Synthese (2016) 193:3383–3395
DOI 10.1007/s11229-015-0937-4

Competing accounts of contrastive coherence

Michael Schippers[1]

Received: 13 May 2015 / Accepted: 25 September 2015 / Published online: 5 October 2015

Abstract The proposition that Tweety is a bird coheres better with the proposition that Tweety has wings than with the proposition that Tweety cannot fly. This relationship of contrastive coherence is the focus of the present paper. Based on recent work in formal epistemology we consider various possibilities to model this relationship by means of probability theory. In a second step we consider different applications of these models. Among others, we offer a coherentist interpretation of the conjunction fallacy.

Keywords Contrastivism · Coherence · Probability · Confirmation · Bayesian epistemology · Conjunction fallacy

1 Introduction

Bayesian coherentism is sometimes characterized by three main assumptions, namely (i) that coherence is the hallmark of epistemic justification, (ii) that the degree of coherence of a set of beliefs is fully determined by the probabilistic features of the propositions contained in the set and (iii) that the binary relation 'being no less coherent than' is an ordering (cf. Bovens and Hartmann 2003, p. 11f.). These latter two assumptions are also the hallmark of a variety of probabilistic approaches that try to capture a set's degree of coherence by means of probability theory (Douven and Meijs 2007; Fitelson 2003; Glass 2002; Olsson 2002; Shogenji 1999). These approaches are based on functions that take as input the probabilistic features of the propositions in the set and give as output a real number representing the set's degree of coherence

✉ Michael Schippers
mi.schippers@uni-oldenburg.de

1 University of Oldenburg, 26111 Oldenburg, Germany

(given some probability distribution Pr). Accordingly, all these approaches satisfy assumption (ii) above. Furthermore, at least with respect to contingent propositions the relation 'being no less coherent than' that is provided by these measures is anti-symmetric, transitive and complete; hence, it is a total order and so these approaches also satisfy assumption (iii).

This latter assumption is a rather strong constraint that in particular requires that each pair of sets (X, Y) be comparable as regards their respective degree of coherence. For example it sometimes seems impossible to compare pairs of propositions with respect to their degrees of coherence when these pairs are completely different as regards content.[1] To illustrate, consider the following pair of situations: in situation 1 a card is drawn at random from a standard deck of 52 cards and the outcome is hidden. Let x be the proposition that the drawn card is a spade and y the proposition that it is black. Given that x entails y, the set containing both these propositions intuitively seems coherent. On the other hand, assume that a dodecahedron is thrown and the outcome is hidden. Let x' be the proposition that the dodecahedron came up a number less than 4 and y' says that it came up an odd number less than 6. Again, the set containing x' and y' intuitively seems coherent. But is it more or less coherent than the set containing x and y? On the one hand, when moving from the first to the second situation, the deductive entailment relation obtaining among the members of the former set is lost. On the other hand, y' is more specific than y. The relative overlap in both situations is the same. Also a comparison of conditional probabilities does not give us a clear-cut ordering: while in situation 1 one conditional probability is equal to 1/2 and the other is equal to 1, in situation 2 both are equal to 2/3.[2]

Thus, even as an ardent defender of a probabilistic approach to coherence one might hesitate to embrace the ordering position that one is implicitly committed to when using probabilistic measures of coherence. A more alleviated position might be one that blurs the clear-cut distinctions between degrees of coherence and instead opts for relations like 'being approximately as coherent as' and 'being considerably more coherent than'. We won't dwell into these fuzzy coherence relations in this paper and instead focus on some well-specified subset of all sets of propositions for which a total ordering demand seems less questionable. These will be pairs of propositions with a common member. It seems that given a certain proposition x we can always assess whether x coheres better with a proposition y or with a another proposition z. This is the relationship of *contrastive coherence* that will introduced in the next section. The following Sect. 3 will then be devoted to two possible applications of the idea of contrastive coherence.

2 Contrastive coherence

Contrastive coherence is a relationship between a proposition x on the one hand and a pair of propositions (or *contrast class*) $(y, z)_x$ on the other such that x coheres better

[1] A different example is given by Bovens and Hartmann in their Tokyo murder case (cf. Bovens and Hartmann 2003, p. 39f.).

[2] An analysis of the coherence measures to be introduced in the next section shows disagreement with respect to this test case (proof omitted).

with either y or z. Alternatively, one might say that one pair of propositions (x, y) is more (or less) coherent than another such pair (x, z). In the literature there are various contrastive accounts of concepts like confirmation (Fitelson 2007; Chandler 2007, 2013), explanation (Lipton 1990) and causation (Hitchcock 1996, 1999).[3] This paper contributes to this debate with a discussion of formal models of contrastive coherence. To this end, let $\mathscr{L}$ be a propositional language such that $x, y, z \in \mathscr{L}$, then $y \succ_x z$ denotes the relationship that x coheres better with y than with z. Alternatively, we will utilize the following notation $(x, y) \succ (x, z)$ to denote the very same relation. In the next section we will consider some straightforward conditions for contrastive coherence orderings. Among these is the idea that x coheres better with y than with z if $\Pr(x|y)$ exceeds $\Pr(x|z)$. This is the well-known likelihood-condition embraced by likelihoodists in confirmation theory (cf. Royal 1997). However, so far there is no analysis of this and related conditions like the one containing only the conditional probability of x given the elements of the contrast class. Therefore, we will broaden the family of candidate conditions for contrastive coherence orderings by taking advantage of the flourishing debate on probabilistic coherence measures. That is, if we have a probabilistic measure of coherence **coh** at hand, i.e. a (partial) function assigning each triple $(x, y, \Pr)$, where $x, y \in \mathscr{L}$ and Pr is a probability function, a real number representing the degree of coherence of x and y under Pr, then the following can be utilized as a bridge principle to determine the relationship of contrastive coherence by means of coherence measures:

(†) $y \succeq_x z$ if and only if $\mathbf{coh}(x, y) \geq \mathbf{coh}(x, z)$

As usual, the corresponding relations $\succ_x$ and $=_x$ are defined as follows: $y \succ_x z$ iff $y \succeq_x z$ and not $z \succeq_x y$ and $y =_x z$ iff $y \succeq_x z$ and $z \succeq_x y$. Finally, $y \prec_x z$ iff $z \succ_x y$.

Recent years have seen the upshot of various probabilistic accounts of measuring coherence.[4] Among these are the following (families of) measures:

(i) $\mathscr{D}(x, y) = \Pr(x, y)/[\Pr(x) \times \Pr(y)]$
(ii) $\mathscr{O}(x, y) = \Pr(x, y)/\Pr(x \vee y)$
(iii) $\mathscr{C}_{\mathbf{supp},\tau}(x, y) = \tau(\{\mathbf{supp}(x, y), \mathbf{supp}(y, x)\})$,

where $\tau(\Delta)$ is some average of the members of Δ. In coherence contexts, especially when the focus is on pairs of propositions, usually the straight average is chosen. We stick to this convention in the sequel of the paper (and hence omit reference to τ). Furthermore, **supp** denotes a probabilistic measure of support (a.k.a. confirmation), i.e. a (partial) function from a pair of propositions (x, y) and a probability function Pr into some well-specified interval of real numbers such that $\mathbf{supp}(x, y, \Pr)$ denotes the degree of support that y provides for x under distribution Pr. In what follows we restrict attention to the following prominent proposals for measuring support (for references see Fitelson 1999; Crupi et al. 2007; Shogenji 2012):

$$d(x, y) = \Pr(x|y) - \Pr(x)$$
$$r(x, y) = \Pr(x|y) \times \Pr(x)^{-1}$$

[3] In a recent anthology, Martin Blauw even speaks of a "contrastivist *movement*" in philosophy (Blauw 2013, p. 1).

[4] For a survey see Schippers (2014, 2015).

$s(x, y) = \Pr(x|y) - \Pr(x|\overline{y})$
$n(x, y) = \Pr(y|x) - \Pr(y|\overline{x})$
$l(x, y) = \log\left[\Pr(y|x) \times \Pr(y|\overline{x})^{-1}\right]$
$J(x, y) = \log[\Pr(x|y) \times \Pr(x)^{-1}] \times [-\log \Pr(x)]^{-1}$
$g(x, y) = \Pr(\overline{x}) \times \Pr(\overline{x}|y)^{-1}$
$z(x, y,) = \min\{\Pr(x|y), \Pr(x)\} \times \Pr(x)^{-1} - \min\{\Pr(\overline{x}|y), \Pr(\overline{x})\} \times \Pr(\overline{x})^{-1}$
$ku(x, y) = \Pr(y|x) \times \Pr(y)^{-1}$
$f(x, y) = \Pr(x|y)$

The account of contrastive coherence obtained by measure $\mathscr{C}_{\mathbf{supp}}$ via (†) will be denoted by ($\dagger_{\mathbf{supp}}$); to illustrate, ($\dagger_d$) denotes the account that utilizes confirmation measure d that is plugged into recipe (iii). Analogously, ($\dagger_{\mathscr{D}}$) and ($\dagger_{\mathscr{O}}$) represent the accounts based on coherence measures $\mathscr{D}$ and $\mathscr{O}$. To eliminate redundancies note that

Observation 1 *(i) ($\dagger_{\mathscr{D}}$), ($\dagger_r$) and ($\dagger_{ku}$) are equivalent; (ii) ($\dagger_s$) and ($\dagger_n$) are equivalent; (iii) all[5] other accounts are pairwise inequivalent.*

Proof The proof of (i) and (ii) is straightforward; a proof of (iii) for all pairs of measures is given by probability distributions $\Pr_1$–$\Pr_8$ in Table 1. □

Now we turn to a first evaluation of these rival accounts of contrastive coherence. To start with, a natural requirement is that all proposals be irreflexive, asymmetric and transitive, i.e.:

(Irr) $y \not\succ_x y$ for all $x, y \in \mathscr{L}$
(As) If $y \succ_x z$ than $z \not\succ_x y$ for all $x, y, z \in \mathscr{L}$
(Tr) If $y \succ_x z$ and $z \succ_x z'$, then $y \succ_x z'$ for all $x, y, z, z' \in \mathscr{L}$

Obviously, these constraints are met by all probabilistic proposals that have been considered so far. Now consider the following qualitative discrimination requirement according to which no propositions coheres better with x than x itself, and no proposition coheres worse with x than its opposite $\overline{x}$:

(QD) $(x, x) \succeq (x, y) \succeq (x, \overline{x})$ for all $x, y \in \mathscr{L}$

Observation 2 *All accounts considered so far satisfy* **(QD)**.[6]

Given these basic properties we will now turn to two applications of contrastive coherence measures that form the main part of the present paper.

3 Applications

In this section we will consider two applications of contrastive coherence models. First of all, we evaluate a conjecture by Siebel (2005) on a contrastive coherence ordering in the well-known Tweety case. The second application picks up on Tversky and Kahneman's study on conjunction fallacies. More precisely, we will reconstruct the comparative assessments in the Linda story in terms of contrastive coherence judgments.

[5] Note that the case for the pair (($\dagger_d$), ($\dagger_J$)) is so far unsettled.

[6] ($\dagger_l$) satisfies **(QD)** provided that division by zero is equated with infinity. Alternatively, an ordinally equivalent measure put forward by Kemeny and Oppenheim (1952) can be used.

Table 1 Probability distributions for observations

Propositions	Probability functions							
	$\Pr_1$	$\Pr_2$	$\Pr_3$	$\Pr_4$	$\Pr_5$	$\Pr_6$	$\Pr_7$	$\Pr_8$
$x \wedge y \wedge z$	17/41	1/258	1/33	1/39	23/60	1/52	13/44	0
$x \wedge y \wedge \overline{z}$	1/34	1/31	1/23	1/24	1/842	2/27	7/50	$1/10^9$
$x \wedge \overline{y} \wedge z$	1/1519	3/52	1/26	1/25	1/26	3/49	4/23	$99/10^6$
$x \wedge \overline{y} \wedge \overline{z}$	1/375	5/23	1/23	1/14	1/91	3/29	13/50	$999/10^9$
$\overline{x} \wedge y \wedge z$	2/31	16/37	1/63	12/61	1/36	17/86	1/18	0
$\overline{x} \wedge y \wedge \overline{z}$	8/39	3/65	1/95	4/43	1/999	7/69	0	0
$\overline{x} \wedge \overline{y} \wedge z$	1/810	9/43	1/175	1/78	1/122	3/49	3/40	$9901/10^6$
$\overline{x} \wedge \overline{y} \wedge \overline{z}$	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8
Account	Verdicts on contrastive coherence							
	$\Pr_1$	$\Pr_2$	$\Pr_3$	$\Pr_4$	$\Pr_5$	$\Pr_6$	$\Pr_7$	$\Pr_8$
$(\dagger_{\mathscr{D}})$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_x$	$\succ_x$	$\succ_z$	$\succ_x$
$(\dagger_{\mathscr{O}})$	$\prec_x$	$\prec_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\succ_x$	$\prec_z$	$\prec_x$
$(\dagger_d)$	$\prec_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_z$	$\succ_x$
$(\dagger_s)$	$\prec_x$	$\succ_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_z$	$\succ_x$
$(\dagger_l)$	$\succ_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_z$	$\succ_x$
$(\dagger_J)$	$\prec_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_z$	$\prec_x$
$(\dagger_g)$	$\succ_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_z$	$\succ_x$
$(\dagger_z)$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\prec_x$	$\succ_x$	$\succ_z$	$\succ_x$
$(\dagger_f)$	$\prec_x$	$\prec_x$	$\succ_x$	$\prec_x$	$\prec_x$	$\succ_x$	$\prec_z$	$\succ_x$

Note that the contrast class in $\Pr_1 - \Pr_6$ and $\Pr_8$ is $(y, z)_x$ while in $\Pr_7$ it is $(x \wedge y, x)_z$. I am grateful to an anonymous reviewer for providing me with $\Pr_8$. Where $\theta_1 = 1 - \Pr_1(x \vee y \vee z) = 39\,263\,874\,907/139\,357\,047\,375$, $\theta_2 = 1 - \Pr_2(x \vee y \vee z) = 790\,807/884\,818\,740$, $\theta_3 = 1 - \Pr_3(x \vee y \vee z) = 159\,872\,017/196\,846\,650$, $\theta_4 = 1 - \Pr_4(x \vee y \vee z) = 74\,285\,843/143\,215\,800$, $\theta_5 = 1 - \Pr_5(x \vee y \vee z) = 6\,175\,725\,319/11\,673\,170\,145$, $\theta_6 = 1 - \Pr_6(x \vee y \vee z) = 753\,095\,911/1\,973\,138\,076$, $\theta_7 = 1 - \Pr_6(x \vee y \vee z) = 7/91\,080$, $\theta_8 = 1 - \Pr_8(x \vee y \vee z) = 989\,999/10^6$

3.1 Tweety: Siebel's conjecture

In a short remark on the well-known Tweety case, Siebel (2005) puts forward the following conjecture: "Why does the proposition 'Tweety is a bird' fit 'Tweety has wings' much better than 'Tweety cannot fly'? Because, one might argue, the probability that Tweety has wings, given that it is a bird, strongly exceeds the probability that Tweety cannot fly, given that it is a bird" (p. 335). If one agrees with Siebel's conjecture, the following might be considered a reasonable constraint for models of contrastive coherence:

(D1) If $\Pr(y|x) > \Pr(z|x)$, then $y \succ_x z$.

Thus, according to **(D1)**, a comparison of contrastive coherence should focus exclusively on the conditional probabilities of each element of the contrast class. However, this constraint is not in line with any of the accounts considered beforehand:

Observation 3 *None of the accounts* $(\dagger_{\mathscr{D}})$–$(\dagger_f)$ *satisfies* **(D1)**.

Proof A proof is given by probability distribution $\Pr_4$ in Table 1, where $\Pr_4(y|x) \approx 0.377 > 0.367 \approx \Pr_4(z|x)$. □

Alternatively, one might be inclined to reverse the positions within each conditional probability. In terms of the Tweety case this would amount to the following rationale: 'Tweety is a bird' fits better with 'Tweety has wings' than with 'Tweety cannot fly' because the probability that Tweety is a bird is higher when conditioned on the information that it has wings than when conditioned on the information that it cannot fly. The corresponding constraint reads as follows[7]:

(D2) If $\Pr(x|y) > \Pr(x|z)$, then $y \succ_x z$.

The evaluation of this constraint yields the following

Observation 4 *Of all considered accounts only* $(\dagger_{\mathscr{D}})$ *satisfies* **(D2)**.

Proof The proof for $(\dagger_{\mathscr{D}})$ is straightforward; for all other accounts see distribution $\Pr_5$ in Table 1, where $\Pr_5(x|y) \approx 0.930 > 0.921 \approx \Pr_5(x|z)$. □

Should this be regarded as a vindication of the deviation measure $\mathscr{D}$-based approach to contrastive coherence (or its equivalent counterparts $\mathscr{C}_r$ and $\mathscr{C}_{ku}$)? This might be considered premature. Both constraints **(D1)** and **(D2)** are restricted to *one* conditional probability while entirely neglecting other relevant pieces of information. For one, it might be the case that the imbalance of conditional probabilities that either **(D1)** or **(D2)** focuses on is reversed when we consider the counterpart in each case. That is, it might be the case that $\Pr(x|y) > \Pr(x|z)$ and $\Pr(y|x) < \Pr(z|x)$ so that there is disagreement among the conclusions drawn according to **(D1)** and **(D2)**: according to **(D1)**, $z \succ_x y$, while according to **(D2)**, $y \succ_x z$. Therefore, in what follows we consider a combined version that takes into account both conditional probabilities.[8]

(D3) If $\Pr(y|x) > \Pr(z|x)$ and $\Pr(x|y) > \Pr(x|z)$, then $y \succ_x z$.

In the light of this combined version, x coheres better with y than with z if both y's probability conditional on x exceeds the corresponding probability of z *and* the likelihood of x is higher conditional on y than it is on z. In short, according to this conservative account x coheres better with y than with z if $y \succ_x z$ according to both **(D1)** and (**D2**).

Observation 5 *Of all accounts considered so far only* $(\dagger_{\mathscr{D}})$, $(\dagger_{\mathscr{O}})$ *and* $(\dagger_f)$ *satisfy* **(D3)**.[9]

[7] This constraint is know as the *law of likelihood* (cf. Royal 1997; Fitelson 2007).

[8] Condition **(D3)** is similar to the Bovens–Olsson condition that is well-known in the literature on measuring coherence (cf. Bovens and Olsson 2000, p. 688). In contrast to **(D3)**, the Bovens–Olsson condition considers *one* pair of propositions with respect to *two* different probability distributions where (**D3**) only pertains to cases where *two* different pairs of propositions are assessed with respect to *one* probability distribution. For an assessment of (a generalized form of) the Bovens–Olsson condition with respect to probabilistic measures of coherence see Schippers (2015). **(D3)** itself is discussed by Glass (2007) as a condition for ranking different explanations.

[9] There is a small caveat in this observation: so far the case of $(\dagger)_z$ is unsettled, i.e. I have neither been able to prove that it *does* satisfy (**D3**) nor have I been able to find a counterexample using Branden Fitelson's *PrSAT* (see Fitelson 2008).

Proof As regards ($\dagger_{\mathscr{D}}$), observation 5 is a corollary of observation 4. For all support-based accounts see $\Pr_6$ in table 1, where $\Pr_6(x|y) \approx 0.238 > 0.237 \approx \Pr_6(x|z)$ and $\Pr_6(y|x) \approx 0.362 > 0.312 \approx \Pr_6(z|x)$. The proofs for $\mathscr{O}$ and $\mathscr{C}_f$ are straightforward. □

Thus, if one sticks to this conservative explication then the accounts based ($\dagger_{\mathscr{O}}$) and ($\dagger_f$) are on a par with ($\dagger_{\mathscr{D}}$) (and so are ($\dagger_r$) and ($\dagger_{ku}$)).

3.2 Linda: a contrastive coherentist account

Judgments under uncertainty are fraud with difficulty. One salient example that clearly reveals the pitfalls within this kind of reasoning is due to Tversky and Kahneman (1982). Consider the following personality sketch:

> Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which of these two alternatives is more probable?
> (b) Linda is a bank teller.
> (bf) Linda is a bank teller and is active in the feminist movement.

What was demonstrated in a number of studies (cf. Tversky and Kahneman 1982; Hertwig and Chase 1998) is that a majority of both statistically naive and sophisticated respondents judge ($b \wedge f$) more probable than (b), thereby committing what is called the "conjunction fallacy": given that ($b \wedge f$) logically entails (b), it is a simple theorem of the probability calculus that (b) *must be* at least as probable as ($b \wedge f$).

The Linda case is only one out of a number of studies on reasoning fallacies that triggered an extensive debate on human rationality and its limitations. However, scholars did not unanimously comply with Tversky and Kahneman's conclusion about human irrationality. Many critics draw attention to possible misinterpretations on the respondent's side: as outlined by Hertwig and Gigerenzer (1999), the word 'probability' is polysemous; for one, there are Bayesian and frequentist interpretations of probability. According to the former, probabilities amount to subjective degrees of belief (cf. Ramsey 1926; de Finetti 1937) while the latter interpret probabilities in terms of relative frequencies (Reichenbach 1949; von Mises 1957). To elicit judgment in the Linda case in terms of a frequentist interpretation of probability is problematic since it involves probabilities for single events with no clear reference class given (see Gigerenzer 1994, 2001). More mundane interpretations equate 'probable' with 'plausible' or 'supported by the evidence'. If participants understand the Linda story in either of these ways, it is far from clear whether they should be deemed irrational (cf. Gigerenzer 1996, 2001; Sides et al. 2002; Crupi et al. 2008). Hertwig and Gigerenzer (1999, p.278) even argue that subjects are urged to choose a non-mathematical interpretation because the personality sketch includes a lot of redundant information that is not necessary to solve the task. This is a clear violation of Paul Grice's *relevance maxim* (cf. Grice 1975, p. 27). Other research has focused on the (polysemous) meaning of 'and' (see Mellers et al. 2001; Tentori and Crupi 2012). However, so far there seems to be no fully satisfactory explanation of conjunction effects and similar reasoning fallacies.

In what follows I want to draw attention to an alternative account that has been proposed by Siebel (2003). His *coherentist* account is based on the idea that "participants' judgments can be seen as resulting from inferences to the maximally coherent whole, where explanatory relations are a central factor in increasing coherence" (p. 2). Thus, in a nutshell, the idea is that what respondents try to do in the Linda case is to maximize coherence: instead of choosing the alternative that is more probable, they opt for the alternative that coheres better with the personality sketch.

Siebel stresses the importance of explanatory relations for coherence (2005, 2011).[10] According to his view, when comparing the two hypotheses in the Linda case, "the crucial point is how many explanatory relations there are and how strong they are" (2003, p. 8). This is similar to what Chart (2001) proposes in his explanationist account: according to him, what subjects tend to do is to make an *Inference to the best explanation*:

> In this case, we assess the inference [from the background story to the conjunctive hypothesis] by looking at the explanatory relations between the evidence and the hypothesis. These relations must run in both directions. A theory is supported by evidence if it provides the best explanation for that evidence. On the other hand, a theory predicts that the evidence will be such that the theory can provide a good explanation. Thus, a theory does not predict things that it cannot explain."

Applied to the Linda case he argues as follows:

> Linda's background, as given, provides absolutely no explanation for her becoming a bank teller. It is completely unexpected, nothing in her background seems prone to cause it, and her background and the career are not at all unified. On the other hand, it does provide a partial explanation for her becoming a feminist bank teller: it explains why she is a feminist. Those political views are expected on the basis of her background, could be caused by several elements of it, and the background and politics are quite well unified.

Given the close proximity of concepts like 'coherence' and 'unification', it seems that Chart's account can also be interpreted as 'coherentist'. What is crucial is that the evidence e, which is Linda's background, can better be accounted for by means of the conjunctive hypothesis $(b \wedge f)$ than by means of (b) alone. That is, $(b \wedge f)$ provides a better explanation for e than b does.

[10] Based on the close relationship between the concepts of coherence and explanation, Siebel even argues in these papers for the impossibility of probabilistically measuring coherence. Starting from the observation that the concept of explanation cannot be reduced to probability, Siebel concludes that "if probabilistic accounts cannot cope with explanation, they will hardly be able to deal with coherence because, as BonJour (1985) and many others have pointed out, coherence is a function of explanation." For a rebuttal see Roche and Schippers (2013).

In probabilistic terms, this amounts to the following[11]:

$$\Pr(e|b \wedge f) > \Pr(e|b) \tag{1}$$

Furthermore, b seems not to provide any explanation for e at all; accordingly, in probabilistic terms we can add the following constraint:

$$\Pr(e|b) \leq \Pr(e) \tag{2}$$

In what follows we will draw attention to the connection between this explanationist account and an account based on the notion of confirmation (a.k.a. support). As was outlined above, the majority of probabilistic proposals for measuring coherence are based on the concept of confirmation. So before we dwell on these coherence measures, we will focus on a confirmation-theoretic account that has been suggested by Crupi et al. (2008). Given a probabilistic support measure **supp**, their approach to the conjunction fallacy utilizes the following two conditions that seem beyond reasonable doubt in light of the experimental findings:

$$\mathbf{supp}(b, e) \leq \delta \tag{3}$$

That is, the background story e and the bank teller hypothesis are (if at all) negatively correlated (where δ is the threshold separating confirmation and disconfirmation). The second condition reads as follows:

$$\mathbf{supp}(b \wedge f, e|b) > \delta \tag{4}$$

This means that the background is positively correlated with the conjunctive hypothesis $b \wedge f$ (even) conditionally on b. Based on these principles, Crupi et al. are able to prove the following theorem for virtually all confirmation measures[12]:

Theorem 1 (Crupi et al. 2008) *Given (3) and (4),* $\mathbf{supp}(b \wedge f, e) > \mathbf{supp}(b, e)$.

What is interesting to note is that the same conclusion can be drawn by means of our explanationist principles (1) and (2). The rationale for this coincidence is the following

Observation 6 *(1) and (2) are logically equivalent to (3) and (4).*

[11] Note that this inequality is the common core of various measures of explanatory power (Good 1960; McGrew 2003; Schupbach and Sprenger 2011; Crupi and Tentori 2012), that is, each of these measures $\mathcal{E}_{\Pr}(e, h)$ that quantify the degree of explanatory power that h provides for e (given Pr) satisfies the following principle for any contingent e, h_1, h_2 and any regular probability Pr (cf. Crupi and Tentori 2012):

$$\mathcal{E}_{\Pr}(e, h_1) \gtreqless \mathcal{E}_{\Pr}(e, h_2) \quad \text{iff} \quad \Pr(e|h_1) \gtreqless \Pr(e|h_2)$$

[12] More precisely, this relationship holds for all probabilistic confirmation measures that satisfy the "weak law of likelihood" (cf. Joyce 2004; Fitelson 2007).

Proof $\mathbf{supp}(x, y) \leq \delta$ iff $\Pr(x|y) \leq \Pr(x)$, where the latter is equivalent to $\Pr(y|x) \leq \Pr(y)$. Furthermore, $\mathbf{supp}(y, z|x) > \delta$ iff $\Pr(z|x \wedge y) > \Pr(z|x)$ is easy to prove. □

Thus, the explanationist and the confirmation-theoretic accounts are in fact equivalent. What is more, consider the following constraint on confirmation measures (cf. Brössel 2013; Crupi et al. 2010):

$$(\ddagger)\ \mathbf{supp}(x, y) \gtreqless \mathbf{supp}(x, z) \quad \text{iff} \quad \Pr(x|y) \gtreqless \Pr(x|z)$$

According to Eells and Fitelson (2000, p. 670), "it is not an exaggeration to say that most Bayesian confirmation theorists would accept (‡) as a desideratum for Bayesian confirmation measures." If **supp** satisfies (‡) we can even prove the following claim on comparative confirmation[13]:

Observation 7 *Given (1) and (2), bf* $\succ_e$ *b for all* ($\dagger_{\mathbf{supp}}$) *such that* **supp** *satisfies* (‡).

Proof Assuming (1) and (2) we can infer via observation 6 and Theorem 1 that $\mathbf{supp}(b \wedge f, e) > \mathbf{supp}(b, e)$. Furthermore, (1) entails $\mathbf{supp}(e, b \wedge f) > \mathbf{supp}(e, b)$ for all measures satisfying (‡). Accordingly, we get:

$$\tau(\mathbf{supp}(b \wedge f, e), \mathbf{supp}(e, b \wedge f)) > \tau(\mathbf{supp}(b, e), \mathbf{supp}(e, b))$$

and thus $\mathscr{C}_{\mathbf{supp},\tau}(b \wedge f, e) > \mathscr{C}_{\mathbf{supp},\tau}(b, e)$ and hence $b \wedge f \succ_e b$ for all accounts ($\dagger_{\mathbf{supp}}$) such that **supp** satisfies (‡). □

The same holds for the deviation measure-based approach:

Observation 8 *Given (1) and (2),* $b \wedge f \succ_e b$ *for* ($\dagger_{\mathscr{D}}$).

Proof Straightforward. □

This, however, is not the case if we let our comparative coherence assessments be based on the overlap measure of coherence.

Observation 9 *Even if (1) and (2), it may be that* $b \wedge f \not\succ_e b$ *for* ($\dagger_{\mathscr{O}}$) *and* ($\dagger_f$).

Proof A proof is given by probability distribution $\Pr_7$ in Table 1. This distribution satisfies both $\Pr_7(z|x \wedge y) > \Pr_7(z|x)$ and $\Pr_7(z|x) \leq \Pr_7(z)$, i.e. conditions (1) and (2) with respect to x, y and z instead of b, f and e but nonetheless $x \wedge y \prec_z x$ for both ($\dagger_{\mathscr{O}}$) and ($\dagger_f$). □

Thus, the majority of probabilistic accounts of contrastive coherence agree in that given some reasonable constraints the conjunctive hypothesis $b \wedge f$ coheres better with the background story e than the conjunct b does. Siebel even concludes:

[13] Measures that satisfy (‡) are (among others) d, l, r, z, ku, Shogenji's justification measure J and Kemeny and Oppenheim's (1952) measure k.

> The coherentist account thus nicely explains subjects' judgments in the Linda study. They do not understand 'probable' in the sense of probability theory but rank the given alternatives with respect to the amount of coherence they provide. If they are offered only the two alternatives 'Linda is a bank teller' and 'Linda is a feminist bank teller', they choose the latter because the explanatory relations resulting from the 'feminist' component lead to more coherence. (2003, p. 9)

My interpretation is a more modest one. I do not think that the coherentist results offer a rival or even superior account of what people actually do in the Linda case. On the contrary, I think that the results match perfectly with other proposals that in combination with the coherentist results offer a plausible interpretation of the results interpreting subjects as rational beings equipped with a high enough degree of sophistication in order to draw reasonable conclusions from puzzling questionnaires.[14] Besides the explanation- and the confirmation-based accounts, the coherentist assessment is also in line with Tversky and Kahneman's approach in terms of the "representativeness heuristic" (cf. Kahneman and Tversky 1972; Tversky and Kahneman 1982 and Shafir et al. 1990), according to which assessments of 'typicality' drive subjects' judgments in uncertain environments. Although a detailed comparison of both approaches is beyond the scope of the present paper, it seems reasonable to conclude that typicality-comparisons may well go hand in hand with coherence assessments in the sense that if feature F is more typical for B's than feature F', then the propositions "x is a B" coheres better with the proposition "x has F" than with the proposition "x has F'" (cf. Siebel 2003, p. 10f.).

4 Conclusion

In this paper various models of contrastive coherence based on probabilistic coherence measures have been considered. Two illuminating case studies illustrated the merits of such models. As has been shown, even if one remains skeptical about general coherence orderings and therefore rejects probabilistic coherence measures, one can nonetheless gain insights into various contrastive coherence orderings by utilizing these measures.

So far we have sticked to a dichotomous relation of contrastive coherence, i.e., we have concentrated on whether a proposition x coheres better with a proposition y or with another proposition z. A next step might be to consider *degrees* of contrastive coherence. Granted that x indeed coheres better with y than with z, it might be interesting to know how much better. For example, in the context of the conjunction fallacy we might say that it is not only the case that the conjunctive hypothesis $b \wedge f$ coheres better with the background story than b alone, but that there is a considerable difference in the degree of contrastive coherence as spelled out, for example, by the difference between or the ratio of $\mathbf{coh}(x, y)$ and $\mathbf{coh}(x, z)$.

[14] Another interpretation of the conjunction fallacy is given by Shogenji (2012). There, Shogenji shows that conditions (3) and (4) also imply that the conjunctive hypothesis $b \wedge f$ is more justified by e than b, as measured by his justification measure J. Given that J also satisfies (‡), we also know that the degree justification-based coherence as measured by $\mathscr{C}_J$ is higher for the conjunctive hypothesis.

Furthermore, the results presented in this paper may be generalized so as to include not only a comparison between a proposition x on the one hand and a contrast class of propositions (y, z) on the other but sets of such propositions. So we might ask whether the set X coheres better with y or with z (yielding the contrastive coherence relation $y \succ_X z$), whether the proposition z coheres better with set X or with set Y (yielding the relation $X \succ_z Y$) or even whether the set X coheres better with the set Y or with the set Z (yielding $Y \succ_X Z$). We leave these issues for future research.

Acknowledgments This work was supported by Grant SI 1731/1-1 to Mark Siebel from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program "New Frameworks of Rationality" (SPP 1516).

References

Blauw, M. (Ed.). (2013). *Contrastivism in philosophy*. New York: Routledge.
BonJour, L. (1985). *The structure of empirical knowledge*. Cambridge: Harvard University Press.
Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.
Bovens, L., & Olsson, E. J. (2000). Coherentism, reliability and Bayesian networks. *Mind*, *109*, 685–719.
Brössel, P. (2013). The problem of measure sensitivity redux. *Philosophy of Science*, *80*, 378–397.
Chandler, J. (2007). Solving the tacking problem with contrast classes. *British Journal for the Philosophy of Science*, *58*, 489–502.
Chandler, J. (2013). Contrastive confirmation: Some competing accounts. *Synthese*, *190*, 129–138.
Chart, D. (2001). Inference to the best explanation, Bayesianism, and feminist bank tellers. Online-paper. http://philsci-archice.pitt.edu/documents/disk0/00/00/03/22/.
Crupi, V., Festa, R., & Buttasi, C. (2010). Towards a grammar of Bayesian confirmation. In M. Surez, M. Dorato, & M. Rdei (Eds.), *Epistemology and methodology of science* (pp. 73–93). Berlin: Springer.
Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Thinking and Reasoning*, *14*, 182–199.
Crupi, V., & Tentori, K. (2012). A second look at the logic of explanatory power (with two new representation theorems). *Philosophy of Science*, *79*, 365–385.
Crupi, V., Tentori, K., & Gonzalez, M. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*, *74*, 229–252.
de Finetti, B. ([1937] 1980). Foresight: Its logical laws, its subjective sources. In H. E. Kyburg, Jr. & H. E. Smokler (Eds.), Studies in subjective probability (2nd ed., pp. 53–118). Huntington, NY: Robert E. Krieger.
Douven, I., & Meijs, W. (2007). Measuring coherence. *Synthese*, *156*, 405–425.
Eells, E., & Fitelson, B. (2000). Measuring confirmation and evidence. *Journal of Philosophy*, *97*, 663–672.
Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, *66*, 362–378.
Fitelson, B. (2003). *A probabilistic theory of coherence. Analysis*, *63*, 194–199.
Fitelson, B. (2007). Likelihoodism, Bayesianism, and relational confirmation. *Synthese*, *156*, 473–489.
Fitelson, B. (2008). A decision procedure for probability calculus with applications. *The Review of Symbolic Logic*, *1*, 111–125.
Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York: Wiley.
Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky. *Psychological Review*, *103*, 592–596.
Gigerenzer, G. (2001). Content-blind norms, no norms, or good norms? A reply to Vranas. *Cognition*, *81*, 93–103.
Glass, D. H. (2002). Coherence, explanation, and Bayesian networks. In M. ONeill, et al. (Eds.), Artificial intelligence and cognitive science (pp. 177–182). Berlin: Springer.
Glass, D. H. (2007). Coherence measures and inference to the best explanation. *Synthese*, *157*, 275–296.

Good, I. J. (1960). Weight of evidence, corroboration, explanatory power, information and the utility of experiments. *Journal of the Royal Statistical Society (Series B)*, *22*, 319–331.
Grice, H. P. (1975). Logic and conversation. Studies in the way of words (pp. 22–40). Cambridge, MA: Harvard University Press.
Hertwig, R., & Chase, V. M. (1998). Many reasons or just one: How response mode affects reasoning in the conjunction problem. *Thinking and Reasoning*, *4*, 319–352.
Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*, 275–305.
Hitchcock, C. R. (1996). The role of contrast in causal and explanatory claims. *Synthese*, *107*, 395–419.
Hitchcock, C. D. (1999). Contrastive explanation and the demons of determinism. *British Journal for the Philosophy of Science*, *50*, 585–612.
Joyce, J. (2004). Bayes's theorem. In E. N. Zalta (Ed.). The Stanford encyclopedia of philosophy (Summer 2004 ed.). http://plato.stanford.edu/archives/sum2004/entries/bayes-theorem/.
Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
Kemeny, J., & Oppenheim, P. (1952). Degrees of factual support. *Philosophy of Science*, *19*, 307–324.
Lipton, P. (1990). Contrastive explanation. In D. Knowles (Ed.), *Explanation and its limits* (pp. 247–266). Cambridge: Cambridge University Press.
McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *British Journal for the Philosophy of Science*, *54*, 553–567.
Mellers, A., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*, 269–275.
Olsson, E. J. (2002). What is the problem of coherence and truth? *The Journal of Philosophy*, *99*, 246–272.
Ramsey, F. P. (1926). Truth and probability. In H. E. Kyburg Jr & H. E. Smokler (Eds.), Studies in subjective probability (2nd ed., pp. 23–52). Huntington, NY: R. E. Krieger. 1980.
Reichenbach, H. (1949). *The theory of probability*. Berkeley: University of California Press.
Roche, W., & Schippers, M. (2013). Coherence, probability and explanation. *Erkenntnis*, *79*, 821–828.
Royal, R. (1997). *Statistical evidence: A likelihood paradigm*. London: Chapman and Hall.
Schippers, M. (2014). Probabilistic measures of coherence. *From adequacy constraints towards pluralism. Synthese*, *191*, 3821–3845.
Schippers, M. (2015). The grammar of Bayesian coherentism. *Studia Logica*, *103*, 955–984.
Schupbach, J. N., & Sprenger, J. (2011). The logic of explanatory power. *Philosophy of Science*, *78*, 105–127.
Shafir, E., Smith, E., & Osherson, D. (1990). Typicality and reasoning fallacies. *Memory and Cognition*, *18*, 229–239.
Shogenji, T. (1999). Is coherence truth conducive? *Analysis*, *59*, 338–345.
Shogenji, T. (2012). The degree of epistemic justification and the conjunction fallacy. *Synthese*, 184, 29–48.
Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory and Cognition*, *30*, 191–198.
Siebel, M. (2003). There's something about Linda: Probability, coherence and rationality. Unpublished manuscript.
Siebel, M. (2005). Against probabilistic measures of coherence. *Erkenntnis*, *63*, 335–360.
Siebel, M. (2011). Why explanation and thus coherence cannot be reduced to probability. *Analysis*, *71*, 264–266.
Tentori, K., & Crupi, V. (2012). On the conjunction fallacy and the meaning of and yet again: A reply to Hertwig, Benz, and Krauss (2008). *Cognition*, *122*, 123–134.
Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). Cambridge: Cambridge University Press.
von Mises, R. (1957). *Probability, statistics and truth*. New York: Macmillan.