Noise in Reasoning as a cause of the Conjunction Fallacy

Rita Howe, Fintan Costello

University College Dublin, Dublin, Ireland
Rita.Howe@ucdconnect.ie
Fintan.Costello@ucd.ie

Abstract. The conjunction fallacy occurs when people judge a conjunction A&B as more likely than a constituent A, contrary to the rules of probability theory. We describe a model where this fallacy arises purely as a consequence of noise and random error in the probability estimation process. We describe an experiment testing this proposal by assessing the relationship between fallacy rates and the average difference between conjunction and constituent estimates (in the model, the smaller this difference, the more likely it is that the conjunction fallacy can occur due to random error), and by assessing the degree of inconsistency in people's conjunction fallacy responses for repeated presentations of the same probability questions (in the model these responses should tend to be inconsistent due to random error, especially in cases where the average difference in estimates is low). Experimental results support both these predictions.

Keywords: probability estimation; conjunction fallacy; cognitive science

1 Introduction

Research into human reasoning under uncertainty has uncovered many surprising results about how humans estimate the probabilities of uncertain events. One of the most surprising of these findings is the conjunction fallacy, which arises when subjects judge some conjunction of events A&B to be more likely (more probable) than one of the constituent events of that conjunction, A. This constitutes a violation of the conjunction rule of probability theory, which requires that $P(A\&B) \leq P(A)$ and $P(A\&B) \leq P(B)$ must always hold (simply because A&B cannot occur with A or B themselves occurring). The conjunction, A&B, under the probabilistic laws, cannot be more likely than the single constituent A, thus when a participant chooses the conjunction A&B as more probable, they are committing a fundamental violation of rational probabilistic reasoning.

This finding was first uncovered by Tversky and Kahneman [1]. They found that when presented with the now famous Linda problem, over 80% of their participants made an erroneous judgement. The Linda problem is as follows:

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.

A. Linda is a bank teller

B. Linda is a bank teller and is active in the feminist movement

The conjunction fallacy is ubiquitous in this field of cognitive science. Tversky and Kahneman's widely replicated results were taken as an indication that humans do not reason in a normative fashion – that is, they don't apply probabilistic rules to real-life contexts. Instead, they reasoned that people employ a "representativeness heuristic" when they encounter probability problems such as Linda. Under this theory, the fallacy occurs as the person described in the conjunction is more representative of the information presented in the character sketch.

However, research has called the validity of the representativeness account into question. Experiments that manipulated class inclusion, for instance, demonstrated that the fallacy occurs regardless of whether the conjunction is representative or not [2]. Formal probabilistic models have sought to show that a range of biases can be explained as a function of probabilistic reasoning instead of a heuristic process. The majority of these look towards quantum probability theory and standard probability theory [3, 4].

In this paper we describe an experiment testing two predictions of one of these accounts: the probability theory plus noise model [4]. This model assumes that people estimate probabilities using a fundamentally rational process which is, however, subject to the systematic biasing effects of noise in the reasoning process. This model makes two predictions about the conjunction fallacy. First, this model predicts that fallacy rates for a conjunction A&B should be reliably related to the average difference between estimates A&B and A (the closer these estimates are, the greater the chance that random noise will move one individual estimate for A&B above that for A, producing a conjunction fallacy response). Experimental results strongly support this prediction. Second, this model assumes that conjunction fallacy responses will be inconsistent in certain cases, due to random noise (with a given participant sometimes producing the fallacy but sometimes not producing the fallacy, for the same conjunction A&B). Experimental results confirm that this inconsistency occurs in just the cases predicted by the model.

1.1 Background

Many of the language based experimental manipulations of the fallacy focus on response mode or framing effects. Response modes typically examine the difference in fallacy rates when the question is posed in a probability or frequency format, and much research has shown that the probability format inflates the fallacy rates in comparison to frequency formats, sometimes quite dramatically [5].

Other manipulations focused on the role of the "frame" in the scenario. The frame is the character sketch or additional information supplied to the participant. Framing descriptions are usually designed so there is a strong confirmation link between one of the constituents in the conjunction and the frame. Participants in experiments that have framing descriptions tend to have a high rate of the conjunction errors [1], [6], [7]. Multiple studies have found that the conjunction fallacy occurs reliably even in studies where the participants saw no character sketch or description of an occurrence, though fallacy rates tend to be somewhat lower in these studies than in studies using a framing scenario [7, 8, 9]. While the framing description is the not critical component for eliciting the fallacy, it does provide a mediating factor for the extent to which the fallacy occurs. Consistently, across the research, higher rates of the conjunction fallacy occur when there is a framing description. Not only do participants commit the fallacy at higher rates when they are provided with framing information, they also commit the fallacy at higher rates if they are given analogous probability based information instead of a character sketch. In both those cases, the participants committed the fallacy at about twice the rate as when there was no frame [8]. The inclusion of "filler items" in frames increases both the standard error and the likelihood that the fallacy will be committed [10]. Fallacy rates will decrease when the frame is decreased but to our knowledge, no research has specifically investigated whether the opposite effect will occur.

1.2 Noise

However, while fallacy rates are generally quite high, a frequent observation among this research is that a small number of participants do not seem overly susceptible to the fallacy. In addition, over a number of conjunction problems, participants rarely have 100% error rates. Stanovich and West recognized that individuals can differ greatly on their performances on cognitive bias eliciting tasks. They found that subjects with higher cognitive ability were disproportionally likely to avoid committing a number of cognitive biases including the conjunction fallacy [11]. Hilbert proposed a theoretical framework of noisy information processing to explain a number of cognitive biases [12]. Under this framework, memory based processes convert observations stored in memory into decisions. By assuming that these processes are subject to noisy deviations and that the noisy deviations are a generative mechanism for fallacious decision making, the framework provides an explanation for a number of well-known cognitive biases including the conjunction fallacy. Previously, weighted models were popular as a means to explain the range of results that were consistently observed in fallacy research [13, 14]. However, these were limited in the scope of results that they could predict. A more successful iteration of these weighted models is the configural weighted average (CWA) model [15]. This sophisticated weighted model includes a "noise component" that randomly disturbs probability judgements. Fisher and Wolfe tested the predictions of a number of "noisy" models of reasoning – both normative and non-normative - and found that many of the normative noisy model predictions were good fits for a range of experimental results on cognitive biases [16].

The Probability Theory Plus Noise model. A survey of the conjunction fallacy literature reveals that for standard conjunction problems, the fallacy rate is highly variable (ranging from ~10% to ~90% for different problem scenarios). Costello and Watts [15], [17] proposed an alternative account (the probability theory plus noise model) which can account for this variability in occurrence of the conjunction fallacy. In this model people do reason in a normative fashion according to probability theory, but are subject to random error in the reasoning process. The reasoner's decision making processes, which are memory based, reliably apply the conjunction rule during the probability estimation process, but random noise causes fluctuations in judgement that sometimes lead to the subjective probability of a conjunction exceeding the subjective probability of the constituent. Costello and Watts showed that a simulation implementing this model produced a wide range of fallacy rates (from less than 10% to close to 70%) and produced conjunction fallacy rates for a given set of materials that closely matched those seen in experimental studies for the same materials [17].

The model is built on the assumption of a rational reasoner with a functionally normal long-term memory who is subject to a minimal amount of transient random noise. Long-term memory contains m number of episodes where each recorded episode, i, has a flag that is set to 1 if i contains event A and is set to 0 if it does not. The random noise is represented by d, the small probability that when the flag is read an incorrect value for it is returned. Each event is assumed to have a minimal representation in that it is represented by a flag. Each flag has the same probability of being read wrong. Given this, a randomly sampled item could be counted as an instance of A in two different ways; if it is an instance of A and is counted correctly or if it isn't an instance of A and is counted incorrectly. The probability of an item being an instance of A and being counted correctly, then, is P(A)(1- d), where P(A) is the chance of an item being an instance of A and (1- d) the is chance of any item being counted correctly. The probability of an item being not A and being incorrectly read as A is (1 – P(A))d, where (1-P(A)) is the chance of an item not being an item of A and d is the chance of an item being read incorrectly. Combining these two expressions, the expected value or average for a noisy estimate of P(A) is

$$\langle P_E(A) \rangle = P(A)(1 - d) + (1 - P(A)) = (1 - 2d)P(A) + d \quad (1)$$

With individual estimates $P_E(A)$ being expected to vary independently around this expected value. The model predicts this bias for noise for all types of events, however, it will be more pronounced for complex events. This is due to the process of counting instances of A&B being more complex than counting instances of A as there is more scope for error. To reflect this increased chance of error, the model assumes a random error rate of d + Δd for complex events. The expected value for conjunctive estimates is:

$$\langle P_E(A\&B) \rangle = (1 - 2[d + \Delta d])P(A\&B) + [d + \Delta d] \quad (2)$$

The model proposes that the conjunction fallacy occurs as the individual estimates for A and A&B vary around their expected values as a result of transient noise. This random variation produces a situation where some PE(A) < PE(A&B) and the conjunction fallacy will result. The model predicts that the closer PE(A) and PE(A&B) are to each other, the higher the likelihood of a fallacy occurring – it predicts that the rate of fallacious responses will increase with the difference between the average estimates, which is given by the expression:

$$\langle P_E(A\&B)\rangle - \langle P_E(A)\rangle$$
$$= (1 - 2[d + \Delta d])P(A\&B) + [d + \Delta d] - (1 - 2d)P(A) - d$$
$$= (1 - 2d)[P(A\&B) - P(A)] + \Delta d[1 - 2P(A\&B)] \quad (3)$$

When this difference is negative (and so the expected value for estimates of A, $\langle P_E(A)\rangle$, is greater than the expected value for estimates of A&B, $\langle P_E(A\&B)\rangle$), the fallacy rate is predicted to be less than 50% (because random error will 'move' the estimate for A&B above that for A less than 50% of the time. When this difference is positive, however (and so the expected value for estimates of A, $\langle P_E(A)\rangle$, is less than the expected value for estimates of A&B, $\langle P_E(A\&B)\rangle$), the fallacy rate is predicted to be greater than 50% because random error will 'move' the estimate for A&B below that for A less than 50% of the time. This difference will be positive when

$$\Delta d[1 - 2P(A\&B)] > (1 - 2d)[P(A) - P(A\&B)] \quad (4)$$

From this expression, we see that the average estimate for A&B can be higher than the average estimate for constituent A when Δd is positive; when 1 - 2P(A&B) > 0 (that is, when P(A&B) < 0.5); and when P(A) - P(A&B) is small (that is, when P(A&B) is close to P(A)). When these three requirements hold, the conjunction fallacy should occur more than 50% of the time for individual estimates. In addition to this, the more positive the difference between the terms, P_E(A&B) and P_E(A), the greater the occurrence of the conjunction fallacy. When the difference between these terms is negative, the model predicts a conjunction fallacy rate of less than 50%.

Note that, since in this model the conjunction fallacy is produced by random variation in probability estimates for P_E(A&B) and P_E(A), a fundamental prediction of this model is that the occurrence the conjunction fallacy will be inconsistent. Specifically, if we ask the same participant to estimate the same probabilities P_E(A&B) and P_E(A) at two separate times, this model predicts that they may produce the conjunction fallacy at one but not at the other time (due to random variation in estimates). The more negative the difference $\langle P_E(A\&B)\rangle - \langle P_E(A)\rangle$, the more likely it is that the same participant will not produce the fallacy at either time (random error being unlikely to ‘move’ estimates enough to produce the fallacy). The more positive the difference $\langle P_E(A\&B)\rangle - \langle P_E(A)\rangle$, the more likely it is that the same participant will produce the fallacy at both times. The closer the difference $\langle P_E(A\&B)\rangle - \langle P_E(A)\rangle$ is to 0, the more likely inconsistent fallacy responses will be (random error causing a fallacy at one time, but not at the other).

Because this model assumes a frequency representation for people’s judgments of probability, it gives a natural account for the observed pattern of lower fallacy rates when materials are presented in terms of frequencies rather than probabilities (the frequency representation is closer to the representation used in people’s probability estimates, and so is less subject to random error and less susceptible to the conjunction fallacy). The model also gives a natural account for the higher fallacy rates seen when probabilities are estimated relative to a framing scenario: that scenario introduces more information to the probability estimation process, and so increases the chance of random error (and so the chance of the conjunction fallacy). More generally, in the noise model, probability estimates for longer statements should be more variable than estimates for shorter statements (simply because the long statements contain more information and so more opportunity for random error to occur). If people's probability estimation is subject to the type of noise assumed in the noise model, we should see this difference in variability for probability estimates for statements of different lengths.

1.3 Predictions

We can summarize the probability theory plus noise model's predictions as follows: First, the rate of conjunction fallacy occurrence relative to A (that is, the rate at which people judge P(A&B)>P(A)) should follow the average difference in estimates for P(A&B) and P(A), and should be greater than 50% when that difference is positive. Second, participants should show inconsistency in conjunction fallacy production, with the same participant producing the fallacy at one time T1 but not at another time T2, for some conjunctions A&B. This inconsistency in fallacy responses should be most frequent when the difference P(A&B)-P(A) is small. When the difference is negative (when P(A&B)-P(A)<0) the model predicts consistent avoidance of the fallacy (no fallacy at either times); when the difference is positive (when P(A&B)-P(A)>0) the model predicts consistent production of the fallacy (fallacy occurrence at both times).

The model also makes some predictions about variation in probability estimates: it predicts that the degree of variation in probability estimates for conjunctive statements A&B should tend to be greater than the degree of variation in estimates for single statements A or B. It also predicts that the degree of variance in estimates for long statements should be higher than the degree of variance for short statements. We test these predictions below.

2 Experiment

2.1 Design

The experimental materials consisted of 8 probability estimation problems, presented to participants twice at two time periods, T1 and T2. Each of the problems was a simple scenario describing an individual (in the style of the Linda problem) followed by four statements regarding that person. There were three different blocks of problems. The first block consisted of 4 scenarios with P(A), P(B), P(A&B), and P(AorB) estimates, the second block consisted of 2 scenarios with P(A), P(A&B), P(C&D), and P(AorC) problems. The final block consistent of 2 scenarios with P(B), P(A&B), P(C&D), and P(BorC) estimates. There were four "fillers" which asked the participants to estimate the probabilities of P(A), P(B), P(C), and P(D). The filler scenarios also asked for probability judgements but the statements contained no conjunctions or disjunctions and each of the four were independent of the others.

In addition to this, an extra manipulation was introduced: some of the probability statements were presented in "short" versions and some in "long" versions. The short versions mirrored the classic materials used in past research e.g. participants estimated the probability that "Linda is a boxer and works in an art gallery", while the long versions keep the semantic content as close to the short versions as possible, while increasing the overall sentence length in the probabilities to be estimated e.g. "Linda is a very successful boxer in the amateur welterweight division and a curator responsible for post minimalism modern art in an art gallery". No claim is being made that the long and short versions were semantically identical. Rather, they were kept as close as possible to allow for comparison: the rationale for the manipulation being that the longer statements should increase the variability of participant's responses. Each participant saw exactly the same materials twice, at T1 and T2.

In total, the participants gave probability estimates for 128 items. To control for a memory effect, the order of the scenarios was randomly generated for both T1 and T2 testing. Within the scenarios, the order of the statements was also randomly generated for each trial. There was a 50/50 split for short and long scenarios for each participant.

2.2 Method

40 female and male students were recruited from the student body. They were given a brief description of their task (a two-part estimation task on novel events) and informed that there was no time limit on task completion. The participants were asked to provide probability judgements for statements following a description of a fictional person. Once they had completed part one of the task, they were given part two to complete. At no stage did they have access to their responses to the previous part.

2.3 Results

Of the 40 participants who completed the task, 3 were excluded from the final analysis as they did not complete a sufficient number of questions (missing more than 10% of responses). As in other experiments, many of the participants had high conjunction fallacy rates.

Time. For T1, 97% of the participants had at least one conjunction fallacy in their responses. The mean fallacy rate was 45% for the group, while the short condition had 39% mean fallacy rate, the long condition had a higher rate of 51%. In T2, 92% of the participants committed the conjunction fallacy at least once with a mean fallacy rate of 44%. In the short condition, the mean fallacy rate was 41% while the fallacy occurrences were slightly higher for the long condition, with a mean fallacy rate of 47%.

Difference in Probability Estimates. The conjunction fallacy plus noise model predicts conjunction fallacy rates for A&B versus A to be greater than 50% for cases where the P(A&B) - P(A) values are greater than zero. P(A&B) – P(A) values less than zero should produce fallacy rates that are less than 50% and P(A&B) – P(A) values that are close to zero should produce fallacy rates around 50%. The same predictions should also hold for P(A&B) – P(B).

We tested these predictions by calculating the average values for each of the constituents and complex events across time and compared these values to the observed fallacy rates for the scenarios. Table 1 above displays the results of the analysis for the short scenarios. The prediction was tested the same way for the long scenarios. These results are displayed in Table 2. As predicted by the model, the cases where a positive difference between P(A&B) – P(A) or P(A&B) – P(B) was observed, the average fallacy rate was above 50%. The estimate differences that were zero or close to it, produced fallacy rates around 50%. Cases that had negative differences produced low fallacy rates, ranging from 2% to 20%.

There was a strong positive correlation between difference in estimates and fallacy rates, $r = 0.93$, $p < 0.0001$. The strong positive correlations between estimate difference and fallacy rate held for both long and short versions of the materials. The estimate difference from the short condition had a strong positive correlation with fallacy rates, $r = 0.95$, $p < 0.0001$ as did the difference in estimates for the long condition, $r = 0.95$, $p < 0.0001$.

Table 1. Observed conjunction fallacy rates for difference in conjunction and constituent estimate for short scenarios. The average difference between the constituent and the conjunction was calculated for each of the short scenarios. The results are sorted in descending order. Overall, there was a significant positive correlation between fallacy rates and average difference. Note: Participants did not give P(A) estimates for scenario 7 and P(B) estimates for scenario 8

Scenario	P(A^B) vs P(A)		Scenario	P(A^B) vs P(B)	
	Difference	Fallacy rate		Difference	Fallacy rate
8	-0.121	20.3%	1	-0.423	5.3%
2	0.088	52.8%	3	-0.395	2.6%
4	0.141	69.4%	2	-0.389	2.8%
1	0.166	68.4%	4	-0.279	8.3%
3	0.189	78.9%	7	-0.167	2.7%

Table 2. Observed conjunction fallacy estimates for difference in conjunction fallacy estimates for long scenarios. The average fallacy rate had a significant positive correlation with average difference. Note: Participants did not give P(A) estimates for scenario 5 and P(B) estimates for scenario 6.

Scenario	P(A^B) - P(A)		Scenario	P(A^B) - P(B)	
	Difference	Fallacy rate		Difference	Fallacy rate
6	0.056	44.6%	2	-0.378	2.6%
4	0.102	55.3%	3	-0.374	2.8%
1	0.102	61.1%	4	-0.352	10.5%
2	0.129	73.7%	1	-0.306	2.8%
3	0.159	66.7%	5	-0.058	17.6%

Inconsistent Fallacy Production. One of the marked observations of these results is the variable performance of participants on the fallacy problems. While some participants were consistent in their responses to a particular scenario, either reliably reproducing the fallacy on both occasions or not producing the fallacy at all for that problem, there was a large number of inconsistent responses where the fallacy was produced by a participant on one presentation of a given scenario (at T1 or T2), but not produced by that participant on the other presentation of exactly the same scenario. These inconsistent fallacy responses made up 21% of the total responses to conjunctions. To test the noise model's predictions about this inconsistency, the rate of occurrence of these inconsistent fallacy responses was compared to the average difference in conjunction and constituent estimates for that scenario. Recall that the probability theory plus noise model predicts that these inconsistent fallacy responses should be most frequent when this difference in estimates is close to 0. Figure 1 graphs the frequency of occurrence of this inconsistency in the conjunction fallacy. Of the inconsistent fallacy responses, nearly half (45%) had average differences in estimates that fell between -0.05 and +0.05, as predicted by the model. As the graph shows, these inconsistent fallacy responses were distributed approximately symmetrically around 0; cases with no fallacy responses overwhelmingly had negative average differences (fell on the left side of the graph) and cases with where two fallacy responses overwhelmingly had positive average differences (fell on the right side of the graph). Again, this supports the predictions of the noise model.

Variability. One of the predictions of the model is that complex events should be more variable than single events. Table 3 below displays the overall standard deviations (SDs) in the participants' estimates for $P_E(A)$, $P_E(B)$, and $P_E(A\&B)$ for each of the scenarios regardless of length type. Comparisons of the variability in the conjunctive events versus the simple events showed higher variability in the complex events in 88% of the cases. A similar pattern was observed in the short and long scenarios where the estimates for constituents A and B tended to be less variable than the conjunction, A& B. Comparisons showed that for both types, the conjunction was more variable in 83% of the occasions. F-tests for equality of variance were then carried

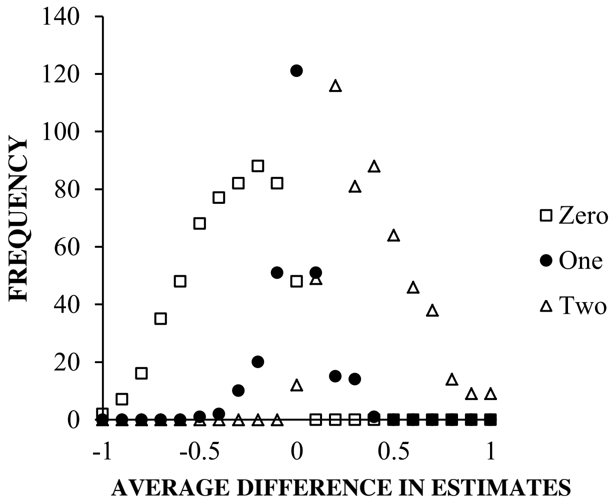

Fig. 1. Participants' consistency in their production of the conjunction in T1 and T2 were calculated. There were 296 possible fallacy responses over T1 and T2, recorded as 'zero' if a participant produced no fallacy at either time, 'one' if the produced the fallacy at one time but not the other (an inconsistent response) and 'two' if they produced the fallacy at both T1 and T2. For each fallacy response the corresponding average difference in estimates, $P_E(A\&B) - P_E(A)$, was calculated; fallacy responses were placed into 'bins' from -0.95 to +0.95 in steps of 0.1 based on this difference. The graph shows the frequency of each conjunction fallacy response in each bin. The model predicts that cases of zero fallacy should be associated with a negative difference in estimates, cases of inconsistent fallacy production should produce differences close to 0 and cases of the fallacy occurring on both occasions should be associated with a positive difference. The graph supports the model's predictions.

out to compare the conjunctive and constituent variability.

Overall, the F-test for equality of variance ($df1 = df2 = 36$) was statistically significant in 33% of the occasions ($p < 0.05$) with another 17% of the sample approaching significance ($p = 0.10$).

For the short scenarios ($df1 = df2 = 36$), the F-test for equality of variance was statistically significant in 40% of the cases ($p < 0.05$) with a further 10% approaching significance. For the long scenarios ($df1 = df2 = 36$), the F-test for equality of variance was statistically significant in 30% of the cases ($p < 0.05$) with a further 20% of comparisons were approaching significance ($p = 0.10$).

A final prediction for the model is that the degree of variance should be higher in the estimates for the long statements than the shorter statements. We carried out a similar analysis to see if there was any difference in variance between the corresponding long and short statements.

For this, the data from scenarios 1 to 4 was used as the participants has either seen a long or short version of each and their probability estimates for the long and short

Table 3. Standard deviations in estimates for long and short scenarios for constituents A, B and for the conjunction A&B. SDs for the constituents were reliably lower than SDs for the conjunctions. F-tests for equality of variance found a number of constituents that had a statistically significant difference in variability relative to their conjunction.

Scenario	Long			Short		
	SD of $P_E(A)$	SD of $P_E(B)$	SD of $P_E(A \wedge B)$	SD of $P_E(A)$	SD of $P_E(B)$	SD of $P_E(A \wedge B)$
1	19.34	22.40	18.41	17.74*	16.49*	27.47
2	19.73*	19.00*	25.06	17.40	18.48	18.59
3	20.09	19.53	20.32	20.02	18.10*	22.60
4	22.75	28.45	23.44	22.12	24.96	23.49
5	-	19.86*	25.29	-	-	-
6	22.54	-	23.50	-	-	-
7	-	-	-	-	20.64	20.84
8	-	-	-	27.20*	-	21.44

* $p < 0.05$

versions of each statement could be compared with each other. Comparisons of the average variance for $P_E(A)$, $P_E(B)$, and $P_E(A\&B)$ for those scenarios showed that the long versions were more variable than the corresponding short version in 75% of the cases. F-tests of equality of variance (df1 = df2 = 36) was statistically significant in 33% of the comparisons between the long and short versions of the scenarios.

3 Discussion

The results of this experiment support the probability theory plus noise model's proposal that the conjunction fallacy is a consequence of random error in probability estimation (with greater random variation in conjunctions than constituent events). First, conjunction fallacy rates were strongly predicted by the average difference between conjunction and constituent probability estimates (as we would expect if the fallacy occurs due to random variation in these estimates: the smaller the difference between estimates, the greater the chance that random variation would 'move' individual estimates to produce a conjunction fallacy response). Second, variability was typically higher for estimates for a conjunctive probability than in estimates for a constituent probability, as assumed in the probability theory plus noise model. Thirdly, estimates for the longer probability statements were typically more variable than estimates for the corresponding short statements, as predicted by the model. With a small sample size it is difficult to produce a clear picture of the variability of the estimates but future research will look towards addressing this issue. Finally, conjunction fallacy responses were inconsistent (with the same participant producing the fallacy for one presentation of a given scenario but not in another presentation), just as we would expect if the fallacy is a consequence of random error. This inconsistency was most frequent when the average difference between estimates was small, just as predicted by the probability theory plus noise model.

References

1. Tversky, A., & Kahneman, D. (1983). Extensional Versus Intuitive Reasoning - the Conjunction Fallacy in Probability Judgment. *Psychological Review*, *90*(4), 293–315.
2. Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology*, *61*(2), 181.
3. Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological Review*, *118*(2), 193.
4. Costello, & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*(3), 463–480.
5. Hertwig, R., & Chase, V. M. (1998). Many reasons or just one: How response mode affects reasoning in the conjunction problem. *Thinking & Reasoning*, *4*(4), 319–352.
6. Hertwig, R., & Gigerenzer, G. (1999). The ‘conjunction fallacy’ revisited: how intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*(4), 275
7. Wedell, D. H., & Moro, R. (2008). Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type. *Cognition*, *107*(1), 105–136.
8. Stolarz-Fantino, S., Fantino, E., & Kulik, J. (1996). The conjunction fallacy: Differential incidence as a function of descriptive frames and educational context. *Contemporary Educational Psychology*, *21*(2), 208–218.
9. Stolarz-Fantino, S., Fantino, E., Zizzo, D. J., & Wen, J. (2003). The conjunction effect: New evidence for robustness. *The American Journal of Psychology*.
10. Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*(4), 269–275.
11. Stanovich, K. E., & West, R. F. (1998). Individual Differences in Framing and Conjunction Effects. *Thinking & Reasoning*, *4*(4), 289–317.
12. Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin*, *138*(2).
13. Thüring, M., & Jungermann, H. (1990). The conjunction fallacy: Causality vs. event probability. *Journal of Behavioral Decision Making*, *3*(1), 61–74.
14. Yates, J. F., & Carlson, B. W. (1986). Conjunction errors: Evidence for multiple judgment procedures, including ‘signed summation’. *Organizational Behavior and Human Decision Processes*, *37*(2), 230–253
15. Nilsson, 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*(4), 517–534.
16. Fisher, C. R., & Wolfe, C. R. (2014). Are People Naïve Probability Theorists? A Further Examination of the Probability Theory + Variation Model. *Journal of Behavioral Decision Making*, *27*(5), 433–443.
17. Costello, F., & Watts, P. (2016). Explaining High Conjunction Fallacy Rates: The Probability Theory Plus Noise Account. *Journal of Behavioral Decision Making*.