Synthese (2012) 184:3–12
DOI 10.1007/s11229-009-9701-y

How the conjunction fallacy is tied to probabilistic confirmation: Some remarks on Schupbach (2009)

Katya Tentori · Vincenzo Crupi

Received: 1 July 2009 / Accepted: 9 November 2009 / Published online: 2 December 2009

Abstract Crupi et al. (Think Reason 14:182–199, 2008) have recently advocated and partially worked out an account of the conjunction fallacy phenomenon based on the Bayesian notion of confirmation. In response, Schupbach (2009) presented a critical discussion as following from some novel experimental results. After providing a brief restatement and clarification of the meaning and scope of our original proposal, we will outline Schupbach's results and discuss his interpretation thereof arguing that they do not actually undermine our point of view if properly construed. Finally, we will foster such a claim by means of some novel data.

Keywords Confirmation · Conjunction fallacy · Judgmental biases · Probability · Rationality · Reasoning

1 Introduction

Starting from seminal inquiries by Amos Tversky and Daniel Kahneman in the psychology of judgment under uncertainty (Tversky and Kahneman 1982, 1983), a remarkable body of empirical evidence has pointed to the existence of a reasoning error known as the "conjunction fallacy". Such a fallacy amounts to ranking the probability of a conjunctive statement $h_1 \wedge h_2$ over the probability of one of its conjuncts (e.g., h_1), contrary to standard and sound probabilistic principles.

K. Tentori
University of Trento, Trento, Italy

K. Tentori
University College London, London, UK
e-mail: katya.tentori@unitn.it

V. Crupi (✉)
University of Turin, Turin, Italy
e-mail: vincenzo.crupi@unito.it

Among scholars concerned with rationality and reasoning, a strong interest in fallacies traces back to the origins of the discipline—and the conjunction fallacy is no exception. Its occurrence has been widely discussed (see Bovens and Hartmann 2003, pp. 85–88; Gigerenzer 1996; Hintikka 2004; Kahneman and Tversky 1996; Levi 1985; Moro 2008; Stich 1990, pp. 6–7, among others), and yet not fully and satisfactorily explained so far.

On the background of a few earlier suggestions (most notably Sides et al. 2002; also see Tenenbaum and Griffiths 2001), Crupi et al. (2008) (CFT from now on) have recently advocated and partially worked out an account of the conjunction fallacy phenomenon based on the notion of *confirmation* as meant in Bayesian terms (Crupi et al. 2007; Earman 1992; Festa 1999; Fitelson 1999, 2006; Howson and Urbach 2006). In response, Schupbach's (2009) presented a critical discussion of CFT as following from some novel experimental results.

From our standpoint, Schupbach's work has a number of virtues. For one, it is—to the best of our knowledge—the first experimental investigation which explicitly addresses confirmation relations as possible determinants of the conjunction fallacy effect, thus taking seriously a confirmation-theoretically based account of the phenomenon as empirically relevant and testable. On the other hand, however, we also find Schupbach's conclusions questionable in several respects.

In what follows we will display some comments and novel data addressing the main points raised by Schupbach's work and their implications. After providing a brief restatement and clarification of the meaning and scope of CFT's original contribution towards an explanatory account of the conjunction fallacy (Sect. 2), we will outline Schupbach's results and discuss his interpretation thereof (Sects. 3 and 4), arguing that they do not actually undermine CFT's point of view if properly construed. Finally, we will foster such a claim by means of some novel data (Sect. 5).

2 Bayesian confirmation and the conjunction fallacy: a restatement

First of all, CFT meant to advocate a *general* framework for the explanation of the conjunction fallacy phenomenon based on a similarly general working hypothesis, i.e., that in experimental conjunction scenarios participants' responses might be accounted for by taking into consideration the confirmation relations among specific information provided to participants, or otherwise available to them, and the two conjuncts. As argued in CFT's paper (pp. 184–187), such a working hypothesis gains some indirect empirical support from results obtained with judgment tasks akin to (but not strictly included in) psychological inquiries on the conjunction fallacy (e.g., Osherson et al. 1990; Lagnado and Shanks 2002). On CFT's proposed reading, such results document how confirmation relations may indeed affect probability judgments in ways which make them depart from relevant normative standards of reasoning.

It is also very clear from the literature, however, that conjunction scenarios and conjunction effects come in different forms. Indeed, ever since Tversky and Kahneman (1983) extensive investigation, commonly used conjunction scenarios could be roughly split into a few subsets of cases in terms of the material employed. Tversky and Kahneman themselves referred to *two* distinct "paradigms" in which conjunction fallacy effects could be expected. Based on their discussion, the first paradigm (eminently instantiated by the Linda problem, and labelled the "$M \rightarrow A$ paradigm" in Tversky and Kahneman (1983), p. 305), a state of affairs (e) is presented (e.g., Linda's description) of which one of the conjuncts (h_1) is "unrepresentative" (e.g., being a bank teller), the other (h_2) being instead highly "representative" (e.g., being a feminist activist). In the second paradigm (labelled the "$A \rightarrow B$ paradigm" in Tversky and Kahneman (1983, p. 305), on the contrary, Tversky and Kahneman postulated that conjunction fallacy effects may occur depending on a positive association between the two conjuncts h_1 and h_2 (e.g., in the presence of a causal relationship between the corresponding events).

In their contribution, CFT suggested that the relevant structural features of *those conjunction scenarios akin to the Linda problem* (thus presumably belonging to Tversky and Kahneman's "$M \rightarrow A$ paradigm") could be neatly captured in confirmation-theoretic terms on the basis of the following three conditions:

(i) e (e.g., Linda's description) does not confirm h_1 (bank teller);
(ii) e does confirm h_2 (feminist), even conditionally on h_1;
(iii) h_1 and h_2 mildly (if at all) disconfirm each other, even conditionally on e.[1]

This fact may appear quite clear for anyone with some familiarity with Bayesian confirmation theory. Still, CFT suggested, its potential explanatory role does not seem to have been generally and thoroughly appreciated so far. CFT endorse the thesis that people's sensitivity to conditions (i)–(iii) plays a crucial role in leading their intuitive judgments astray from relevant probabilistic principles and in determining the prevalence of the conjunction fallacy effect in the concerned class of experimental problems. In *psychological* (*descriptive*) terms, the idea is the following (see CFT, p. 188). As long as the relevant confirmation relations between h_1 and h_2 are slightly negative or negligible (condition (iii)), participants might have a tendency to make a probabilistic judgment about the conjunction $h_1 \wedge h_2$ reflecting some averaging integration of the different confirmation-theoretic impact of e on h_1 and h_2, respectively. The latter being clearly higher than the former (conditions (i) and (ii)), this would lead to the mistaken judgment of $h_1 \wedge h_2$ as more probable than h_1. In a more formal vein, CFT have also proven that (i) and (ii) are sufficient to yield $c(h_1 \wedge h_2, e) > c(h_1, e)$ for a number of well-known Bayesian confirmation measures, thus implying that on such conditions $h_1 \wedge h_2$ is in fact better confirmed by e than h_1 is.[2]

The outline above represents a first elaboration of CFT's general working hypothesis concerning the relevance of confirmation relations for the conjunction fallacy phenomenon, still requiring quantitative refinements and extensions. The ultimate goal would be the identification of the list of candidate confirmation-theoretic deter-

[1] As a matter of fact, the latter clause ("even conditionally on e") was not included in CFT's original presentation of conditions (i)–(iii). This refinement is rather subtle and, in any event, immaterial for the present discussion. Still, further reflection and ongoing research convinced us that it can indeed be relevant in some cases.

[2] This clearly leaves room for the relationship $c(h_1 \wedge h_2, e) > c(h_1, e)$ obtaining by routes other than the fulfilment of conditions (i) and (ii). Possibilities of this sort are illustrated by Schupbach's (2009) himself (Appendix 1). Atkinson et al. (2009) explored the issue in a more systematically fashion with intriguing results.

minants of the conjunction fallacy along with the quantitative dependencies among such determinants and the size of the effect. Although much work needs to be done to achieve a complete model, some relevant tenets of it can be already presented much along the lines of CFT's original paper. In particular, CFT conjectured "the difference between mean ratings of $c(h_2, e)$ and of $c(h_1, e)$ to be positively correlated with the percentage of conjunction errors" (p. 195). This empirically testable conjecture amounts to assuming the size of the conjunction fallacy effect to yield a *positive* dependency on the perceived degree of confirmation $c(h_2, e)$ and a *negative* dependency on the perceived degree of confirmation $c(h_1, e)$. In hindsight, this conjecture already gains some support from Tversky and Kahneman's findings. In one of their studies, they restricted Linda's description to simply "a 31 years old woman", thus reducing the evidence provided to almost completely irrelevant information for both h_1 ("bank teller") and h_2 ("feminist"). With this modified scenario "almost all respondents obeyed the conjunction rule" (Tversky and Kahneman 1983, p. 305). Notably, Tversky and Kahneman's results also suggest how confirmation relations involving h_1 and h_2 might be psychologically crucial for the conjunction fallacy effect. A case in point involves their character "Bill": $e =$ "34 years old, intelligent but unimaginative, compulsive, and generally lifeless; when in school, strong in maths but weak in the humanities"; $h_1 =$ "plays jazz for a hobby"; $h_2 =$ "is bored by music". No conjunction fallacy effect was observed with this material (Tversky and Kahneman 1983, p. 305). Notice that here quite clearly e does disconfirm h_1 and confirm h_2 (so that conditions (i) and (ii) are both satisfied). Apparently, however, this is compensated by h_1 and h_2 being "highly incompatible" (*ibid.*), i.e., in our terms, *strongly* disconfirming each other (contrary to condition (iii)).

Schupbach's (2009) inquiry on confirmation and the conjunction fallacy is meant to test the tenability of the thesis that conditions (i)–(iii) above are either necessary or sufficient for the occurrence of the conjunction fallacy. In addition, challenging conclusions are drawn concerning the existence and strength of a causal relationship between CFT's conditions and the occurrence of the phenomenon. We will present and discuss these points in turn, starting from the issue about "necessity".

3 On the necessity issue

Schupbach's (2009) experiment 1 was explicitly conceived to investigate whether "CFT's specific conditions are necessary for the conjunction fallacy" (Schupbach's 2009). The following scenario was employed: $e =$ "Linda participates in anti-war protests, votes Democrat, and subscribes to a popular liberal magazine"; $h_1 =$ "Linda is a poet"; $h_2 =$ "Linda is a feminist". Fifty-six participants had to choose the most likely statement between "Linda is a poet" and "Linda is a feminist poet". After that, they expressed some additional probabilistic judgments (e.g., whether h_1 was made more, equally or less likely by coming to know that e) by which relevant qualitative confirmation relations concerning e, h_1 and h_2 were inferred (e.g., $c(h_1, e) >/=/< 0$). On the basis of the latter, participants were then classified as "meeting" or "not meeting" CFT's conditions (i)–(iii).

In his overall sample, Schupbach found a 52% conjunction fallacy rate. Such fallacious judgments, however, were not equally split into the "conditions met" vs. "conditions not met" groups (24 vs. 32 participants, respectively): the fallacy was committed by 70.8% of the former vs. 37.5% of the latter. Schupbach readily recognizes that the significantly higher conjunction fallacy rate in the "conditions met" group does provide prima facie supporting evidence to CFT's conditions as capturing relevant determinants of the conjunction fallacy effect (2009). Beyond that, however, since the conjunction fallacy rate in the "conditions not met" group is higher than 0% (i.e., 37.5%), Schupbach also remarks that "CFT's conditions are not necessary for the conjunction fallacy" (2009). He also suggests that this contributes to point out significant "limitations" in CFT's position (2009). On closer scrutiny, however, this suggestion seems hasty, as we will now argue.

First and foremost, no claim about "necessity" can be detected wherever in CFT's paper. This is because, as already explained (see Sect. 2), conditions (i)–(iii) were meant to capture *one* major class of scenarios (roughly matching the structure of the *original* Linda problem) among those in which a substantial rate of conjunction errors had already been documented in the literature. This focus of analysis was motivated by the fact that no fully satisfactory explanation had been achieved so far even for this limited class of cases, despite it having been extensively investigated (see CFT, pp. 189–194).

The class of Linda-like scenarios was thus the only one explicitly addressed by CFT. Notice that, based on their own responses, participants in Schupbach's "conditions met" group appeared to see the scenario presented as yet another instance of that class and, accordingly, exhibited a high conjunction fallacy rate.

As to participants in Schupbach's "conditions not met" group, their behaviour need not remain unexplained in a confirmation-theoretical approach to the conjunction fallacy when properly extended. Surely a thorough interpretation of these data is somewhat hindered by "conditions not met" being a rather coarse description, for obviously conditions (i)–(iii) may have been violated in a variety of ways whose analysis is not included in Schupbach's (2009) treatment. Still, despite the lack of relevant details about these participants' departures from conditions (i)–(iii), we think that some reasonable speculations can be advanced along the following lines.

To begin with, consider Schupbach's "conditions not met" participants *not* committing the fallacy. We have already mentioned that Tversky and Kahneman themselves found no conjunction fallacy effect with an impoverished description of Linda ("a 31 old woman") which presumably set the perceived values of both $c(h_1, e)$ and of $c(h_2, e)$ to negligible levels. We also noticed how this fits with CFT's hypothesis about the tendency to commit the fallacy being associated with the perceived difference between those two confirmation values (in the presence of mildly negative or null confirmation relations between the conjuncts—condition (iii)). Now it doesn't seem implausible that a substantial proportion of "conditions not met" participants avoiding the fallacy in Schupbach's exp. 1 might have found themselves in a similar situation, i.e., judging Linda's given description (e), "poet" (h_1) and "feminist" (h_2) as simply confirmationally irrelevant to each other (which would amount to the violation of CFT's condition (ii)).

Secondly, consider Schupbach's "conditions not met" participants who *did* commit the fallacy. In the above discussion, and again by reference to Tversky and Kahneman's findings, we have also noticed that h_1 and h_2 being *strongly* at odds may cause the conjunction fallacy rate to drop to zero. Symmetrically, significant relations of positive confirmation between h_1 and h_2 may foster a quite large conjunction fallacy effect even when $c(h_1, e)$ and $c(h_2, e)$ seem to be comparable (so that their difference is negligible). Indeed, a neat demonstration of the conjunction fallacy effect was obtained by Tentori et al. (2004) with their "Scandinavia" scenario, which seems to be precisely of this sort, i.e., $e =$ "x is a (randomly selected) Scandinavian individual", $h_1 =$ "x has blonde hair" and $h_2 =$ "x has blue eyes". It doesn't seem implausible that a substantial proportion of "conditions not met" participants committing the fallacy in Schupbach's exp. 1 might have found themselves in a similar situation, i.e., judging Linda being a poet (h_1) and her being a feminist (h_2) as confirming each other (which would amount to the violation of CFT's condition (iii)).

4 On sufficiency and "causal strength"

A second major point discussed by Schupbach concerns the "sufficiency" of conditions (i)–(iii) for the conjunction fallacy to occur. Schupbach's experiments 2 and 3 were designed to address this issue. In what follows we will focus on the "lottery" scenario (exp. 3), in which $e =$ "Jim is an occasional purchaser of a mega millions lottery ticket", $h_1 =$ "Jim is a scientist", and $h_2 =$ "Jim will win the multi-million dollar jackpot in the lottery". Nothing in the discussion below, however, crucially depends on the reference to these data as compared to those obtained on Schupbach's American Idol problem (exp. 2).

In an overall sample of 21 participants, results showed complete absence (0%) of the conjunction fallacy effect in the "conditions not met" group (7 participants) as compared to a modest effect (14.3%) in the "conditions met" group (14 participants). Schupbach concluded that "CFT's conditions are not sufficient for the conjunction fallacy" (2009) and, moreover, that they only exhibit a "very weak causal relationship" with the phenomenon (2009). For these reasons, he sees the data as quite strongly challenging CFT's (and others') confirmation-theoretic account. Again, we would like to question the latter conclusion.

The intended study design in the lottery case was the same as above, with "conditions met" vs. "conditions not met" as a dichotomic between-subjects predictor variable. Here again, in order to discuss the results, it seems very instructive to consider *how* participants might have departed from conditions (i)–(iii). To our mind, almost the only conceivable way in which this might have occurred is by the neglect of an actual confirmatory relevance of e (occasionally purchasing a ticket for a presumably gigantic lottery) to h_2 (win the lottery). And indeed, *all* of Schupbach's participants in the "conditions not met" group did precisely this (and a strictly analogous pattern occurred in the "American Idol" scenario) (Schupbach, personal communication). As a consequence, the variable which is actually at work here simply amounts to participants' judgments indicating either a null or positive perceived value of $c(h_2, e)$. As we have seen, Schupbach observed a modest difference in conjunction fallacy rates across his two groups of participants (0 vs. 14.3%). Notably, modest effects (say, being only slightly drunk) may depend on at least two very different states of affairs: either a causal factor which is only weakly relevant (say, cider) is largely present (e.g., one bottle drunk), or a quite strongly relevant causal factor (say, wodka) is present only to a small extent (a half glass). Apparently, Schupbach postulated the former kind of situation and concluded that $c(h_2, e) > 0$ causally affects the occurrence of the conjunction fallacy "only to a very weak degree" (2009). Yet the alternative reading can hardly be disregarded: quite simply, in Schupbach's experiment, the quantity c("win a mega millions lottery", "occasionally purchase a ticket") might not have been perceived as sufficiently different from 0, thus not high enough to let the conjunction fallacy rate attain its upper levels. In fact, CFT explicitly made the quantitative prediction that, other things being equal, the higher the assessment of $c(h_2, e)$, "the higher the percentage of conjunction errors should be in the corresponding standard probabilistic task" (p. 195). Schupbach's results, then, do display a tendency which is in line with CFT's prediction, suggesting that the difference between a null and a positive (and presumably low) perceived value of $c(h_2, e)$ actually yields a (limited) increase of the probability of committing the fallacy.

It should be noticed that the size of samples in the lottery experiment was small and that the crucial *quantitative* extent of assessments concerning $c(h_2, e)$ could not be measured or otherwise controlled for by means of the procedure employed. So the finding should be taken with caution. A more robust pattern of results and a more telling test of CFT's prediction would involve a systematic manipulation of $c(h_2, e)$ across a number of quantitative levels carried out in somewhat larger samples of participants. By exploiting and adapting Schupbach's ingenious lottery scenario, we did perform such an experiment, as illustrated in the next section.

5 A direct quantitative test of CFT's prediction

The present experiment was meant to systematically investigate whether the conjunction fallacy rate increases as the confirmatory impact of the evidence provided, e, on the added conjunct, h_2, is raised. Participants read a brief scenario involving a character (Mark) and a small-scale local lottery counting 100 tickets. They were presented with two hypotheses: the single conjunct "Mark is a scientist" (h_1) and the conjunction "Mark is a scientist and will win the lottery" ($h_1 \wedge h_2$). Across subjects, the order of presentation of the two hypotheses (the single conjunct vs. the conjunction) was balanced. The hypothesis which appeared first was labelled "A", the other one "B". Participants' task was then to indicate which (if any) was more probable among A and B. The degree of confirmation on the added conjunct h_2 ("Mark will win the lottery") was employed as an independent variable and set at six increasing levels as determined by the evidence provided (e) concerning the number of tickets out of 100 [*none*, *1*, *20*, *50*, *80*, *all*] that Mark was said to have succeeded to buy. (See the Appendix for the complete scenario and the experimental stimulus.)

The design was between subjects, so we had six groups of participants overall, one for each level of the independent variable $c(h_2, e)$. The dependent variable was the

Table 1 Percentages of participants choosing each of the possible rankings of h_1 ("Mark is a scientist") and $h_1 \wedge h_2$ ("Mark is a scientist and will win the lottery"). Figures in the third column represent conjunction fallacy responses. The number of participants in each experimental group appears in the last column

No. of tickets	$P(h_1 \mid e) > P(h_1 \wedge h_2 \mid e)$, %	$P(h_1 \mid e) < P(h_1 \wedge h_2 \mid e)$, %	$P(h_1 \mid e) = P(h_1 \wedge h_2 \mid e)$, %	N
0	89	0	11	36
1	66	16	18	38
20	70	20	10	40
50	50	28	22	40
80	29	47	24	38
100	3	55	42	40

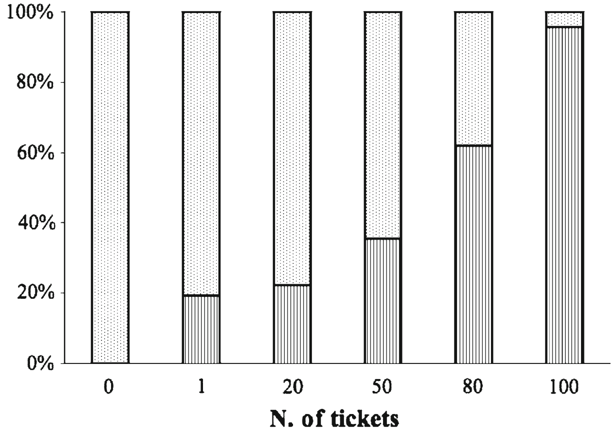

Fig. 1 Relative proportion of responses indicating the conjunction $h_1 \wedge h_2$ (striped areas) vs. the conjunct h_1 (dotted areas) as the more probable hypothesis in each experimental condition

judged probability ranking between h_1 and $h_1 \wedge h_2$, i.e.: $P(h_1|e) > P(h_1 \wedge h_2|e)$ vs. $P(h_1|e) = P(h_1 \wedge h_2|e)$ vs. $P(h_1|e) < P(h_1 \wedge h_2|e)$. Results are reported in Table 1.

As it can be seen, the data clearly display the expected pattern: the percentage of rankings amounting to the conjunction fallacy—i.e., $P(h_1|e) < P(h_1 \wedge h_2|e)$—turned out to be strictly increasing with the increase of $c(h_2, e)$. To appreciate the relative proportion of responses indicating the conjunct h_1 vs. the conjunction $h_1 \wedge h_2$ as the more probable hypothesis, consider Fig. 1, in which equal probability judgments—i.e., $P(h_1|e) = P(h_1 \wedge h_2|e)$—have been left aside. As the figure shows, with $c(h_2, e)$ at its minimum (corresponding to $e =$ "0 tickets bought", which implies h_2 being false) no one judged $P(h_1 \wedge h_2|e)$ as higher than $P(h_1|e)$. Furthermore, the relative amount of preferences for $P(h_1 \wedge h_2|e)$) over $P(h_1|e)$ steadily increases as $c(h_2, e)$ increases, approaching 100% with the latter value at its maximum level (corresponding to $e =$ "all 100 tickets bought", which implies h_2 being true).

Conclusion

In our view, the above results importantly extend Schupbach's findings from his own lottery experiment, and concurrently foster our interpretation of those findings. Participants' responses on both Schupbach's and our experiments lend support to CFT's quantitative prediction concerning a positive relationship between the degree of confirmation $c(h_2, e)$ involving the added conjunct and the conjunction fallacy rate. Such a relationship, while naturally emerging from CFT's reconstruction and analysis of a major class of experimental scenarios from the literature, provides a first contribution towards a more detailed characterisation of the possible confirmation-theoretic determinants of conjunction fallacy effects.

Acknowledgements Research supported by a grant from the SMC/Fondazione Cassa di Risparmio di Trento e Rovereto for the CIMeC (University of Trento) research project on *Inductive reasoning* and by a grant from the Spanish Department of Science and Innovation (FFI2008-01169/FISO). Parts of this work have been presented at the workshop *Probability, confirmation and fallacies*, Center for Logic and Analytical Philosophy, University of Leuven. We would like to especially thank Branden Fitelson, Jonah Schupbach and Tomoji Shogenji for useful discussions.

Appendix

Each year at the beginning of summer a party takes place in the town where Mark lives. The party ends with a fundraising lottery counting 100 tickets and raffles off a dinner with a famous guest.

In 2008, the dinner guest is Mark's favourite Italian female writer. Mark succeeded to buy [*no*, *1*, *20*, *50*, *80*, *all*] lottery ticket[s].

Consider the following hypotheses:

A: Mark is a scientist
B: Mark is a scientist and will win the lottery

Which of the following statements do you think is correct?
(Please indicate your answer with a cross [X] in the corresponding box below.)
[ ] A is more probable than B
[ ] B is more probable than A
[ ] A and B are equally probable

References

Atkinson, D., Peijnenburg, J., & Kuipers, T. (2009). How to confirm the disconfirmed: On conjunction fallacies and robust confirmation. *Philosophy of Science, 76*, 1–21.

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford, UK: Oxford University Press.

Crupi, V., Tentori, K., & Gonzalez, M. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science, 74*, 229–252.

Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation and the conjunction fallacy. *Thinking and Reasoning, 14*, 182–199.

Earman, J. (1992). *Bayes or bust?* Cambridge, MA: MIT Press.

Festa, R. (1999). Bayesian confirmation. In M. Galavotti & A. Pagnini (Eds.), *Experience, reality, and scientific explanation* (pp. 55–87). Dordrecht: Kluwer.

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science, 66*, S362–S378.

Fitelson, B. (2006). Logical foundations of evidential support. *Philosophy of Science, 73*, 500–512.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky. *Psychological Review, 103*, 592–596.

Hintikka, J. (2004). A fallacious fallacy? *Synthese, 140*, 25–35.

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach*. Chicago/La Salle, IL: Open Court.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582–591.

Lagnado, D. A., & Shanks, D. R. (2002). Probability judgment in hierarchical learning: A conflict between predictiveness and coherence. *Cognition, 83*, 81–112.

Levi, I. (1985). Illusions about uncertainty. *British Journal for the Philosophy of Science, 36*, 331–340.

Moro, R. (2008, forthcoming). On the nature of the conjunction fallacy. *Synthese*.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185–200.

Schupbach, J. N. (2009). Is the conjunction fallacy tied to probabilistic confirmation? *Synthese*. doi:10.1007/s11229-009-9698-2.

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory and Cognition, 30*, 191–198.

Stich, S (1990). *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. Cambridge, MA: MIT Press.

Tenenbaum, J. B., & Griffiths, T. L. (2001). The rational basis of representativeness. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 1036–1041). Hillsdale: Erlbaum.

Tentori, K., Bonini, N., & Osherson, D. (2004). The conjunction fallacy: A misunderstanding about conjunction?. *Cognitive Science, 28*, 467–477.

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, UK: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional *vs*. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.