Focused Correlation, Confirmation, and the Jigsaw Puzzle of Variable Evidence*

Maximilian Schlosshauer and Gregory Wheeler†‡

Focused correlation compares the degree of association within an evidence set to the degree of association in that evidence set given that some hypothesis is true. Wheeler and Scheines have shown that a difference in incremental confirmation of two evidence sets is robustly tracked by a difference in their focus correlation. In this essay, we generalize that tracking result by allowing for evidence having unequal relevance to the hypothesis. Our result is robust as well, and we retain conditions for bidirectional tracking between incremental confirmation measures and focused correlation.

1. Introduction. Auditors have found an irregularity in the books for Acme, Inc. Bungled paperwork explains most anomalies of this kind, but fraud cannot be ruled out, so teams are assembled to investigate. Team 1 learns that Carson, a longtime Acme bookkeeper, is divorcing his wife, which is sad news for the Carson family but of little interest to Team 1. Most people manage a divorce without defrauding their employer, after all. Team 2 learns that Carson has a Brazilian lover, but they merely pause to think, lucky Carson. Team 3 finds out that Carson took a flight to Rio de Janeiro seated beside a stunning but dodgy Brazilian financier.

Considered in isolation, each team's evidence signals little more than middle-life boredom. Yet, when taken together, the evidence suggests fraud. It is not the facts alone that warrant hauling Carson in for questioning but rather how those pieces of evidence fit together into a compelling case against him.

But what does it mean for evidence to 'fit together', and what is it

*Received April 2010; revised August 2010.

†To contact the authors, please write to: Gregory Wheeler, Center for Artificial Intelligence, New University of Lisbon, Portugal; e-mail: grw@fct.unl.pt.

‡This work was supported by the European Science Foundation, the Portuguese Science and Technology Foundation, and the Danish Natural Science Research Council.

Philosophy of Science, 78 (July 2011) pp. 376–392. 0031-8248/2011/7803-0003$10.00

about combined evidence that lends more support to a hypothesis than each piece of evidence on its own? One answer is to compare the likelihood of evidence occurring together to its co-occurrence given a cause or some unifying explanation. The events of the story are unusual on their own, but a conspiracy to defraud Acme involving Carson provides a unifying explanation for those events, one that would be alarming enough for Acme investigators to want a word with Mr. Carson.

The core idea animating this example is encapsulated in a measure called focused correlation (Myrvold 1996; Wheeler 2009), which is defined for n binary evidence variables and a single binary hypothesis variable as follows:

$$\mathrm{For}_{\boldsymbol{H}}(\boldsymbol{E}_1, \ldots, \boldsymbol{E}_n) = \frac{P(\boldsymbol{H}|\boldsymbol{E}_1 \cap \cdots \cap \boldsymbol{E}_n)P(\boldsymbol{H})^{n-1}}{P(\boldsymbol{H}|\boldsymbol{E}_1) \times \cdots \times P(H|\boldsymbol{E}_n)}$$

$$= \frac{P(\boldsymbol{E}_1 \cap \cdots \cap \boldsymbol{E}_n|\boldsymbol{H})/P(\boldsymbol{E}_1|\boldsymbol{H}) \times \cdots \times P(\boldsymbol{E}_n|\boldsymbol{H})}{P(\boldsymbol{E}_1 \cap \cdots \cap \boldsymbol{E}_n)/P(\boldsymbol{E}_1) \times \cdots \times P(\boldsymbol{E}_n)}. \tag{1}$$

Focused correlation has been given a variety of interpretations: to capture the effect that diversity of evidence has on confirmation of some hypothesis (Myrvold 1996); to provide a partial explication of the relationship between epistemic coherence and incremental confirmation of a hypothesis (Wheeler 2009); and, in normalized form, to provide an account of unified evidence for a hypothesis (Myrvold 2003). But, the basic idea is that by comparing the degree of association within an evidence set to the degree of association in that evidence set given some hypothesis, in certain circumstances several questions about confirmation relationships between evidence and that hypothesis can be answered.

The conditions under which focused correlation tracks confirmation, pace Myrvold (2003), depend on what type of logical or causal structure governs the relationships between evidence and hypothesis in a problem (Wheeler and Scheines 2011). For example, Erik Olsson (2005) and Luc Bovens and Stephan Hartmann (2006) have drawn attention to problematic features of models in which all binary evidence variables $\boldsymbol{E}_i$ of a problem are conditionally independent of the binary hypothesis variable $\boldsymbol{H}$; that is, $\forall i \neq j, \boldsymbol{E}_i \perp\!\!\!\perp \boldsymbol{E}_j|\boldsymbol{H}$. The correlation between any distinct pair of evidence variables $\boldsymbol{E}_i$ and $\boldsymbol{E}_j$ in this class of models is the product of the correlations between each $\boldsymbol{E}$ and $\boldsymbol{H}$ (Danks and Glymour 2001). But this fact yields counterintuitive results if one interprets association as epistemic coherence since then association among the evidence is strictly less than the (nondeterministic) confirmation that each piece of evidence lends to the hypothesis.

In the case of focused correlation, it is easy to see from equation (1) that the focused correlation of an evidence set with respect to some hypothesis under this conditional independence assumption is strictly less than one, and thus is negative, whereas the confirmation of the hypothesis by that evidence may be positive. This conditional independence condition is central to the witness testimony models of Bovens and Hartmann (2003) and Olsson (2005), and it is the condition that drives their noted negative results concerning the construction of a probabilistic measure of coherence (Wheeler 2009).

But, contra Olsson (2005, 135), there are interesting cases outside of this class of models where a measure like focused correlation can be useful, and, pace Luc Bovens and Stephan Hartmann (2006), there are conditions under which a difference in confirmation lent to some hypothesis by two evidence sets is tracked by a difference in their focused correlation, even within this obstinate class of models.[1] To illustrate, consider again *L'affaire de Carson*. Suppose that a performance review of the Acme Investigation Department involves comparing the evidence set **E**, consisting of a divorcing employee who traveled with a dubious associate, to the evidence set **E***, consisting of a divorcing employee who is having an affair. Acme is a big company with many cases on file, and their models suggest that **E** is a better indicator of fraud than **E*** because fraud is a slightly 'better' explanation for the co-occurrence of divorce and suspicious travel than it is for divorce and an adulterous affair. Under what conditions can Acme management draw this inference?

In Wheeler and Scheines (2011), sufficient conditions for a wide variety of incremental confirmation measures are given, under which a difference in focused correlation entails a difference in confirmation lent to that hypothesis and vice versa. But their results are only trivially true in this problematic class of models because one of the conditions for the comparison result is that all individual items of evidence have equal relevance to the hypothesis. In effect, this restrictive condition, when combined with the conditional independence assumption mentioned above, deals with the problematic cases driving Bovens and Hartmann's counterexamples (2003, 2006), by ensuring that at least one of the antecedent conditions is false. The motivation for this strong assumption is clear in their project, for they address whether a difference in association of two evidence sets induced by some hypothesis can track confirmation of that hypothesis.

1. Observe that this problematic class of models may be characterized by those in which this conditional independence condition holds. Hence, that class of models—provided that the assumptions needed to activate the theory of graphical causal models are satisfied—can be identified nonparametrically and handled by other means. So, while it is true that no probabilistic measure of coherence can operate correctly within this class, that is only one way to look at it; another is to observe that there is no need for a coherence measure in this class of models.

But including equal relevance as a ceteris paribus condition will not do for practical comparisons of evidence sets; it is much too restrictive.

In this article, we show that the equal evidence assumption can indeed be relaxed to identifiable cases involving unequal relevance. Our result is robust as well, and we retain conditions for bidirectional tracking between the class of incremental confirmation measures and focused correlation. The cases in which tracking fails are instructive, suggesting that focused correlation is a stronger measure of evidential support than classical incremental confirmation measures, and we illustrate this point through some examples.

The structure of the article is as follows. In section 2, we set up the machinery by making explicit our notation and the confirmation measures we consider. Section 3 presents our results, and in section 4, we explain why two common nonincremental confirmation measures are not tracked by focused correlation. This analysis offers further evidence for distinguishing between classical, incremental confirmation measures and other notions of confirmation. Proofs are set in the appendix.

2. Setup. In this section we specify notation and list the confirmation functions we will consider.

Define a *probability space* $(\Omega, \mathcal{F}, P)$ such that $\mathcal{F}$ is a σ-algebra over a set Ω, and $P{:}\mathcal{F} \rightarrow [0,1]$ is a probability measure defined on the space $(\Omega, \mathcal{F}, P)$ satisfying

P1. $P(\Omega) = 1$.

P2. $P(\cup_{i=1}^{\infty} X_i) = \sum_{i=i}^{\infty} P(X_i)$, when X_i are countable, pairwise disjoint elements of $\mathcal{F}$.

A *probability structure* is a tuple $M = (\Omega, \mathcal{F}, P, V)$, where $(\Omega, \mathcal{F}, P)$ is a probability space and V is an interpretation function associating each element $\omega \in \Omega$ with a truth assignment on the primitive propositions in $A, B, \ldots \in \Phi$ such that $V(\omega)(A) \in \{1, 0\}$ for each $\omega \in \Omega$ and for every proposition in Φ.

For each primitive proposition in Φ, we define $M, \omega \models A$ if and only if (iff) $V(\omega)(A) = 1$ and proceed by induction on the structure of propositional formulas. Since P is defined on events in $\mathcal{F}$ rather than propositions, let $[[A]]_M$ denote the set of outcomes within Ω in M where A is true, which will correspond to a subset of $\mathcal{F}$. The following equations make explicit the relationship between propositions and events for arbitrary propositional formulas A and B:

i) $[[A \wedge B]]_M = [[A]]_M \cap [[B]]_M$,

ii) $[[A \vee B]]_M = [[A]]_M \cup [[B]]_M$,

iii) $[[\neg A]]_M = \overline{[[A]]}_M$.

An evidence set is a set of propositions, written $\mathbf{E} = \{E_1, E_2\}$. We defined focused correlation in equation (1) with respect to random variables $\boldsymbol{E}_i$ and $\boldsymbol{H}$ each taking values 0 or 1. We use the abbreviation E for $\boldsymbol{E} = 1$ and $\neg E$ for $\boldsymbol{E} = 0$ and likewise will use the abbreviation $P(E_1 \wedge \neg E_2|H)$ for $P(\boldsymbol{E}_1 = 1 \cap \boldsymbol{E}_2 = 0|\boldsymbol{H} = 1)$. We will relax notation by expressing focused correlation of propositions as well as of variables, using the conventions here for distinguishing between propositions and variables to signal which is which.

2.1. Confirmation Measures. A statement E is confirmation to hypothesis H with respect to a classical probability model specifying the measure P just when E and H are positively correlated under P; that is, when $P(H|E) > P(H)$. A confirmation function measures the degree to which evidence confirms a hypothesis, and there are several proposals:[2]

$$inc_1(H, E_1, E_2) := P(H|E_1 \wedge E_2) - P(H|E_1).$$

$$inc_2(H, E_1, E_2) := \frac{P(H|E_1 \wedge E_2) - P(H|E_1)}{1 - P(H|E_1)}.$$

$$ko(H, \mathbf{E}) := \frac{P(\mathbf{E}|H) - P(\mathbf{E}|\neg H)}{P(\mathbf{E}|H) + P(\mathbf{E}|\neg H)}.$$

$$r(H, \mathbf{E}) := \log \frac{P(H|\mathbf{E})}{P(H)}.$$

$$l(H, \mathbf{E}) := \log \frac{P(\mathbf{E}|H)}{P(\mathbf{E}|\neg H)}.$$

2.2. Comparing Equal-Relevance Evidence Sets. In our discussion throughout, we assume that $P(\mathbf{D})$, a probability distribution defined over a domain of propositions $\langle H, \mathbf{E} \rangle$, is positive. In addition, we may appeal to two conditions.

2. Measures inc_1 and inc_2 are both variants of L. Jonathan Cohen's notion of *incremental convergence* (Cohen 1977): inc_1 reports the contribution that E_2 simpliciter makes to H, whereas inc_2 reports the contribution E_2 makes with respect to the possible available evidence; inc_1 and inc_2 are the two cases that comprise measure Z (Crupi, Tentori, and Gonzalez 2007), where $inc_2(H, E)$ is used if $P(H|E) > P(H)$, and $inc_1(H, E)$ is used otherwise. The Kemeny and Oppenheim fitness measure is *ko* (1952); *r* is a generic relevance measure, versions of which have been endorsed from Keynes (1921) to Milne (1997), among others; *l* is ordinally equivalent to *ko*; and *ko*, *r*, and *l* are typically discussed for evidence E representing an evidence set $\mathbf{E}$ of arbitrary size by a single conjunction of the propositions in $\mathbf{E}$. We will discuss cases where $|\mathbf{E}| = 2$. See Kyburg (1983) for an overview of confirmation measures and Eells and Fitelson (2002) for a recent discussion.

(A1) **Positive Relevance**: All propositions in **E** are positively relevant to H, just in case $\forall E_i \in \mathbf{E}, P(H|E_i) > P(H) > P(H|\neg E_i)$.
(A2) **Equal Relevance**: All propositions in a set of evidence **E** are equally confirmatory, just in case $\forall E_i, E_j \in \mathbf{E}, P(H|E_i) = P(H|E_j)$, and $P(H|\neg E_i) = P(H|\neg E_j)$.

First, evidence and hypothesis may be correlated or anticorrelated, unless independent. Thus, evidence may be either positively relevant or negatively relevant to a hypothesis, if it is relevant at all. Condition A1 restricts attention to cases where evidence is positively relevant. Second, strength of relevance need not be the same, but condition A2 restricts attention to just those cases where all pieces of evidence are equally relevant to the hypothesis.[3] This is the condition we shall explore how to weaken. But first, let's see the benefit from a positive distribution over **D** that satisfies both A1 and A2.

Proposition 1 (Wheeler and Scheines 2011). If $\mathbf{E} = \{E_1, E_2\}$ and $\mathbf{E}^* = \{E_1, E_3\}$, and $\mathbf{E} \cup \mathbf{E}^*$ satisfies A1 and A2 with respect to H, then all of the following inequalities are equivalent:[4]

- $\mathrm{For}_H(\mathbf{E}) > \mathrm{For}_H(\mathbf{E}^*)$
- $r(H, \mathbf{E}) > r(H, \mathbf{E}^*)$
- $l(H, \mathbf{E}) > l(H, \mathbf{E}^*)$
- $ko(H, \mathbf{E}) > ko(H, \mathbf{E}^*)$
- $inc_1(H, \mathbf{E}) > inc_1(H, \mathbf{E}^*)$
- $inc_2(H, \mathbf{E}) > inc_2(H, \mathbf{E}^*)$

Proposition 1 tells us that focused correlation tracks incremental confirmation and vice versa, but the proof leans on A2 to isolate the effect of focused correlation on incremental confirmation. This condition is too restrictive to exploit in an application, and, as observed in the introduction, it deals with a problematic class of models by effectively excluding them. In the next section, we consider how to preserve these tracking properties of focused correlation without condition A2.

3. Unequal Relevance. We consider, as before, two evidence sets $\mathbf{E} = \{E_1, E_2\}$ and $\mathbf{E}^* = \{E_1, E_3\}$, but now we shall relax assumption A2: we no longer require that all propositions E_i, $i = 1, 2, 3$, are equally confir-

3. See lemma 2 and remarks in the appendix about the strength of these two assumptions when combined.

4. For the proposition to hold for inc_1, we stipulate that $inc_1(H, \mathbf{E}) > inc_1(H, \mathbf{E}^*)$ stands for $inc_1(H, E_1, E_2) > inc_1(H, E_1, E_3)$. A similar remark applies for interpreting inc_2.

matory. That is, we allow for the degree of positive relevance $P(H|E_i)$ of the individual pieces of evidence to vary.

What we would like to check now is whether, in this generalized situation, focused correlation continues to track confirmation, and vice versa. That is, we would like to explore under what conditions proposition 1 still holds. 'Confirmation' is here understood as quantified by the measures defined in section 2.1. We shall denote these measures collectively by $c(H, \mathbf{E})$ in what follows; that is, $c(H, \mathbf{E})$ jointly stands for any of the functions $inc_1(H, \mathbf{E})$, $inc_2(H, \mathbf{E})$, $ko(H, \mathbf{E})$, $r(H, \mathbf{E})$, and $l(H, \mathbf{E})$.

A quick look at the definitions of these confirmation measures (sec. 2.1) shows that they are based only on the probabilities of H conditional on the full evidence set and on the shared evidence piece E_1 but not on the probabilities of H conditional on the distinct individual pieces E_2 and E_3, that is, not on $P(H|E_2)$ and $P(H|E_3)$. The appearance of $P(H|E_2)$ and $P(H|E_3)$ in equation (1) is therefore a particular feature of focused correlation. This feature exemplifies the additional input to the measure, which takes into account the conditional probabilities $P(H|E_i)$ for all $E_i \in \mathbf{E}$.

As an aside, we note that the relevance $P(H|E_1)$ of the evidence E_1 will not enter into the following considerations, not withstanding its role in inc_1 and inc_2. This is so because E_1 is shared among both evidence sets $\mathbf{E}$ and $\mathbf{E}^*$, and we are here concerned only with a comparison of the degrees of focused correlation and confirmatory power of these two sets.

3.1. From Focused Correlation to Confirmation. Let us first describe the condition under which a larger degree of focused correlation for evidence set $\{E_1, E_2\}$ than for $\{E_1, E_3\}$ implies larger confirmation of the hypothesis by $\{E_1, E_2\}$ than by $\{E_1, E_3\}$.

Proposition 2. If $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3)$ and

$$\frac{P(H|E_2)}{P(H|E_3)} > \frac{\mathrm{For}_H(E_1, E_3)}{\mathrm{For}_H(E_1, E_2)}, \tag{2}$$

then also $c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$.

Thus, the confirmation measures $c(H, \{E_1, E_2\})$ track focused correlation if the bound (2) is fulfilled. In particular, this will be the case whenever $P(H|E_2) \geq P(H|E_3)$, but the implication $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3) \Rightarrow c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$ will also hold if $P(H|E_2) < P(H|E_3)$ by an amount that is a function of the difference in the focused correlations of the two sets $\mathbf{E}$ and $\mathbf{E}^*$, as specified by (2). However, if $P(H|E_2)$ drops below this bound, then the tracking fails, and the confirmation measures $c(H, \{E_1, E_2\})$ will indicate less confirmatory power for $\mathbf{E}$ than $\mathbf{E}^*$, despite the fact that $\mathbf{E}$ has larger focused correlation than $\mathbf{E}^*$.

This observation has an intuitive explanation, as $c(H, \{E_1, E_2\}) < c(H, \{E_1, E_3\})$ is equivalent to $P(H|E_1 \wedge E_2) < P(H|E_1 \wedge E_3)$ when both A1 and A2 hold. Now, as mentioned above, focused correlation goes a step further than the measures $c(H, \{E_1, E_2\})$ because it weighs the joint relevance $P(H|\mathbf{E})$ by how much each individual piece of evidence confirms H. If E_1 and E_2 individually confirm H very little compared to the confirmation of H provided by the conjunction $E_1 \wedge E_2$, while the probability of H given $E_1 \wedge E_3$ does not exceed by much the product of the probabilities of H given E_1 and E_3 alone, then the set $\{E_1, E_2\}$ may have larger focused correlation than the set $\{E_1, E_3\}$, even if $P(H|E_1 \wedge E_2) < P(H|E_1 \wedge E_3)$. In other words, $\text{For}_H(E_1, E_2) \geq \text{For}_H(E_1, E_3)$ may arise from a situation in which $P(H|E_2)$ is so much smaller than $P(H|E_3)$ as to be able to counteract the greater confirmatory power of $\mathbf{E}^*$ over $\mathbf{E}$ as quantified by the measures $c(H, \{E_1, E_2\})$.

We will illustrate our results by presenting alternative versions of Mr. Carson's scandal. In each of our examples, we shall let E_1 represent the fact that Carson is divorcing his wife, E_2 = Carson's trip abroad seated next to a shady financier, E_3 = Carson's affair with a Brazilian lover, and H = Carson swindled Acme. We begin by considering the case in which the confirmation measures $c(H, \mathbf{E})$ fail to track focused correlation.

Example 1. Suppose that the evidence set $\{E_1$ = divorce, E_2 = travel$\}$ has larger focused correlation than $\{E_1$ = divorce, E_3 = affair$\}$, with respect to the hypothesis H. We interpret this relation as saying that evidence of both Carson's divorce proceedings and his travel cohere better toward supporting the hypothesis of fraud than evidence of Carson's divorce and affair.

But this fact does not necessarily mean that the coincidence of divorce and travel is also more confirmatory of the hypothesis of fraud than the divorce-plus-affair scenario. Why not?

Imagine a backstory such that Carson's affair, considered in isolation, supports fraud much more strongly than a trip to Brazil, considered in isolation. Perhaps poor Carson is in over his head: his affair is a surprise to all and an expensive one, at that. His trip to Rio, although unusual, pales by comparison, and on its own there is no reason to think that his having a shady seat companion is anything but a coincidence. Relative to this background knowledge, A, $P_A(H|E_3) > P_A(H|E_2)$. Now add the news of Carson's divorce proceedings. This would not dent the confirmatory power of the affair scenario since adding a divorce to that affair only piles on more expenses for Carson. But learning of the divorce would also do little to alter the confirmatory force of the travel scenario on judging whether Carson is complicit in defrauding Acme. Thus, $P_A(H|\{E_1, E_3\}) > P_A(H|\{E_1, E_2\})$.

But then there is little reason to believe that we need the coincidence of a divorce and an affair to send up a red flag in Acme's fraud department, as opposed to just knowing of Carson's liaison. The coherence of divorce and affair may therefore be less than the coherence of divorce and travel with respect to confirming the hypothesis of fraud, and still the conjunction of divorce and affair may be more confirmatory of fraud than the conjunction of divorce and travel. As we have seen, this happens when the confirmatory power of the affair (considered by itself) becomes sufficiently large in comparison with the confirmatory power of the travel (considered by itself). When coherence is measured by focused correlation, what qualifies as 'sufficiently large' is then precisely quantified by the bound (2).

3.2. From Confirmation to Focused Correlation. Let us now tackle the converse relation, namely, the implication from a difference in confirmation to a difference in focused correlation.

Proposition 3. If $c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$ and

$$P(H|E_2) \leq P(H|E_3), \tag{3}$$

then also $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3)$.

The bound (3) is a catchall condition that holds for all confirmation measures $c(H, \mathbf{E})$ considered in section 2.1. It may be tightened by focusing on a particular measure. For example, for $inc_1(H, \mathbf{E})$ and $inc_2(H, \mathbf{E})$, the bound is

$$\frac{P(H|E_2)}{P(H|E_3)} < \frac{inc_k(H, E_1, E_2)}{inc_k(H, E_1, E_3)}, \qquad k = 1, 2, \tag{4}$$

which allows for $P(H|E_2) > P(H|E_3)$ within certain limits.

The existence of a bound on the ratio of $P(H|E_2)$ to $P(H|E_3)$ has a similar explanation as the bound (2). If $P(H|E_2) > P(H|E_3)$ by a sufficiently large amount, then the larger confirmatory power of $\mathbf{E}$ over $\mathbf{E}^*$ may be overshadowed by the fact that the confirmation of H by E_3 alone, with respect to the confirmation by the conjunction $E_1 \wedge E_3$, is much less than the confirmation of H by E_2 alone, with respect to the confirmation by the conjunction $E_1 \wedge E_2$. Or, put more plainly, the beliefs in the set $\mathbf{E}^*$ cohere better toward supporting H than do the beliefs in the set $\mathbf{E}$.

Example 2. In the spirit of the introduction, suppose that Carson's affair is viewed as an ordinary disaster, and his travel plans are more suspicious—it is known that he booked both his seat and his companion's, let's say. So, suppose that the conjunction of divorce and travel is more confirmatory of the hypothesis of fraud than the conjunction of divorce and an affair on this background knowledge, labeled B: $c_B(H, \{E_1, E_2\}) > c_B(H, \{E_1, E_3\})$, and $P_B(H|E_1 \wedge E_2) > P_B(H|E_1 \wedge E_3)$.

Now, what happens when we consider focused correlation instead of confirmation? Equation (1) weighs $P_B(H|E_1 \wedge E_2)$ and $P_B(H|E_1 \wedge E_3)$ by how much each of the individual pieces of evidence differing between the two evidence sets—that is, Carson's travel (E_2) and affair (E_3)—confirms the suspicion of fraud. If condition (3) is violated, that is, if $P_B(H|E_2) > P_B(H|E_3)$, this would mean that isolated knowledge of Carson's travel makes us more inclined to think of the possibility that he may be engaged in fraudulent business than if we simply heard that Carson is embroiled in an affair. It is the merit of focused correlation to now put up a cautionary flag when it comes to considering the implications for the relative degrees of coherence of the two evidence sets $\{E_1 = \text{divorce}, E_2 = \text{travel}\}$ and $\{E_1 = \text{divorce}, E_3 = \text{affair}\}$. How so?

The argument is similar to the discussion of the previous example in section 3.2. Since according to background B the evidence of travel, considered in isolation, is already more supportive of the hypothesis of fraud than the evidence of an affair, also considered in isolation, and since the only other piece of evidence—namely, the fact that Carson is in the midst of divorce proceedings—is common to both evidence sets, then one should not automatically conclude that it is the coincidence of divorce and travel that boosts our confidence in fraud. Rather, it may be the case that the evidence of travel alone is what makes the hypothesis so eminently plausible, not the fact that it coincides with Carson's divorce. Conversely, while Carson's affair alone may barely raise suspicions of fraud, its coincidence with Carson's being in the midst of a messy—and potentially financially threatening—divorce may boost the likelihood of fraud by a large margin. In other words, although the coincidence of Carson's divorce and travel is more confirmatory of fraud than the coincidence of his divorce and affair, the latter evidence set may well be judged to have a larger degree of coherence. Focused correlation captures this subtlety.

Conversely, if condition (3) is obeyed, that is, if $P_B(H|E_2) \leq P_B(H|E_3)$, focused correlation and confirmation will gravitate toward favoring the same evidence set. Together with the fact that $P_B(H|E_1 \wedge E_2) > P_B(H|E_1 \wedge E_3)$, we can now safely conclude that divorce and travel cohere better toward supporting the hypothesis of fraud than the conjunction of divorce and affair: although evidence of travel alone is less (or equally) indicative of fraud compared with the evidence of an affair, when taken together with the evidence of an ongoing divorce, Carson's travel will ring Acme's alarm bells more readily than evidence of Carson's involvement with his Brazilian sweetheart.

3.3. Bidirectional Tracking. Propositions 2 and 3 show that when the ratio of $P(H|E_2)$ to $P(H|E_3)$ is within a certain range, focused correlation will be tracked by confirmation, and when it is within another range, focused correlation will track confirmation. It is important to note that these ranges overlap. Propositions 2 and 3 readily quantify this overlap, that is, the range of values of $P(H|E_2)/P(H|E_3)$, for which bidirectional tracking holds. Combining propositions 2 and 3 yields:

Corollary 1. If

$$\frac{\text{For}_H(E_1, E_3)}{\text{For}_H(E_1, E_2)} < \frac{P(H|E_2)}{P(H|E_3)} \leq 1, \tag{5}$$

then $\text{For}_H(E_1, E_2) > \text{For}_H(E_1, E_3) \Leftrightarrow c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$.

That is, ensuring that the tracking goes both ways does not require us to impose the rather strong assumption A2 of equal relevance of all bits of evidence.

Choosing a particular confirmation measure allows us to extend even further the range over which the bidirectional tracking holds. For example, for $inc_1(H, \mathbf{E})$ or $inc_2(H, \mathbf{E})$, we can use the bound (4) instead of (3), which gives a 'relaxed' version of corollary 1:

Corollary 2. If

$$\frac{\text{For}_H(E_1, E_3)}{\text{For}_H(E_1, E_2)} < \frac{P(H|E_2)}{P(H|E_3)} < \frac{inc_k(H, E_1, E_2)}{inc_k(H, E_1, E_3)}, \qquad k = 1, 2, \tag{6}$$

then $\text{For}_H(E_1, E_2) > \text{For}_H(E_1, E_3) \Leftrightarrow inc_k(H, \{E_1, E_2\}) > inc_k(H, \{E_1, E_3\})$.

Figure 1 illustrates the different ranges of $P(H|E_2)/P(H|E_3)$ for which uni- and bidirectional tracking between focused correlation and confirmation obtains.

3.4. Equality of Focused Correlation and Confirmation. For the sake of completeness, let us also explicitly consider the connection between equal values of focused correlation and equal values of confirmation:

Proposition 4. $\text{For}_H(E_1, E_2) = \text{For}_H(E_1, E_3) \Leftrightarrow c(H, \{E_1, E_2\}) = c(H, \{E_1, E_3\})$ iff

$$P(H|E_2) = P(H|E_3). \tag{7}$$

In other words, equal values of focused correlation for the two sets $\mathbf{E}$ and $\mathbf{E}^*$ translate into equal degrees $c(H, \{E_1, E_2\})$ and $c(H, \{E_1, E_3\})$ of confirmation—and vice versa—if and only if the equal-relevance assumption A2 is made. The reason for this result should now be clear from the above discussion of propositions 2 and 3.

4. Discussion. In this article, we have demonstrated that the equal-rele-

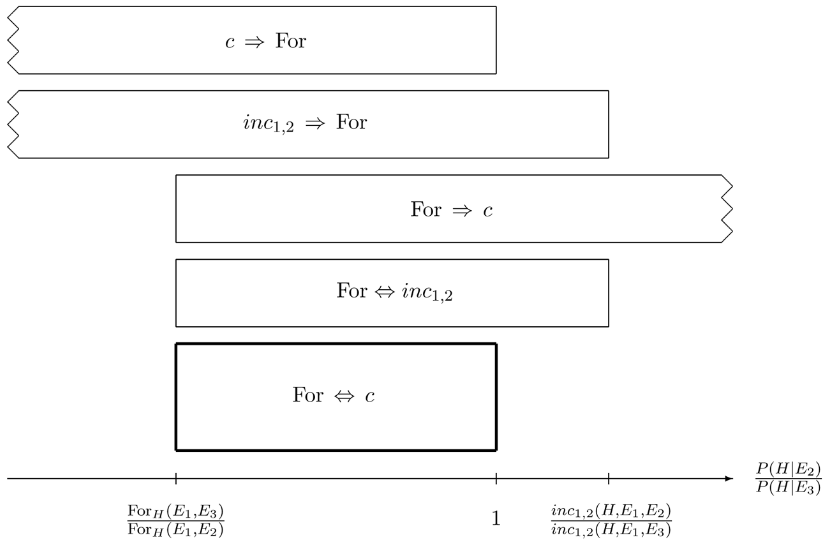

Figure 1. Regimes of $P(H|E_2)/P(H|E_3)$ for which focused correlation tracks or is tracked by confirmation measures. The relevance of a piece of evidence E_i for a given hypothesis H is $P(H|E_i)$, and 'For $\Rightarrow c$' is shorthand for the implication $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3) \Rightarrow c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$, which represents a tracking of focused correlation by the confirmation measures $c(H, \mathbf{E})$. Similarly for '$c \Rightarrow$ For'. Over a range of values for $P(H|E_2)/P(H|E_3)$, bidirectional tracking obtains (*bold box*); that is, $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3) \Leftrightarrow c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$. For specific choices of $c(H, \mathbf{E})$, for example, for the measures $inc_1(H, \mathbf{E})$ or $inc_2(H, \mathbf{E})$ shown here, the tracking range is extended further.

vance assumption can be relaxed within certain limits without upsetting the bidirectional tracking between focused correlation and several incremental confirmation measures; if only unidirectional tracking is required, these limits become even less stringent. However, if the degrees of relevance diverge too strongly, the tracking fails. We have shown that the ratio of $P(H|E_2)$ to $P(H|E_3)$ is crucial in assessing, both qualitatively and quantitatively, when and where the tracking fails.

However, this observation should not be viewed as reflecting a failure or deficiency of the measure of focused correlation. Rather, it may be considered as illuminating a shortcoming of incremental confirmation measures since they fail to take into account the additional information provided by the relevance of each individual bit of evidence. As we have seen, this information is important in flagging a situation in which a large boost in confirmation caused by adding a piece of evidence has nothing or only little to do with the coincidence of this evidence with the existing body of evidence. One merit of focused correlation is to detect such cases.

In one class of cases that has received a lot of attention, namely, the single-factor common-cause model mentioned in the introduction, any measure linking positive association to positive confirmation breaks down, and comparisons (assuming A2) are never false because they cannot be different. But these facts take nothing away from focused correlation, for two reasons. First, because this class of models is determined nonparametrically, we can learn whether one hypothesis is a common cause (Silva et al. 2006) and then attack the problem by other methods—namely, by looking at the relative strengths of the evidence. We do not need the sign of focused correlation in single-factor common-cause models to indicate the sign of confirmation if we understand clearly when the two do not align. In the case of single-factor common-cause models, the correlation of the evidence is completely determined by the relevant strengths of the evidence (Danks and Glymour 2001). Faulting focused correlation for failing to indicate confirmation in this class of models is akin to faulting screwdrivers for failing to drive nails: it is a true claim that is entirely beside the point. Second, it should be clear that the failure of focused correlation to track differences in confirmation of evidence in common-cause models in proposition 1 is driven by A2 rather than a deficiency in focused correlation. But, again, although our results here allow for some tracking in common-cause models, at bottom a difference in confirmation in this class of models boils down to a difference in the individual strengths, $P(H|E_i)$.

Focused correlation is defined generally for n evidence variables, but we only discuss evidence sets of size two, and difficulties loom for attempts to make comparisons of larger evidence sets (Bovens and Hartmann 2006). Even though focused correlation is defined for arbitrary-sized information sets, the relationships of confirmation, covariance, and correlation are binary—or three-place in conditional form. Thus, the expansion of evidence sets beyond size two will require a partition of the evidence set since there are many incremental confirmation questions that are compatible with one focused correlation problem involving an evidence set of a size greater than two. To expect otherwise is a category mistake, and negative results should be no surprise. That said, if there is not an interest in linking focused correlation for large evidence sets—which, technically, is not a measure of correlation but is instead a distance-from-independence measure—to confirmation, then one may explore extending the results presented here to facilitate direct comparisons of large evidence sets.

Finally, two commonly discussed confirmation measures have been omitted from discussion, Carnap's (1962) relevance measure r and the old evidence measure *oe* (Christensen 1999; Joyce 1999), which are defined as

$$\tau(H, E)P(H \wedge E) - P(H)P(E)$$

and

$$oe(H, E)P(H|E) - P(H|\neg E).$$

Neither measure is tracked by focused correlation. The reason is that both measures are less constrained than incremental confirmation measures. In the case of Carnap's measure of relevance, observe that $\tau(H, E)$ is the covariance of the binary variables H and E:

$$\mathrm{Cov}(H, E) = P(H \wedge E) - P(H)P(E) = P(H)(P(E|H) - P(E)).$$

Although positive covariance for H and a single evidence variable E is ensured by the positive relevance condition A1, conditions A1 and A2 and equation (9a) do not constrain the sign of $\tau(H, \mathbf{E})$ for evidence sets $\mathbf{E}$ of size two. In the case of *oe*, there is some recognition already of the difference between incremental confirmation measures and measures for novelty of evidence like *oe* (Joyce 1999). The failure to compare evidence sets by their degree of surprise, particularly under the restrictive conditions of positivity, A1, A2, and equation (9a), is an additional reason to sharply distinguish between these two types of measures. Countermodels for *oe* are discussed in the appendix.

Appendix

The following proofs of propositions 2–4 are similar in spirit and rely heavily on lemma 1, which follows from the definition of $\mathrm{For}_H(E_1, E_2)$ in equation (1).

Lemma 1. If $\mathrm{For}_H(E_1, E_2) = \lambda \mathrm{For}_H(E_1, E_3)$ with $\lambda > 0$, then

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} = \lambda \frac{P(H|E_2)}{P(H|E_3)}. \tag{8}$$

Proof of Lemma 1.

$$\frac{P(H|E_1 \wedge E_2)P(H)}{P(H|E_1)P(H|E_2)} = \lambda \frac{P(H|E_1 \wedge E_3)P(H)}{P(H|E_1)P(H|E_3)}$$

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_2)} \times \frac{P(H)}{P(H|E_1)} = \lambda \frac{P(H|E_1 \wedge E_3)}{P(H|E_3)} \times \frac{P(H)}{P(H|E_1)}$$

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} = \lambda \frac{P(H|E_2)}{P(H|E_3)}.$$

QED

Proof of Proposition 2. We begin with two observations, assuming A1 and A2:

$$P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3) \Leftrightarrow c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\}), \quad \text{(9a)}$$
$$P(H|E_1 \wedge E_2) = P(H|E_1 \wedge E_3) \Leftrightarrow c(H, \{E_1, E_2\}) = c(H, \{E_1, E_3\}). \quad \text{(9b)}$$

Turning to $\mathrm{For}_H(\mathbf{E})$, by A2, $P(H|E_2)/P(H|E_3) = 1$. So, by lemma 1, the left-hand side of equations (9a) and (9b) may be expressed by equation (10) when $\lambda > 1$ and $\lambda = 1$, respectively.

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} = \lambda. \quad (10)$$

Without A2, observe that

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} = \lambda \frac{P(H|E_2)}{P(H|E_3)},$$

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} \times \frac{P(H|E_3)}{P(H|E_2)} = \lambda,$$

$$\frac{P(H|E_1 \wedge E_2)P(H)}{P(H|E_1)P(H|E_2)} \times \frac{P(H|E_1)P(H|E_3)}{P(H|E_1 \wedge E_3)P(H)} = \lambda, \quad (11)$$

$$\frac{\mathrm{For}_H(E_1, E_2)}{\mathrm{For}_H(E_1, E_3)} = \lambda.$$

So $\lambda > 1$, given (11), and $\mathrm{For}_H(E_1, E_2) > \mathrm{For}_H(E_1, E_3)$, by hypothesis. Therefore, $P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3)$, and thus $c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$ iff

$$\frac{P(H|E_2)}{P(H|E_3)} > \frac{1}{\lambda} = \frac{\mathrm{For}_H(E_1, E_3)}{\mathrm{For}_H(E_1, E_2)}. \quad (12)$$

QED

Proof of Proposition 3. By assumption, $c(H, \{E_1, E_2\}) > c(H, \{E_1, E_3\})$, and thus $P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3)$ from equations (9a) and (9b). Then by lemma 1

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_1 \wedge E_3)} > \frac{P(H|E_2)}{P(H|E_3)},$$

$$\frac{P(H|E_1 \wedge E_2)}{P(H|E_2)} > \frac{P(H|E_1 \wedge E_3)}{P(H|E_3)}.$$

We can tighten the bound if we choose a particular confirmation measure. For example,

$$inc_k(H, E_1, E_2) = \gamma\, inc_k(H, E_1, E_3)$$

$$\Leftrightarrow\ P(H|E_1 \wedge E_2) = \gamma P(H|E_1 \wedge E_3), \qquad k = 1, 2,$$

and thus the bound is

$$\frac{P(H|E_2)}{P(H|E_3)} < \gamma = \frac{inc_k(H, E_1, E_2)}{inc_k(H, E_1, E_3)}, \tag{13}$$

which proves equation (4). QED

Proof of Proposition 4. If $\mathrm{For}_H(E_1, E_2) = \mathrm{For}_H(E_1, E_3)$, then definition (1) implies that $P(H|E_1 \wedge E_2) = P(H|E_1 \wedge E_3)$ iff $P(H|E_2) = P(H|E_3)$. Conversely, if $c(H, \{E_1, E_2\}) = c(H, \{E_1, E_3\})$, then equation (9b) implies that $P(H|E_1 \wedge E_2) = P(H|E_1 \wedge E_3)$, and definition (1) then shows that $\mathrm{For}_H(E_1, E_2) = \mathrm{For}_H(E_1, E_3)$ iff $P(H|E_2) = P(H|E_3)$. QED

Countermodels for Measure Tracking *oe*. To show that $P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3)$ does not entail $oe(H, \{E_1, E_2\}) > oe(H, \{E_1, E_3\})$ under positivity, A1, and A2, it suffices to provide a model in which

$$P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3) \tag{14}$$

but

$$\begin{aligned} &P(H|E_1 \wedge E_2) - P(H|\neg E_1 \vee \neg E_2) \\ \ngtr\ &P(H|E_1 \wedge E_3) - P(H|\neg E_1 \vee \neg E_3). \end{aligned} \tag{15}$$

The constraints A1 and A2 are strong by design, and we can simplify equation (15) by observing the following lemma.

Lemma 2. If $P(\mathbf{D})$ is a positive probability distribution over $\langle H, \mathbf{E}\rangle$ satisfying A1 and A2, then

i) $P(E_1) = P(E_2) = P(E_3)$, and
ii) $P(H|\neg E_1) = P(H|\neg E_2) = P(H|\neg E_3) = \alpha$.

Proof of Lemma 2. From A2, it follows that the covariance of each piece of evidence with the hypothesis is identical, $P(E_i|H)/P(E_i)$ for $i = 1, 2, 3$, which can be rewritten as

$$P(E_i)[P(H|E_i) - P(H)].$$

Assumption A2 guarantees lemma 2*i*, and A1 ensures that the covariance is positive.

From total probability, A2, and lemma 2*i*,

$$P(H) = P(H|E_i)P(E_i) + P(H|\neg E_i)1 - P(E_i).$$

QED

Lemma 2 allows for the reduction of (15) to

$$P(H|E_1 \wedge E_2) - [2\alpha - P(H|\neg E_1 \wedge \neg E_2)] \not> P(H|E_1 \wedge E_3) - [2\alpha - P(H|\neg E_1 \wedge \neg E_3)]. \tag{16}$$

But while $P(H|E_1 \wedge E_2) > P(H|E_1 \wedge E_3)$ by hypothesis, the remaining terms are unconstrained. Countermodels exist that violate the inequality. Analogously, there are countermodels in which $P(H|E_1 \wedge E_2) = P(H|E_1 \wedge E_3)$ but $oe(H, E_1 \wedge E_2) \neq oe(E_1 \wedge E_3)$. In sum, more incremental confirmation does not ensure more evidential surprise, and it should be clear from the discussion that the converse does not hold either.

REFERENCES

Bovens, L. and S. Hartmann. 2003. *Bayesian Epistemology*. Oxford: Oxford University Press.

———. 2006. "An Impossibility Result for Coherence Rankings." *Philosophical Studies* 128: 77–91.

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

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

Cohen, L. J. 1977. *The Probable and the Provable*. Oxford: Clarendon.

Crupi, V., K. Tentori, and M. Gonzalez. 2007. "On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues." *Philosophy of Science* 74 (2): 229–52.

Danks, D., and C. Glymour. 2001. "Linearity Properties of Bayes Nets with Binary Variables." In *Uncertainty in Artificial Intelligence: Proceedings of the 17th Conference*, ed. J. Breese and D. Koller, 88–104. San Francisco: Kaufmann.

Eells, E., and B. Fitelson. 2002. "Symmetries and Asymmetries in Evidential Support." *Philosophical Studies* 107 (2): 129–42.

Joyce, J. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

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

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

Kyburg, H. E., Jr. 1983. "Recent Work in Inductive Logic." In *Recent Work in Philosophy*, ed. T. Machan and K. Lucey, 87–150. Totowa, NJ: Rowman & Allanheld.

Milne, P. 1997. "log $[P(h/eb)/P(h/b)]$ Is the One True Measure of Confirmation." *Philosophy of Science* 63:21–26.

Myrvold, W. 1996. "Bayesianism and Diverse Evidence: A Reply to Andrew Wayne." *Philosophy of Science* 63:661–65.

———. 2003. "A Bayesian Account of the Virtue of Unification." *Philosophy of Science* 70:399–423.

Olsson, E. 2005. *Against Coherence: Truth, Probability and Justification*. Oxford: Oxford University Press.

Silva, R., C. Glymour, R. Scheines, and P. Spirtes. 2006. "Learning the Structure of Latent Linear Structure Models." *Journal of Machine Learning Research* 7:191–246.

Wheeler, G. 2009. "Focused Correlation and Confirmation." *British Journal for the Philosophy of Science* 60 (1): 79–100.

Wheeler, G., and R. Scheines. 2011. "Causation, Association, and Confirmation." In *Explanation, Prediction, and Confirmation: New Trends and Old Ones Reconsidered*. The Philosophy of Science in a European Perspective. Dordrecht: Springer.