < y < 1$, by which $xy/(xy + (1 - y)) = P(h|e) \leq 1$, so that $P(h \wedge e) \leq P(e)$, and thus $P(\neg h \wedge e) \geq 0$;

(iv) $(z \leq y) \vee (x \leq (1 - y)/(z - y))$, by which $xz/(xy + (1 - y)) = P(e|h) \leq 1$, so that $P(h \wedge e) \leq P(h)$, and thus $P(h \wedge \neg e) \geq 0$;

(v) $x \geq (y + z - 1)/y$, by which $z/(xy + (1 - y)) = P(e|\neg h) \leq 1$, so that $P(\neg h \wedge e) \leq 1 - P(h)$, and thus $P(h \wedge e) + P(\neg h \wedge e) + P(h \wedge \neg e) \leq 1$.

We thus posit j: $\{(x, y, z) \in \{\{\Re^+ \cup \{0\}\} \times (0, 1)^2 \,|\, (y + z - 1)/y \leq x$ and $((z \leq y) \vee (x \leq (1 - y)/(z - y)))\} \rightarrow \Re$ and denote the domain of j as D_j.

Lemma 2.1: For any x,y,z_1,z_2 such that $x \in \Re^+ \cup \{0\}$, y, z_1, $z_2 \in (0, 1)$, $(y + z_1 - 1)/y$, $(y + z_2 - 1)/y \leq x$, $((z_1 \leq y) \vee (x \leq (1 - y)/(z_1 - y)))$ and $((z_2 \leq y) \vee (x \leq (1 - y)/(z_2 - y)))$, there exist h, e_1, $e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $P'(e_1|h)/P'(e_1|\neg h) = P'(e_2|h)/P'(e_2|\neg h) = x$, $P'(h) = y$, $P'(e_1) = z_1$, and $P'(e_2) = z_2$.

Proof: The equalities in Lemma 2.1 arise from the following scheme of probability assignments:

$P'(h \wedge e_1 \wedge e_2) = \left(\frac{z_1 x}{yx+(1-y)}\right)\left(\frac{z_2 x}{yx+(1-y)}\right)y$; $P'(\neg h \wedge e_1 \wedge e_2) = \left(\frac{z_1}{yx+(1-y)}\right)\left(\frac{z_2}{yx+(1-y)}\right)(1-y)$;

$P'(h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1 x}{yx+(1-y)}\right)\left(1-\frac{z_2 x}{yx+(1-y)}\right)y$; $P'(\neg h \wedge \neg e_1 \wedge e_2) = \left(1-\frac{z_1}{yx+(1-y)}\right)\left(\frac{z_2}{yx+(1-y)}\right)(1-y)$;

$P'(h \wedge \neg e_1 \wedge e_2) = \left(1-\frac{z_1 x}{yx+(1-y)}\right)\left(\frac{z_2 x}{yx+(1-y)}\right)y$; $P'(\neg h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1}{yx+(1-y)}\right)\left(1-\frac{z_2}{yx+(1-y)}\right)(1-y)$;

$P'(h \wedge \neg e_1 \wedge \neg e_2) = \left(1-\frac{z_1 x}{yx+(1-y)}\right)\left(1-\frac{z_2 x}{yx+(1-y)}\right)y$; $P'(\neg h \wedge \neg e_1 \wedge \neg e_2) = \left(1-\frac{z_1}{yx+(1-y)}\right)\left(1-\frac{z_2}{yx+(1-y)}\right)(1-y)$.

Suppose there exist $(x, y, z_1), (x, y, z_2) \in D_j$ such that $j(x, y, z_1) \neq j(x, y, z_2)$. Then, by Lemma 2.1 and the definition of D_j, there exist $h, e_1, e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $P'(e_1|h)/P'(e_1|\neg h) = P'(e_2|h)/P'(e_2|\neg h) = x$, $P'(h) = y$, $P'(e_1) = z_1$, and $P'(e_2) = z_2$. By the probability calculus, if the latter equalities hold, then $P'(h|e_1) = P'(h|e_2)$. Thus, there exist $h, e_1, e_2 \in L_c$ and $P' \in \boldsymbol{P}$ such that $C_{P'}(h, e_1) = j(x, y, z_1) \neq j(x, y, z_2) = C_{P'}(h, e_2)$ even if $P'(h|e_1) = P'(h|e_2)$, contradicting A1. Conversely, A1 implies that, for any $(x, y, z_1), (x, y, z_2) \in D_j$, $j(x, y, z_1) = j(x, y, z_2)$. So, for A1 to hold, there must exist k such that, for any $h, e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = k[P(e|h)/P(e|\neg h), P(h)]$ and $k(x, y) = j(x, y, z)$. We thus posit $k: \{(x, y) \in \{\Re^+ \cup \{0\}\} \times (0, 1)\} \rightarrow \Re$ and denote the domain of k as D_k.

Lemma 2.2: For any x, y_1, y_2 such that $x \in \Re^+ \cup \{0\}$ and $y_1, y_2 \in (0, 1)$, there exist $h, e_1, e_2 \in L_c$ and $P'' \in \boldsymbol{P}$ such that $P''(e_1|h)/P''(e_1|\neg h) = x$, $P''(e_1|h) = P''(e_1|h \wedge e_2)$, $P''(e_1|\neg h) = P''(e_1|\neg h \wedge e_2)$, $P''(h) = y_1$, and $P''(h|e_2) = y_2$.

Proof: Let $z_1, z_2 \in (0, 1)$ be given so that $z_1 \leq y_1, xy_1 + (1-y_1)$ (as the latter quantities must all be positive, z_1 exists) and $z_2 \leq y_1/y_2, (1-y_1)/(1-y_2)$ (as the latter quantities must all be positive, z_2 exists). The equalities in Lemma 2.2 arise from the following scheme of probability assignments:

$P''(h \wedge e_1 \wedge e_2) = \left(\frac{z_1 x}{y_1 x+(1-y_1)}\right)y_2 z_2$; $P''(\neg h \wedge e_1 \wedge e_2) = \left(\frac{z_1}{y_1 x+(1-y_1)}\right)(1-y_2)z_2$;

$P''(h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1 x}{y_1 x+(1-y_1)}\right)(y_1 - y_2 z_2)$; $P''(\neg h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1}{y_1 x+(1-y_1)}\right)\left(1-\frac{y_1-y_2 z_2}{(1-z_2)}\right)(1-z_2)$;

$P''(h \wedge \neg e_1 \wedge e_2) = \left(1-\frac{z_1 x}{y_1 x+(1-y_1)}\right)y_2 z_2$; $P''(\neg h \wedge \neg e_1 \wedge e_2) = \left(1-\frac{z_1}{y_1 x+(1-y_1)}\right)(1-y_2)z_2$;

$P''(h \wedge \neg e_1 \wedge \neg e_2) = \left(1-\frac{z_1 x}{y_1 x+(1-y_1)}\right)(y_1 - y_2 z_2)$; $P''(\neg h \wedge \neg e_1 \wedge \neg e_2) = \left(1-\frac{z_1}{y_1 x+(1-y_1)}\right)\left(1-\frac{y_1-y_2 z_2}{(1-z_2)}\right)(1-z_2)$.

Suppose there exist $(x, y_1), (x, y_2) \in D_k$ such that $k(x, y_1) \neq k(x, y_2)$. Then, by Lemma 2.2 and the definition of D_k, there exists $h, e_1, e_2 \in L_c$ and $P'' \in \boldsymbol{P}$ such that $P''(e_1|h)/P''(e_1|\neg h) = x$, $P''(e_1|h) = P''(e_1|h \wedge e_2)$, $P''(e_1|\neg h) = P''(e_1|\neg h \wedge e_2)$, $P''(h) = y_1$, and $P''(h|e_2) = y_2$. Clearly, if the latter equalities hold, then $P''(e_1|\pm h) = P''(e_1|\pm h \wedge e_2)$. Thus, there exist $h, e_1, e_2 \in L_c$ and $P'' \in \boldsymbol{P}$ such that $C_{P''}(h, e_1) = k(x, y_1) \neq k(x, y_2) = C_{P''}(h, e_1|e_2)$ even if $P''(e_1|\pm h) = P''(e_1|\pm h \wedge e_2)$, contradicting A2*. Conversely, A2* implies that, for any

$(x, y_1), (x, y_2) \in D_k$, $k(x, y_1) = k(x, y_2)$. So, for A2* to hold, there must exist f such that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = f(P(e|h)/P(e|\neg h))$ and $f(x) = k(x, y)$. We thus posit f: $\{\Re^+ \cup \{0\}\} \rightarrow \Re$ and denote the domain of f as D_f.

Lemma 2.3: For any $x_1, x_2 \in \Re^+ \cup \{0\}$, there exist h, e_1, $e_2 \in L_c$ and $P''' \in \boldsymbol{P}$ such that $P'''(e_1|h)/P'''(e_1|\neg h) = x_1$ and $P'''(e_2|h)/P'''(e_2|\neg h) = x_2$.

Proof: Let y, z_1, $z_2 \in (0, 1)$ be given so that $z_1 \leq y$, $x_1 y + (1-y)$ (as the latter quantities must all be positive, z_1 exists) and $z_2 \leq y$, $x_2 y + (1-y)$ (as the latter quantities must all be positive, z_2 exists). The equalities in Lemma 2.3 arise from the following scheme of probability assignments:

$$P'''(h \wedge e_1 \wedge e_2) = \left(\frac{z_1 x_1}{yx_1 + (1-y)}\right)\left(\frac{z_2 x_2}{yx_2 + (1-y)}\right)y;$$
$$P'''(h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1 x_1}{yx_1 + (1-y)}\right)\left(1 - \frac{z_2 x_2}{yx_2 + (1-y)}\right)y;$$
$$P'''(h \wedge \neg e_1 \wedge e_2) = \left(1 - \frac{z_1 x_1}{yx_1 + (1-y)}\right)\left(\frac{z_2 x_2}{yx_2 + (1-y)}\right)y;$$
$$P'''(h \wedge \neg e_1 \wedge \neg e_2) = \left(1 - \frac{z_1 x_1}{yx_1 + (1-y)}\right)\left(1 - \frac{z_2 x_2}{yx_2 + (1-y)}\right)y;$$
$$P'''(\neg h \wedge e_1 \wedge e_2) = \left(\frac{z_1}{yx_1 + (1-y)}\right)\left(\frac{z_2}{yx_2 + (1-y)}\right)(1-y);$$
$$P'''(\neg h \wedge e_1 \wedge \neg e_2) = \left(\frac{z_1}{yx_1 + (1-y)}\right)\left(1 - \frac{z_2}{yx_2 + (1-y)}\right)(1-y);$$
$$P'''(\neg h \wedge \neg e_1 \wedge e_2) = \left(1 - \frac{z_1}{yx_1 + (1-y)}\right)\left(\frac{z_2}{yx_2 + (1-y)}\right)(1-y);$$
$$P'''(\neg h \wedge \neg e_1 \wedge \neg e_2) = \left(1 - \frac{z_1}{yx_1 + (1-y)}\right)\left(1 - \frac{z_2}{yx_2 + (1-y)}\right)(1-y).$$

Suppose there exist $x_1, x_2 \in D_f$ such that $x_1 > x_2$ and $f(x_1) \leq f(x_2)$. Then, by Lemma 2.3 and the definition of D_f, there exist h, e_1, $e_2 \in L_c$ and $P''' \in \boldsymbol{P}$ such that $P'''(e_1|h)/P'''(e_1|\neg h) = x_1$ and $P'''(e_2|h)/P'''(e_2|\neg h) = x_2$. By the probability calculus, if the latter equalities hold, then $P'''(h|e_1) > P'''(h|e_2)$. Thus, there exist h, e_1, $e_2 \in L_c$ and $P''' \in \boldsymbol{P}$ such that $C_{P'''}(h, e_1) = f(x_1) \leq f(x_2) = C_{P'''}(h, e_2)$ even if $P'''(h|e_1) > P'''(h|e_2)$, contradicting A1. Conversely, A1 implies that, for any $x_1, x_2 \in \Re^+ \cup \{0\}$, if $x_1 > x_2$ then $f(x_1) > f(x_2)$. By a similar argument, A1 also implies that, for any $x_1, x_2 \in \Re^+ \cup \{0\}$, if $x_1 = x_2$ then $f(x_1) = f(x_2)$. So, for A1 to hold, it must be that, for any h, $e \in L_c$ and any $P \in \boldsymbol{P}$, $C_P(h, e) = f(P(e|h)/P(e|\neg h))$ and f is a strictly increasing function. □

References

Aczel, J. [1966]: *Lectures on Functional Equations and their Applications*, New York: Academic Press.

Barratt, A., Wyer, P. C., Hatala, R., McGinn, T., Dans, A. L., Keitz, S., Moyer, V. and Guyatt, G. [2004]: 'Tips for Learners of Evidence-Based Medicine: Relative Risk Reduction, Absolute Risk Reduction, and Number Needed to Treat', *Canadian Medical Association Journal*, **171**, pp. 353–8.

Chater, N. and Oaksford, M. [2008, unpublished]: 'Formal Measures of Confirmation and Argument Strength', Pre-conference workshop of the *Sixth International Conference of Thinking* (Venice) on *The Cognitive Science of Induction and Confirmation*.

Cox, R. T. [1946]: 'Probability, Frequency, and Reasonable Expectation', *American Journal of Physics*, **14**, pp. 1–13.

Crupi, V. and Tentori, K. [2010]: 'Irrelevant Conjunction: Statement and Solution of a New Paradox', *Philosophy of Science*, **77**, pp. 1–13.

Crupi, V. and Tentori, K. [forthcoming (a)]: 'Confirmation Theory', in A. Hájek and C. Hitchcock (*eds*), *Oxford Handbook of Probability and Philosophy*, Oxford: Oxford University Press.

Crupi, V. and Tentori, K. [forthcoming (b)]: 'Confirmation as Partial Entailment: A Representation Theorem in Inductive Logic', *Journal of Applied Logic*.

Crupi, V. and Tentori, K. [2012]: 'A Second Look at the Logic of Explanatory Power (with Two Novel Representation Theorems)', *Philosophy of Science*, **79**, pp. 365–85.

Crupi, V., Festa, R. and Buttasi, C. [2010]: 'Towards a Grammar of Bayesian Confirmation', in M. Suárez, M. Dorato and M. Rédei (*eds*), *Epistemology and Methodology of Science*, Dordrecht: Springer, pp. 73–93.

Crupi, V., Tentori, K. and Gonzalez, M. [2007]: 'On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues', *Philosophy of Science*, **74**, pp. 229–52.

Deeks, J. J. and Altman, D. G. [2004]: 'Diagnostic Tests 4: Likelihood Ratios', *British Medical Journal*, **329**, pp. 168–9.

Dembczynski, K., Greco, S., Kotlowski, W. and Slowinski, R. [2007]: 'Relationship between Loss Functions and Confirmation Measures', *Lecture Notes in Computer Science*, **4482**, pp. 338–45.

Duda, R. O., Hart, P. E. and Nilsson, N. J. [1976]: 'Subjective Bayesian Methods for Rule-Based Inference Systems', in B. L. Webber and N. J. Nilsson (*eds*), *Readings in Artificial Intelligence*, Palo Alto, CA: Tioga, 1981, pp. 192–9.

Edwards, A. W. F. [1972]: *Likelihood*, Cambridge: Cambridge University Press.

Eells, E. [1991]: *Probabilistic Causality*, Cambridge: Cambridge University Press.

Eells, E. and Fitelson, B. [2000]: 'Measuring Confirmation and Evidence', *Journal of Philosophy*, **97**, pp. 663–72.

Festa, R. [1999]: 'Bayesian Confirmation', in M. Galavotti and A. Pagnini (*eds*), *Experience, Reality, and Scientific Explanation*, Dordrecht: Kluwer, pp. 55–87.

Fitelson, B. [2001]: 'A Bayesian Account of Independent Evidence with Applications', *Philosophy of Science*, **68**, pp. S123–40.

Fitelson, B. [2006]: 'Logical Foundations of Evidential Support', *Philosophy of Science*, **73**, pp. 500–12.

Fitelson, B. [2007]: 'Likelihoodism, Bayesianism, and Relational Confirmation', *Synthese*, **156**, pp. 473–89.

Fitelson, B. and Hitchcock, C. [2011]: 'Probabilistic Measures of Causal Strength', in P. McKay Illari, F. Russo and J. Williamson (*eds*), *Causality in the Sciences*, Oxford: Oxford University Press, pp. 600–27.

Good, I. J. [1950]: *Probability and the Weighing of Evidence*, London: Griffin.

Good, I. J. [1960]: 'Weight of Evidence, Corroboration, Explanatory Power, Information and the Utility of Experiments', *Journal of the Royal Statistical Society: Series B*, **22**, pp. 319–31.

Good, I. J. [1968]: 'Corroboration, Explanation, Evolving Probability, Simplicity, and a Sharpened Razor', *British Journal for the Philosophy of Science*, **19**, pp. 123–43.

Good, I. J. [1984]: 'The Best Explicatum for Weight of Evidence', *Journal of Statistical Computation and Simulation*, **19**, pp. 294–9.

Good, I. J. [1985]: 'Weight of Evidence: A Brief Survey', in J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith (*eds*), *Bayesian Statistics* 2, Amsterdam: North-Holland, pp. 249–70.

Hacking, I. [1965]: *Logic of Statistical Inference*, Cambridge, MA: Cambridge University Press.

Hájek, A. and Joyce, J. [2008]: 'Confirmation', in S. Psillos and M. Curd (*eds*), *Routledge Companion to the Philosophy of Science*, New York: Routledge, pp. 115–29.

Halpern, J. [1999]: 'A Counterexample to Theorems of Cox and Fine', *Journal of Artificial Intelligence Research*, **10**, pp. 67–85.

Heckerman, D. [1986]: 'Probabilistic Interpretation for MYCIN's Certainty Factors', in L. N. Kanal and J. F. Lemmer (*eds*), *Uncertainty in Artificial Intelligence*, Amsterdam: North-Holland, pp. 167–96.

Heckerman, D. [1988]: 'An Axiomatic Framework for Belief Updates', in J. F. Lemmer and L. N. Kanal (*eds*), *Uncertainty in Artificial Intelligence 2*, Amsterdam: North-Holland, pp. 11–22.

Heckerman, D. and Shortliffe, E. [1992]: 'From Certainty Factors to Belief Networks', *Artificial Intelligence in Medicine*, **4**, pp. 35–52.

Hintikka, J. [1968]: 'The Varieties of Information and Scientific Explanation', in B. van Rootselaar and J. F. Staal (*eds*), *Logic, Methodology, and Philosophy of Science* III: *Proceedings of the 1967 International Congress*, Amsterdam: North-Holland, pp. 151–71.

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

Howson, C. [2000]: *Hume's Problem: Induction and the Justification of Belief*, New York: Oxford University Press.

Huber, F. [2008]: 'Milne's Argument for the Log-Ratio Measure', *Philosophy of Science*, **75**, pp. 413–20.

Kaye, D. H. and Koehler, J. J. [2003]: 'The Misquantification of Probative Value', *Law and Human Behavior*, **27**, pp. 645–59.

Kemeny, J. and Oppenheim, P. [1952]: 'Degrees of Factual Support', *Philosophy of Science*, **19**, pp. 307–24.

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

Kuipers, T. [2000]: *From Instrumentalism to Constructive Realism*, Dordrecht: Reidel.

Lewis, D. [1986]: 'Postscript to *Causation*', in his *Philosophical Papers* II, Oxford: Oxford University Press, pp. 173–213.

Lo, Y., Sides, A., Rozelle, J. and Osherson, D. [2002]: 'Evidential Diversity and Premise Probability in Young Children's Inductive Judgment', *Cognitive Science*, **26**, pp. 181–206.

McGrew, T. [2003]: 'Confirmation, Heuristics, and Explanatory Reasoning', *British Journal for the Philosophy of Science*, **54**, pp. 553–67.

Milne, P. [1996]: 'Log[p(h/eb)/p(h/b)] Is the One True Measure of Confirmation', *Philosophy of Science*, **63**, pp. 21–6.

Minsky, M. and Selfridge, O. G. [1961]: 'Learning in Random Nets', in E. C. Cherry (*ed.*), *Information Theory*, London: Butterworth, pp. 335–47.

Mortimer, H. [1988]: *The Logic of Induction*, Prentice Hall: Paramus.

Nozick, R. [1981]: *Philosophical Explanations*, Oxford: Clarendon Press.

Oaksford, M. and Hahn, U. [2007]: 'Induction, Deduction, and Argument Strength in Human Reasoning and Argumentation', in A. Feeney and E. Heit (*eds*), *Inductive Reasoning: Experimental, Developmental, and Computational Approaches*, New York: Cambridge University Press, pp. 269–301.

Popper, K. R. [1954]: 'Degree of Confirmation', *British Journal for the Philosophy of Science*, **5**, pp. 143–9.

Popper, K. R. [1963]: *Conjectures and Refutations*, London: Routledge & Keegan.

Schupbach, J. N. [2011]: 'New Hope for Shogenji's Coherence Measure', *British Journal for the Philosophy of Science*, **62**, pp. 125–42.

Schupbach, J. N. and Sprenger, J. [2011]: 'The Logic of Explanatory Power', *Philosophy of Science*, **78**, pp. 105–27.

Shogenji, T. [1999]: 'Is Coherence Truth-Conducive?' *Analysis*, **59**, pp. 338–45.

Sober, E. [2008]: *Evidence and Evolution*, New York: Cambridge University Press.

Steel, D. [2003]: 'A Bayesian Way to Make Stopping Rules Matter', *Erkenntnis*, **58**, pp. 213–22.

Steel, D. [2007]: 'Bayesian Confirmation Theory and the Likelihood Principle', *Synthese*, **156**, pp. 55–77.

Suppes, P. [1970]: *A Probabilistic Theory of Causation*, Amsterdam: North-Holland Publishing Company.

Taroni, F., Aitken, C., Garbolino, P. and Biedermann, A. [2006]: *Bayesian Networks and Probabilistic Inference in Forensic Science*, Chichester: Wiley.

Tenenbaum, J. B. and Griffiths, T. L. [2001]: 'The Rational Basis of Representativeness', in J. D. Moore and K. Stennings (*eds*); *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, Hillsdale: Erlbaum, pp. 1036–41.

Tentori, K., Crupi, V. and Osherson, D. [2007]: 'Determinants of Confirmation', *Psychonomic Bulletin & Review*, **14**, pp. 877–83.

Tentori, K., Crupi, V. and Osherson, D. [2010]: 'Second-Order Probability Affects Hypothesis Confirmation', *Psychonomic Bulletin & Review*, **17**, pp. 129–34.

Viale, R. and Osherson, D. [2000]: 'The Diversity Principle and the Little Scientist Hypothesis', *Foundations of Science*, **5**, pp. 239–53.

Watanabe, S. [1969]: *Knowing and Guessing: A Quantitative Study of Inference and Information*, New York: Wiley.