

## COMMENTARY

# Noisy Probability Judgment, the Conjunction Fallacy, and Rationality: Comment on Costello and Watts (2014)

Vincenzo Crupi  
University of Turin

Katya Tentori  
University of Trento

According to Costello and Watts (2014), probability theory can account for key findings in human judgment research provided that random noise is embedded in the model. We concur with a number of Costello and Watts's remarks, but challenge the empirical adequacy of their model in one of their key illustrations (the conjunction fallacy) on the basis of recent experimental findings. We also discuss how our argument bears on heuristic and rational thinking.

*Keywords:* conjunction fallacy, rationality, heuristics, inductive confirmation

An assumption shared by competing approaches to the study of cognition is that human judgment under uncertainty is governed by so-called heuristics rather than by the principles of the probability calculus (see Gilovich, Griffin, & Kahneman, 2002; Gigerenzer, Hertwig, & Pachur, 2011). Costello and Watts (2014) leveled an interesting challenge against this popular view. In their opinion, a probabilistic model of human reasoning is able to account for observed behavior provided that it embeds the role of random noise in the judgment process. They illustrated their point with analyses of some key examples, and concluded that, in none of them, is the appeal to heuristics required to explain the findings.

According to Costello and Watts (2014), the “surprising rationality” of human judgment is primarily supported by the close agreement between people's probability estimates and the requirements of probability theory for expressions such as,  $P(x \wedge y) + P(x \vee y) - P(x) - P(y) = 0$ , for which—Costello and Watts submitted—the effect of noise is cancelled out. Costello and Watts then proceeded to argue that their model explains a number of well-known biases in probabilistic reasoning. In what follows, we focus on the latter claim, particularly on the cornerstone case of the conjunction fallacy, to which a good deal of Costello and Watts's efforts were devoted. We start with some clarification remarks as to how Costello and Watts's idea of “surprising rationality” might square with the occurrence of biases like the conjunction fallacy. We then question directly their claim that the conjunction fallacy can be explained by combining probability theory with noisy reasoning processes in the way that they have proposed. This is not a minor point in their argument, because, for Costello and Watts, showing “that observed biases cannot be explained as the result of random noise” would “demonstrate conclusively that people are

using heuristics” (p. 478). Accordingly, we conclude with a brief discussion of heuristic and rational thinking.

### The Conjunction Fallacy: Neither Rational, nor Explained by Probability and Noise

Costello and Watts clearly accept that compliance of graded subjective credences with the probability calculus is an adequate norm of rationality. Although they do not elaborate explicitly, this view is popular and backed by traditional strategies of justification, such as Dutch book arguments (see Hahn, 2014; Osherson, 1995; Vineberg, 2011) or results concerning accuracy dominance avoidance (D'Agostino & Sinigaglia, 2010; Leitgeb & Pettigrew, 2010a, 2010b; Pettigrew, 2013; Predd et al., 2009). Costello and Watts also have made clear that the conjunction fallacy is real, that is, that the phenomenon is documented by experimental investigations that are methodologically compelling. Here, we will take both of the foregoing assumptions for granted (see Tentori, Bonini, & Osherson, 2004; Tentori & Crupi 2012a; Crupi & Girotto, 2014; Cruz, Baratgin, Oaksford, & Over, 2015, as regard the latter, in particular). Against this background, Costello and Watts also have claimed that their probabilistic model accounts for major biases, including the conjunction fallacy, while insisting that human judgment is “surprisingly rational,” as the title of their article states. This seems to call for some clarifications as to what Costello and Watts's idea of surprising rationality might or might not imply. Because Costello and Watts's indications seem rather scant, we will briefly address this point before taking issue with whether their model in fact explains the conjunction fallacy.

Let us grant for the moment that Costello and Watts's model satisfactorily explains the conjunction fallacy. If so, then although human reasoning would not arise from heuristic rules, which depart systematically from probabilistic principles, it would still be true that systematically biased judgments ensue. That is, even if “conjunction fallacy responses” are not “systematically influenced” by any factor other than the “systematic distorting influence of noise” (Costello & Watts, 2014, p. 477), they would be systematically biased nonethe-

---

Vincenzo Crupi, Center for Logic, Language, and Cognition, University of Turin; and Katya Tentori, Center for Mind/Brain Science, University of Trento.

Correspondence concerning this article should be addressed to Vincenzo Crupi, Center for Logic, Language, and Cognition, University of Turin, via Sant'Ottavio 20, 10124 Turin, Italy. E-mail: [vincenzo.crupi@unito.it](mailto:vincenzo.crupi@unito.it)

less. To illustrate, participants have been shown to be willing to gamble on a conjunction  $x \wedge y$  when they could just as easily gamble on  $x$  alone (as in Bonini, Tentori, & Osherson, 2004; Nilsson & Andersson, 2010), a tendency that seems beyond any attempt at rational accommodation. In this respect, granting the normativity of probability theory and the reality of the conjunction fallacy, the alleged explanatory success of Costello and Watts's model would still not justify the conclusion that people are behaving rationally in the experimental tasks at issue. Here, we do not mean to ascribe such conclusion to Costello and Watts, yet we suggest that its rejection would better be spelled out explicitly, because it makes it clear that Costello and Watts's "surprising rationality" still falls short of quite basic aspects of Bayesian rationality.

Let us now consider our main target of discussion, namely, whether Costello and Watts's attempted explanation of the conjunction fallacy succeeds. Indeed, in Costello and Watts's article, this is a point of major interest in itself. Once again, we subscribe without reservation to a key starting point of Costello and Watts's approach, to wit, that a satisfactory explanation of the conjunction fallacy should be able to account both for when it occurs and for when it does *not* occur. To illustrate, Costello and Watts (2014) considered a naive averaging model where the judged probability of a conjunction of hypotheses  $x \wedge y$  amounts to a weighted average of the judged probabilities of the conjuncts. This account predicts that the fallacious assessment of  $P(x \wedge y)$  as being higher than either  $P(x)$  or  $P(y)$  is ubiquitous, no matter what  $x$  and  $y$  are and what prior evidence is available (p. 469). But surely, as Costello and Watts (2014) have rightly pointed out, the fallacy "does not occur for all or even most conjunctions" (p. 467). When does it occur, then? What are its determinants?

In Costello and Watts's (2014) model, a reasoner would have her "true" subjective probabilities  $P(x)$  and  $P(x \wedge y)$  satisfying the standard axioms, so that  $P(x) \geq P(x \wedge y)$ , according to the conjunction rule of probability theory. Compliance with the probability axioms is guaranteed because these values arise from an underlying memory store of instances making each of  $x$  and  $y$  true or false. However, as the two values  $P(x)$  and  $P(x \wedge y)$  approach each other, the chances increase that occasional estimates of those values—labeled  $P_E(x)$  and  $P_E(x \wedge y)$  in Costello and Watts's notation—violate the conjunction rule, for such occasional judgments would fluctuate randomly around the (probabilistically coherent) true credences. In fact, it is clear from Costello and Watts's (2014) analysis that the expected occurrence of the conjunction fallacy about statements  $x$  and  $x \wedge y$  should be inversely proportional to the difference  $P(x) - P(x \wedge y)$  (see pp. 467–468). We take it as a virtue of their proposal that this key assumption is quite directly testable.

Suppose that, given some evidence  $e$ , the most likely statement has to be chosen among three, that is, a single hypothesis  $h_1$  and two conjunctions  $h_1 \wedge h_2$  and  $h_1 \wedge h_3$ , as in the scenario below.

O. Received a violin degree. [ $e$ ]

Which of the following hypotheses do you think is the most probable?

- O. Is an expert mountaineer. [ $h_1$ ]
- O. Is an expert mountaineer and gives music lessons. [ $h_1 \wedge h_2$ ]
- O. Is an expert mountaineer and owns an umbrella. [ $h_1 \wedge h_3$ ]

In one of their studies, Tentori, Crupi, and Russo (2013) observed that mean estimates were largely and reliably lower for  $P(h_2|e \wedge h_1)$  than for  $P(h_3|e \wedge h_1)$  (35% vs. 67%). So, again, following Costello and Watts's (2014) notation,

$$P_E(h_2|e \wedge h_1) < P_E(h_3|e \wedge h_1)$$

These are conditional probability judgments, thus lying outside the scope of Costello and Watts's (2014) original model (see p. 464). However, one can postulate the true probabilistic credence  $P(x|y)$  to amount to the proportion of  $x$ -instances among  $y$ -instances in stored memory, and the estimate  $P_E(x|y)$  itself to be subject to random variation. This natural way to extend Costello and Watts's (2014) model of "simple" probabilities to conditional probabilities is explicitly pursued by Costello and Watts in subsequent work (Costello & Watts, 2015). Because by assumption the error is random, noisy probability estimates will preserve, on average, the rank order of the corresponding "true" probabilities. Thus, the above ranking of mean estimates  $P_E(h_2|e \wedge h_1) < P_E(h_3|e \wedge h_1)$  must reflect the corresponding ranking  $P(h_2|e \wedge h_1) < P(h_3|e \wedge h_1)$ , so that, by the probability calculus, we then have

$$P(h_2|e \wedge h_1)P(h_1|e) < P(h_3|e \wedge h_1)P(h_1|e)$$

$$P(h_1 \wedge h_2|e) < P(h_1 \wedge h_3|e)$$

$$P(h_1|e) - P(h_1 \wedge h_2|e) > P(h_1|e) - P(h_1 \wedge h_3|e)$$

Because Costello and Watts have maintained that the chances of a fallacious ranking of  $P(x \wedge y)$  over  $P(x)$  are inversely proportional to  $P(x) - P(x \wedge y)$ , a neat prediction now follows for the violin scenario: given evidence  $e$ , fallacious choices of  $h_1 \wedge h_2$  as the most probable statement must be fewer than those of  $h_1 \wedge h_3$ . However, in contrast with this prediction, a large majority (83%) of the fallacious responses (which were 24 overall, from a total of 30 judgments) concerned  $h_1 \wedge h_2$  rather than  $h_1 \wedge h_3$  (Tentori et al., 2013, Experiment 2).

The violin scenario was meant to convey some key features of Tversky and Kahneman's (1983, p. 305) *M-A* paradigm of conjunction problems, mostly known through the Linda case, where some psychologically salient connection is assumed to exist between a relevant "model"  $M$  (i.e., Linda's description) and the added conjunct  $A$  (being a feminist activist; also see below). Other scenarios employed by Tentori et al. (2013) are instead inspired by so-called *A-B* paradigm, with no specific information conveyed at the outset to describe or evoke a "model," but rather an added conjunct  $A$  providing a "plausible cause or motive" for  $B$  (i.e., the "basic" hypothesis of interest, which is displayed both in isolation and within the conjunctive statement). Here is a simple example:

Do you think it is most probable that a person

- Is Swiss. [ $h_1$ ]
- Is Swiss and can ski. [ $h_1 \wedge h_2$ ]
- Is Swiss and has a driving license. [ $h_1 \wedge h_3$ ]

With this material, 79% of the fallacies concerned  $h_1 \wedge h_2$  rather than  $h_1 \wedge h_3$  (in this case, fallacious responses were 19 overall, from a total of 40 judgments). However, mean estimates of  $P(h_2|h_1)$  and  $P(h_3|h_1)$  were statistically undistinguish-

able (76% vs. 75%). Therefore, we have  $P(h_1) - P(h_2|h_1) = P(h_1) - P(h_2|h_1)P(h_1) = P(h_1) - P(h_3|h_1)P(h_1) = P(h_1) - P(h_3|h_1)$ , so that Costello and Watts's model again fails to account for the observed imbalance in conjunction fallacy responses.

The foregoing findings come from a series of four experimental studies (varying in elicitation procedures, experimental designs, classes of problems, and content) consistently challenging the widespread idea that the occurrence of the conjunction fallacy crucially depends on how likely the added conjunct is judged to be. The results contradict this idea, which—as Tentori et al. (2013) argued—is a shared implication of a varied collection of accounts of the conjunction fallacy, including weighted average (Fantino, Kulik, Stolarz-Fantino, & Wright, 1997), configural weighted average (Nilsson, Winman, Juslin, & Hansson, 2009), multiplicative combination rules either with configural weights (Einhorn, 1985) or without (Birnbaum, Anderson, & Hynan, 1990), signed summation (Yates & Carlson, 1986), and Bayesian networks with source reliability (Bovens & Hartmann, 2003). Tentori et al. (2013) observed that the same holds for an earlier version of Costello and Watts's noisy recall model (see Costello, 2009). Because Costello and Watts (2014) did not address these results, pointing out their negative import seems justified and important.

To sum up, even assuming the explanatory adequacy of Costello and Watts's noisy recall model for the cornerstone case of the conjunction fallacy, conjunction fallacy findings themselves clearly show that human judgment systematically violates probabilistic principles of rationality. Moreover, as we have argued, the idea that the noisy recall model does account for the conjunction fallacy phenomenon does not stand scrutiny once relevant and recent experimental results are considered.

### Irrationality Explained by Subtle Heuristic Thinking

Tentori et al. (2013) devised their experiments to test major extant accounts of the conjunction fallacy against their proposal that *inductive confirmation* (or *evidential impact*) is the key determinant for its occurrence. In Bayesian confirmation theory,  $x$  is said to be confirmed by  $y$ , formally  $\text{conf}(x,y) > 0$ , if  $y$  increases the initial probability of  $x$ , that is, if  $P(x|y) > P(x)$ . In the violin scenario, although overall less probable than  $h_3$  (owning an umbrella),  $h_2$  (giving music lessons) is more strongly confirmed by the evidence,  $e$  (having a violin degree). More precisely, as explained above, people judged  $P(h_2|e|h_1) < P(h_3|e|h_1)$ . However, they also judged  $\text{conf}(h_2,e|h_1) > \text{conf}(h_3,e|h_1)$  (see Tentori, Crupi, & Osherson, 2007; Tentori, Chater, & Crupi, 2015, for the elicitation of confirmation judgments).<sup>1</sup> Accordingly, Tentori et al.'s (2013) account predicted the observed pattern of conjunction fallacy responses in this and other instances of the *M-A* paradigm, that is, a majority of fallacious responses for  $h_1 \wedge h_2$  rather than  $h_1 \wedge h_3$ .

In the *A-B* paradigm, on the other hand, the key explanatory factor is the strength of the confirmation relation from the isolated conjunct to the added conjunct. In the Swiss man scenario, for instance, judgments implied that  $P(h_2|h_1)$  and  $P(h_3|h_1)$  were indistinguishable, but also that  $\text{conf}(h_2,h_1) > \text{conf}(h_3,h_1)$ , thus explaining more fallacies concerning  $h_1 \wedge h_2$  rather than  $h_1 \wedge h_3$ .<sup>2</sup> Notably, conjunction fallacy responses in

Costello and Watts's (2014) experiments fit the *A-B* paradigm and therefore are amenable to the same kind of analysis: they arise with statements such as “an ordinary day in Ireland will be rainy and cloudy,” where clearly a positive relation of confirmation occurs between the conjuncts (see p. 473).

If inductive confirmation rather than probability plus random noise explains the conjunction fallacy, what can be said about another main target of Costello and Watts's discussion, that is, the role of heuristics in cognition? A fresh look at two classical examples may clarify matters. First, consider the Linda scenario, in which the following points appear compelling.

- (i) Linda's description ( $e$ ) does not confirm “bank teller” ( $h_1$ ). More precisely,  $P(h_1|e) \leq P(h_1)$  so that  $\text{conf}(h_1,e) \leq 0$ .
- (ii) Linda's description  $e$  does confirm “feminist activist” ( $h_2$ ), even conditionally on  $h_1$ . More precisely,  $P(h_2|e|h_1) > P(h_2|h_1)$ , so that  $\text{conf}(h_2,e|h_1) > 0$ .
- (iii)  $h_1$  and  $h_2$  mildly (if at all) disconfirm each other, even conditionally on  $e$ . More precisely,  $P(h_2|e|h_1)$  may be lower than  $P(h_2|e)$ , but only slightly so, thus  $\text{conf}(h_2,h_1|e)$  is possibly negative, but relatively small.

As shown by Tentori et al.'s (2013) studies on the *M-A* paradigm, as long as the relevant confirmation relations between  $h_1$  and  $h_2$  are slightly negative or negligible (Condition iii), participants may mistakenly rank the conjunction  $h_1 \wedge h_2$  as more probable than  $h_1$  because of the perception that the added conjunct,  $h_2$ , obtains some positive degree of inductive confirmation in the scenario (Condition ii), whereas the isolated conjunct,  $h_1$ , does not (Condition i). This of course recalls Tversky and Kahneman's (1983) idea that Linda's description is “highly representative” of a feminist activist, quite unrepresentative of a bank teller, and partly representative of a feminist bank teller. However, unlike the latter largely informal account, Bayesian confirmation theory delivers a formal proof that, on Conditions i–iii above, the following holds<sup>3</sup>:

$$\text{conf}(h_1 \wedge h_2, e) > \text{conf}(h_1, e) \quad (1)$$

<sup>1</sup> The confirmatory impact of  $e$  on, say,  $h_2$  conditional on  $h_1$  is formally and conceptually distinct from how  $h_2$  is affected by  $e$  and  $h_1$  taken as a joint item of evidence. The former quantity— $\text{conf}(h_2,e|h_1)$ —reflects the relationship between  $Pr(h_2|e|h_1)$  and  $Pr(h_2|h_1)$ ; for the latter—denoted as  $\text{conf}(h_2,e \wedge h_1)$ —the relevant probabilistic values are  $Pr(h_2|e \wedge h_1)$  and  $Pr(h_2)$ , instead. This distinction is entirely standard in Bayesian confirmation theory and is more extensively discussed in Tentori et al. (2013).

<sup>2</sup> One should note that, theoretically, there are many nonequivalent ways to formalize the confirmation function  $\text{conf}$  (see Crupi & Tentori, 2013; Crupi, Chater, & Tentori, 2013; Festa, 2012; Brössel, 2013; Glass, 2013; Roche & Shogenji, 2014). This choice is left open in Tentori et al. (2013), for their conclusions are invariant across the main options available.

<sup>3</sup> The confirmation-theoretic account of the *M-A* paradigm is not only more precise, but also more comprehensive than the representativeness heuristic in several ways. For instance, Gavanski and Roskos-Ewoldsen (1991) suggested that conjunction fallacies for statements such as “Linda is a bank teller and Jason is an artist” are problematic, because no single “representative model” exists for two distinct characters. However, these scenarios naturally fit our framework, because Gavanski and Roskos-Ewoldsen (1991) had a cover story about both Linda and Jason clearly providing supporting evidence for the added conjunct “Jason is an artist” (and not for “Linda is a bank teller”).



That is, in the Linda problem, the conjunctive hypothesis  $h_1 \wedge h_2$  demonstrably happens to be more strongly confirmed than  $h_1$  (see Crupi, Fitelson, & Tentori, 2008, for the details and references to earlier contributions in a similar vein).<sup>4</sup> This result allows for a sharp rendition of Tversky and Kahneman's (1983) remark that "feminist bank teller is a better hypothesis about Linda than bank teller" (p. 311)—indeed, the former, compared with the latter, is better confirmed by Linda's description (see the Appendix for a numerical illustration).<sup>5</sup> Accordingly, it makes sense to say that "the answer to a question [probability] can be biased by the availability of an answer to a cognate question [confirmation]" (p. 312, square brackets added).

Not only does this analysis readily extend to all instances of the *M-A* paradigm as defined by Conditions i–iii above (e.g., the violin scenario), but it also yields a parallel reconstruction of the *A-B* paradigm, which would lie beyond the scope of the representativeness heuristic reading in its original form, thus achieving a higher level of theoretical unity within Tversky and Kahneman's "attribute-substitution" overarching scheme (see Kahneman & Frederick, 2002).

Consider Tversky and Kahneman's (1983) health survey scenario, with two conjuncts stating that a randomly selected adult male, Mr. F., has had one or more heart attacks ( $h_1$ ) and is over 55 years of age ( $h_2$ ). Unlike Linda, Mr. F. purposely lacks any specific characterization at the outset, so that, in our current notation, the explicit evidence  $e$  would be empty (i.e., tautological). In Tversky and Kahneman's (1983) view, the conjunction fallacy may arise here because the added conjunct  $h_2$  provides a "plausible cause or motive" for  $h_1$  (p. 305). A counterpart confirmation-theoretic rendition rests on the consideration that the observation of effects provides inductive confirmation for the occurrence of their plausible causes. Accordingly, the key feature of the *A-B* paradigm turns out to be the confirmation-theoretic connection between the two conjuncts themselves,  $h_1$  and  $h_2$ , as implied by participants' own background knowledge and beliefs. As a useful modeling exercise, we can then isolate the relevant piece of background knowledge involved and represent it as  $r$ . To emphasize,  $r$  is now assumed to convey the participant's view of how  $h_1$  (e.g., having had one or more heart attacks) and  $h_2$  (being over age 55) are related in the real world. Obviously, it is only against the background of  $r$  that  $h_1$  confirms  $h_2$ . Thus, with  $r$  separately represented, we have  $\text{conf}(h_2, h_1 | r) > 0$ , while  $\text{conf}(h_2, h_1) = 0$ . On the other hand, as by definition  $r$  only concerns the relationship between the conjuncts, it cannot affect the credibility of any of them in and by itself, so that  $\text{conf}(h_1, r) = \text{conf}(h_2, r) = 0$ . Notably, these conditions are sufficient to derive a strict *A-B* analog of (1) above, that is,<sup>6</sup>

$$\text{conf}(h_1 \wedge h_2, r) > \text{conf}(h_1, r) \quad (2)$$

Hence, in this reconstruction, the exemplars of the *M-A* and *A-B* paradigms are effectively brought back to a unified pattern. In both kinds of cases, some piece of given information—either explicitly submitted to participants (as in the Linda scenario) or otherwise antecedently available to them (as in the health survey scenario)—inductively confirms the conjunction of hypotheses,  $h_1 \wedge h_2$ , to a greater extent than one of those hypotheses, that is,  $h_1$ .

The idea that people rely on detecting relations of confirmation in order to assess likelihoods under uncertainty is not only well supported by conjunction fallacy research, but also consistent with other related findings (Crupi et al., 2008; Tentori et al., 2013, provide more discussion). Furthermore, as suggested in Tentori et al. (2015), this arrangement is broadly effective, because in ordinary circumstances hypotheses that are strongly supported by salient evidence also tend to be quite probable overall, so that the two kinds of assessments tend to converge. Much as is commonly said of traditional heuristics, judging probabilities by confirmation often yields accurate results. However, it can lead to systematic errors when probability and confirmation are dissociated, as happens in several influential experimental paradigms concerning uncertain judgment. When this is the case, people prove to be prone to judgmental biases—some of which are now widely known—whose direction and magnitude depend on the relevant confirmation-theoretic patterns.

## Conclusion

Costello and Watts's argument against heuristics is stimulating but ultimately unconvincing, for their nonheuristic model is disproven by available evidence, as we explained above. According to some influential advocates, on the other hand, heuristic thinking achieves high degrees of accuracy in an adaptive way, despite exploiting a very limited array of basic cognitive tools (see Gigerenzer, Todd, & the ABC Research Group, 1999, for a now classic statement; see Juslin, Nilsson, & Winman, 2009, for a consonant view). Yet heuristics need not be unsophisticated. As far as the conjunction fallacy is concerned, we suggest that a somewhat opposite situation prevails, which is, in our view, no less interesting. That is, even when heuristic reasoning is defective in its final outcome, it may still be tracking subtle notions of high importance, like sound relations of evidential impact. Hence, subtly irrational, yet still irrational.

<sup>4</sup> The relationship  $\text{conf}(h_1 \wedge h_2, e) > \text{conf}(h_1, e)$  can also obtain by routes other than the fulfillment of the conditions considered here. Possibilities of this sort are illustrated by Schupbach (2012, Appendix 1; also see Tentori & Crupi, 2012b, for discussion) and explored in a more systematic fashion by Atkinson, Peijnenburg, and Kuipers (2009), with intriguing results.

<sup>5</sup> See Cevolani, Crupi, and Festa (2010), Hartmann and Meijs (2012), and Shogenji (2012) for partly alternative ways to flesh out this remarkable statement.

<sup>6</sup> For a proof, note that, by hypothesis,  $\text{conf}(h_1, r) = \text{conf}(h_2, r) = 0$  and  $\text{conf}(h_2, h_1 | r) > 0$ , while  $\text{conf}(h_2, h_1) = 0$ . Thus,  $P(h_1 | r) = P(h_1)$  and  $P(h_2 | h_1 \wedge r) > P(h_2 | r) = P(h_2) = P(h_2 | h_1)$ . Then we have  $P(h_2 \wedge h_1 | r) = P(h_2 | h_1 \wedge r)P(h_1 | r) = P(h_2 | h_1 \wedge r)P(h_1) > P(h_2)P(h_1) = P(h_2 \wedge h_1)$ , thus  $\text{conf}(h_2 \wedge h_1, r) > 0 = \text{conf}(h_1, r)$ . The original derivation of this result is credited to Tomoji Shogenji (personal communication, 2008).

## References

- Atkinson, D., Peijnenburg, J., & Kuipers, T. (2009). How to confirm the conjunction of two disconfirmed hypotheses. *Philosophy of Science*, 76, 1–21. <http://dx.doi.org/10.1086/598164>
- Birnbaum, M. H., Anderson, C. J., & Hynan, L. G. (1990). Theories of bias in probability judgment. In J. P. Caverni, J. M. Fabre, & M. Gonzalez (Eds.), *Advances in psychology. Vol. 68: Cognitive biases* (pp. 477–499). North-Holland, The Netherlands: Elsevier. [http://dx.doi.org/10.1016/S0166-4115\(08\)61338-2](http://dx.doi.org/10.1016/S0166-4115(08)61338-2)

- Bonini, N., Tentori, K., & Osherson, D. (2004). A different conjunction fallacy. *Mind & Language*, *19*, 199–210. <http://dx.doi.org/10.1111/j.1468-0017.2004.00254.x>
- Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. New York, NY: Oxford University Press.
- Brössel, P. (2013). The problem of measure sensitivity redux. *Philosophy of Science*, *80*, 378–397. <http://dx.doi.org/10.1086/671172>
- Cevolani, G., Crupi, V., & Festa, R. (2010). The whole truth about Linda: Probability, verisimilitude, and a paradox of conjunction. In M. D'Agostino, G. Giorello, F. Laudisa, T. Pievani, & C. Sinigaglia (Eds.), *New trends in logic and philosophy of science* (pp. 603–615). London, England: College Publications.
- Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*, 463–480. <http://dx.doi.org/10.1037/a0037010>
- Costello, F., & Watts, P. (2015). *People's conditional probability judgments follow probability theory (plus noise)*. Manuscript submitted for publication.
- Costello, F. J. (2009). How probability theory explains the conjunction fallacy. *Journal of Behavioral Decision Making*, *22*, 213–234. <http://dx.doi.org/10.1002/bdm.618>
- Crupi, V., Chater, N., & Tentori, K. (2013). New axioms for probability and likelihood ratio measures. *The British Journal for the Philosophy of Science*, *64*, 189–204. <http://dx.doi.org/10.1093/bjps/axs018>
- Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation and the conjunction fallacy. *Thinking & Reasoning*, *14*, 182–199. <http://dx.doi.org/10.1080/13546780701643406>
- Crupi, V., & Girotto, V. (2014). From is to ought, and back: How normative concerns foster progress in reasoning research. *Frontiers in Psychology*, *5*, 219. <http://dx.doi.org/10.3389/fpsyg.2014.00219>
- Crupi, V., & Tentori, K. (2013). Confirmation as partial entailment: A representation theorem in inductive logic [Erratum in *Journal of Applied Logic*, *12*, 230–231]. *Journal of Applied Logic*, *11*, 364–372. <http://dx.doi.org/10.1016/j.jal.2013.03.002>
- Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology*, *6*, 192.
- D'Agostino, M., & Sinigaglia, C. (2010). Epistemic accuracy and subjective probability. In M. Suárez, M. Dorato, & M. Rédei (Eds.), *Epistemology and methodology of science* (pp. 95–105). Berlin, Germany: Springer.
- Einhorn, J. H. (1985). *A model of the conjunction fallacy*. Chicago, IL: University of Chicago, Center for Decision Research.
- Fantino, E., Kulik, J., Stolarz-Fantino, S., & Wright, W. (1997). The conjunction fallacy: A test of averaging hypotheses. *Psychonomic Bulletin & Review*, *4*, 96–101. <http://dx.doi.org/10.3758/BF03210779>
- Festa, R. (2012). For unto every one that hath shall be given: Matthew properties for incremental confirmation. *Synthese*, *184*, 89–100. <http://dx.doi.org/10.1007/s11229-009-9695-5>
- Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology*, *61*, 181–194. <http://dx.doi.org/10.1037/0022-3514.61.2.181>
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199744282.001.0001>
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511808098>
- Glass, D. H. (2013). Confirmation measures of association rule interestingness. *Knowledge-Based Systems*, *44*, 65–77. <http://dx.doi.org/10.1016/j.knsys.2013.01.021>
- Hahn, U. (2014). The Bayesian boom: Good thing or bad? *Frontiers in Psychology*, *5*, 765. <http://dx.doi.org/10.3389/fpsyg.2014.00765>
- Hartmann, S., & Meijs, W. (2012). Walter the banker: The conjunction fallacy reconsidered. *Synthese*, *184*, 73–87. <http://dx.doi.org/10.1007/s11229-009-9694-6>
- Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review*, *116*, 856–874. <http://dx.doi.org/10.1037/a0016979>
- Kahneman, D., & Frederick, S. (2002). Representativeness revised: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511808098.004>
- Leitgeb, H., & Pettigrew, R. (2010a). An objective justification of Bayesianism I: Measuring inaccuracy. *Philosophy of Science*, *77*, 201–235. <http://dx.doi.org/10.1086/651317>
- Leitgeb, H., & Pettigrew, R. (2010b). An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science*, *77*, 236–272. <http://dx.doi.org/10.1086/651318>
- Nilsson, H., & Andersson, P. (2010). Making the seemingly impossible appear possible: Effects of conjunction fallacies in evaluation of bets on football games. *Journal of Economic Psychology*, *31*, 172–180. <http://dx.doi.org/10.1016/j.joep.2009.07.003>
- Nilsson, H., Winman, A., Juslin, P., & Hansson, G. (2009). Linda is not a bearded lady: Configural weighting and adding as the cause of extension errors. *Journal of Experimental Psychology: General*, *138*, 517–534. <http://dx.doi.org/10.1037/a0017351>
- Osherson, D. (1995). Probability judgment. In E. E. Smith & D. Osherson (Eds.), *An invitation to cognitive science. Vol. 3: Thinking* (pp. 35–75). Cambridge, MA: MIT Press.
- Pettigrew, R. (2013). Epistemic utility and norms for credences. *Philosophy Compass*, *8*, 897–908. <http://dx.doi.org/10.1111/phc3.12079>
- Predd, J. B., Seiringer, R., Lieb, E. J., Osherson, D., Poor, H. V., & Kulkarni, S. R. (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory*, *55*, 4786–4792. <http://dx.doi.org/10.1109/TIT.2009.2027573>
- Roche, W., & Shogenji, T. (2014). Dwindling confirmation. *Philosophy of Science*, *81*, 114–137. <http://dx.doi.org/10.1086/674346>
- Schubach, J. N. (2012). Is the conjunction fallacy tied to probabilistic confirmation? *Synthese*, *184*, 13–27. <http://dx.doi.org/10.1007/s11229-009-9698-2>
- Shogenji, T. (2012). The degree of epistemic justification and the conjunction fallacy. *Synthese*, *184*, 29–48. <http://dx.doi.org/10.1007/s11229-009-9699-1>
- Tentori, K., Bonini, N., & Osherson, D. N. (2004). The conjunction fallacy: A misunderstanding about conjunction? *Cognitive Science*, *28*, 467–477. [http://dx.doi.org/10.1207/s15516709cog2803\\_8](http://dx.doi.org/10.1207/s15516709cog2803_8)
- Tentori, K., Chater, N., & Crupi, V. (2015). Judging the probability of hypotheses versus the impact of evidence: Which form of inductive inference is more accurate and time-consistent? *Cognitive Science*. Advance online publication. <http://dx.doi.org/10.1111/cogs.12259>
- Tentori, K., & Crupi, V. (2012a). On the conjunction fallacy and the meaning of *and*, yet again: A reply to Hertwig, Benz, and Krauss (2008). *Cognition*, *122*, 123–134. <http://dx.doi.org/10.1016/j.cognition.2011.09.002>
- Tentori, K., & Crupi, V. (2012b). How the conjunction fallacy is tied to probabilistic confirmation: Some remarks on Schubach (2012). *Synthese*, *184*, 3–12. <http://dx.doi.org/10.1007/s11229-009-9701-y>
- Tentori, K., Crupi, V., & Osherson, D. (2007). Determinants of confirmation. *Psychonomic Bulletin & Review*, *14*, 877–883. <http://dx.doi.org/10.3758/BF03194115>

Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: Probability *versus* inductive confirmation. *Journal of Experimental Psychology: General*, *142*, 235–255. <http://dx.doi.org/10.1037/a0028770>

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

Vineberg, S. (2011). Dutch book arguments. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Summer 2011 Ed.). Retrieved from <http://plato.stanford.edu/archives/sum2011/entries/dutch-book>

Yates, J. F., & Carlson, B. W. (1986). Conjunction errors: Evidence for multiple judgment procedures, including “signed summation”. *Organizational Behavior and Human Decision Processes*, *37*, 230–253. [http://dx.doi.org/10.1016/0749-5978\(86\)90053-1](http://dx.doi.org/10.1016/0749-5978(86)90053-1)

## Appendix

### Linda: Feminist Bank Teller Can Be More Confirmed Than Bank Teller

Consider the following probability assignments ( $L$  = Linda’s description,  $b$  = bank teller,  $f$  = feminist activist):

$$P(L) = 1\%$$

$$P(b) = 2\%$$

$$P(f) = 10\%$$

$$P(b|L) = 1\%$$

$$P(f|b) = 7\%$$

$$P(f|L) = 75\%$$

$$P(f|L \wedge b) = 50\%$$

These values are probabilistically coherent. In fact, they determine a single complete probability distribution over the algebra generated by the statements  $L$ ,  $b$ , and  $f$ , namely

$$P(L \wedge b \wedge f) = 1/20,000$$

$$P(L \wedge b \wedge \neg f) = 1/20,000$$

$$P(L \wedge \neg b \wedge f) = 149/20,000$$

$$P(L \wedge \neg b \wedge \neg f) = 49/20,000$$

$$P(\neg L \wedge b \wedge f) = 27/20,000$$

$$P(\neg L \wedge b \wedge \neg f) = 371/20,000$$

$$P(\neg L \wedge \neg b \wedge f) = 1,479/20,000$$

$$P(\neg L \wedge \neg b \wedge \neg f) = 17,923/20,000$$

Note that the confirmation-theoretic defining conditions of the  $M$ - $A$  paradigm are all satisfied in this illustration, because (i)  $L$  disconfirms  $b$ , that is,  $P(b|L) = 1\% < 2\% = P(b)$ , (ii)  $L$  confirms  $f$  even on the background assumption  $b$ , that is,  $P(f|L \wedge b) = 50\% > 7\% = P(f|b)$ , and (iii)  $b$  and  $f$  mildly disconfirm each other, that is,  $P(f|b) = 7\% < 10\% = P(f)$ —similarly, one has  $P(b|f) = 1,4\% < 2\% = P(b)$ . Of course, the conjunction rule of standard probability theory is fulfilled. In particular, one can compute  $P(b \wedge f|L) = 0.5\%$ , so that

$$P(b \wedge f|L) < P(b|L)$$

However, calculations also yield  $P(b \wedge f|L) = 0.5\% > P(b \wedge f) = 0.14\%$ , so that  $\text{conf}(b \wedge f, L) > 0$ , while  $P(b|L) = 1\% < 2\% = P(b)$ , so that  $\text{conf}(b, L) < 0$ . As a consequence, feminist bank teller is indeed more strongly confirmed by Linda’s description than bank teller, that is,

$$\text{conf}(b \wedge f, L) > \text{conf}(b, L)$$

Received December 2, 2014

Revision received April 17, 2015

Accepted April 27, 2015 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!