Synthese
DOI 10.1007/s11229-010-9845-9

Contrastive confirmation: some competing accounts

Jake Chandler

Received: 4 August 2010 / Accepted: 20 October 2010

Abstract I outline four competing probabilistic accounts of contrastive evidential support and consider various considerations that might help arbitrate between these. The upshot of the discussion is that the so-called 'Law of Likelihood' is to be preferred to any of the alternatives considered.

Keywords Law of Likelihood · Contrastive confirmation · Bayesianism · Probability

1 Introduction

Our battery of everyday epistemic concepts includes a notion that has so far been comparatively under-discussed in the confirmation-theoretic literature: the concept of what one might call contrastive evidential support, of an item of evidence's *favouring* a hypothesis over one of its specific alternatives. One might claim, for instance, that the pollen traces on the victim's garments support the hypothesis that he spent the summer in Poland over the hypothesis that he spent it in Provence. And of course, one could consistently do so whilst maintaining that the very same evidence fails to discriminate between the victim's having spent the time in Poland and his having spent it in East Germany.

Although literature on this topic is somewhat scarce, a probabilistic analysis of the relation has been on offer since the 1960s. Sometimes known as the 'Law of Likelihood' (LL), the proposal has enjoyed a list of distinguished supporters, including most

J. Chandler (✉)
Center for Logic and Analytic Philosophy, HIW, KU Leuven,
Kardinaal Mercierplein 2, 3000 Leuven, Belgium
e-mail: Jacob.Chandler@hiw.kuleuven.be

Published online: 11 November 2010

notably Hacking (1964, 1965) and Royall (1997). [1] However, it is fair to say that LL has received little by way of justification and, owing to a theoretical environment so far void of competing proposals, has remained subject to very little critical scrutiny.

In an interesting article recently published in this journal, Fitelson (2007) has offered a principle that immediately brings a significant number of competing accounts to the table, calling for an examination of some adjudicating considerations. His suggestion is that the following principle can be used to generate an analysis of the relation of contrastive evidential support from an appropriate measure of degree of (non-contrastive) evidential support:

(†) $H_1 \succ_E H_2$ iff $c(H_1, E) > c(H_2, E)$

Here, '$H_1 \succ_E H_2$' stands in for 'E favours H_1 over H_2' and '$c(H, E)$' denotes the degree of evidential support conferred on H by E. In other words, Fitelson's suggestion is that E favours H_1 over H_2 iff E is stronger evidence for H_1 than it is for H_2.

Now notoriously, there are many ordinally inequivalent proposals regarding $c(H, E)$ (see for instance Fitelson (1999) for relevant references). I shall restrict my attention here to what are arguably the six most popular:[2]

(d) $c(H, E) = \Pr(H \mid E) - \Pr(H)$
(s) $c(H, E) = \Pr(H \mid E) - \Pr(H \mid \overline{E})$
(c) $c(H, E) = \Pr(H \cap E) - \Pr(H)P(E)$
(n) $c(H, E) = \Pr(E \mid H) - \Pr(E \mid \overline{H})$
(l) $c(H, E) = \log\left[\frac{\Pr(E|H)}{\Pr(E|\overline{H})}\right]$
(r) $c(H, E) = \log\left[\frac{\Pr(H|E)}{\Pr(H)}\right]$

Following Fitelson, let us denote the corresponding accounts of contrastive confirmation, obtained via (†), by $(\dagger_d)$, $(\dagger_s)$, etc. We can immediately note the following:

Observation 1 (i) $(\dagger_s)$, $(\dagger_c)$ and $(\dagger_d)$ are equivalent. (ii) $(\dagger_d)$, $(\dagger_n)$, $(\dagger_l)$ and $(\dagger_r)$ are pairwise inequivalent.

We are therefore left with the following distinct options:

$(\dagger_d)$ $H_1 \succ_E H_2$ iff $\Pr(H_1 \mid E) - \Pr(H_1) > \Pr(H_2 \mid E) - \Pr(H_2)$
$(\dagger_n)$ $H_1 \succ_E H_2$ iff $\Pr(E \mid H_1) - \Pr(E \mid \overline{H_1}) > \Pr(E \mid H_2) - \Pr(E \mid \overline{H_2})$
$(\dagger_l)$ $H_1 \succ_E H_2$ iff $\Pr(E \mid H_1)\Pr(E \mid \overline{H_2}) > \Pr(E \mid H_2)\Pr(E \mid \overline{H_1})$
$(\dagger_r)$ $H_1 \succ_E H_2$ iff $\Pr(E \mid H_1) > \Pr(E \mid H_2)$

The last of these, $(\dagger_r)$, is a familiar figure, being none other than LL. To the best of my knowledge, the remaining accounts have yet to be endorsed, with the possible exception of $(\dagger_l)$, tentatively defended by Fitelson (2003, 2007). All four suggestions share

[1] Note that Royall also endorses a *quantitative* analogue of LL, a measure of the degree to which an item of evidence favours one hypothesis over another. Forster and Sober (2004) have suggested that this measure was also endorsed by Hacking. This however would seem to be incorrect, as Hacking himself has informed me in correspondence. I shall not be discussing the issue of degree of contrastive support in what follows.

[2] In what follows, it will be assumed that E, H_1 and H_2 have *non-extremal probabilities*. This requirement is typically simply imposed to insure that the relevant conditional probabilities are well defined.

some attractive features. They all yield the intuitive result that $\succ_E$ is both transitive and asymmetric. They also turn out to be alike in validating an attractive principle, once adduced by Royall (1997) in support of ($\dagger_r$), according to which, if the first hypothesis is conclusive evidence for the data and the second is conclusive evidence for the negation of the data, then the data supports the former over the latter:

Observation 2 Each of ($\dagger_d$) to ($\dagger_r$) entails (P1).

where

(P1) If (i) $\Pr(E \mid H_1) = 1$ and (ii) $\Pr(E \mid H_2) = 0$, then $H_1 \succ_E H_2$. [3]

In what follows, I shall attempt to evaluate the comparative merits and lacunae of these proposals. I shall however steer clear of both (i) the general plausibility of (†) and (ii) the comparative merits of the various measures of degree of evidential support (d)–(r). Instead I will *directly* focus on the issue of how ($\dagger_d$)–($\dagger_r$) fare with respect to various desiderata for an account of contrastive evidential support. I will however briefly return to issues (i) and (ii) in the final section. Proofs of the various observations made are collected in the Appendix.

2 Mutual exclusiveness and Fitelson's case against ($\dagger_r$)

In (Fitelson 2007), we are asked to consider the following scenario:

Cards 1: observe that a spade has been drawn from a standard pack of cards (E). The two hypotheses under consideration are the hypothesis that the card is the ace of spades (H_1) and the hypothesis that the card is black (H_2).

Fitelson argues that we have a clear intuition here that E favours H_2 over H_1. If he is right, we have grounds to rule out a number of our candidates. Indeed, whilst ($\dagger_d$) and ($\dagger_l$) both yield the result in question, ($\dagger_r$) and ($\dagger_n$) do not, yielding the result that E favours H_1 over H_2. See Case 1 of Table 1 for a summary of these facts alongside the relevant probability function.

Fitelson puts our alleged intuitions concerning this case down to an application of a general "highly intuitive" principle (called '(*)' in the original article), according to which "If E provides conclusive evidence for H_1, but non-conclusive evidence for H_2…then E favours H_1 over H_2". In formal terms, the principle appears to be understood as follows:

(P2) If (i) $\Pr(H_1 \mid E) = 1$ and (ii) $\Pr(H_2 \mid E) < 1$, then $H_1 \succ_E H_2$.

But now, as Fitelson remarks (2007, p. 488, fn. 25):

Observation 3 (i) ($\dagger_l$) satisfies (P2) but (ii) ($\dagger_r$), ($\dagger_d$) and ($\dagger_n$) do not.[4]

[3] In response to Royall's argument for ($\dagger_r$) from (P1), Forster and Sober (2004) speculated that various competing accounts of contrastive support might also have this principle as a consequence. This result confirms their suspicions.

[4] Fitelson does not in fact discuss ($\dagger_n$) in his paper, restricting his attention to ($\dagger_r$), ($\dagger_d$) and ($\dagger_l$). He also omits the proof, which is included here in the appendix.

Table 1 Probability models and corresponding verdicts on contrastive support

Propositions	Probability functions						
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
$E \cap H_1 \cap H_2$	1/52	0	0	0	1/5	0	0
$E \cap H_1 \cap \overline{H_2}$	0	1/52	1/100	4/31	1/5	3/100	3/10
$E \cap \overline{H_1} \cap H_2$	12/52	2/52	0	9/28	0	2/100	2/10
$E \cap \overline{H_1} \cap \overline{H_2}$	0	1/52	8/100	5/27	0	0	0
$\overline{E} \cap H_1 \cap H_2$	0	0	0	0	5/100	0	0
$\overline{E} \cap H_1 \cap \overline{H_2}$	0	0	9/10	1/810	5/10	3/100	3/10
$\overline{E} \cap \overline{H_1} \cap H_2$	1/4	0	1/100	2/21	0	2/100	2/10
$\overline{E} \cap \overline{H_1} \cap \overline{H_2}$	1/2	48/52	0	13453/50220	5/100	9/10	0
Account	**Verdicts on contrastive support**						
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
$(\dagger_d)$	$\prec_E$	$\prec_E$	$\prec_E$	$\prec_E$	$\prec_E$	$\succ_E$	$\prec_E$
$(\dagger_n)$	$\succ_E$	$\prec_E$	$\prec_E$	$\succ_E$	$\prec_E$	$\succ_E$	$\prec_E$
$(\dagger_l)$	$\prec_E$	$\prec_E$	$\succ_E$	$\succ_E$	$\succ_E$	$\succ_E$	$\prec_E$
$(\dagger_r)$	$\succ_E$	$\sim_E$	$\succ_E$	$\prec_E$	$\prec_E$	$\prec_E$	$\prec_E$

If Fitelson is right about the plausibility of (P2), it would appear that we have a consideration that quite remarkably favours $(\dagger_l)$ over *all* competitors considered here.

But something is amiss. Indeed, the claim that a favouring relation obtains between E, H_1 and H_2 plausibly presupposes that H_1 and H_2 are *competing* hypotheses, at least in the weak sense that the probability of their intersection is nil.[5] And indeed, this is the very reason why, Fitelson's claims notwithstanding, it is far from clear that it even makes any sense, with respect to Cards 1, to ask oneself whether or not the fact that the card drawn is a spade favours the hypothesis that it was an ace of spades over the hypothesis that it was a black card. That the card was an ace of spades and that the card was black are simply not alternative hypotheses. Cards 1 thereby fails to provide a suitable test case with which to assess the proposals on offer. Furthermore, in the presence of the assumption that H_1 and H_2 are mutually exclusive in the sense specified, (P2) fails to have any relevant probative force, since:

Observation 4 On the assumption that $\Pr(H_1 \cap H_2) = 0$, $(\dagger_l)$, $(\dagger_r)$, $(\dagger_d)$ and $(\dagger_n)$ all satisfy (P2).

The requirement that $\Pr(H_1 \cap H_2) = 0$, which incidentally, was satisfied by the countermodels used to establish the inequivalences reported in Observation 1, will be tacitly assumed throughout the remainder of this paper. Before moving on however,

[5] See (Chandler 2007). It is worth noting that a similar diagnosis has been given, in the context of discussions of the notions of contrastive *explanation*, of the unintelligibility of utterances such as 'That'll explain why he left on Sunday rather than on Sunday morning'. There is of course a certain amount of controversy here. Whilst the mutual exclusiveness requirement is endorsed by the likes of Garfinkel (1981), Ruben (1987) and Temple (1988), Lipton (1990), for instance, claims that this requirement is too strong, adducing an alleged counterexample. In (Chandler 2007), I argued that Lipton's argument fails.

two brief comments regarding this assumption are in order here, the first to dispel a potential worry that it might raise, the second to report an equivalence that it generates.

First of all, as a number of people have reminded me, Sober and Forster (2004) claim that, in the context of discussions of model selection, scientists sometimes say of an item of evidence that it 'favours' a given model over a logically weaker one (where a model is the—typically infinite—disjunction of a number of competing hypotheses regarding the probability distribution of some random variable). Although this practice may prima facie seem to be at odds with what I have just claimed, the tension vanishes upon closer inspection. Indeed it turns out that, as Sober and Forster use the term, 'E favours model M_1 over model M_2' is actually shorthand for 'E favours the likeliest (in the technical sense) disjunct $L(M_1)$ of model M_1 over the likeliest disjunct $L(M_2)$ of model M_2', with $L(M_1) \cap L(M_2) = \varnothing$.

Secondly, in (Chandler 2007), borrowing from Hitchcock's account of contrastive explanation (Hitchcock 1996, 1999), I suggested the following:

(CB) $H_1 \succ_E H_2$ iff (1) $\Pr(H_1 \mid H_2) = 0$ and (2) $\Pr(H_1 \mid E \cap (H_1 \cup H_2)) > \Pr(H_1 \mid \overline{E} \cap (H_1 \cup H_2))$.

This analysis was then put to work in providing a contrastivist resolution of the so called Tacking Problem for Bayesian confirmation theory. Interestingly, however, Fitelson has pointed out to me that the following noteworthy result can easily be established:

Observation 5 On the assumption that $\Pr(H_1 \cap H_2) = 0$, (CB) is equivalent to ($\dagger_r$).[6]

3 Leeds' case

In an unpublished paper, Leeds (ms) argues that ($\dagger_r$) yields some counter-intuitive judgments. He offers something analogous to the following scenario:

Cards 2 I observe that an ace has been drawn from a standard pack of cards (E). The two hypotheses under consideration are the hypothesis (H_1) that the card is the ace of hearts and the hypothesis that (H_2) the card is either the ace of spades or the ace of clubs.

[6] It is interesting to note that there would be a straightforward way of adapting (CB) to provide an analysis of a *quaternary* evidential relation between two competing items of evidence and two competing hypotheses ('Its being the case that E_1 rather than E_2 provides support for the hypothesis that H_1 over the hypothesis that H_2'). Indeed, the obvious analyses would be

(1) $\Pr(H_1 \mid H_2) = 0$, (2) $\Pr(E_1 \mid E_2) = 0$ and (3) $\Pr(H_1 \mid E_1 \cap (H_1 \cup H_2) \cap (E_1 \cup E_2)) = \Pr(H_1 \mid E_1 \cap (H_1 \cup H_2)) > \Pr(H_1 \mid E_2 \cap (H_1 \cup H_2) \cap (E_1 \cup E_2)) = \Pr(H_1 \mid E_2 \cap (H_1 \cup H_2))$

Interestingly, given (1) and (2), (3) is equivalent to

$$\frac{\Pr(H_1 \cap E_1)\Pr(H_2 \cap E_2)}{\Pr(H_1 \cap E_2)\Pr(H_2 \cap E_1)} > 0$$

The quantity on the left hand side of this last inequality is of course rather reminiscent of the well known odds ratio measure of correlation between binary variables. Note that by setting $E_2 = \overline{E_1}$, we would recover LL as a special case.

See Case 2 of Table 1 for the relevant probability distribution. Now it turns out that the following holds:

Observation 6 In Cards 2, (i) $H_1 \succ_E H_2$ according to $(\dagger_d)$, $(\dagger_n)$ and $(\dagger_l)$, (ii) $H_1 \sim_E H_2$ according to $(\dagger_r)$.[7]

This is an important test case. In contrast to the situation in Cards 1, the hypotheses here are genuine alternatives: the mutual exclusiveness condition is satisfied. If our intuitions square with the verdict yielded by $(\dagger_r)$, and we take those intuitions seriously, we have in effect eliminated a huge amount of competition. According to Leeds, however, our intuitions do not square with $(\dagger_r)$'s verdict: he maintains that, in this situation, intuitions dictate that E favours H_2 over H_1. And Sober (2005), an erstwhile defender of $(\dagger_r)$, appears to agree: Cards 2 speaks against $(\dagger_r)$.

I find these alleged intuitions quite puzzling: it seems quite clear to me that $(\dagger_r)$ gets things *right* here and E does *not* discriminate between H_1 and H_2. And indeed, in his discussion of Cards 2, Fitelson concedes that some will be unlikely to 'be swayed by such examples' and will reject Leeds' and Sober's intuitions here. But if Leeds and Sober are wrong, what exactly is it that drives their claim? As Fitelson notes, Sober appears to make an appeal to the relative values of posterior probabilities of H_1 and H_2, following up his claim that $H_2 \succ_E H_1$ with the observation that $\Pr(H_2 \mid E) > \Pr(H_1 \mid E)$. This would suggest an endorsement of an account of favouring according to which $H_1 \succ_E H_2$ iff $\Pr(H_1 \mid E) > \Pr(H_2 \mid E)$. But this is an unpromising suggestion, since, as Fitelson points out, it counterintuitively allows for E to favour H_1 over H_2 even when E simultaneously raises the probability of H_2 and lowers the probability of H_1, a situation precluded by the other accounts discussed here.

4 Behaviour under conditionalisation

Suppose that you find out that James prefers pasta to pizza. Given your background knowledge, this information plausibly favours James' ordering pasta at the restaurant tonight over his ordering pizza (I am assuming that James will muster enough self-restraint to order only one dish). Now suppose you now find out both that James will definitely order a main dish this evening and that the restaurant will only be serving pasta or pizza today, due to a shortage of risotto rice. In other words, you find out that James will be ordering either pasta or pizza. Given this new data, does the information that you have regarding James' culinary preferences still favour his ordering pasta over his ordering pizza? The answer is clearly that it does. This suggests the following attractive principle:

(P3) If $H_1 \succ_E H_2$, then $H_1 \succ_{E|H_1 \cup H_2} H_2$.

In other words: if an item of evidence E favours a hypothesis H_1 over a rival hypothesis H_2, then this would remain the case were one to find out that one or another of H_1 or H_2 is true, assuming that one updates on new knowledge by strict conditionalisation. I take this to be intuitively compelling.

[7] Where $H_1 \sim_E H_2$ iff neither $H_1 \succ_E H_2$ nor $H_2 \succ_E H_1$.

In fact, it would seem that (P3) can be strengthened: the direction of entailment would intuitively appear to go both ways. We should endorse the following principle, which is somewhat reminiscent of the so-called 'Context-Free Ordering Assumption' in the preference literature (see McLennen 1990, p. 29):

(P3*) $H_1 \succ_E H_2$ iff $H_1 \succ_{E|H_1 \cup H_2} H_2$.

But now it turns out that that the following is true, strongly favouring ($\dagger_r$) over its competitors:

Observation 7 (i) Both (P3) and (P3*) are true according to ($\dagger_r$) and (ii) neither (P3) nor (P3*) are true according to ($\dagger_d$), ($\dagger_n$) or ($\dagger_l$).

5 Decisional tie-breaking

There is a dog race coming up this Sunday and you are offered a choice between two dogs to put your money on: Slip the Lark and Kungfu Panda. The odds of the bets are such that, assuming that your preferences go by expected monetary gain, you are indifferent between the two options presented. I now offer you information that favours the hypothesis of a win for Slip over the hypothesis of a win for Panda, say the information that Sunday will be a rainy day and that Slip has a preference for a muddy racecourse, whilst Panda does not. It seems clear that your indifference ought to give way to a strict preference for the bet on Slip. It also seems clear that your indifference would have persisted had the information given failed to favour any one of the dogs over the other.

More generally, the following would appear to hold true. Let us denote by $\langle H, Q, k\rangle$ a bet on H such that the bettor agree to forfeit a quantity Q of money ($Q \geq 0$) in the event that H is false in exchange for kQ ($k \geq 0$) in the event that H is true. We can now offer the following: If (i) the utility of money is linear and rational choice goes by relative expected utility, (ii) one is indifferent between a bet $B_1 = \langle H_1, Q, k_1\rangle$ and a bet $B_2 = \langle H_2, Q, k_2\rangle$, and (iii) E is irrelevant to the utilities of the hypothesis/act pairs, then upon learning E, (iv) one should strictly prefer the bet on H_1 to the bet on H_2 iff (v) E favours H_1 over H_2. More formally, where $1 \leq i \leq 2$, $1 \leq j \leq 3$, H_3 denotes $\overline{H_1 \cup H_2}$ and $EU_E(B_i)$ stands for the expected utility of act B_i after conditionalising on E:

(P4) If (i) $B_1 = \langle H_1, Q, k_1\rangle$ and $B_2 = \langle H_2, Q, k_2\rangle$ with $Q, k \geq 0$, (ii) $EU(B_1) = EU(B_2)$, and (iii) $U(B_i \cap H_j \cap E) = U(B_i \cap H_j)$ then (iv) $EU_E(B_1) > EU_E(B_2)$ iff (v) $H_1 \succ_E H_2$.

It is however easy to show that the following holds:

Observation 8 (P4) is true iff ($\dagger_r$) is so too.

This again provides what I take to be a strong consideration weighing in favour of ($\dagger_r$), this time not only to the detriment of ($\dagger_d$), ($\dagger_n$) and ($\dagger_l$), but to that of any further alternative option that one may wish to envisage.

6 Concluding comments

I have argued, in Sects. 2 and 3, that existing considerations purporting to arbitrate between some of the proposals discussed here are inconclusive at best. On the heels of this, I have also offered, in Sects. 4 and 5, two new arguments favouring $(\dagger_r)$, aka the Law of Likelihood, over the alternatives considered.

This raises the immediate question of whether or not the present discussion has any significant repercussion on the vexed 'measures of confirmation' debate. Indeed, although, in the presence of (†), the correctness of $(\dagger_r)$ would fail to single out a *unique* measure of confirmation (since (†) generates a mapping from measures of confirmation to analyses of contrastive confirmation that is, as we have seen in Sect. 1, many-to-one), it would rule out a number of the most popular alternatives. All of this, however, hinges on the status of (†). Unfortunately, however, in contrast with (Fitelson 2007), I must admit that I fail to see the intuitive appeal of the principle. But then again, neither did Fitelson's former self, as he has himself recently pointed out to me. Indeed, in his earlier PhD dissertation, he appeared to view (†) with a certain amount of suspicion, arguing that its truth is 'far from obvious' (Fitelson 2001, p. 29).

Of course, this is not to say that (†) is *false*. It may indeed turn out to be true, but, in view of the considerations adduced in the present paper, this matter will hinge on the fate of measure (r). To co-opt a remark made by Fitelson in relation to the analysis of the favouring relation (Fitelson 2007, fn. 22), the battle for (†) is one that now needs to be fought at the level of nonrelational confirmation.

Acknowledgments I am grateful to Luc Bovens, Richard Bradley, Igor Douven, Chris Kelp and Adam Rieger for useful feedback on various parts of this paper, as well as to the CPNSS, London School of Economics, and the Formal Epistemology Project, KU Leuven, for providing the stimulating research environments in which the article was written. Special credit is due to Branden Fitelson, for both his engaging and helpful correspondence, and the invaluable PrSAT Mathematica package (written in collaboration with Jason Alexander and Ben Blum; see Fitelson (2008) for further discussion) used to generate the models in Table 1 of the Appendix.

Appendix

Proof of Observation 1 (i) Conjoining (c) and (†) gives us

$(\dagger_c)$: $H_1 \succ_E H_2$ iff $\Pr(H_1 \cap E) - \Pr(H_1)\Pr(E) > \Pr(H_1 \cap E) - \Pr(H_1)\Pr(E)$.

Dividing both sides by $\Pr(E)$ immediately yields $(\dagger_d)$. Furthermore, conjoining (s) and (†) gives us

$(\dagger_s)$: $H_1 \succ_E H_2$ iff $\Pr(H_1 \mid E) - \Pr(H_1 \mid \overline{E}) > \Pr(H_2 \mid E) - \Pr(H_2 \mid \overline{E})$.

By the definition of conditional probability, $\Pr(H_1 \mid E) - \Pr(H_1 \mid \overline{E}) > \Pr(H_2 \mid E) - \Pr(H_2 \mid \overline{E})$ iff $\Pr(H_1 \mid E) - \Pr(H_1 \cap \overline{E})/\Pr(\overline{E}) > \Pr(H_2 \mid E) - \Pr(H_2 \cap \overline{E})/\Pr(\overline{E})$. By elementary algebra and the law of total probability, the RHS is equivalent to $[\Pr(H_1 \mid E)(1-\Pr(E)) - \Pr(H_1 \cap \overline{E})]/\Pr(\overline{E}) > [\Pr(H_2 \mid E)(1-\Pr(E)) - \Pr(H_2 \cap \overline{E})]/\Pr(\overline{E})$. By the definition of conditional probability, this is itself equivalent to $\Pr(H_1 \mid E) - \Pr(H_1 \cap E) - \Pr(H_1 \cap \overline{E}) > \Pr(H_2 \mid E) - \Pr(H_2 \cap E) - \Pr(H_2 \cap \overline{E})$, and hence, by the law of total probability, to $\Pr(H_1 \mid E) - \Pr(H_1) > \Pr(H_2 \mid E) - \Pr(H_2)$. □

Proof of Observation 1 (ii) See Cases 3 and 4 of Table 1. □

Proof of Observation 2 It is well known that each of (d) to (r) are such that

$$c(H, E) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0 \quad \text{iff}\, \Pr(H \mid E) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} \Pr(H).$$

In order to obtain the desired result, we can therefore simply establish that, given the antecedent of (P1), it follows that $\Pr(H_1 \mid E) > \Pr(H_1)$ and $\Pr(H_2 \mid E) \leq \Pr(H_2)$. Assuming that $\Pr(E \mid H_1) = 1$, it follows that $\Pr(H_1 \mid \overline{E}) = 0$. Now, by the law of total probability and the definition of conditional probability, $\Pr(H_1) = \Pr(H_1 \mid E) \Pr(E) + \Pr(H_1 \mid \overline{E}) \Pr(\overline{E})$, so assuming furthermore that $1 > \Pr(E) > 0$, it then also follows that $\Pr(H_1 \mid E) > \Pr(H_1)$. Assuming that $\Pr(E \mid H_2) = 0$, it follows that $\Pr(H_2 \mid E) = 0$. So finally, assuming that $\Pr(H_2) > 0$, we have $\Pr(H_2 \mid E) < \Pr(H_2)$. □

Proof of Observation 3 (i) Assume that $\Pr(H_1 \mid E) = 1$. Since, as it is easy to demonstrate, $\Pr(H_1 \mid E) = 1 - [(\Pr(E \mid \overline{H_1}) \Pr(\overline{H_1}))/ \Pr(E)]$, it follows that $\Pr(E \mid \overline{H_1}) = 0$. It also follows, from the same assumption, that $\Pr(E \cap H_1) > 0$, and hence that $\Pr(E \mid H_1) > 0$. Therefore $\Pr(E \mid H_1) \Pr(E \mid \overline{H_2}) > \Pr(E \mid H_2) \Pr(E \mid \overline{H_1})$ iff $\Pr(E \mid \overline{H_2}) > 0$. In other words, according to ($\dagger_l$), $H_1 \succ_E H_2$ iff $\Pr(E \mid \overline{H_2}) > 0$. Assume furthermore that $\Pr(H_2 \mid E) < 1$. It quickly follows that $\Pr(E) > \Pr(E \cap \overline{H_2})$. Since, as it is easy to demonstrate, $\Pr(E \mid \overline{H_2}) = (\Pr(E) - \Pr(E \cap \overline{H_2}))/ \Pr(\overline{H_2})$, it follows, in turn, that $\Pr(E \mid \overline{H_2}) > 0$ and hence that $H_1 \succ_E H_2$ according to ($\dagger_l$). □

Proof of Observation 3 (ii) For ($\dagger_r$) and ($\dagger_n$), see Case 1 of Table 1 (i.e. Cards 1), where according to both accounts, and contrary to (P1*), $H_1 \succ_E H_2$. For ($\dagger_d$), see Case 5 of the same table, where, according to this proposal and again contrary to (P1*), $H_1 \succ_E H_2$. □

Proof of Observation 4 It follows from the opening remarks made regarding (d) to (r) in the proof of Observation 2 above, that if $\Pr(H_1 \mid E) > \Pr(H_1)$ and $\Pr(H_2 \mid E) < \Pr(H_2)$, then $H_1 \succ_E H_2$ according to each of ($\dagger_d$) to ($\dagger_r$). Now, provided $\Pr(H_1 \mid E) = 1$ and $\Pr(H_1) < 1$, it is obviously the case that $\Pr(H_1 \mid E) > \Pr(H_1)$. Provided furthermore that $\Pr(H_1 \cap H_2) = 0$ and $\Pr(H_2) > 0$, then it is also the case that $\Pr(H_2 \mid E) = 0 < \Pr(H_2)$. □

Proof of Observation 5 By the definition of conditional probability, $\Pr(E \mid H_1) > \Pr(E \mid H_2)$ is equivalent to $\Pr(E \cap H_1)/ \Pr(H_1) > \Pr(E \cap H_2)/ \Pr(H_2)$. Providing $\Pr(H_1 \cap H_2) = 0$, this is equivalent to $\Pr(E \cap H_1 \cap \overline{H_2})/((\Pr(E \cap H_1 \cap \overline{H_2}) + \Pr(\overline{E} \cap H_1 \cap \overline{H_2})) > \Pr(E \cap \overline{H_1} \cap H_2)/((\Pr(E \cap \overline{H_1} \cap H_2) + \Pr(\overline{E} \cap \overline{H_1} \cap H_2))$, that is: $\Pr(E \cap H_1 \cap \overline{H_2})/((\Pr(E \cap H_1 \cap \overline{H_2}) + \Pr(E \cap \overline{H_1} \cap H_2)) > \Pr(\overline{E} \cap H_1 \cap \overline{H_2})/((\Pr(\overline{E} \cap H_1 \cap \overline{H_2}) + \Pr(\overline{E} \cap \overline{H_1} \cap H_2))$. Again, providing $\Pr(H_1 \cap H_2) = 0$, this is equivalent to $\Pr(E \cap H_1)/ \Pr(E \cap (H_1 \cup H_2)) > \Pr(\overline{E} \cap H_1)/ \Pr(\overline{E} \cap (H_1 \cup H_2)$, in other words, again by the definition of conditional probability: $\Pr(H_1 \mid E \cap (H_1 \cup H_2)) > \Pr(H_1 \mid \overline{E} \cap (H_1 \cup H_2))$. □

Proof of Observation 6 See Case 2 of Table 1. □

Proof of Observation 7(i) Trivial, since $\Pr(E \mid H_1) = \Pr(E \mid H_1 \cap (H_1 \cup H_2))$ and $\Pr(E \mid H_2) = \Pr(E \mid H_2 \cap (H_1 \cup H_2))$. □

Proof of Observation 7(i) See Cases 6 and 7 of Table 1, where Case 7 is obtained from Case 6 by strict conditionalisation on $H_1 \cup H_2$. The proposition favoured changes from one case to the other on all accounts but $(\dagger_r)$. □

Proof of Observation 8 Assume $EU(B_1) = EU(B_2)$, i.e. $\Pr(H_1)k_1Q - \Pr(H_2)Q - \Pr(H_3)Q = \Pr(H_2)k_2Q - \Pr(H_1)Q - \Pr(H_3)Q$. Rearranging things a little, we can see that this is equivalent to $\Pr(H_1)/\Pr(H_2) = (k_2Q + Q)/(k_1Q + Q)$. Since $k_1, Q \geq 0$, it follows that $k_1Q + Q \geq 0$. Given this and the fact that $U(B_i \cap H_j \cap E) = U(B_i \cap H_j)$ for $1 \leq i, j \leq 2$, we similarly find that $EU_E(B_1) > EU_E(B_2)$ iff $\Pr(H_1 \mid E)/\Pr(H_2 \mid E) > (k_2Q + Q)/(k_1Q + Q)$. We can now see that $EU_E(B_1) > EU_E(B_2)$ iff $\Pr(H_1 \mid E)/\Pr(H_2 \mid E) > \Pr(H_1)/\Pr(H_2)$, i.e. iff $\Pr(E \mid H_1) > \Pr(E \mid H_2)$, i.e. iff E favours H_1 over H_2 according to $(\dagger_r)$. □

References

Chandler, J. (2007). Solving the tacking problem with contrast classes. *British Journal for the Philosophy of Science, 58*(3), 489–502.
Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science, 66*, 362–378.
Fitelson, B. (2001). *Studies in Bayesian confirmation theory*. PhD Dissertation, University of Wisconsin, Madison, WI.
Fitelson, B. (2003). Review of Richard Swinburne (Ed.) *Bayes's theorem. Notre Dame philosophical reviews*.
Fitelson, B. (2007). Likelihoodism, Bayesianism, and relational confirmation. *Synthese, 156*(3), 473–489.
Fitelson, B. (2008). A decision procedure for probability calculus with applications. *The Review of Symbolic Logic, 1*, 111–125.
Forster, M. R., & Sober, E. (2004). Why likelihood. In M. Taper & S. Lele (Eds.), *The nature of scientific evidence.* (pp. 153–165). Chicago, IL: Chicago University Press.
Garfinkel, A. (1981). *Forms of explanation: Rethinking the questions in social theory*. New Haven, CT: Yale University Press.
Hacking, I. (1964). On the foundations of statistics. *British Journal for the Philosophy of Science, 15*(57), 1–26.
Hacking, I. (1965). *Logic of statistical inference*. Cambridge: Cambridge University Press.
Hitchcock, C. R. (1996). The role of contrast in causal and explanatory claims. *Synthese, 107*, 395–419.
Hitchcock, C. R. (1999). Contrastive explanation and the demons of determinism. *British Journal for the Philosophy of Science, 50*, 585–612.
Leeds, D. (ms). *Other minds, support and likelihoods*.
Lipton, P. (1990). Contrastive explanation. In D. Knowles (Ed.), *Explanation and its limits.* (pp. 247–266). Cambridge: Cambridge University Press.
McLennen, E. F. (1990). *Rationality and dynamic choice*. Cambridge: Cambridge University Press.
Royall, R. (1997). *Statistical evidence: A likelihood paradigm*. New York: Chapman and Hall.
Ruben, D. (1987). Explaining contrastive facts. *Analysis, 47*, 35–37.
Sober, E. (2005). Is drift a serious alternative to natural selection as an explanation of complex adaptive traits? In A. O'Hear (Ed.), *Philosophy, biology and life.* Cambridge: Cambridge University Press
Temple, D. (1988). Discussion: The contrast theory of why-questions. *Philosophy of Science, 55*, 141–151.