CONTRASTIVE BAYESIANISM

BRANDEN FITELSON

ABSTRACT. Bayesianism provides a rich theoretical framework, which lends itself rather naturally to the explication of various "contrastive" and "non-contrastive" concepts. In this (brief) discussion, I will focus on issues involving "contrastivism", as they arise in some of the recent philosophy of science, epistemology, and cognitive science literature surrounding Bayesian confirmation theory.

1. What *is* "Bayesianism"?

I.J. Good [19] once estimated that there are 46,656 varieties of Bayesianism. He based his estimate on a number of "dimensions" along which different sorts of Bayesianism could be characterized. From the perspective of this volume, Good's is probably an *under*-estimate, since he was talking mainly about applications of "Bayesianism" to problems involving statistical inference. In contemporary analytic philosophy, there are still further "dimensions" along which (many) additional "Bayesianisms" might be distinguished. What all "Bayesianisms" have in common is that they all make essential use of *probability* as their main theoretical tool. For the most part, disagreements among different kinds of "Bayesians" will involve differing *interpretations* of probability. There are many interpretations of probability in the philosophical universe (see [22] for an excellent survey). In this article, I will try to remain as neutral as possible on the various interpretive disputes that arise among the myriad of "Bayesians" one encounters in the philosophical literature. Of course, certain applications will most naturally be associated with certain kinds of "Bayesianism" (in contrast with others). But, I will not dwell on such differences, unless they are essential to the "contrastivist" character of the accounts in question. Because my focus will be rather narrow (I'll be focusing on issues that arise in recent applications of Bayesian confirmation theory), I will be able to sidestep many (but, as we will see, not quite all) of these intramural interpretive disputes.

2. What *is* "Contrastivism"?

This volume is about "contrastivism". So, it is natural to wonder what distinguishes "contrastive" philosophical approaches or accounts from "non-contrastive" ones. I won't attempt a demarcation. Indeed, I won't even say very much (in any *general* way) about this distinction. Instead, I will try to illustrate how "contrastivist" thinking arises in some recent applications of "Bayesian" techniques.

Date: 07/14/10. Draft: final version to appear in *Contrastivism in Philosophy* (M. Blaauw *ed.*).

Key words and phrases. Bayesianism, probability, confirmation, contrastivism, likelihoodism, evidence, support, formal epistemology, conjunction fallacy, irrelevant conjunction.

Thanks to Fabrizio Cariani, Jake Chandler, Vincenzo Crupi, Kenny Easwaran, Hartry Field, Clark Glymour, Jim Hawthorne, Chris Hitchcock, Jim Joyce, Jim Pryor, Susanna Rinard, Teddy Seidenfeld, Elliott Sober, Dan Steel, Katya Tentori, and Mike Titelbaum for useful discussions.

Hopefully, this will give some sense of how "Bayesian philosophers" (broadly construed) think about "contrastivism" and its philosophical significance. To this end, I will examine several recent case studies from the contemporary literature on Bayesian confirmation theory, which, as we shall see, is implicated in Bayesian philosophy of science, Bayesian epistemology, and Bayesian cognitive science.

3. Likelihoodism, Bayesianism, and Contrastive Confirmation

It is useful to begin with a discussion of a prominent "contrastivist" probabilistic account that has appeared in contemporary philosophy of science. This will simultaneously set the theoretical stage for subsequent sections, and illustrate a concrete example of "contrastivism" in (broadly) Bayesian philosophy of science.

Elliott Sober has been defending what he calls "contrastive empiricism" (CE) for over 20 years. In his original statement and defense of (CE), Sober [44] explains:

> ...theory testing is a contrastive activity. If you want to test a theory T, you must specify a range of alternatives — you must say what you want to test T *against*. There is a trivial reading of this thesis that I do not intend. To find out if T is plausible is simply to find out if T is more plausible than *not*-T. I have something more in mind: there are various contrasting alternatives that might be considered. If T is to be tested against T', one set of observations [E] may be needed, but if T is to be tested against T'' a different set of observations [E'] may be needed. By varying the contrasting alternatives, we formulate genuinely different testing problems.

Here, Sober intends to be contrasting his view of the testing of scientific theories [(CE)] with what he takes to be its main philosophical rival: *Bayesian confirmation theory* (BCT). As Carnap [4] explains, there are two distinct probabilistic notions of "support" or "confirmation" that "Bayesians" (broadly construed) must be careful to distinguish (I will be returning to this crucial distinction in subsequent sections):

- **Confirmation as firmness** (confirms$_f$). E confirms$_f$ H (relative to background corpus K) iff $\Pr(H \mid E \,\&\, K) > t$, where t is some (possibly contextually determined) threshold value. And, the *degree* to which E confirms$_f$ H, relative to background corpus K [$\mathfrak{c}_f(H, E \mid K)$] is given by $\Pr(H \mid E \,\&\, K)$.
- **Confirmation as increase in firmness** (confirms$_i$). E confirms$_i$ H (relative to background corpus K) iff $\Pr(H \mid E \,\&\, K) > \Pr(H \mid K)$ And, the *degree* to which E confirms$_i$ H, relative to background corpus K [$\mathfrak{c}_i(H, E \mid K)$] is given by some function [$\mathfrak{c}_i$] of $\Pr(H \mid E \,\&\, K)$ and $\Pr(H \mid K)$, where (informally) $\mathfrak{c}_i$ is some measure of "the degree to which evidence E increases the firmness/probability of H (relative to background corpus K)".[1]

[1]Here, we are assuming that all bodies of evidence are *propositional*. This is a typical assumption made by Bayesian epistemologists and philosophers of science. It is not universally accepted (in epistemology generally) that all evidence is propositional. But, this is not an unpopular view either. See [52, ch. 9] and [35] for recent discussions. Also, we're understanding conditional probability as "probability *on an indicative supposition*". This is pretty standard in the present context. See [28] for more on this "probability on an indicative supposition" conception, and how it differs from a *subjunctive*-suppositional conception of $\Pr(\cdot \mid \cdot)$, which may be more appropriate for causal or explanatory applications. Finally, following [21], we assume that *all* probabilities are *inherently conditional* in nature. When we talk about the "prior probability" $\Pr(p)$ of a proposition p, this is really just shorthand for a conditional probability $\Pr(p \mid K)$, relative to some unspecified background corpus K.

Intuitively, $\mathfrak{c}_f(H, E)$ is a measure of *how probable H is, on the supposition of evidence E*.[2] This is often called the *posterior* probability of H *on evidence* E. But, $\mathfrak{c}_i(H, E)$ is something different. It's meant to be a measure of the degree to which E is *evidentially relevant* to H, where this is explicated in terms of (some measure of) the degree to which E is *probabilistically relevant* to H. To gauge $\mathfrak{c}_i(H, E)$, we must *compare* the *posterior* probability of H *on* E with the *prior* probability of H.

On closer inspection, we can see that both kinds of confirmation-theoretic notions are implicated in Sober's quotation, above. Sober talks about a hypothesis H being "plausible", which he takes to be synonymous with H's being "more plausible than *not-H*." Here, Sober has in mind the *firmness* concept $\mathfrak{c}_f$. He is talking about a hypothesis H being "more probable than not" (given E). Formally, he has in mind cases in which $\Pr(H \mid E) > \Pr(\sim H \mid E)$, which is equivalent to $\Pr(H \mid E) > 1/2$, *i.e.*, $\mathfrak{c}_f(H, E) > 1/2$. Sober wants to distinguish this "non-contrastive" plausibility claim with a "contrastive" claim to the effect that evidence E *favors* one hypothesis H_1 over a *concrete alternative* hypothesis H_2 (where H_2 is *not* equivalent to *not*-H_1).

When it comes to the *contrastive* "favoring" relation, Sober [44] is a *Likelihoodist*. That is, he accepts the following *Law of Likelihood* (for all relations of favoring).

> (LL) Evidence E favors hypothesis H_1 over hypothesis H_2 *if and only if* H_1 confers greater probability on E than H_2 does.
> [Formally, E favors H_1 over H_2 iff $\Pr(E \mid H_1) > \Pr(E \mid H_2)$.]

This is called the "Law of Likelihood" because it says that the relation "E favors H_1 over H_2" boils down to a comparison of the *likelihoods* $\Pr(E \mid H_1)$ and $\Pr(E \mid H_2)$ of the two alternative hypotheses H_1 and H_2, respectively — *relative to evidence E*.

Interestingly, Sober's Likelihoodist *favoring* relation is intimately related to Carnap's *confirmation as increase in firmness* concept $\mathfrak{c}_i$. To see this, we need to delve a bit deeper into $\mathfrak{c}_i$-theory. First, note that various $\mathfrak{c}_i$-measures have been proposed and defended by (BCT)-ers. Here are a few of the most popular $\mathfrak{c}_i$-measures.[3]

- *Difference*: $d(H, E) \overset{\text{def}}{=} \Pr(H \mid E) - \Pr(H)$.
- *Ratio*: $r(H, E) \overset{\text{def}}{=} \log\left[\dfrac{\Pr(H \mid E)}{\Pr(H)}\right]$.
- *Likelihood-Ratio*: $l(H, E) \overset{\text{def}}{=} \log\left[\dfrac{\Pr(E \mid H)}{\Pr(E \mid \sim H)}\right] = \log\left[\dfrac{\Pr(H \mid E) \cdot [1 - \Pr(H)]}{[1 - \Pr(H \mid E)] \cdot \Pr(H)}\right]$.

[2] From now on, I will (for simplicity) suppress the background corpus K, unless we need to be explicit about its content. But, as I implied in *fn.* 1, all confirmation claims and quantities are (implicitly) *relativized to background corpora.* So, when we write $\mathfrak{c}(H, E)$ this is really just shorthand for $\mathfrak{c}(H, E \mid K)$, for some unspecified background corpus K (where, formally, $\mathfrak{c}(H, E \mid K)$ is obtained from $\mathfrak{c}(H, E)$ by *conditionalizing* $\Pr(\cdot)$ in $\mathfrak{c}(H, E)$ on K). The behavior of *confirmational contrasts* involving different background corpora has been under-discussed in the literature. See [13] for a notable exception. I will not have the space here to discuss contrasts involving alternative background corpora. But, toward the end of the paper, we will see an example where the content of K becomes important.

[3] See [14], [29], and [8] for contemporary discussions of the various measures of $\mathfrak{c}_i$ that have been proposed and defended in the Bayesian confirmation-theoretic literature. My definition of $\mathfrak{c}_i$ — as a *function* of the *posterior* and *prior* probabilities of H — is intentionally restrictive (*viz.*, intentionally less than maximally general), so as to avoid various technical subtleties that are not central to the issues I'm discussing in this article. Moreover, we take *logarithms* of the ratio measures to ensure that they are *positive* when E confirms$_i$ H (relative to K), *negative* when E *dis*confirms$_i$ H (relative to K), and *zero* when E is neither confirms$_i$ nor disconfirms$_i$ H (relative to K). This is merely a useful convention, which does not affect the comparative/contrastive structure imposed by the measures.

Second, consider the following *bridge principle*, which connects the *favoring* relation and certain comparative claims involving *confirmation as increase in firmness*:

> ($\dagger_{\mathfrak{c}_i}$) Evidence E favors hypothesis H_1 over hypothesis H_2, according to measure $\mathfrak{c}_i$, *if and only if* $\mathfrak{c}_i(H_1, E) > \mathfrak{c}_i(H_2, E)$.

What ($\dagger_{\mathfrak{c}_i}$) says is that *favoring* relations *supervene on* $\mathfrak{c}_i$*-relations*. This is a natural bridge principle for a (BCT)-theorist to accept. After all, it just says that E favors H_1 over H_2 iff E confirms$_i$ H_1 more strongly than E confirms$_i$ H_2. I will discuss the philosophical case for (and against) such a bridge principle, below. But, first, note that ($\dagger_{\mathfrak{c}_i}$) forges the following intimate connection between (LL) and (BCT):

> (1) ($\dagger_r$) entails (LL).[4]

What (1) says is that if one adopts *both* (*i*) r as one's $\mathfrak{c}_i$-measure, *and* (*ii*) the bridge principle ($\dagger_{\mathfrak{c}_i}$) connecting favoring and comparative-$\mathfrak{c}_i$, then one's (Bayesian) confirmation$_i$-theory *entails* the Law of Likelihood (LL). Moreover, it can be shown that r is the *only* choice of $\mathfrak{c}_i$-measure (among contemporary $\mathfrak{c}_i$'s) that entails [*via* ($\dagger_{\mathfrak{c}_i}$)] the (LL). In this sense, it is somewhat misleading for Sober to represent (BCT) and (LL) as *mutually exclusive* alternative approaches to "favoring". In fact, what (1) reveals is that Sober's (LL) is a *consequence of a particular way of being a Bayesian confirmation-theorist.* And, as it happens, one prominent Bayesian (Peter Milne [34]) has *used* ($\dagger_{\mathfrak{c}_i}$) and (LL) to argue in favor of r as "the one true measure of $\mathfrak{c}_i$".

Thus, we find ourselves in a somewhat uncomfortable dialectical position. We have both Likelihoodists and (some) Bayesians accepting the Law of Likelihood (LL), but (all) Likelihoodists seem to think that their approach is (somehow) *incompatible* with *any* Bayesianism. I suspect that the problem here has to do with the *bridge principle* ($\dagger_{\mathfrak{c}_i}$). While it would be natural for a *Bayesian* (such as Milne) to accept ($\dagger_{\mathfrak{c}_i}$), it is not at all clear why a *Likelihoodist* (*per se*) should accept ($\dagger_{\mathfrak{c}_i}$). Indeed, it seems to me that Likelihoodists must *reject* ($\dagger_{\mathfrak{c}_i}$), *if* they are to claim that their theory of favoring is somehow *incompatible* with a (BCT) that accepts ($\dagger_r$).

But, this is a delicate matter. After all, a proponent of the ($\dagger_r$)-flavor of (BCT) will *agree* with Likelihoodists on *all favoring claims.* That is, there won't be any concrete examples involving hypothesis testing on which the Likelihoodist and the ($\dagger_r$)-theorist will disagree about any of the favoring relations. So, what *can* they disagree about? It seems that all they *can* disagree about is the *existence* of the "non-contrastive" confirmational quantities — $r(H_1, E)$ and $r(H_2, E)$ — that feature in the Bayesian bridge principle ($\dagger_r$). But, why would Likelihoodists doubt the *existence* of $r(H_1, E)$ and $r(H_2, E)$? Likelihoodists often complain that the existence of such quantities presupposes the existence of *prior probabilities* $\Pr(H_1)$ and $\Pr(H_2)$ of the hypotheses H_1 and H_2. And, Likelihoodists are skeptical about the existence of prior probabilities of hypotheses. This has lead some Likelihoodists to view (LL) as a "*ceteris paribus* law". For instance, Sober [45] now says:

> The Law of Likelihood should be restricted to cases in which the probabilities of hypotheses are not under consideration (perhaps because they are not known or are not even "well-defined") and one is limited to information about the probability of the observations given different hypotheses.

[4] For the sake of brevity, I have omitted all proofs of technical claims from this article.

The worry here seems to trade on the fact that $r(H,E)$ is a function $\left[\frac{\Pr(H|E)}{\Pr(H)}\right]$, which featutes the prior probability of H, $\Pr(H)$, as a term (and Likelihoodists are skeptical about the existence of such priors). But, this is potentially misleading, since $r(H,E)$ is *numerically identical* to a function $\left[\frac{\Pr(E|H)}{\Pr(E)}\right]$, which does *not* feature the prior probability of H as a term — it features only the *likelihood* of H [$\Pr(E \mid H)$] and the prior probability of the *evidence* E [$\Pr(E)$]. So, being more careful now, I suppose the Likelihoodist (*e.g.*, Sober) should say that they are *only* willing to countenance *likelihoods* of hypotheses $\Pr(E \mid H)$, and that they are skeptical about the existence of *prior probabilities of both hypotheses and bodies of evidence.*

What do we *mean* when we say that Likelihoodists are "skeptical about the existence of some conditional (see *fn.* 1) probabilities (*e.g.*, priors), but not others (*e.g.*, likelihoods)"? This is where we enter into vexed controversies involving various *interpretations* of the salient confirmation-theoretic conditional probability functions $\Pr(\cdot \mid \cdot)$. A Bayesian will typically interpret such conditional probabilities either as the *actual* (conditional) *degrees of belief* of an agent ([10], [38]) *or* as *justified* or *epistemically rational* (conditional) degrees of belief, relative to some body of evidence ([30], [4], [52, ch. 10]). Likelihoodists, on the other hand, are (mainly) talking about *statistical* probabilities, which are implied by *statistical hypotheses* [41]. And, when it comes to statistical probabilities, it is typically assumed that these are (paradigmatically) *likelihoods of statistical hypotheses.* Prior probabilities (of *either* hypotheses *or* bodies of evidence) are usually *not* determined by the sorts of statistical models Likelihoodists have in mind. As such, Likelihoodists seem to be assuming that *only* probabilities that are implied by certain sorts of statistical models are "fair game" for use in theories of favoring (or confirmation). In other words, Likelihoodists might be willing to grant that there is a *sense* in which (*e.g.*) subjective Bayesian probabilities "exist" (*viz.*, *in the minds of Bayesians*), but Likelihoodists will maintain that such probabilities lack *epistemological probative value* in hypothesis testing (specifically, in the testing of *statistical* hypotheses).[5]

This is the best explanation of the motivation behind Sober's proposed *restriction* of (LL) that I have been able to come up with. But, I remain unconvinced. First, the restriction seems rather *ad hoc*. And, second, it seems to presuppose that — in whatever contexts the Likelihoodists have in mind — there aren't any *objective* (and probative) *epistemological constraints* on "initial credences" in hypotheses (and/or bodies of evidence). I'm not sure why we should believe that.[6] Having said that, I want to avoid getting bogged-down in disputes involving competing interpretations of confirmation-theoretic conditional probabilities. For this reason, I won't delve any further into this particular dialectical labyrinth (but, see *fn.* 6 for references). Instead, I want to return to the *material adequacy* of the so-called "Law of Likelihood". I think the (LL) — in its original formulation — is simply *false*.

[5] Some philosophers have urged that *both* sorts of probabilities are needed. For instance, Jim Hawthorne [23] has argued that Bayesians need *both* "degrees of belief" (which are *subjective*) *and* "degrees of support" (which are *objective*, and *implied by concrete statistical hypotheses/models*).

[6] There is a rather vast literature on so-called "objective Bayesianism", which tries to identify certain features of *objectively reasonable/rational* initial/prior credence functions. I won't discuss that literature here (since I don't want to dwell on issues surrounding the interpretation of confirmation-theoretic probabilities). But, "objective Bayesianism" has a rather long (and rather tortuous) history, which features many notable figures, *e.g.*, Leibniz [20], Keynes [30], Carnap [4], de Finetti [9], and Maher [33]. For a state-of-the art defence of "objective Bayesianism", see Jon Williamson's [51].

Here is what I take to be a rather clear counterexample to (LL). The experimental set-up for my counterexample to (LL) involves a standard, well-shuffled deck of playing cards, from which we are going to sample a single card. Let $E \overset{\text{def}}{=}$ the card is a spade, $H_1 \overset{\text{def}}{=}$ the card is the ace of spades, and $H_2 \overset{\text{def}}{=}$ the card is black. In this example (assuming the standard statistical model of card draws), we have $\Pr(E \mid H_1) = 1 > \Pr(E \mid H_2) = \frac{1}{2}$. So, according to (LL), E favors H_1 over H_2. But, this seems absurd. After all, the truth of E *guarantees* the truth of H_2, but the truth of E does *not* guarantee the truth of H_1. In this sense, E constitutes *conclusive evidence* for H_2, and *less than* conclusive evidence for H_1. If *that* doesn't imply that E *favors* H_2 over H_1, then what would? This suggests the following principle:

(CE) If E constitutes conclusive evidence for H_1, but E constitutes less than conclusive evidence for H_2 (where it is assumed that E, H_1, and H_2 are all contingent), then E favors H_1 over H_2.

It seems to me that (CE) should be a *desideratum* for any adequate explication of "favoring". And, because (LL) implies the existence of counterexamples to (CE), this seems to *refute* (LL). Moreover, (CE) has strong ramifications for any Bayesian theory of confirmation$_i$ which accepts the bridge principle ($\dagger_{\mathfrak{c}_i}$). Once we accept (CE) and ($\dagger_{\mathfrak{c}_i}$), then we must *not* adopt r as our measure of $\mathfrak{c}_i$, since ($\dagger_r$) entails (LL).

One way to respond to this counterexample would be to adopt Sober's restriction of (LL) to contexts in which "only likelihoods are available, and no priors are available". However, it is unclear whether the intuitive verdict about this example even *depends on* the "availability of priors". It seems to me that there is a simple *logical asymmetry* that explains the intuitive verdict that E favors H_2 over H_1 in the example. I don't think this requires any appeal to "verboten priors". As a result, I don't think the Soberian "*ceteris paribus* law" reading of (LL) is probative here.

Another way to respond to my counterexample to (LL) would be to find a different strategy for "restricting the scope" of (LL) — one which is motivated in some other way. Recently, Dan Steel [46] and Jake Chandler [6] have (independently) responded to (CE) and my putative counterexample to (LL) in just such a way. Both Steel and Chandler argue that "favoring" is inherently *contrastive* in nature. And, as a result, that the hypotheses H_1 and H_2 that appear in (LL) and (CE) must be *mutually exclusive*. This rules-out my example, since in my example H_1 *entails* H_2.[7]

Following Hitchcock's approach to contrastive probabilistic explanation ([26], [27]) — which presupposes that members of contrast classes (of *explananda*) are *mututally exclusive* — Chandler ([5], [6]) proposes the following Hitchcock-style account of favoring that *builds-in* mutual exclusivity of the alternative hypotheses:

(HC) E favors H_1 over H_2 iff (*i*) H_1 and H_2 are mutually exclusive, and (*ii*) $\Pr(H_1 \mid E \,\&\, (H_1 \vee H_2)) > \Pr(H_1 \mid {\sim}E \,\&\, (H_1 \vee H_2))$.

Interestingly, it turns out that (HC) is *logically equivalent* to the following:

(LL*) E favors H_1 over H_2 iff (*i*) H_1 and H_2 are mutually exclusive, and (*ii*) $\Pr(E \mid H_1) > \Pr(E \mid H_2)$.

[7] It is no accident that the logical relations in my example are precisely the way they are. It is important that the intuitive verdict be explicable solely on the basis of *logical asymmetries* between E, H_1, and H_2. And, this sort of structure is (more or less) the *only* one that will do the trick. See my [16] and Chandler's [6] for further discussion of this and other putative counterexamples to (LL).

That is, assuming mutual exclusivity of H_1 and H_2, the Hitchcock-Chandler approach to favoring (HC) *is equivalent to the Likelihoodist theory of favoring* (LL).[8]

This suggests a natural revision/restriction of the original Law of Likelihood (LL), which is embodied in (LL*), above. The idea is that (LL) is true — *provided that the alternative hypotheses H_1 and H_2 are mutually exclusive.* This restriction of (LL) is *not ad hoc* in the way that Sober's restriction was. Indeed, the restriction does have some intuitive plausibility. When we make contrastive claims, we *often* presuppose that the members of the salient contrast class are mutually exclusive. But, it is natural to ask whether such a presupposition is *always* present.

It seems clear to me that such a presupposition is *not always* present in cases of favoring/contrastive confirmation.[9] For instance, in the context of statistical hypothesis testing, Likelihoodists (and other "anti-Bayesians") are quick to *criticize* Bayesian approaches that *do* (generally) presuppose the mutual exclusivity of alternative hypotheses. Indeed, in the context of statistical model selection, it is supposed to be one of the relative *strengths* of Likelihoodism — *as opposed to* certain flavors of Bayesianism — that it *is* capable of testing *nested models* (*i.e.*, models that bear *containment or entailment* relations to each other, as my H_1 and H_2, above, do) against each other. So, this "mutual exclusivity requirement" reply to my counterexample to (LL) is simply *not available* to traditional, *statistical* Likelihoodists. Structurally analogous cases involving statistical model selection will be among the examples *used by* Likelihoodists as "evidence" that is supposed to "favor" their own approach to model selection over certain Bayesian approaches. See [17] for a thorough discussion of this debate among philosophers of statistics.

Another worry I have about the "mutual exclusivity presupposition" maneuver is that it seems to make it *too easy* to refute the bridge principle ($\dagger_{\mathfrak{c}_i}$). A Bayesian confirmation theorist who accepts a bridge principle ($\dagger_{\mathfrak{c}_i}$) will hear "E favors H_1 over H_2" as *synonymous* with "E evidentially supports H_1 more strongly than E evidentially supports H_2". And, it seems clear that this latter *comparative confirmation* claim does *not* (always) presuppose that the alternative hypotheses (H_1 and H_2) are mutually exclusive. Surely, evidence E can sometimes support a logically stronger (or logically weaker) hypothesis more strongly than E supports a logically weaker (or logically stronger) one. In the next two sections, I will discuss two well-known illustrations of such phenomena. This will provide further test cases for the various accounts of "favoring" (and confirmation) that we've been discussing.

4. The Problem of Irrelevant Conjunction

This section has three parts. In the first part, I will rehearse (in some detail) the historical dialectic concerning the "problem of irrelevant conjunction" (PIC). In the second part, I will explain how this problem (and our proposed resolution of it) can shed some light on the questions about Likelihoodism, Bayesianism, "favoring",

[8]This also means that Hitchcock's account of contrastive probabilistic explanation ([26], [27]) — for two contrasted *explananda*, relative to a single *explanans* — *reduces to a simple comparison of the likelihoods* of the contrasted *explananda*, relative to the *explanans.* This is an interesting and important theoretical connection (and unification). It reveals that something like (LL) is presupposed in various contemporary "contrastive probabilistic explications" of *both* explanation *and* confirmation.

[9]Peter Lipton [31] voices an analogous complaint about Hitchcock-style approaches to contrastive explanation. Lipton thought that the members of explanatory contrast classes did *not* always have to be mutually exclusive. I won't be able to discuss that dialectic here. But, the parallel is worth noting. In the final section, we'll see some psychological data that support our Liptonian complaint.

and "constrastive confirmation" that were discussed in the previous section. In the third part of this section, I will discuss an objection to our approach to (PIC) that was recently raised by Patrick Maher, and an alternative, "contrastivist" account of (PIC) due to Jake Chandler, which was inspired by Maher's objection.

4.1. **Part I: The Historical Dialectic of the Problem of Irrelevant Conjunction.** One of the traditional (deductive) accounts of confirmation that features prominently in the history of confrimation theory is the so-called Hypothetico-Deductive (or HD) account of confirmation (I will call this confirmation relation "confirms$_h$"):

(HD) E confirms$_h$ H iff H entails E.

Due to the monotonicity of entailment, confirmation$_h$ has the following property:

(2) If E confirms H, then E confirms $H \,\&\, X$ (for *any* X).

Clark Glymour [18] raises two worries in connection with property (2):

(2a) If H entails E, then so will $H \,\&\, X$, where X is any sentence whatsoever. But, we cannot admit, generally, that E will lend any plausibility to an arbitrary X. One might, of course, deny what Hempel calls the special consequence condition, namely, that if E confirms a hypothesis, then E will confirm every logical consequence of that hypothesis. But this is hardly satisfactory. Sometimes, anyway, confirmation does follow entailment, at least over some paths.

(2b) As evidence accumulates, we may come to accept a hypotheses ... and when we accept a hypothesis we commit ourselves to accepting all of its logical consequences. So, if a body of evidence could bring us to accept hypothesis H, and whatever confirms H confirms $H \,\&\, X$, where X is any irrelevant hypothesis, then the same evidence that brings us to accept H, ought, presumably, to bring us to accept X.

Both of these worries about (2) have to do with (some of) the evidential support provided by E (for H) somehow "rubbing off" onto an *irrelevant conjunct* X. But, the two worries seem to involve *two different notions* of evidential support:

- E **supports$_1$** H iff E is (positively) evidentially relevant to H.
- E **supports$_2$** H iff E warrants/justifes belief/acceptance of H.

Glymour's worry (2b) involves support$_2$, and Glymour's worry (2a) involves support$_1$. As I mentioned above, Carnap cautioned us not to conflate confirms$_f$ and confirms$_i$. Similarly, we need to be careful not to conflate these two notions of "evidential support". For the sake of the present discussion, we will follow Carnap, who thought of confirms$_i$ as an *explicatum* for supports$_1$, and confirms$_f$ as an *explicatum* for supports$_2$. Thus, we will say that Glymour's worry (2a) involves confirms$_i$, and his worry (2b) involves confirms$_f$. Finally, *nobody* thinks that confirms$_h$ is a good *explicatum* for supports$_2$.[10] So, I won't bother to discuss Glymour's worry (2b).

Glymour's (2a) is more interesting. This worry has become known as "the problem of irrelevant conjunction" (PIC). In worry (2a), Glymour mentions Hempel's *special consequence condition* [25], which entails the following condition:

[10]Any hypothesis H will entail/predict *many* observational consequences. It just *can't* be the case that verifying *any one* of these (myriad) predictions would be *sufficient to warrant belief in H*. It is also useful to note that confirms$_f$ does *not* satisfy (2) — not even in the case where H entails E (unlike confirms$_i$). For these reasons, I'm focusing on confirms$_i$ for the remainder of this section.

(SCC) If E confirms $H \,\&\, X$, then E confirms X (for *any* X).

Clearly, no adequate account of confirmation can satisfy *both* (2) *and* (SCC). Any such theory would entail that *any* evidence E (if it confirms *any* hypothesis) will confirm *every* proposition X. This is why Glymour suggests that one "cheap" way out of the problem of irrelevant conjunction is to (merely) *deny* (SCC). But, as Glymour also suggests, we want *more than a mere denial of* (SCC) here. We want a response to "the problem of irrelevant conjunction" that also involves a *principled* (and intuitively plausible) way of determining when (SCC) holds and when it fails.

This is where Bayesian confirmation theory (*viz.*, confirmation$_i$-theory) comes into the picture. First, note that confirmation$_i$-theory does *not* entail (2). That is:

(3) E confirms$_i$ $H \nRightarrow E$ confirms$_i$ $H \,\&\, X$.

But, confirmation$_i$-theory *does* entail the following *deductive special case* of (2):

(4) If H *entails* E, then E confirms$_i$ $H \,\&\, X$ (for *any* X).

Fact (4) has inspired several confirmation$_i$-theorists to offer "resolutions" of (PIC).

John Earman [12] appeals to the following to try to "soften the impact" of (4):

(4.1) If H entails E, then $\mathfrak{c}_i(H \,\&\, X, E) < \mathfrak{c}_i(H, E)$.

The idea behind Earman's (4.1) is that — while it is true that confirms$_i$ entails the (PIC)-like property (4) in the (HD) case where H entails E — tacking irrelevant conjuncts onto such [(HD)-confirmed] hypotheses will always *lower the degree to which E confirms$_i$ them.* That is, one will *pay a confirmation$_i$-theoretic price* for tacking irrelevant conjuncts onto an [(HD)-confirmed] hypotheses. Closer scrutiny of Earman's response to (4) reveals the following three worrisome features:

(a) The "irrelevance" of the conjuncts X is *irrelevant* to the decrease in $\mathfrak{c}_i$. After all, (4.1) is true for *all* X — irrelevant or otherwise. It would be preferable if the irrelevance of X was (in some sense) playing an explanatory role.

(b) (4.1) is *not true for all* measures $\mathfrak{c}_i$ of confirmation$_i$. For instance, (4.1) fails to hold for the ratio measure $r(H, E)$ that we discussed above, in connection with Likelihoodism. As such, proponents of ($\dagger_r$), *e.g.*, Milne, won't be in a position to avail themselves of Earman's approach to (PIC).

(c) (4.1) only applies to cases of *deductive* evidence (*i.e.*, cases in which E confirms$_h$ H). As we'll see shortly, confirmation$_i$ faces a *more general* (PIC)-type problem. And, Earman's approach won't be applicable to it.

Other Bayesians have offered similar responses to (PIC)/(4). Rosenkrantz [40] appeals to the following, rather similar result:

(4.2) If H entails E, then $d(H \,\&\, X, E) = \Pr(X \mid H) \cdot d(H, E)$.

Using (4.2), Rosenkrantz tries to address *some* of the problems we raised for Earman's (4.1)-approach. Rosenkrantz explains the rationale behind (4.2) as follows:

> ...I hope you will agree that the two extreme positions on this issue are equally unpalatable, (i) that a consequence E of H confirms $H \,\&\, X$ not at all, and (ii) that E confirms $H \,\&\, X$ just as strongly as it confirms H alone. ...In general, intuition expects intermediate degrees of confirmation that depend on the degree of compatibility of H with X.

Basically, what Rosenkrantz is doing here is: (*i*) adopting $\Pr(X \mid H)$ as a measure of the "degree of compatibility of H with X", and (*ii*) adopting the difference measure d as his measure of confirmation$_i$. Does this ultimately lead to an *improvement* on Earman's (4.1)? This depends on whether Rosenkrantz really has adequately addressed worries (a)–(c), above. Unfortunately, I don't think he has.

In a way, Rosenkrantz is *trying* to address (a) here. He seems to be thinking of $\Pr(X \mid H)$ as a kind of measure of "the degree of (ir)relevance" of X — *qua conjunct* in $H \,\&\, X$. But, this is a *peculiar* way for a *Bayesian* to explicate "(ir)relevance". Normally, Bayesians use *probabilistic (in)dependence* relations to explicate *(ir)relevance* relations. Moreover, because it is a relation involving only X and H, $\Pr(X \mid H)$ can tell us nothing about "degrees of relevance" involving X, H — *and* E. And, in general, this "irrelevant conjunct" relation (whatever it turns out to be) must be *evidence relative.* That is, what we want is an explication of "X is an irrelevant conjunct *to* H, *relative to evidence* E." When it comes to (b), Rosenkrantz is in even worse shape than Earman. Rosenkrantz's approach works *only* for confirmation$_i$-measures that are *very similar* to the difference measure d. In this sense, Earman's approach is strictly more general. Finally, Rosenkrantz is still only addressing the *deductive* case. So, like Earman's approach, Rosenkrantz's approach will not be useful for more general, inductive varieties of (PIC), which we will see shortly.

Ultimately, confirmation$_i$-theorists need to *re-think* the problem of irrelevant conjunction, and its possible resolution(s). To that end, let's think about how $\mathfrak{c}_i$-theory handles *irrelevant* conjunctions, in the general, *inductive* case. First, we need to say what it *means* for X to be an *irrelevant conjunct* to a hypothesis H, *with respect to evidence* E. Here, we can adopt stronger or weaker explications of "irrelevant conjunct". The strongest, natural Bayesian explication is:

- X is a *strongly irrelevant conjunct to* H*, with respect to evidence* E, just in case X is *probabilistically independent* of H, E, *and* $H \,\&\, E$.[11]

One could also adopt the following weaker explication of "irrelevant conjunct":

- X is a *weakly irrelevant conjunct to* H*, with respect to evidence* E, just in case $\Pr(E \mid H \,\&\, X) = \Pr(E \mid H)$ [*i.e., if* $X \perp\!\!\!\perp E \mid H$].

The idea behind the weak explication of "irrelevant conjunct" is that a weakly irrelevant conjunct X *does not affect the likelihood of the hypothesis* H (on evidence E). In this sense, X does not add anything to H — insofar as its predictions about (the probability of) the evidence are concerned. This is a natural (weak, Bayesian) way of capturing the idea that X is an "irrelevant conjunct" to H, with respect to evidence E. The strong explication entails the weak explication, but it also entails *much more.* While the weak explication appeals *only to likelihoods* of hypotheses, the strong explication trades in *priors of both hypotheses and evidence*, and so will, presumably, not be something that Likelihoodists will (generally) find kosher.[12]

[11] It follows from this strong explication of "irrelevant conjunct" that X is probabilistically independent of *all logical combinations* of H and E. That's why it's the *strongest* Bayesian explication.

[12] In my [15], I adopted the strong explication of "irrelevant conjunct". In a more recent paper [24], Jim Hawthorne and I have shown that the weaker explication suffices for the Earman-style approach to (PIC) that we favor. That being said, there are certain dialectical advantages (for a "full-blown Bayesian") to using the stronger explication of "irrelevant conjunct". See *fn.* 13 for further discussion.

Now that we have both strong and weak explications of "irrelevant conjunct", we are in a position to investigate how tacking on irrelevant conjuncts affects confirmation relations (both qualitative and quantitative). Here are two key results:

(5) If E confirms$_i$ H, and X is an (either strongly or weakly) irrelevant conjunct to H, with respect to evidence E, then E *also* confirms$_i$ $H \,\&\, X$.

(6) If E confirms$_i$ H, and X is an (either strongly or weakly) irrelevant conjunct to H, with respect to evidence E, then $\mathfrak{c}_i(H \,\&\, X, E) < \mathfrak{c}_i(H, E)$, for all measures of $\mathfrak{c}_i$ (under consideration), *except the ratio measure* r.

What (5) tells us is that confirmation$_i$-theory *does suffer from* a *general* problem of *irrelevant* conjunction, which is analogous to the problem faced by confirmation$_h$-theory (*i.e.*, HD-confirmation). That is, tacking *irrelevant* conjuncts onto a hypothesis that is confirmed$_i$ by E yields *conjunctions that are also confirmed$_i$ by* E. Result (6) is a generalization of Earman's (4.1), which avoids our criticisms (a) and (c) of Earman's account. Regarding (a), our (6) *makes essential use of* the *irrelevance* of the conjunct X. Regarding (c), we have generalized both the problem and its (Earman-style) resolution far beyond the deductive case in which H entails E. Indeed, we now see that the deductive case is just an *extreme, limiting-case* in which H *screens-off* X from E. The fact that *weak* irrelevance (*viz.*, the screening-off of X from E, by H) is sufficient for both (5) and (6) provides a *unified explanation* of *why* the deductive *and* inductive cases of (PIC) behave the way they do. Regarding our criticism (b) of Earman's account, we can unfortunately do no better than Earman did. Alas, our results only go through for $\mathfrak{c}_i$ measures *other than* r. This means that Bayesians (like Milne) who accept r will not be able to avail themselves of our approach to (PIC). And, those (like Milne) who accept the bridge principle ($\dagger_r$) will not be able to say that E *favors* H over $H \,\&\, X$ when X is an irrelevant conjunct to H, with respect to evidence E. This connection to ($\dagger_r$) — and therefore (LL) — is an interesting and important one. I will return to this dialectical thread in Part II of this section. Before moving on to Part II of this section, however, it is useful to illustrate our account of (PIC) with the following simple concrete example:

- Suppose we'll be sampling a card at random from a standard deck. Let $\mathcal{E}$ be the proposition that the card is black. Let $\mathcal{X}$ be the hypothesis that the card is an ace, and let $\mathcal{H}$ be the hypothesis that the card is a spade.

In this example (assuming, as above, the standard probability model for random draws from a standard deck of playing cards), we have the following facts:

(7) $\mathcal{E}$ confirms$_i$ $\mathcal{H}$.

(8) $\Pr(\mathcal{E} \mid \mathcal{H} \,\&\, \mathcal{X}) = \Pr(\mathcal{E} \mid \mathcal{H})$.

Thus, the preconditions of our (5) and (6) are met in the example. Therefore:

(9) $\mathcal{E}$ confirms$_i$ $\mathcal{H} \,\&\, \mathcal{X}$.

(10) $\mathfrak{c}_i(\mathcal{H} \,\&\, \mathcal{X}, \mathcal{E}) < \mathfrak{c}_i(\mathcal{H}, \mathcal{E})$, for all $\mathfrak{c}_i$, *except the ratio measure* r.

Finally, we *also* have the following important fact:

(11) $\mathcal{E}$ does *not* confirm$_i$ $\mathcal{X}$.

In light of (9) and (11), this also constitutes a counterexample to (SCC) for confirms$_i$. But, Bayesian confirmation$_i$ theory doesn't *merely reject* (SCC) here. Rather, it provides a principled and illuminating account of when (SCC) fails (and when it holds). This seems to satisfy Glymour's desire for an account of confirmation that (*i*) has something interesting and illuminating to say about (PIC), and (*ii*) simultaneously provides a principled and explanatory rejection of (SCC).[13] In Part II of this section, I return to Likelihoodism, favoring, and contrastive confirmation, in light of (PIC).

4.2. **Part II: (PIC), Favoring, and Contrastive Confirmation.** At the end of section 3, I mentioned that there were clear-cut examples in which evidence E supports H more strongly than E supports $H \,\&\, X$. In Part I of this section, I explained how (PIC) provides an interesting class of examples of precisely this kind, and how Bayesian confirmation$_i$ theory provides the (intuitively) correct verdicts concerning such cases. I also mentioned that advocates of the $\mathfrak{c}_i$ measure r are unable to reproduce these verdicts. Here is a result that furnishes a more precise explanation:

(12) If E confirms$_i$ H, and X is an (either strongly or weakly) irrelevant conjunct to H, with respect to evidence E, then $r(H \,\&\, X, E) = r(H, E)$.

This is not surprising, since (*i*) "weak irrelevance" entails that the *likelihoods* of H and $H \,\&\, X$ are equal (relative to E), and (*ii*) a comparison of $r(H, E)$ and $r(H \,\&\, X, E)$ boilds down to a comparison of the *likelihoods* of H and $H \,\&\, X$ (relative to E). Moreover, since ($\dagger_r$) entails (LL), advocates of ($\dagger_r$) will have to say that there can be *no favoring* of H over $H \,\&\, X$ by E, whenever X is an "irrelevant conjunct" (in either our weak or strong senses). What should a Likelihoodist say about this situation?

There are two ways that advocates of (LL) typically respond to these facts about (PIC). The first way [exemplified by Milne, and other "Bayesian-Likelihoodists" who adopt the bridge principle ($\dagger_r$)] is to *bite the bullet*, and insist that E *evidentially supports* H *and* $H \,\&\, X$ *equally strongly* in examples of "irrelevant conjunction" (like our concrete example involving the deck of cards in §4.1, above). I won't address this strategy here (except to say that I don't find it intuitively compelling). Rather, I will focus on the second kind of response given by advocates of (LL).

The second type of response by defenders of (LL) comes from "contrastivist-Likelihoodists" (such as Chandler) who accept (LL), but do *not* accept the bridge principle ($\dagger_r$). Such advocates of (LL) will insist that *it doesn't make sense* to talk about "favoring" in cases like (PIC), where the alternative hypotheses are *not* mutually exclusive. In other words, the second response is an instance of the *mutual exclusivity requirement* (on alternative hypotheses involved in favoring relations), which we discussed in section 3 above. Ultimately, I don't think this response obviates the need for something very much like Milne's bullet-biting response. This is because, as Susanna Rinard [39] has pointed out, requiring mutual exclusivity of alternative hypotheses doesn't manage to avoid all of the problems raised by (PIC)-type considerations. We can see Rinard's point clearly with this simple modification of our card-sampling example from §4.1 (which is similar to her example):

[13] Glymour (personal communication) criticized the "weak irrelevance" versions of our results on the grounds that our weak explication of irrelevance may classify conjuncts that are "intuitively irrelevant" as "relevant". We can avoid this worry (as well as worries about distinguishing *redundant* conjuncts and *irrelevant* conjuncts) by using the strong explication of irrelevance instead. But, this is probably not a move that a Likelihoodist would be inclined to make. See *fn.* 12 for further discussion.

- Let $\mathcal{E}$ be the proposition that the card is black. Let $\mathcal{X}$ be the hypothesis that the card is an ace, let $\mathcal{H}_1$ be the hypothesis that the card is a spade, and let $\mathcal{H}_2$ be the hypothesis that the card is a club.

"Contrastivist-Likelihoodists" (*e.g.*, Chandler) will complain that claims such as:

(12) $\mathcal{E}$ *favors* $\mathcal{H}_1$ over $\mathcal{H}_1$ & $\mathcal{X}$.

[which, owing to (5) and (6), are implied by all bridge principles ($\dagger_{\mathfrak{c}_i}$) *except for* ($\dagger_r$)] are *infelicitous* on the grounds that "favoring" claims presuppose that the alternative hypotheses involved are mutually exclusive. However, mutual exclusivity is not really the issue here. To see this, consider the following *analogous* claim:

(13) $\mathcal{E}$ *favors* $\mathcal{H}_2$ over $\mathcal{H}_1$ & $\mathcal{X}$.

Here, the alternative hypotheses *are* mutually exclusive. So, it seems that there can be no "constrastivist-presuppositional" grounds for claiming that (13) is infelicitous. Therefore, it seems that *all* advocates of (LL) must say that (13) is (felicitous and) *false*. After all, the *likelihoods* of $\mathcal{H}_2$ and $\mathcal{H}_1$ & $\mathcal{X}$ are *equal* (they are both equal to 1). Moreover, it seems clear that — from a probabilistic point of view — claims $\mathcal{H}_1$ and $\mathcal{H}_2$ *on a par with respect to the evidential relations they bear to* $\mathcal{E}$ *and* $\mathcal{X}$ in our example. This seems to imply that something very much like Milne's *bullet-biting* response still needs to be embraced by defenders of (LL) — even those who reject the bridge principle ($\dagger_r$) on "contrastivist-presuppositional" grounds. Because I think it's implausible to claim *either* that $\mathcal{E}$ *evidentially supports* $\mathcal{H}_2$ *and* $\mathcal{H}_1$ & $\mathcal{X}$ *equally strongly or* that $\mathcal{E}$ does *not* favor $\mathcal{H}_2$ over $\mathcal{H}_1$ & $\mathcal{X}$ in the present example, I don't hold out much hope for Likelihoodists (of either stripe) to tell a compelling, general story about (PIC). That said, I will briefly discuss a recent "contrastivist-Likelihoodist" alternative approach to (PIC), due to Chandler.

4.3. **Maher's Objection and Chandler's Alternative "Contrastivist" (PIC)-Account.** Patrick Maher [32] has recently criticized our [(5)&(6)-based] approach to (PIC) [24]. He complains that our approach doesn't even *address* (PIC), because he thinks the (PIC) is *motivated essentially by the following (faulty) intuition about support*$_1$:

($\star$) If X is an irrelevant conjunct to H, with respect to evidence E, then E *does not support*$_1$ H & X.

Maher's approach to (PIC) is to simply explain why this is a *false intuition*. He does so by discussing concrete counterexamples to ($\star$), much like our simple card examples in sections (4.1) & (4.2) above. Our response to Maher is quite simple. Of course, we agree that ($\star$) is false (indeed — *that's part* of *our* story about (PIC), too!). But, we *disagree* with the claim that (PIC) is motivated (*essentially*) by "intuition ($\star$)". We would say that (PIC) arises because of the *truth* of (5), above. It is (5) that implies the *existence* of a "problem of irrelevant conjunction" for supports$_1$/confirms$_i$. And, pointing out the *falsity* of ($\star$) — which, of course, we *also* do — doesn't do anything to address (or "soften the impact") of the truth of (5). This is why we think (6) is an essential part of any complete Bayesian confirmation$_i$-theoretic story about (PIC). Thus, we think Maher has the wrong diagnosis here.[14]

[14]We also think that Maher's claim about the centrality of ($\star$) as an intuition that drives people to think there *is* a (PIC) in the first place is, at best, dubious — as a matter of *historical* fact. For instance, we seriously doubt that ($\star$) was essential (or even *significant*) in *Glymour's* [18] thinking about (PIC).

Chandler [5] has picked-up on Maher's ($\star$)-line on (PIC). He has taken Maher's line as a point of departure for his own "contrastivist-Likelihoodist" alternative to our (non-Likelihoodist) approach to (PIC). Chandler starts with his "contrastivist-Likelihoodist" approach to favoring: (HC)/(LL*), which (as we explained above) is just the Law of Likelihood, *plus* the requirement that alternative hypotheses in favoring relations must always be mutually exclusive. Then, he combines (HC)/(LL*) with Maher's claim that intuition ($\star$) is what's driving (PIC). This culminates in the suggestion that people (falsely) *think* that the *non-contrastive* supports$_1$/confirms$_i$ claim ($\star$) is true, because they are *conflating* it with a *true contrastive-favoring* claim. Specifically, Chandler suggests that such people are conflating ($\star$) with

($\star\star$) If X is an irrelevant conjunct to H, with respect to evidence E, then E does not *favor* H & X over H & $\sim X$,

which, according to Chandler's contrastivst-favoring theory (HC)/(LL*), *is* true.

I'll just make a couple of brief remarks about Chandler's approach. First, as I've already explained, I think Maher's diagnosis of "intuition-($\star$)" as *the* (or even *a primary*) source of (PIC) is off the mark. But, let's bracket that and just focus on Chandler's discussion concerning ($\star\star$). Here, I think his discussion is misleading in several ways. First, he presents his theory of favoring in its (HC)-form, which obscures the fact that it is equivalent to (LL*).[15] Second, once we realize that Chandler's theory is a *Likelihoodist* theory of favoring, it is a *trivial* matter that ($\star\star$) comes out true on his theory of favoring, since he is assuming our weak explication of "irrelevant conjunction", which just *is* the salient Likelihood *identity*. Third, because (LL) is rather controversial, Chandler's claim that ($\star\star$) is true is *also* rather controversial. Basically, *only* those who accept (LL) will accept ($\star\star$). As a result, it is somewhat misleading of Chandler to represent his account as providing a "charitable reconstruction" of how someone might come to accept "something like" ($\star$). Finally, it is often *psychologically implausible* to suggest that people attend to *contrastive* claims in contexts where they are confronted with *non-contrastive* questions (which is what Chandler's "contrastivist-error-theory" seems to presuppose).

In the final section of this paper, I will discuss "The Conjunction Fallacy". There, we will see that psychological hypotheses analogous to Chandler's "contrastivist-error-theory" hypothesis [regarding the conflation of ($\star$) and ($\star\star$) by *actual subjects*] are not borne out by the data. We'll also see another example with a similar confrimation-theoretic structure to (PIC), except, this time, it will be the logically *stronger* of two hypotheses that is confirmed$_i$ more strongly than the logically *weaker* alternative. And, this time, the examples won't be *merely* theoretical/philosophical in nature — we'll have lots of psychological data to draw upon.

5. The Conjunction Fallacy

Tversky & Kahneman [50] presented subjects with the (now infamous) "Linda example". In this example, subjects are presented with the following *evidence*:

[15]To be fair to Chandler, he didn't realize that (HC) was equivalent to (LL*) at that time (and neither did Hitchcock). That's something I pointed out to him (and Hitchcock) after his paper was published. In his more recent article [6], Chandler incorporates this insight into his discussion.

(e) Linda is 31, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and she also participated in antinuclear demonstrations.

Then, subjects are asked which of the following two hypotheses about Linda is *more probable, given the above evidence e about Linda.*

(h_1) Linda is a bank teller.

(h_2) Linda is a bank teller *and* an active feminist.

Most subjects answer that h_2 is more probable than h_1, given evidence e. That is, *prima facie*, most subjects report that their conditional credences are such that:

(14) $\Pr(h_2 \mid e) > \Pr(h_1 \mid e)$.

Unfortunately, (14) *contradicts the probability calculus*, according to which:

(15) If p entails q, then $\Pr(p \mid e) \leq \Pr(q \mid e)$, for *any* e.

Since Tversky & Kahneman's paper was published, there has been a great deal of ink spilled (in both cognitive science and philosophy) about what might be going on with subjects who (*prima facie*) report that their degrees of belief satisfy (14).

I won't attempt to address even a small fraction of this literature here. Rather, I will consider two possible explanations of what might be going on with such subjects that have been proposed in the recent literature. This will tie-in nicely with the dialectic we've been discussing. The first explanation, which has been proposed by various cognitive scientists ([11], [36]), is that when subjects are asked the question about h_1 and h_2, they are inclined to presuppose that — because h_1 and h_2 are meant to be *alternative* hypotheses about Linda — h_1 and h_2 are (or should conversationally be treated as if they are) *mutually exclusive*. To better understand this proposal, it helps to work with a finer-grained representation of the content of h_1 and h_2. To this end, I'll now work with the following notation:

(b) Linda is a bank teller.

(f) Linda is an active feminist.

With this notation in hand, we can clearly express the "contrastivist" explanation of "the conjunction fallacy" that is now on the table. The proposal is that, when subjects are asked to contrast h_1 and h_2, what they *actually* end-up contrasting are the following *mutually exclusive* alternative hypotheses about Linda:

($b \,\&\, {\sim}f$) Linda is a bank teller *and not* an active feminist.

($b \,\&\, f$) Linda is a bank teller *and* an active feminist.

Thus, or so this "contrastivist" proposal goes, when subjects report their answer to the question, they are *actually* indicating that their credences are such that:

(16) $\Pr(b \,\&\, f \mid e) > \Pr(b \,\&\, {\sim}f \mid e)$.

And, because (16) does *not* imply probabilistic *incoherence*, it was unfair (and premature) of Tversky & Kahneman (and others) to conclude that typical responses to the Linda question reveal any (Bayesian) *irrationality* in actual subjects.

Various experiments have been performed in recent years, which are designed to explicitly test this "contrastivist explanation" of "the conjunction fallacy". My favorite sets of experiments are reported by Tentori *et. al.* [47] and Bonini *et. al.* [2]. In these experiments, subjects are asked to *bet* on the truth of various logical combinations of b and f (and similar "conjunction fallacy" conjuncts). And, *in the very same contexts*, subjects are *also* asked to perform basic *logical inferences* (*e.g., conjunction elimination*) involving various logical combinations of b and f. These experiments show quite clearly that — even when subjects are clearly presupposing that h_2 and h_1 have the *logical forms* $b \,\&\, f$ and b, respectively — they tend (in proportions not too dissimilar to those seen in the original Tversky & Kahneman experiments) to *bet more money on* the truth of h_2 ($b \,\&\, f$) than on the truth of h_1 (b). To my mind, these experiments show rather definitively that the "contrastivist" explanation of subjects' responses to the Linda question is *inadequate* (and that subjects' responses are in violation of Bayesian rational requirements after all).

In light of these recent psychological experiments involving betting and logical inference, I prefer an alternative way of "explaining" what might be going on in the Linda case. I favor "explanations" that begin by *conceding* that subjects *are* violating Bayesian rational requirements in the Linda case. What I'm more interested in is *why* subjects might tend to make the sorts of errors they make in the Linda case. That is, I'm interested in the question of whether (and to what extent) these mistakes are "understandable", from a (broadly) Bayesian point of view.

In recent joint work with Crupi and Tentori [7], I have endorsed one possible way of understanding why subjects tend to report things like (14) in the Linda case. Our main idea is to begin (as Carnap cautioned us to do) by distinguishing $\mathfrak{c}_f$, which is just *conditional probability*, and $\mathfrak{c}_i$, which gauges *degree of evidential relevance*. Because (as we have already seen) confirmation$_i$ does *not* entail (SCC), this makes it *possible* (*i.e., probabilistically coherent*) for subjects' credences to be such that:

(16) $\mathfrak{c}_i(b \,\&\, f, e) > \mathfrak{c}_i(b, e)$.

for various measures $\mathfrak{c}_i$ of degree of confirmation as increase in firmness. That is, while it is impossible for e to *confirm*$_f$ $b \,\&\, f$ *more strongly* than e *confirms*$_f$ b, it is *not* impossible for e to *confirm*$_i$ $b \& f$ *more strongly* than e *confirms*$_i$ b. Our strategy was to try to identify confirmation$_i$-theoretic conditions which (*i*) are sufficient to entail (16), and (*ii*) would be accepted by most subjects (in the context of the Linda experiments). This lead to the following central confirmation$_i$-theoretic result [7]:

(17) *If* the following two conditions are satisfied:

(17.1) $\mathfrak{c}_i(b, e \mid f) \leq 0$, *and*

(17.2) $\mathfrak{c}_i(f, e) > 0$

then $\mathfrak{c}_i(b \,\&\, f, e) > \mathfrak{c}_i(b, e)$, for *all* measures $\mathfrak{c}_i$ (that have been proposed).[16]

What condition (17.1) expresses is that the claim that the evidence e about Linda is *not positively relevant* to the claim that Linda is a bank teller — *even if* it is presupposed that Linda is an active feminist. And, what condition (17.2) says is that the evidence e about Linda *is* positively evidentially relevant to the claim that Linda is

[16] We add the parenthetical caveat "that have been proposed" in (17), because it is theoretically possible to gerrymander bizarre relevance measures that violate (17). However, none of these bizarre measures is on the table in the contemporary dialectic. For present purposes, all that really matters is that result holds for *both* the *Likelihoodist* measure r, and all of the *non-Likelihoodist* measures (*e.g.,* d and l) that appear in the literature. This ensures that the Bayesian debates about the "conjunction fallacy" are (in an important sense) *orthogonal* to the Likelihoodism debate we've been discussing.

an active feminist (making *no* presuppositions about Linda). It seems that conditions (17.1) and (17.2) are (intuitively) *true* in the Linda case (and, in our experience, the vast majority of subjects are inclined to agree with this assessment). Thus, it follows from (17) that (16) must also be true in the Linda case — for *all* measures of confirmation$_i$ that have been proposed in the literature. We think this robust fact about the confirms$_i$-relations in the Linda case sheds light on why subjects may be caused to give incorrect answers to questions about the confirms$_f$-relations in the Linda case.[17] As the historical literature from philosophy of science reveals, the distinction between confirms$_f$ and confirms$_i$ is rather subtle.[18] As such, it wouldn't be very surprising if people had a tendency to make mistakes in cases where the two concepts come apart (in surprising ways).

Let's return, finally, to the philosophical dialectic concerning our various accounts of favoring and contrastive/comparative confirmation. The Linda case provides a nice illustration of the fact that evidence (e) can sometimes constitute *stronger evidence* for one hypothesis (h_2) than another (h_1), *even though* h_2 *is logically stronger than* h_1. And, unlike "irrelevant conjunction" cases, this assessment does *not* depend on one's choice of $\mathfrak{c}_i$-measure. *All* Bayesian confirmation-theorists should accept that e *confirms*$_i$/supports$_1$ h_2 *more strongly than* h_1 in the Linda case. Moreover, the betting/logical inference experiments ([47], [2]) indicate that actual subjects are *not* hearing the probability question as a "contrastive" one — *if* this requires a presupposition that the alternative hypotheses about Linda are *mutually exclusive*. This casts doubt on the Chandler-Hitchcock-style strategy (as applied to "irrelevant conjunction" cases) of substituting a contrastive question for a non-contrastive one in cases where there is a (*prima facie*) logical dependence between alternative hypotheses. In closing, I wonder whether subjects would be inclined to judge that e *favors* h_2 over h_1 in the Linda case. I conjecture that that actual subjects would *not* balk at such a claim (at least, not on grounds of non-mutual-exclusivity of the alternatives). Indeed, it seems to me quite natural to say that e *favors* h_2 over h_1 in the Linda case. Having said that, I should note that this is a burgeoning area of philosophical and psychological research. As such, much work remains to be done here — both on the philosophical side, and on the empirical side. The good news is that this provides some exciting opportunities for future collaborative research between philosophers and cognitive scientists.[19]

References

[1] D. Atkinson, J. Peijnenburg, and T. Kuipers. How to Confirm the Conjunction of Disconfirmed Hypotheses. *Philosophy of Science*, 76:1–21, 2009.

[17] It is worth noting that there are (equally plausible) *alternative* sets of sufficient conditions for our desired conclusion (16) that involve *only likelihoods* (relative to e) of different logical combinations of b and f [7, *fn.* 1]. So, even *Likelihoodists* should accept the conclusion that $\Pr(e \mid h_2) > \Pr(e \mid h_1)$ — *even if* they refuse to accept the claim that e *favors* h_2 over h_1, on the "constrastivist-presuppositional" grounds that such claims presuppose that h_1 and h_2 are mutually exclusive.

[18] Indeed, Popper's critique [37] of the first edition Carnap's *Logical Foundations of Probability* [3] was that *Carnap himself* had conflated confirms$_f$ and confirms$_i$. In the second edition of *LFP* [4], Carnap basically conceded this point to Popper. This is what led Carnap [4, new preface] to implore his readers to be careful about distinguishing confirms$_f$ and confirms$_i$ in the first place.

[19] In the last few years, there has been a flurry of both philosophical and psychological work on confirmation/support judgments and their relation to probability judgments. See [49], [43], [1], [42], and [48] for more on this recent strand of philosophical and cognitive-scientific research.

[2] N. Bonini, K. Tentori, and D. Osherson. A different conjunction fallacy. *Mind & Language*, 19(2):199–210, 2004.
[3] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, Chicago, first edition, 1950.
[4] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, Chicago, second edition, 1962.
[5] J. Chandler. Solving the tacking problem with contrast classes. *The British Journal for the Philosophy of Science*, 58(3):489–502, 2007.
[6] J. Chandler. Contrastive Confirmation: some Competing Accounts. To appear in *Synthese*, 2010.
[7] V. Crupi, B. Fitelson, and K. Tentori. Probability, confirmation, and the conjunction fallacy. *Thinking & Reasoning*, 14(2):182–199, 2008.
[8] V. Crupi, K. Tentori, and M. Gonzalez. On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues*. *Philosophy of Science*, 74:229–252, 2007.
[9] B. de Finetti. Initial probabilities: A prerequisite for any valid induction. *Synthese*, 20(1):2–16, 1969.
[10] B. de Finetti. Probabilism. *Erkenntnis*, 31(2):169–223, 1989.
[11] D. Dulany and D. Hilton. Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition*, 9(1):85–110, 1991.
[12] J. Earman. *Bayes or bust?* MIT Press, Cambridge, MA, 1992.
[13] B. Fitelson. A Bayesian account of independent evidence with applications. *Philosophy of Science*, 68(3):123–140, 2001.
[14] B. Fitelson. *Studies in Bayesian Confirmation Theory*. PhD thesis, University of Wisconsin–Madison, 2001.
[15] B. Fitelson. Putting the irrelevance back into the problem of irrelevant conjunction. *Philosophy of Science*, 69:611–622, 2002.
[16] B. Fitelson. Likelihoodism, Bayesianism, and relational confirmation. *Synthese*, 156(3):473–489, 2007.
[17] M. Forster. Bayes and Bust: Simplicity as a Problem for a Probabilist's Approach to Confirmation. *British Journal for the Philosophy of Science*, 46:399–424, 1995.
[18] C. Glymour. *Theory and evidence*. Princeton University Press, Princeton, 1980.
[19] I.J. Good. 46656 varieties of Bayesians. *The American Statistician*, 25(5):56–63, 1971.
[20] I. Hacking. The Leibniz-Carnap program for inductive logic. *The Journal of Philosophy*, 68:597–610, 1971.
[21] A. Hájek. What conditional probability could not be. *Synthese*, 137(3):273–323, 2003.
[22] A. Hájek. Interpretations of Probability. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2010 edition, 2010.
[23] J. Hawthorne. Degree-of-belief and degree-of-support: why Bayesians need both notions. *Mind*, 114(454):277–320, 2005.
[24] J. Hawthorne and B. Fitelson. Discussion: Re-solving Irrelevant Conjunction with Probabilistic Independence. *Philosophy of Science*, 71(4):505–514, 2004.
[25] C. Hempel. Studies in the logic of confirmation. *Mind*, 54:1–26, 97–121, 1945.
[26] C. Hitchcock. The role of contrast in causal and explanatory claims. *Synthese*, 107(3):395–419, 1996.
[27] C. Hitchcock. Contrastive explanation and the demons of determinism. *The British Journal for the Philosophy of Science*, 50(4):585–612, 1999.
[28] J. Joyce. *The Foundations of Causal Decision Theory*. Cambridge University Press, Cambridge, 1999.
[29] J. Joyce. On the Plurality of Probabilist Measures of Evidential Relevance. Presented at the *Bayesian Epistemology Workshop of the 26th International Wittgenstein Symposium*, Kirchberg, Austria (August 2003), unpublished manuscript, 2003.
[30] J. Keynes. *A Treatise on Probability*. Macmillan, London, 1921.
[31] P. Lipton. Contrastive explanation. *Royal Institute of Philosophy Supplement*, 27:247–266, 1990.
[32] P. Maher. Bayesianism and Irrelevant Conjunction. *Philosophy of Science*, 71:515–520, 2004.
[33] P. Maher. Bayesian probability. *Synthese*, 172(1):119–127, 2010.
[34] P. Milne. $\log[p(h/eb)/p(h/b)]$ is the one true measure of confirmation. *Philosophy of Science*, 63:21–26, 1996.
[35] R. Neta. What Evidence Do You Have? *The British Journal for the Philosophy of Science*, 59(1):89–119, 2008.

[36] G. Politzer and I. Noveck. Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research*, 20(2):83–103, 1991.
[37] K. Popper. Degree of confirmation. *The British Journal for the Philosophy of Science*, 5:143–149, 1954.
[38] F. Ramsey. Truth and probability. In R. Braithwaite, editor, *The Foundations of Mathematics and Other Logical Essays*, pages 156–198. Kegan, Paul, London, 1931.
[39] S. Rinard. Comments on Dan Steel's "Bayesian confirmation theory and the likelihood principle". Presented at *FEW 2005*, UT–Austin, 2005.
[40] R. Rosenkrantz. Bayesian confirmation: Paradise regained. *The British Journal for the Philosophy of Science*, 45:467–476, 1994.
[41] R. Royall. *Statistical Evidence: A Likelihood Paradigm*. Chapman & Hall, London, 1997.
[42] J. Schupbach. Is the conjunction fallacy tied to probabilistic confirmation? To appear in *Synthese*, 2010.
[43] T. Shogenji. The degree of epistemic justification and the conjunction fallacy. To appear in *Synthese*, 2010.
[44] E. Sober. Contrastive empiricism. In *From A Biological Point of View: Essays in Evolutionary Philosophy*, pages 114–135. Cambridge University Press, Cambridge, 1994.
[45] E. Sober. Is Drift a Serious Alternative to Natural Selection as an Explanation of Complex Adaptive Traits? *Royal Institute of Philosophy Supplement*, 80(56):10–11, 2006.
[46] D. Steel. Bayesian confirmation theory and the likelihood principle. *Synthese*, 156(1):53–77, 2007.
[47] K. Tentori, N. Bonini, and D. Osherson. The conjunction fallacy: a misunderstanding about conjunction? *Cognitive Science*, 28(3):467–477, 2004.
[48] K. Tentori and V. Crupi. How the conjunction fallacy is tied to probabilistic confirmation: Some remarks on Schupbach. To appear in *Synthese*, 2010.
[49] K. Tentori, V. Crupi, N. Bonini, and D. Osherson. Comparison of confirmation measures. *Cognition*, 103(1):107–119, 2007.
[50] A. Tversky and D. Kahneman. Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 91:293–315, 1983.
[51] J. Williamson. *In Defence of Objective Bayesianism*. Oxford University Press, Oxford, 2010.
[52] T. Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2002.

Philosophy Dept., Rutgers University, 1 Seminary Place, New Brunswick, NJ 08901-1107.
E-mail address: branden@fitelson.org
URL: http://fitelson.org/