Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type

Douglas H. Wedell [*], Rodrigo Moro

Department of Psychology, Barnwell College, University of South Carolina, Columbia, SC 29208, United States

Abstract

Two experiments used within-subject designs to examine how conjunction errors depend on the use of (1) choice versus estimation tasks, (2) probability versus frequency language, and (3) conjunctions of two likely events versus conjunctions of likely and unlikely events. All problems included a three-option format verified to minimize misinterpretation of the base event. In both experiments, conjunction errors were reduced when likely events were conjoined. Conjunction errors were also reduced for estimations compared with choices, with this reduction greater for likely conjuncts, an interaction effect. Shifting conceptual focus from probabilities to frequencies did not affect conjunction error rates. Analyses of numerical estimates for a subset of the problems provided support for the use of three general models by participants for generating estimates. Strikingly, the order in which the two tasks were carried out did not affect the pattern of results, supporting the idea that the mode of responding strongly determines the mode of thinking about conjunctions and hence the occurrence of the conjunction fallacy. These findings were evaluated in terms of implications for rationality of human judgment and reasoning.

[*] Corresponding author. Tel.: +1 803 777 4258; fax: +1 803 777 9558.
E-mail address: Wedell@sc.edu (D.H. Wedell).

1. Introduction

In a classic article, Tversky and Kahneman (1983) presented compelling evidence that people often incorrectly perceive the probability of a conjunction of events as more likely than the probability of one of the constituent events, referred to as the conjunction fallacy. In that original report, they explored a variety of different problems, response methods, incentives, and subject variables and argued for the robust nature of this violation of extensional reasoning. They explained conjunction errors as arising from a tendency to use a similarity based heuristic, called the representativeness heuristic, as a proxy for calculating probabilities. In the ensuing years, numerous researchers have challenged the basic interpretation put forth by Tversky and Kahneman (1983) and, consequently, the generality of this effect. The research we report in this article attempted to systematically study manipulations along three dimensions, conceptual focus, response mode, and type of problem, in order to present a clearer picture of the generality of the conjunction fallacy and thereby promote a clearer understanding of the mechanisms behind it.

One basis for challenging the conjunction fallacy has centered on the idea that people simply misinterpret the alternatives presented to them, which we will refer to as the misinterpretation hypothesis. Dulany and Hilton (1991) delineated this objection by including questions that asked participants to explain how they interpreted the different options. For example, in the classic Linda problem, the option *"Linda is a bank teller"* when placed in the context of an option *"Linda is a bank teller and is active in the feminist movement"* could possibly be interpreted as *"Linda is a bank teller and is not active in the feminist movement."* Indeed, they found evidence that this was the case for many of their participants. Further, they suggested this type of interpretation might follow from Gricean maxims for conversational implicature. However, although the misinterpretation hypothesis has garnered some support (Dulany & Hilton, 1991; Fiedler, 1988), there is also evidence against it. One way to test for this type of misinterpretation is to simply include in the set of alternatives the conjunction of the base event with the complement of the added event, i.e., include the statement *"Linda is a bank teller and is not active in the feminist movement"* as one of the options. Gricean rules no longer would support an interpretation of the base event as reflecting the conjunction of the base event and the complement of the added event in this situation, because that option is overtly presented in the set. Recent research in which this type of third option has been included in conjunction problems has demonstrated robust conjunction errors, thereby casting strong doubt upon the misinterpretation hypothesis (Bonini, Tentori, & Osherson, 2004; Tentori, Bonini, & Osherson, 2004). In the research we report in this article, we include this third type of option (base event with complement of the added event) in our problems

because we feel it provides a simple but elegant guard against this particular type of misinterpretation.[1]

Perhaps the strongest critique of the work of Kahneman and Tversky on the conjunction fallacy has been presented by Gigerenzer and colleagues (Gigerenzer, 1994, 1996; Gigerenzer, Hell, & Blank, 1988; Hertwig & Gigerenzer, 1999). Their objections have largely revolved around a lack of ecological validity surrounding the basic experimental paradigm typically used, along with the concomitant idea that probability principles are being misapplied in most of these problems. The issue of ecological validity argues that the tasks and response formats used to explore the conjunction fallacy are not representative of those people typically encounter in daily life. Central to this idea is that the natural way to process information in the environment is based on coding events in terms of frequencies and not in terms of probabilities. much of the confusion seemingly exhibited by individuals in this task may result from performing a task that they do not typically encounter and being asked to evaluate options in a manner that is quite artificial. Along this line of reasoning, Gigerenzer has proposed that people generally interpret the word probability as it occurs in the Linda problem differently from the technical meaning that researchers impute to it. In addition, Gigerenzer and colleagues argue that the ways in which probability is applied to these situations are questionable. For instance, applying probabilities to the Linda problem only makes sense from a Bayesian point of view in which probabilities can be applied to degrees of subjective belief. Gigerenzer and colleagues further argue that rules of probability must recognize the importance of context, and thus extensionality may not always apply. In support of this perspective, this group of researchers has demonstrated time and again that the conjunction fallacy is minimized when problems are posed in terms of sampling from a population and responses are based on estimating the frequency of each type of event occurring in the drawn sample (Hertwig & Gigerenzer, 1999).

The large reduction in conjunction errors found when participants are asked to estimate frequency of events from a sample of events is an important limiting condition, which was actually recognized in the original report by Tversky and Kahneman (1983, p. 309). One question that arises from this finding, however, is the extent to which this strong reduction is due to the focus on frequencies rather than probabilities or to the use of an estimation procedure rather than a choice or ranking procedure. In their original report, Tversky and Kahneman addressed this issue by considering the results for the same conjunction problem using different tasks: choices based on probabilities, estimations of probabilities, and estimations of frequencies. The problem they studied asked people to assess which was more likely in a survey sample, that a person suffers from one or more heart attacks or that the person suffers from one or more heart attacks

---

[1] Several techniques have been used to avoid this type of misunderstanding. One method is to rephrase the conjunct option, for example, by changing "*Linda is a bank teller*" to "*Linda is a bank teller whether or not she is active in the feminist movement*" (Tversky & Kahneman, 1983). A second method is to change the set of options, for example, by including in that set the disjunction "*Linda is a bank teller or is active in the feminist movement* (Morier & Borgida, 1984). A third method is to avoid the word "and," for example, by choosing "*Lendl will win the finals*" instead of "*Lendl will play the finals and win*" (Politzer & Noveck, 1991). Results obtained by the use of these techniques are conflicting and controversial. For a detailed analysis of these techniques, results, and derived controversies, see Moro (2007), chapter 2.

and is over 55 years old. In the original choice format that asked participants to check the option they thought more probable, 58% of 115 participants committed the conjunction fallacy. When given the same two alternatives and asked to estimate percentages in the sample that matched each of the two descriptions, 65% of 147 respondents committed the conjunction fallacy. When participants estimated the percentages for three descriptions, the two constituent events and their conjunction, the percentage of committing the fallacy was significantly reduced to 31% of 159 respondents, a strong decrease. When the task was changed to estimating the frequencies fitting each description out of a sample of 100, the number of conjunction fallacies was further reduced to only 25% of 117 respondents for the two-description case and 11% of 360 respondents for the three-description case. This pattern suggests effects of both response mode (choice or ranking versus estimation) and conceptual focus (probability or percentages versus actual frequencies). Subsequent research has demonstrated similar systematic effects of ranking versus estimation (Hertwig & Chase, 1998; Hertwig & Gigerenzer, 1999; Morier & Borgida, 1984).[2]

From the work cited above, it appears that estimation typically leads to a reduction of conjunction errors as compared to choice or ranking procedures and that frequency focus tends to lead to a reduction compared to a probability or percentage focus. However, there are clearly exceptions to each of these effects. For example, the data from Tversky and Kahneman (1983) cited above indicate that conjunction errors were just as likely for choices based on probability and estimation of probabilities in the two alternative case of the heart attack problem. A recent study Tentori et al. (2004) provides another example of finding no difference between frequency focus and probability focus. In their study, participants either chose which of three descriptions was more *probable* for a randomly selected person or which of the three described groups in a random sample of 100 was more *numerous*. They found large proportions of conjunction errors that did not depend on whether the conceptual focus was framed in terms of probability of an event or frequency of groups within a sample.[3]

---

[2] The use of percentages by Tversky and Kahneman (1983) as the method for expressing probabilities has advantages and disadvantages when comparing the results to frequency estimation. These derive from the similarity between expressions of percentages and frequencies out of a sample of 100 as used in their examples. A disadvantage of this test is that the probabilities expressed as percentages may capitalize on the similarity to the frequency format and hence effects may be due in part to invoking frequencies. On the other hand, an advantage of this method is that the differences they observed between frequency and percentage estimation tasks demonstrates that despite the similar response format, these may be conceived as somewhat separable constructs. In the experiments we report, we ask participants to express probabilities as fractions or decimals and consequently make the difference between using probabilities and frequencies somewhat more distinct.

[3] From a developmental perspective, it may seem quite surprising that Tentori et al. (2004) found that adults chose the group of Scandinavians with blond hair and blue eyes to be more numerous than the group of Scandinavian with blond hair, as this choice clearly violates the subset rule. As Hertwig and Chase (1998) have indicated, developmental psychologists have shown that 8-year-old children can recognize and make use of subset relations so that they no longer pick the subset as more numerous than a more inclusive set. How is it possible, then, that the adults in Tentori et al.'s experiment do not able apply the subset rule, even when the task format is similar to that used in developmental studies? This is a question at the heart of much of the research on conjunction errors and is not easily answered. It is clear that adults are able to see and use subset relations, but that under some specific contexts such as those described in the literature on conjunction errors, people's responses do not reflect the importance or relevance of the implied subset relations. One purpose of the studies reported here was to systematically explore conditions under which responses do or do not reflect those relations.

Based on these conflicting results, we believe there is a clear need to study both of these variables (response mode and conceptual focus) in the context of each other to see how they combine. We know of only one study to date that has done so. As part of a larger set of studies focused primarily on base rate neglect, Sloman, Over, Slovak, and Stibel (2003) conducted a study (Experiment 5) that explored the effects of conceptual focus and response mode on the conjunction fallacy. They found that estimation significantly reduced conjunction errors relative to ranking procedures but that there was no effect of framing the problem in terms of probabilities versus frequencies on conjunction errors. We believe two methodological aspects of that study make it difficult to draw clear cut conclusions from their data. First, their response format included seven filler items between the relevant elements, thus making the inclusion relationship extremely difficult for participants to detect. We believe that including only directly relevant options provides a more transparent and general test these effects. Second, they used a relatively small number of participants in a between-subjects design based on the Bill and Linda problems. We manipulated these variables within-subjects, with fairly large samples, to increase power and also provide a clearer comparison of conjunction errors for the same individuals across these conditions.

Before describing our experimental work in detail, we first discuss the issue of how conjunction problems may differ. Tversky and Kahneman (1983) described two simple recipes for creating the conjunction fallacy based on relationships between the causal model, as entailed in the problem description, the base event, which is assumed to be unlikely given the model, and the added event. In one recipe, the added event is positively associated with the model, as in the classic Linda problem. If similarity is guided by summing matching and nonmatching features as described in Tversky's (1977) contrast model of similarity, then adding an event with a high association with the model will tend to increase similarity to the model and hence increase probability estimates based on similarity. The often used Linda problem follows this recipe, with the base event, *"Linda is a bank teller,"* unlikely given the description of her, and the added event, *"Linda is active in the feminist movement,"* likely given the description. In the other recipe, the added event is positively associated with the base event. This method presumably increases the plausibility of the base event occurring, which is then reflected in the overall probability assessment of the conjunction. An example of this type of problem is when people judge *"Mr. F. has one or more heart attacks"* less probable than *"Mr. F. has one or more heart attacks and is over 55 years old."* Both types of problems created strong conjunction errors.

In addition to these types of problems, Wells (1985) investigated problems in which both base and added events are likely or unlikely. He found that conjunction errors were almost entirely eliminated when both events were unlikely, but that there were a substantial though greatly reduced number of conjunction errors when both events were likely. More recently, Tentori et al. (2004) demonstrated strong and robust conjunction errors when both base and added events were likely. In these problems, the two likely events were sometimes strongly associated (e.g., *"having blonde hair and blue eyes"*) but sometimes they were not strongly associated (e.g.,

"*being less than 21 years old and taller than 5′10″*"). Thus, a third way to create substantial conjunction errors would appear to be to combine two likely events (that may or may not also imply one another).

It is not clear how the different response mode and conceptual focus manipulations apply to these different types of conjunction problems. To explore this issue, we included two basic types of problems in our studies. The first used the typical formula of combining a low likelihood event with a high likelihood event, as reflected in the Linda problem. The second used a formula for combining two high likelihood events that may or may not be highly related to one another, as exemplified in the recent studies of Tentori et al. (2004).

2. Experiment 1

In Experiment 1 we created booklets that represented the $2 \times 2 \times 2$ combination of response mode (choice and estimation), conceptual focus (probability and frequency), and problem type (low and high likelihood events and two high likelihood events). These eight combinations were presented to each participant in a within-subject design. Unique topics were created for each of the eight cases judged, with the different topics counterbalanced across response mode and conceptual focus conditions (but nested within problem type). This study then provides a systematic examination of the effects of these different ways of responding to and conceptualizing conjunction problems.

After the eight experimental trials, participants had to solve one of four versions of the Linda problem generated from the 2×2 combination of type of description (full or reduced) and number of options (two or three). The reduced description stated that Linda was 31 years old, single, outspoken, and very bright. The full description added that she was a philosophy major in college and that as student, she was concerned about social discrimination issues and had participated in anti-nuclear demonstrations. The two-option condition had the standard minimal structure, i.e., B, B&F. The three-option condition had the same structure as the rest of our problems: B, B&F, B&~F. After responding to the Linda problem participants were asked to indicate how they interpreted the single event statement, either as "*Linda is a bank teller and she is not active in the feminist movement*" or as "*Linda is a bank teller whether or not she is active in the feminist movement.*" Thus, this last problem and question allows us to (1) examine the extent to which participants are misinterpreting the statements in the problem and (2) see whether this depends on the options presented (cf., Dulany & Hilton, 1991).

2.1. Method

2.1.1. Participants and design

Participants were 96 undergraduate psychology students (24 male and 72 female) at the University of South Carolina who volunteered in exchange for course credit. Because we were looking to maintain strict counterbalancing, whenever we found a

participant who did not fill out a booklet correctly according to instructions (e.g., estimating probabilities when asked to estimate frequencies, etc.), we replaced that individual in that condition until we met our criterion of 96 participants.

The design was a $2 \times 2 \times 2$ within-subject design. The three factors were response mode (choice and estimation), conceptual focus (probability and frequency), and type of conjunct (likely and unlikely).[4] A given problem topic appeared only once within each booklet. Across booklets, problem topics were nested within problem type but crossed with response mode and conceptual focus so that each problem occurred equally often in these four conditions. Problems were blocked by response mode so that half the participants made choices for the first four problems followed by estimates for the next four problems and the other half completed the tasks in reverse order. Each problem type by conceptual focus condition occurred equally often in the four positions within each block. Altogether there were 32 different booklets constructed based on the combination of matching topic to condition and varying order. Three participants were randomly assigned to each of the 32 booklets.

In all booklets, the final problem consisted of one of two versions of the Linda problem (generic or descriptive) with either two alternatives or three alternatives for response options. The generic description simply stated that Linda was 31 years old, single, outspoken, and very bright, and the descriptive statement was the description typically used. This was followed by the question an how the base event was interpreted.

2.1.2. Booklets

The study involved eight different problems (see the Appendix). Each problem, in turn, had four different versions that resulted from the factorial combination of response mode and conceptual focus. For example, one of the problems was based on hair color and eye color of Scandinavians, as derived from the Tentori et al. (2004) study. For this problem, the initial paragraph read:

> *The Scandinavian Peninsula is the European area with the greatest percentage of people with blond hair and blue eyes. This is the case even though every possible combination of hair color and eye color occurs in those countries.*

The four resulting conditions differed in how participants were asked to think about the problem and how they were asked to make a response. These are delineated below for this problem.

2.1.2.1. Probability focus, choice mode. Suppose we choose at random an individual from the Scandinavian population. Which event do you think is most probable? (check your choice)

---

[4] The reason we distinguish between types of problems by the type of conjunct rather than for the type of conjunction is very simple. Given that we added the third option (base event and complement of the added event) to avoid misunderstandings, we now have two conjunctions rather than one. So, it would be confusing to classify problems by the type of conjunction. However, notice that for every topic, only one conjunction will be very tempting. It will be either the one that contains a very likely added event or the one that contains an added event strongly associated with the base event.

– The individual has blond hair.
– The individual has blond hair and blue eyes.
– The individual has blond hair and does not have blue eyes.

2.1.2.2. Frequency focus, choice mode. Suppose we choose at random 100 individuals from the Scandinavian population. Which group do you think is most numerous? (check your choice)

– Individuals who have blond hair.
– Individuals who have blond hair and blue eyes.
– Individuals who have blond hair and do not have blue eyes.

2.1.2.3. Probability focus, estimation mode. Suppose we choose at random an individual from the Scandinavian population. Your task will be to estimate the probability of the events listed below. Express your probability estimate in terms of a number in the range 0–1, where 0 means minimal probability and 1 maximal probability. You are free to use the whole range (including 0 and 1); both decimal estimates (e.g., .10) and fractional estimates (e.g., 1/10) are acceptable. Estimate the probabilities for the following events:

– The individual has blond hair.
– The individual has blond hair and blue eyes.
– The individual has blond hair and does not have blue eyes.

2.1.2.4. Frequency focus, estimation mode. Suppose we choose at random 100 individuals from the Scandinavian population. Estimate the number of people in each of the following groups (use numbers from 0 to 100)

– Individuals who have blond hair.
– Individuals who have blond hair and blue eyes.
– Individuals who have blond hair and do not have blue eyes.

The corresponding unlikely-conjunct conditions are exactly the same except that now the options are based on an unlikely attribute paired with a likely attribute.[5] For example, the options for probability focus would be as follows:

– The individual has green eyes.
– The individual has green eyes and blond hair.
– The individual has green eyes and does not have blond hair.

---

[5] Note the example of green eyes and blonde hair was only used in Experiment 2. However, we describe it in the method section of Experiment 1 because it follows from the example we have been using. Complete descriptions of problems and the experiments in which they appeared are provided in the Appendix.

We also note that the problems we used could be classified into three different types. The first is what we refer to as sampling of discrete categories, of which the Scandinavian problem is an example. We also have problems that sample continuous categories and indicate a specific cutoff. For example, in a problem on height and weight, the options concerned whether the sampled individual was under a given weight cutoff and over a given height cutoff. Finally, a special case of discrete sampling is captured in two problems that represent games of chance. In these problems, one may consider the likelihood of rolling different values on dice or drawing colored balls from different urns. We raise this distinction because in Experiment 1, problems were nested within problem type (likely conjunct or unlikely conjunct). Hence any effect of problem type may be confounded with other aspects of the problem other than the type of conjunction.

2.1.3. Procedure

Participants in groups of 5–50 reported to one of several large rooms and were given an informed consent and brief instructions that they were to complete the booklet, working their way through each page before continuing to the next. Participants took between 10 and 20 min to fill out their booklets before receiving a debriefing.

2.2. Results

2.2.1. Analysis of the factorial design

We tallied conjunction errors as occurring when either of the conjunctions was chosen or estimated as higher in probability or frequency than the constituent event, with ties in estimation reflecting no violation of extensional reasoning. Table 1 presents the percentage of conjunction errors for each problem as a function of experimental condition. The overall level of conjunction errors was 53.9%, which is quite substantial but also clearly lower than the 80–90% often found in the two alternative choice problems.[6] Conjuncts of unlikely and likely events produced more conjunction errors (65.9%) than conjuncts of two likely events (41.9%). Response mode also had a strong effect on the occurrence of conjunction errors, with conjunction errors higher in the choice mode (66.9%) than in the estimation mode (40.9%). However, conceptual focus appeared to have minimal effects on conjunction errors, with

---

[6] One way conjunction errors might occur is if participants simply did not care about the task and responded randomly. In our case, random responding would produce a conjunction error rate of 66.7%, as we had two conjunctions in each response set. Random responding implies an average frequency of 3.33 cases of estimating higher or choosing the conjunct of the base and complement events. We tallied this score for each participant, with a maximum possible score of 8 and minimum of 0 to consider the likelihood of random responding. The mean number endorsements of the base–complement conjunction was 0.677, which was significantly lower than the chance rate, $t(95) = -23.9$, $p < .001$. Indeed, only 3 of 98 participants endorsed this option more than twice. A parallel analysis conducted on the data from Experiment 2 produced similar results, with the mean number of endorsements of the base–complement conjunction begin 1.01, which was significantly lower than the chance rate $t(127) = -26.1$.

Table 1
Percentage of conjunction errors for each problem as a function of response mode, conceptual focus, and type of conjunction (Experiment 1)

Problem type/name	Probability choice	Frequency choice	Probability estimation	Frequency estimation	Mean
Unlikely conjunct					
Dice	66.67	70.83	66.67	75.00	69.79
Urn	70.83	75.00	75.00	66.67	71.88
Football	87.50	75.00	45.83	37.50	61.46
Movies	75.00	70.83	50.00	45.83	60.42
Mean	75.00	72.92	59.38	56.25	65.89
Likely conjunct					
Height	66.67	58.33	33.33	29.17	46.88
NBA	62.50	70.83	41.67	41.67	54.17
Scandinavia	70.83	75.00	12.50	16.67	43.75
India	45.83	29.17	4.17	12.50	22.92
Mean	61.46	58.33	22.92	25.00	41.93

Note. $N = 96$, with $N = 24$ for each problem in a specific condition.

conjunction errors just slightly higher for probability than for frequency focus (54.7% versus 53.1%). The pattern of data also appears to reflect an interaction of response mode and type of problem. The effect of response mode on conjunction errors was substantially larger for likely conjuncts (a 35.9% reduction in the estimation mode) than for unlikely conjuncts (a 16.2% reduction in the estimation mode). Finally, the effect of order of tasks (choice then estimation or vice versa) had surprisingly little effect, with conjunction errors in choice only slightly lower after estimation (55.7% versus 52.1%) and conjunction errors in estimation slightly higher after choice (39.6% versus 42.2%).

To analyze these effects, we conducted a repeated measures categorical data analysis implemented in the CATMOD procedure of the SAS software (Version 9.1). The three repeated measures factors included in the analysis were response mode, conceptual focus, and problem type. In addition, one between-subjects factor reflecting the most important order manipulation (choice first versus estimation first) was included in the analysis. The initial saturated model produced no significant effect of order and no effect of focus or interaction with focus. A reduced model that eliminated the focus variable again showed no main effects or interactions with order, and no significant residual variance. The final model then only included an intercept, the main effects of response mode and problem type and their interaction. All effects were statistically significant: $\chi^2(1) = 55.0$, $p < .001$ for response mode, $\chi^2(1) = 68.3$, $p < .001$ for problem type, and $\chi^2(1) = 19.8$, $p < .001$ for the interaction. The residual variance was nonsignificant, $\chi^2(4) = 0.8$, $p > .05$. Thus the data can be well characterized by a reduction in conjunction errors for estimation and for likely conjuncts, with greater reductions with estimation occurring for likely conjuncts than for unlikely conjuncts.

2.2.2. Analysis of the estimation data

The estimation data provide the basis for additional ways to understand how participants generate their estimates. One way to evaluate these data is to make use of the structure of the three options. Because the added conjunction option in our study was the conjunction of the base with the complement event, then by the rules of extensionality the probabilities for the two conjunctions should add up to the probability of the base event. For each problem we tallied the number of participants whose estimates satisfied this constraint, which would naturally prevent the conjunction error from occurring. These percentages were as follows: 14.6% for probability estimates of unlikely conjuncts, 25.0% for probability estimates of likely conjuncts, 28.2% for frequency estimates of unlikely conjuncts, and 28.1% for frequency estimates of likely conjuncts. A repeated measures categorical analysis was conducted on these frequencies following the 2×2 design of conceptual focus crossed with problem type. In the saturated model, only the intercept and the main effect of focus were significant. The reduced model confirmed this result, with the effect of focus statistically significant, $\chi^2(1) = 4.63$, $p < .05$, and no significant residual, $\chi^2(2) = 3.71$, $p > .05$. Thus, there was a small but significant tendency to apply the correct distributional model to these problems in the frequency focus compared to the probability focus conditions.

Because the dice and urn problems provided precise probabilities of events, they allowed us to consider more precisely how different participants' response patterns fit different models of the estimation process. Each participant estimated either frequencies or probabilities for just one of these two problems, resulting in 24 participants in each of the four estimation conditions (probability or frequency estimation for dice or urn problem). We classified participants' data into one of three basic models that we postulated either a priori or after viewing the data. One a priori model was the distributional model, which refers to the normative way to calculate the conjunction frequencies or probabilities. Predicted probabilities (frequencies) from this model were 0.167 (17), 0.139 (14), and 0.028 (3) for the dice problem and 0.10 (10), 0.08 (8), and 0.02 (2) for the urn problem. A second a priori model was the averaging model, which refers to evaluation of the conjunction of probabilities or frequencies by averaging the values for each event. This model may be loosely related to aspects Tversky's (1977) contrast model of similarity, or it can be conceived as a heuristic or simple rule for making estimates. The averaging model has also been shown to fit participants' data well in previous studies (Fantino, Kulik, & Stolarz-Fantino, 1997). Predicted probabilities (frequencies) from this model were 0.167 (17), 0.500 (50), and 0.167 (17) for the dice problem and 0.10 (10), 0.45 (45), and 0.15 (15) for the urn problem. A third model we inferred from viewing the data was the single event model, which refers to evaluating of the probability or frequency of the conjunction by using the value of the higher probability event. One way to think about this model is that it represents a lower bound for the disjunction of the two events. Predicted probabilities (frequencies) from this model were 0.167 (17), 0.833 (83), and 0.167 (17) for the dice problem and 0.10 (10), 0.80 (80), and 0.20 (20) for the urn problem.

To classify a participant's responses into one of these three models, we averaged the absolute deviation of the participant's estimates from those prescribed by each

model. A participant's responses were classified as fitting a given model if that model had the least average absolute deviation from the responses under the following two restrictions. First, we set a minimum average absolute deviation value of 0.15 in the probability condition and 15.0 in the frequency condition for classification into a given model.[7] This resulted in a total of 9 of 96 participants remaining unclassified. Second, the spirit of the distributional model requires that no conjunction error is made, yet this model might have best characterized some participants whose estimates reflected a conjunction error. We thus created a new category, the quasi-distributional model, for participants whose responses failed the extensionality requirement but were still closest to the distributional model. Participants following the quasi-distributional model appear to have some notion that the conjoint probabilities would be low, but fail to make them lower than the base event.

The classification of each participant's pattern of data into different models is shown in Table 2. The mean estimates for the participants classified into each model are shown as well. Note that mean values are close to the values prescribed by the corresponding model, reflecting some face validity for the classification scheme. As shown, a substantial percentage of the participants (27.6%) adhered to the distribution model, with an additional 14.9% of participants classified into the quasi-distributional model, reflecting a tendency to assign low probabilities or frequencies to the conjunction of events, but assigning the more likely conjunct a higher value than the base event. The largest percentage of participants (40.2%) appeared to engage in a simple averaging of values. This type of integration may be more akin to a similarity based assessment, in that conjoining a high similarity event with a low similarity event results in an intermediate value of similarity. Of course, this procedure blatantly ignores the extensionality property of conjunctions. Finally, 17.2% of participants appeared to simply use the highest single event value as a convenient estimate, largely ignoring the conjunction. In some ways, this tendency may reflect an approximation to the disjunction of these events rather than the conjunction.

It is also instructive to see how participants with different inferred models based on the urn or dice estimation task performed on the estimation task for the likely conjuncts and choice tasks for the unlikely and likely conjuncts. The 24 participants classified as using the distributional model produced 12.5%, 47.9%, and 45.8% conjunction errors in these three conditions, respectively. Thus, it appears that this group's normative performance transferred well to the other estimation conditions but not to the choice conditions in which they produced substantial conjunction errors. The 35 participants classified as using the averaging model produced

---

[7] Although our criterion requiring a model to have a minimum average absolute deviation of 0.150 for probabilities (and 15.0 for frequencies) to be considered as characterizing a person's data is somewhat arbitrary, we felt it important to have some way of designating participants as unclassified. When we evaluated the three models using data generated from each model, average absolute deviations of the incorrect models varied from 0.11 to 0.30, with an average of 0.179 for incorrect models. From this analysis we determined that 0.150 would be a reasonable lower bound to attribute a model to a particular pattern of data. Note that upper maximum values of average absolute deviations for these three models varied from 0.722 to 0.933, with an average of 0.830. Thus, our criterion for model fit required a deviation roughly 1/5th of the maximum deviation.

Table 2
Mean estimates from participants classified into each model for dice and urn problems (Experiment 1)

Problem/model	Probability estimation				Frequency estimation			
	N	E_L	$E_L \& E_H$	$E_L \& {\sim} E_H$	N	E_L	$E_L \& E_H$	$E_L \& {\sim} E_H$
Dice								
Distributional	6	0.161	0.152	0.079	5	14	11.8	4
Quasi-distributional	5	0.113	0.238	0.082	4	10	18.8	4.8
Averaging	7	0.160	0.542	0.127	11	13.3	51.5	22.0
Single event	3	0.133	0.767	0.200	3	18.3	76.7	21.7
Urn								
Distributional	5	0.100	0.066	0.016	8	10.6	8.0	5.3
Quasi-distributional	0				4	6.8	18.8	6.0
Averaging	11	0.087	0.473	0.159	6	12.0	50.0	21.0
Single event	5	0.090	0.780	0.266	4	8.8	75.0	18.8

Note. E_L, low probability event; E_H, high probability event; &, conjunction; $\sim$, negation; in the dice problem, $E_L = 1/6 = 0.167$, $E_H = 5/6 = 0.833$, $E_L \& E_H = 5/36 = 0.139$, and $E_L \& {\sim} E_H = 1/36 = 0.028$; in the urn problem, $E_L = 1/10 = 0.100$, $E_H = 8/10 = 0.800$, $E_L \& E_H = 8/100 = 0.080$, and $E_L \& {\sim} E_H = 2/100 = 0.020$; Number of unclassified participants were 3, 1, 3, and 2 in dice-probability, dice-frequency, urn-probability, and urn-frequency conditions, respectively.

22.9%, 80.0%, and 64.3% conjunction errors in these three conditions, respectively, This group managed to commit conjunction errors in estimation of likely conjuncts only about 1/5th of the time, but made abundant conjunction errors in choice (though more so for unlikely conjuncts). The 15 participants classified as using the single event model produced 30.0%, 76.7%, and 70.0% conjunction errors in these three conditions, respectively. This is a pattern similar to the averaging group, but demonstrating higher levels of conjunction errors in estimation.

2.2.3. Analysis of the Linda data

At the end of each booklet, the Linda problem was presented in either full or reduced description form, followed by a question on the interpretation of the alternatives. For the full form, the percentages of conjunction errors were 62.5% with two response options and 66.7% with three response options. Thus, the number of response options did not appear to affect these choices much. The percentages of conjunction errors decreased in the reduced form to 50.0% with two options and 33.3% with three options. The fact that these errors were still substantial suggests that there is a tendency to endorse a richer description as more probable, even if the model description is sparse.

The final question in the booklet was designed to determine whether the three-choice response format that included the conjunction of the base event with the complement of the added event would reduce misinterpretation of the base event as representing a conjunct with the complement of the added event. Misinterpretations were significantly greater in the two option case (27.1%) than in the three-option case (6.2%), $\chi^2(1) = 7.5$, $p < .01$. Hence, the inclusion of the conjunct of base and complement had the intended effect of making it difficult to interpret the single

event as excluding the added event. Note also that even though the number of mis-interpretations was much greater in the two option case, the number of conjunction errors was nearly the same, indicating that this linguistic ambiguity may not have contributed to creating conjunction errors.

2.3. Discussion

The systematic manipulation of response mode, conceptual focus, and problem type used in Experiment 1 provides a clearer picture of how these variables operate and interrelate with one another in producing conjunction errors. The pattern of data we observed largely converged with the patterns reported in other studies that have typically examined a subset of these conditions. Similar to the results reported by Tentori et al. (2004) and by Tversky and Kahneman (1983), Experiment 1 showed that there was no significant effect of conceptual focus on conjunction errors when choosing among alternatives, 68.2% errors for probability-based choices and 65.6% errors for frequency-based choices. Sloman et al. (2003) also showed null result for ranking responses based on probability or frequency. Although Tversky and Kahneman (1983) reported a difference related to conceptual focus in estimation, with probability focus leading to somewhat more conjunction errors than frequency focus, this was not observed in the current experiment (41.2% versus 40.6%, respectively) or in previous work by Sloman et al. (2003).

Also similar to the findings of several researchers (Hertwig & Chase, 1998; Sloman et al., 2003; Tversky & Kahneman, 1983), the estimation procedure produced fewer conjunction errors than the corresponding choice (or ranking) conditions. In Experiment 1, estimation resulted in a 26% reduction of conjunction errors compared to choice. It should be noted that many of the studies comparing frequency to probability methods have tended to conflate this variable with response mode, so that probability rankings are compared with frequency estimates (Fiedler, 1988; Hertwig & Gigerenzer, 1999). This comparison examines the extremes of the continuum, but it does not clarify whether it is estimation or frequency focus that results in this difference. The current results clarify that estimation is a potent way to reduce conjunction errors regardless of whether the focus is on frequencies or probabilities. In the current study, frequency formats did not aid at all in reducing conjunction errors compared with probability formats.

Experiment 1 also shed light on the difference between likely and unlikely conjuncts. The only study that we know which systematically compared unlikely and likely conjuncts was that by Wells (1985). His study used probability estimates and showed that conjunction errors were much more prevalent for unlikely conjuncts (71.9%) than for likely conjuncts (24.0%). The corresponding values in our study were 59.4% and 22.9%, showing a similar advantage of likely conjuncts in reducing conjunction errors. The current study showed that this effect generalizes to frequency estimates and that the effect is somewhat smaller for choice under either a probability or a frequency focus. The greater response mode based reduction of conjunction errors for likely conjuncts may reflect several tendencies. First, it may reflect how ties are broken. In estimation, a tie did not count as a conjunction error,

but in choice a participant cannot indicate ties and so may have resolved them arbitrarily, leading to more conjunction errors. We would argue that ties should be more likely to arise in the likely conjunct case even when participants are using inappropriate models such as averaging or single event models, as discussed below.

The use of precise probability values for two of the problems in Experiment 1 (dice and urn) provided a clear basis for evaluating how participants made estimates. Participants were classified into one for three basic models, based on which model most closely fit their pattern of estimates. The three models were the distributional model, representing the normative approach, the averaging model, a commonly used rule in judgment, and the single event model, which might represent a type of disjunction rule or just a simplified estimation rule in which the maximum constituent event probability was used as an estimate of the joint events. We designated a subset of those classified into the distributional model as following a quasi-distributional model, as their estimates reflected conjunction errors. We believe this categorization scheme is telling in several ways.

First, the scheme shows that participants are truly heterogeneous in their estimation schemes. Those classified into the distributional model provide an entirely different profile of estimates than those classified into either the averaging or single event models. With the quasi-distributional model added, these patterns accounted for all but 6% of the participants' patterns. The mean estimates for each group were close to the a priori values, suggesting these provide a reasonable approximation of the estimation behavior of these individuals.

Second, the performance of these participant groups in the other conditions is telling. As one might expect, those exhibiting a distributional model in the unlikely-conjunct estimation condition appeared to generalize this behavior to the likely conjunct cases and so produced very few conjunction errors. Most telling, however, these self-same participants who rarely made a conjunction error in estimation because the apparently knew and could implement the normative procedure, showed quite substantial conjunction errors in choice. Thus, the use of the within-subjects design clearly demonstrates a different mode of thinking in choice than in estimation. The lack of a significant effect of task order reinforces the general lack of transfer from procedures in the estimation task that produce fewer conjunction errors to the mode of thinking invoked in choice that produces substantial errors. The relative performances of the averaging and single event groups in these other conditions also may have reflected differences in thinking during choice versus estimation, with more conjunction errors in the former. However, because both of these groups had a poorer handle on the correct estimation procedure, they produced fairly large numbers of conjunction errors in estimation.

Third, the modeling suggested a reason why likely conjuncts produce fewer errors than unlikely conjuncts. If we instantiate the similarity based representativeness heuristic in an averaging model (i.e., the similarity between the situation and its model is represented by the average similarity of components), then it follows that the conjunction of two likely events will be close to the similarity of the component events. For example, if in our dice example we compared the single roll that had a 5/6th probability to the conjunction of two rolls, each with a 5/6th probability, the average

of the two events equals the probability of the single event and so this rule would not lead participants to commit the conjunction fallacy (depending on how choices resolve ties). While it is difficult to study this process for choice, it is easier to make inferences concerning processes guiding estimation. We pursue this in Experiment 2.

3. Experiment 2

The large effect of problem type observed in Experiment 1 is interesting, but it is somewhat difficult to interpret because specific problem descriptions were nested within problem type. Thus, it is possible that the key difference between the two types of problems was not in the nature of the conjunct but simply in the differences between the topics used. While we feel this is unlikely given the relatively consistent results across the different problems nested within each type, we undertook to investigate this issue in Experiment 2 by formulating each problem topic as both a likely and unlikely conjunct, thereby more effectively isolating the effects of this manipulation. Thus, we sought to replicate the basic findings of Experiment 1 and clarify the nature of the effect of problem type.

Note that by instantiating dice and urn problems as both unlikely and likely conjuncts, we should be able to better understand how the estimation process differs for these two conjuncts. Our basic hypothesis is that the relative use of the different models will be similar for the two types of problems. Consequently, we predicted a reduction in conjunction errors for participants who use the averaging or the single event model. The reason is that strict application of these strategies will produce conjunction errors under the *unlikely* conjunct condition but not under the *likely* conjunct conditions. For example, under the *unlikely* conjunct condition of the urn problem, the averaging strategy will generate a probability value of 0.10 for the base event and 0.45 for the key conjunction, committing thus the conjunction fallacy. However, under the *likely* conjunct condition of the same problem, averaging will produce a value of 0.90 for the base event and 0.85 for the key conjunction, avoiding thus the conjunction fallacy.

3.1. Method

Participants were a new group of 128 undergraduates sampled from the same population as in Experiment 1. Like Experiment 1, each participant whose responses overtly did not conform to the specified instructions for a given problem was replaced by another randomly sampled participant who was given the same booklet, with four participants randomly assigned to each of the 32 booklet types.

The procedure was identical to Experiment 1. The only differences between Experiments 1 and 2 were in the construction of the booklets. Each booklet consisted of eight different problems, one each of the eight conditions produced by the factorial combination of response mode, conceptual focus and type of problem. Across booklets, each of eight problem topics appeared equally often in each of the eight conditions. Because the India problem produced few conjunction errors in Experiment 1,

it was replaced with the Hockey problem in Experiment 2. The likely and unlikely-conjunct conditions for each problem are displayed in the Appendix. Altogether there were 32 different booklets resulting from counterbalancing of order and counterbalancing of topics across conditions. Again, due to the importance of counterbalancing, each booklet type was used equally often (four times each).

3.2. Results

The data were processed in ways parallel to the procedures described in Experiment 1. Table 3 presents the percentages of conjunction errors observed across conditions, segregated by the specific problem. By and large, the results are quite comparable to those observed in Experiment 1. The total percentage of conjunction errors was 57.0%, with this percentage once again much higher in choice (73.0%) than in estimation (41.0%). There was also a similar tendency to commit more conjunction errors in the unlikely-conjunct condition (67.0%) than in the likely conjunct condition (47.1%). Similar to Experiment 1, this tendency was greater in estimation (a difference of 26.6%) than in choice (a difference of 13.3%). Like Experiment 1, conceptual focus had little effect on conjunction errors, with conjunction errors reduced by only 5.1% in the frequency focus conditions. Finally, as in

Table 3
Percentage of conjunction errors for each problem as a function of conceptual focus, response mode, and type of conjunction (Experiment 2)

Problem type/name	Probability choice	Frequency choice	Probability estimation	Frequency estimation	Mean
Unlikely conjunct					
Dice	93.75	93.75	68.75	56.25	78.13
Urn	87.50	87.50	87.50	75.00	84.38
Football	81.25	75.00	50.00	37.5	60.94
Movies	75.00	68.75	62.50	56.25	65.63
Height	87.50	62.50	56.25	62.50	67.19
NBA	68.75	68.75	31.25	18.75	46.88
Scandinavia	75.00	87.50	62.50	37.50	65.63
Hockey	93.75	68.75	62.50	43.75	67.19
Mean	82.81	76.56	60.16	48.44	66.99
Likely conjunct					
Dice	93.75	68.75	31.25	12.50	51.56
Urn	62.50	68.75	25.00	31.25	46.88
Football	62.50	75.00	37.50	12.50	46.88
Movies	75.00	75.00	56.25	50.00	64.06
Height	75.00	75.00	18.75	50.00	54.69
NBA	62.50	81.25	43.75	31.25	54.69
Scandinavia	43.75	62.50	25.00	18.75	37.50
Hockey	50.00	31.25	0.00	0.00	20.31
Mean	65.63	67.19	29.69	25.78	47.07

Note. $N = 128$, with $N = 16$ for each problem in a specific condition.

Experiment 1, reversing the order of choice and estimation had little overall effect, with conjunction errors at 57.2% for choice followed by estimation and 56.8% for the reverse order. The lack of order effects reinforces the lack of carryover or transfer across response modes.

To analyze these effects, we again conducted a repeated measures categorical data analysis, with the three repeated measures factors being response mode, conceptual focus, and problem type and the one between-subjects factor being task order (choice first versus estimation first). As in Experiment 1, the initial saturated model produced no significant effect of order and no effect of focus or interaction with focus. The reduced model that eliminated the focus variable again showed no main effects or interactions with order, and no significant residual variance. Thus, once again the final model included only an intercept, the main effects of response mode and problem type, and their interaction. All effects were statistically significant: $\chi^2(1) = 120.4$, $p < .001$ for response mode, $\chi^2(1) = 93.5$, $p < .001$ for problem type, and $\chi^2(1) = 6.2$, $p < .05$ for the interaction. The residual variance for this model was nonsignificant, $\chi^2(4) = 6.2$, $p > .05$. Parallel to Experiment 1, the data of Experiment 2 were well characterized by a reduction in conjunction errors for estimation as compared with choice and for likely conjuncts as compared with unlikely conjuncts, with the combination of estimation and likely conjunct conditions creating greater reductions than predicted by an additive model.

3.2.1. Analysis of the estimation data

For each problem we again tallied whether estimates satisfied the normative constraint that the values for the two conjunctions add to the value for the base event. The percentages of participants who satisfied this rule were similar to those found in Experiment 1: 15.6% for probability estimates of unlikely conjuncts, 21.1% for probability estimates of likely conjuncts, 32.0% for frequency estimates of unlikely conjuncts, and 27.3% for frequency estimates of likely conjuncts. A repeated measures categorical analysis was conducted on these frequencies following the 2×2 design of conceptual focus combined with problem type and produced similar results to Experiment 1. A reduced model was derived with the one significant effect being focus, $\chi^2(1) = 8.19$, $p < .01$, and no significant residual effects, $\chi^2(2) = 2.74$, $p > .05$.

We performed the same classification analysis on estimates for the dice and urn problems as in Experiment 1, the one difference being that these problems occurred in both the unlikely and likely conjunct conditions. Therefore, we are able to examine classification across this variable that coincided with fairly large changes in conjunction errors. Each participant estimated either frequencies or probabilities for just one of these two problems, resulting in 16 participants in each of the eight estimation conditions (probability or frequency estimation for dice or urn problem for unlikely or likely conjunct). Classification was carried out in the same way described in Experiment 1. Predicted probabilities for the unlikely conjuncts were the same as described in Experiment 1. Predicted probabilities (frequencies) for the likely conjuncts from the distributional model were 0.833 (83), 0.694 (69), and 0.139 (14) for the dice problem and 0.90 (90), 0.72 (72), and 0.18 (18) for the urn problem. Predicted probabilities (frequencies) for the likely conjuncts from the averaging model

Table 4
Classification of different models participants may have used to estimate probabilities and frequencies for dice and urn problems combined (Experiment 2)

Model	Unlikely conjunct			Likely conjunct		
	Prob.	Freq.	Sum	Prob.	Freq.	Sum
Distributional	7	6	13	6	14	20
Quasi-distributional	4	7	11	1	1	2
Averaging	10	5	15	13	4	17
Single event	7	6	13	4	2	6
Unclassified	4	8	12	8	11	19

Note. Prob., Probability estimation, Freq., Frequency estimation; $N = 32$ for each type of problem by focus condition.

were 0.833 (83), 0.833 (83), and 0.500 (50) for the dice problem and 0.90 (90), 0.85 (85), and 0.55 (55) for the urn problem. Predicted probabilities (frequencies) for the single event model were 0.833 (83), 0.833 (83), and 0.833 (83) for the dice problem and 0.90 (90), 0.90 (90), and 0.90 (90) for the urn problem.

The classification of each participant's pattern of data into different models is shown in Table 4 combined across the dice and urn problems as these yielded similar results.[8] The pattern of model classifications for the unlikely conjuncts is more evenly divided than in Experiment 1, with the mode once again being the averaging model, but only by a small amount. One of the clearest differences between unlikely and likely conjunct conditions is the large reduction of the quasi-distributional model for the likely conjunct. Note that when distributional and quasi-distributional models are combined, both types of conjuncts have similar frequencies (24 and 22 for unlikely and likely conjunct conditions, respectively). However, loose implementation of the distributional model in the unlikely conjunct condition tends to lead to a conjunction error, which may account for the large number classified into the quasi-distributional model. On the other hand, because the base event in the unlikely conjunct condition has a high likelihood, even a loose implementation of the distributional model tends to avoid conjunction errors, and hence avoid classification into the quasi-distributional group.

A further difference in implementing these models in likely and unlikely conjunct conditions lies in the consequences of using the single event and averaging models. In the unlikely-conjunct condition 100% of the 15 participants classified into the averaging model and 100% of the 13 participants classified into the single event model committed conjunction errors. Again, by the nature of the alternatives in the likely conjunct condition, these percentages were reduced in the likely conjunct condition to 37% of the 19 participants in the averaging model and 33% of the 6 participants in the single event model. Thus, these model classification data for estimation help demonstrate why conjunction errors are reduced so much more for the likely

---

[8] The percentage of unclassified participants (24.22%) was somewhat higher in Experiment 2 than in Experiment 1 (9.38%). A number of different patterns characterized the numerical estimates for those left unclassified, so it is difficult to relate these participants to any one alternative strategy.

conjunct condition. Primarily, it is because the same inappropriate model that produces a conjunction error for unlikely conjuncts may not do so for likely conjuncts.

A final way to look at these model classification data is to examine how participants using the different strategies performed in the other estimation condition and in choice conditions. Table 5 presents these data, summing across conceptual focus. First consider those who gave numerical estimates of urn or dice problems for an unlikely conjunct. As shown in Table 5, regardless of the strategy used, these participants show fairly low levels of conjunction errors in the likely-conjunct estimation task. This occurs despite the fact that all participants classified as using the distributional model committed no conjunction errors for the problem from which their model was inferred but all the participants classified as using averaging or single event models committed conjunction errors for the problem from which their model was inferred. Furthermore, note that regardless of inferred model, participants showed very large conjunction errors in the choice conditions. Examining the data in the lower half of Table 5 for participants whose model was derived from the likely conjunct estimation task, we note that the percentage of conjunction errors is moderately high in the unlikely-conjunct estimation condition, despite these participants rarely making conjunction errors for the problem from which their models were inferred. Once again, these participants show high rates of conjunction errors in choice regardless of model.

3.2.2. Analysis of the Linda data

Experiment 2 replicated the administration of the Linda problem and subsequent interpretation question. Results were similar to Experiment 1, but somewhat stronger. For the full form, the percentages of conjunction errors were 81.3% with two response options and 84.4% with three response options and decreased in the reduced form to 34.4% with two options and 56.3% with three options. Also, once again misinterpretations of the base event were significantly greater in the two option case (34.4%) than in the three-option case (12.5%), $\chi^2(1) = 8.54$, $p < .01$.

Table 5
Conjunction error percentages for participants classified into different models from likely and unlikely-conjunct conditions

Model	Condition model inferred from	N	UCE (%)	LCE (%)	UCC (%)	LCC (%)
Distributional	Unlikely conjunct	13	–	26.9	84.6	69.2
Averaging	Unlikely conjunct	15	–	23.3	70.0	63.3
Single event	Unlikely conjunct	13	–	34.6	84.6	92.3
Distributional	Likely conjunct	20	37.5	–	72.5	60.0
Averaging	Likely conjunct	17	47.1	–	91.2	58.8
Single event	Likely conjunct	6	75.0	–	83.3	83.3

Note. UCE, unlikely-conjunct estimation task; LCE, likely conjunct estimation task; UCC, unlikely-conjunct choice task; and LCC, likely conjunct choice task.

3.3. Discussion

A primary aim of Experiment 2 was to replicate the design of Experiment 1 while eliminating the confounding of problem type and problem topic. The basic pattern of conjunction errors was the same across Experiments 1 and 2, demonstrating that the effects attributed to unlikely versus likely conjuncts are not simply due to nesting of different problem topics. In both sets of analyses, the results did not depend on the order in which the tasks were completed, choice first or estimation first. Even though conjunction errors were greatly reduced in the estimation condition, the processes responsible for this reduction did not transfer to the choice task, for which the problem and option forms were identical. As demonstrated by others (Bonini et al., 2004; Sloman et al.'s, 2003; Tentori et al., 2004), shifting the focus from probabilities to frequencies did not significantly reduce conjunction errors in either Experiment 1 or 2 (although it trended in that direction). However, changing the problem from one joining an unlikely event with a likely event to one joining two likely events did significantly reduce conjunction errors in both experiments. The significant interaction of response mode and problem type found in both experiments also demonstrated that the reduction of conjunction errors with estimation was much larger for the likely conjuncts than for the unlikely conjuncts. This appears to be because strategies for estimating probabilities or frequencies that produce conjunction errors for unlikely conjuncts do not do so for likely conjuncts. The lessening of this difference in choice may be due to how ties are resolved or to the use of different strategies in choice than estimation (cf., Hertwig & Chase, 1998). Note also that inclusion of the dice and urn problems in the likely-conjunction condition demonstrated that large conjunction errors can be obtained even when there is no association between the two events (rolling different dice or selecting balls from different urns). Hence, in addition to the two recipes Tversky and Kahneman (1983) give for creating conjunction errors, one may add one that combines two likely events, even if they have no particular association.

A second aim of Experiment 2 was to include a numerically-based problem in each estimation condition so that inferred models could be examined across estimation conditions. In both Experiments 1 and 2, only a minority of participants (25.0% and 20.3%, respectively) were classified as using the correct distributional model in the unlikely conjunct estimation task. An additional small group (13.5% and 17.2%) were classified as following a quasi-distributional model, because their pattern of estimates was closest to that predicted by the distributional model but they still managed to commit conjunction errors. These participants may have an intuitive grasp that conjunctions lead to smaller probabilities but fail to see the subset relationship completely. A rough majority of participants across Experiments 1 and 2 (55.3% and 48.4%, respectively) tended to follow either an averaging model or single event valuation model to make their estimates in the unlikely-conjunct condition, leading to large numbers of conjunction errors. As shown in Table 4, the main change in the strategies inferred from unlikely conjuncts compared with those inferred from likely conjuncts is that in the latter the quasi-dis-

tributional category virtually disappears and the distributional category expands to take these up. This may occur because a loose implementation of the distributional model will generally lead to no conjunction errors in the likely-conjunct case with a high probability base event, but it will generally lead to conjunction errors in the unlikely-conjunct case with a low probability base event.

The only effect of conceptual focus that was observed in both Experiments 1 and 2 concerned the percentage of participants whose estimates followed the subset rule in which the values for the target conjunction and the conjunction with the complementary event add up to that for the base event. In both experiments, frequency estimates were somewhat more likely to follow this rule than probability estimates, perhaps due to the difficulty in dealing with fractions. Although this was the case, it did not lead to statistically significant reductions in conjunction errors with frequency focus.

One other noteworthy effect was the confirmation in both experiments that the base event was rarely considered as representing a conjunction with the complement of the added event when that specific combination was included as one of the options. This was tested using the Linda problem. Averaging across the two experiments, nearly one third of the participants endorsed the interpretation of the base event as implying a conjunction with the complement of the added event (i.e., "*Linda is a bank teller and she is not involved in the feminist movement*") when only the two options were available, but this was reduced to approximately one in ten participants endorsing this interpretation in the three-option case. Because the test problems all appeared in this three-option format, it seems unlikely that conjunction errors found in these conditions are due to this type of misinterpretation of the base event statement.

4. General discussion

Experiments 1 and 2 provided a systematic manipulation of three variables, response mode, conceptual focus, and problem type. The basic pattern of results was replicated across the two experiments and sets some boundary conditions or constraints on when conjunction errors may be expected and how they may be interpreted. We believe these experiments have implications for three different yet related topics: (1) implications for the status of the misinterpretation hypothesis, (2) implications for the understanding the processes underlying conjunction errors, and (3) implications for interpreting the rationality of this behavior. We address each of these in turn.

4.1. The misinterpretation hypothesis

We believe the pattern of effects in our studies provides evidence against the misinterpretation hypothesis being a primary explanation of the conjunction fallacy. The misinterpretation hypothesis postulates that the conjunction fallacy is mainly due to interpreting the base event (e.g., Linda is a bank teller) as the

conjunction of the base event and the complement of the added event (e.g., Linda is a bank teller and not is active in the feminist movement). To help our participants avoid this misinterpretation, we followed the procedure of Tentori et al. (2004) and explicitly included the alternative with which the base event was presumably confused (i.e., base event and complement of added event). Furthermore, at the end of each experiment we tested whether this type of confusion was indeed reduced by including the Linda problem in either the two or three alternative format followed by a probe of how participants interpreted the base event. Consistent with the misinterpretation hypothesis, the two alternative format produced a large minority of participants who misinterpret the base event as including the complement of the added event. However, the percentage misinterpreting the base event was greatly reduced to only about 10% when the third alternative was included. Thus, we have evidence that very few people misinterpret the base event under the three item format, and we find that using this format there were still a very large proportion of conjunction errors in our two experiments, especially in the choice response mode. Therefore, we feel it unlikely that these conjunction errors are primarily the result of misinterpreting the alternatives presented.

Also relevant to this discussion is the percentage of conjunction errors for those who did not misinterpret the base event. One way to examine this is to compare conjunction errors in the Linda problem for the two and three alternative conditions. We replicated the finding presented by Dulany and Hilton (1991) that the two-option condition generates a high rate of misunderstanding of the base event (almost 40% in Experiment 2 in the full description condition). But if the conjunction fallacy is mainly due to this misunderstanding, one would expect that when such misunderstanding is reduced, there would be a similar reduction in conjunction errors. Although the percentage of misinterpretations was significantly reduced by use of the three-option format in both experiments, the number of conjunction errors was not. Thus, a reduction in misunderstanding of the conjunct for the Linda problem did not elicit a similar reduction in conjunction errors for the problem. Taking this approach one step further, we compared the number of conjunction errors in our probability choice and frequency choice conditions across the two experiments for the subset of participants who correctly interpreted the base event of the Linda problem and those who did not. One might expect that those who correctly interpret the base event would show fewer conjunction errors overall in these conditions, but they did not.

Finally, the misinterpretation hypothesis faces the problem of explaining the large reduction in conjunction errors for estimation compared to choice. We believe that the same conversational-pragmatic analysis used to derive the misunderstanding of the conjunct hypothesis in the choice task can be applied to the estimation task. Therefore, if people tend to misinterpret the conjunct in the ranking or choice instruction, why would they not do so as well in the estimation condition? Taken together, we believe the pattern of effects in our experiments is not supportive of the misinterpretation hypothesis being the prime motivator of conjunction errors.

4.2. Underlying processes

A long standing criticism of the heuristics and biases approach is that the heuristics described are too vague to be empirically tested (see Gigerenzer, 1996). For the conjunction fallacy, Tversky and Kahneman (1983) use the representativeness heuristic as an explanatory device. This heuristic is described as incorporating a similarity comparison between the model for generating the events, or the population from which the sample is drawn, and the instance or sample itself. One basic model for similarity comparison that might be used to build a more specific version of representativeness is Tversky's (1977) contrast model, in which similarity is a weighted average of matching and nonmatching features. This model incorporates three weighting parameters, θ, the weighting of common features, α, the weighting of features unique to the referent, and β, the weighting of features unique to the variant. The model was developed to account for violations of metric distance axioms often reflected in similarity judgments (e.g., violations of minimality, symmetry, and triangle inequality). However, even using the simplest form in which weights may be held constant, the model can be interpreted as generally reflecting an averaging rule. More generally, judgment data are often well characterized by averaging models, which might correspond to application of a simple rule or may be the result of a sampling process within a random walk process (Wedell & Senter, 1997). Note that an averaging rule implies the conjunction error and a tendency for it to be reduced in the likely conjunct condition. This rule also appears to reflect the numbers actually generated by several of the participants when attempting to quantify their estimates of probabilities and frequencies in the dice and urn problems (see also Fantino et al., 1997).

It is clear that other rules besides averaging are also used within the estimation task. The single event rule may reflect a simplification of the problem or a minimum expectation one would generate for the probability of a disjunction (a misinterpretation of the conjunction). Other participants appeared to employ something akin to the normative rule in which the distribution is decomposed into subsets reflecting base and conjunction events. However there are three noteworthy aspects to this approach. First, few participants actually achieve it. Second, some participants appear to approximate it but fail to achieve the conditions necessary for extensional reasoning. Finally, regardless of whether one uses this rule in an estimation problem, one still tends to show large conjunction errors in choice.

The clear cut differences in conjunction errors demonstrated by the same participants for choice versus estimation tasks regardless of the order of these tasks argues strongly for the use of different modes of thinking associated with these two response modes (Hertwig & Chase, 1998). Consistent with the compatibility hypothesis (Tversky, Sattath, & Slovic, 1988), different response modes may elicit evaluations that are compatible with the desired response. Hence choice (a qualitative response) elicits qualitative thinking and estimation (a quantitative response) elicits more quantitatively based methods of evaluation. The notion of separable modes of thought permeates the heuristics and biases literature, where people are capable of integrating base rates, for instance, when the problem is formalized,

but when they are given a richer description they are lured into a qualitative, heuristic approach (Kahneman & Tversky, 1996). Others within decision making have proposed similar dichotomies, such as the proposal of gist encoding versus fine-grain encoding in the memory based fuzzy trace theory applications to decision making (Reyna, 2004). Likewise, these distinctions may map onto the difference between associative versus rule-based modes of thinking described by Sloman (1996). As similarity is a basis for association, similarity based modes of thinking reflected in the representativeness heuristic are likely to reflect automatic associative impressions that are consistent with our intuitions and personal experience. This may well be the default mode of thinking unless one is forced to search for a useful rule to combine information, as might be required by estimation tasks.

Regardless of how we classify these different modes of thinking, it would be useful to develop methods for a more fine-grain analysis of these procedures. Estimation using frequency and probability scales provide a clear basis for inferring what types of rules people are using from the pattern of values, as shown in our current set of analyses. The all-or-none nature of choice, however, makes it more difficult to conduct a parallel analysis of these associative, similarity based, or gist extraction processes. One avenue for future research in this regard may be to use graded scales of choice as surrogate continuous measures related to choice that provide the necessary variation to consider greater specification of the processes involved. Such a graded scale would provide more relevant information than simple choice or rankings, which only provide an ordering of alternatives. In addition, the within-subject manipulation of conditions used in the present set of experiments seems quite promising, since the individual differences in the processing of these problems exhibited in the current set of experiments precludes simple averaging across participants. Such analyses would push the field toward the desired goal of explicating the processes underlying the heuristics that are often asserted.

4.3. Implications for rationality

How then do these results relate to the debate on the rationality of human reasoning and decision making? To answer this question requires us to describe our bases for defining rationality. We believe there are several such bases. Within formal systems, rationality is often defined in terms of procedures that lead to consistency. Thus, for example, transitivity is rational because a lack of transitivity leads to inconsistent evaluations that are order dependent. Tversky et al. (1988) have referred to this type of constraint as procedural invariance, in which responses do not differ with irrelevant changes in descriptions or response formats that do not affect the deep structure of the problem. Note that the within-subject design used in Experiments 1 and 2 allows us to highlight violations of this type of invariance. The same participants whose choices indicate that the conjunction is the more probable event systematically provide estimates that are inconsistent with this assertion. Thus, there is clear evidence in our data for the type of irrationality that arises from inconsistency or violations of procedural invariance.

A second basis for evaluations of rationality lay in determining whether behavior conforms to a normative rule. According to this approach, selection of a conjunction as having higher frequency or probability than a base event is a violation of extensional reasoning and irrational. However, one may object to applying the probability calculus to estimates of probability to reflect beliefs about propensities of unique individual, as such estimates are clearly grounded in a subjectivist interpretation of probability rather than a frequentist interpretation (Gigerenzer, 1996). To this end, we avoided problems that considered the propensities for unique circumstances or individuals and relied strictly on problems formulated in sampling terms. Hence, our probability focus considered sampling a given individual out of a population and our frequency focus considered the numbers of individuals within a large random sample that adhered to a given description. We therefore feel that the problems we investigated (with the exception of the Linda problem added at the end) are immune to objections of the potential misapplication of probability to degrees of subjective belief.

But even if people misinterpret the word probability in sampling contexts, this objection does not apply to versions of our problems that were couched in frequency terminology. Roughly two thirds of participants across Experiments 1 and 2 classified the group of people with blonde hair and blue eyes as more *numerous* than the group of people with blonde hair. This is a typical result in the frequency choice condition (see Tables 1 and 3). In the frequency condition, the word probability does not even appear. These results provide strong evidence against the idea that the conjunction fallacy is due to a misunderstanding of the word probability. Clearly, in this condition, there is a flagrant violation of a norm in set theory that a subset can never be bigger than its superset.

Another possible objection to the norm of selecting the base event as more probable or more frequent arises from linguistic contextual considerations. If in the context of the interaction being held between experimenter and participant (through the booklet verbiage) the natural interpretation of the base event is that it implies a conjunction with the complement of the added event, then its selection would not violate normative principles (Dulany & Hilton, 1991). However, we empirically demonstrate that the three options we include in these problems rarely lead to this interpretation and so we reject the hypothesis of linguistic misinterpretation. On these counts then, we believe the experimental demonstrations reinforce the failure to conform to a normative principle demonstrated in other like research and hence reflect a violation of rationality in this sense.

A third interpretation of rationality focuses on the idea of using adequate means to satisfy some goal. In this context, things become murkier because it is very difficult to establish clear cut criteria for rationality. The first problem here is to determine the goals participants have in mind when engaging in these tasks. Is the goal to provide reasonable and coherent evaluations? Or is the goal to spend as little time as necessary to complete the tasks in ways that conform to expectations in order to earn a participation credit? If this last possibility is the case, our experiments do not allow us to question the rationality of our participants. Given the purpose of gaining credit

for participation, just showing up and marking any responses would be rational.[9] On the other hand, research that has used monetary incentives has typically failed to eliminate conjunction errors. If obtaining the presented monetary incentives was an important goal for the participants, then failure to eliminate conjunction errors under these conditions might be irrational under this interpretation.[10]

A related proposal, defended by Hertwig and Gigerenzer (1999) is that participants might be trying to be as *informative* as possible when solving these problems. Let us provisionally accept this idea: The participant's goal is to be informative. Now, given the typical features of the Scandinavian population, for example, checking the blonde hair and blue eye option as most probable or numerous is more informative (even if incorrect) than checking the blonde hair option. But if people just want to be informative in this way, why do they tend to avoid conjunction errors in the estimation condition? Is it the case that they do not want to be informative anymore?[11]

A second difficulty in determining clear standards under this interpretation of rationality is to explicate the cognitive tools participants have at their disposal. If

---

[9] An anonymous reviewer suggested two implications related to the goal of just getting through the experiment. One implication of this goal would be that participants fail to be consistent in applying rules. As a measure consistency we created split halves measures for the data by summing the number of conjunction errors under probability focus and the number under frequency focus and using these to estimate the reliability of the full set of eight items. This estimate was .67 in Experiment 1 and .37 in Experiment 2. One interpretation of these fairly low consistency indices is that participants may have been more focused on completing the experiment than being consistent across problems. A second implication of participants being focused on the goal of just getting through the experiment is that the number of participants who failed to follow instructions might be fairly large. We replaced participants whose responses did not strictly match instructions (for example, if they indicated a probability when a frequency was called for, or a choice when an estimate was called for). Consistent with this implication, nearly 1/3 of participants were replaced on these criteria, suggesting that a fair percentage of our participant pool may have been anxious to get the problems done quickly so that they did not pay attention to some aspects of the problems or instructions.

[10] Several researchers have asked participants to rank options according to their willingness to bet on them, while others have offered participants real betting opportunities on the options. So far, the conjunction fallacy has remained robust under conditions of monetary incentives. Tversky and Kahneman (1983) found that people prefers to bet on ''*Linda is feminist bank teller*'' rather than on ''*Linda is a bank teller*''. This result has been replicated under different conditions and with a variety of examples (see Bar-Hillel & Neter, 1993; for a systematic study, and also Messer & Griggs, 1993; Sides, 2000; and Bonini et al., 2004). In all these cases, researchers have found high percentages (always more than 50%) of conjunction violations. Even if these studies are very suggestive, there is still a possibility that participants have other goals in mind when completing these studies.

[11] An advocate of the informativeness hypothesis might respond as follows. Maybe what the subjects wants to express is that the blonde and blue-eyed group is more numerous than the blonde and not blue-eyed group. In the estimation framework, they have the possibility to express that idea and they usually do it. In the choice framework, the only way to express the same idea is by violating the conjunction rule. This version of the informativeness hypothesis is better but it also runs into similar problems. If this account were correct, one would expect that people follow the conjunction rule in all the estimation conditions. However, this does not happen. In the probability focus unlikely-conjunct condition, most of the people commit the conjunction fallacy. This failure cannot be accounted by pointing the informativeness issue. Furthermore, in the frequency focus unlikely-conjunct condition, the rate of violation was also very substantial (56% and 48% in Experiments 1 and 2, respectively). Again, this result cannot be explained by the second version of the informativeness hypothesis.

their cognitive tools do not contain the conjunction rule or the subset rule, it would seem incorrect to argue that participants are irrational for not using these rules (see Cherniak, 1986). Thus, this line of thought leaves the possibility of rationality open even if there is a norm violation.

A fourth and related interpretation of rationality includes as an essential feature the adaptiveness of the behavior within the relevant context (see Gigerenzer & Selten, 2002). A behavior that leads to clearly negative outcomes when there is a clear basis for choosing a behavior leading to positive outcomes would seem maladaptive and irrational in this regard. In order to explore the adaptiveness of some behavior, this approach closely investigates the environment in which this behavior usually occurs. From this point of view, advocates of human rationality might argue as follows. In typical human environments, people may never face (except in reasoning experiments!) a set of options such that one of the options is a subset of another option. Thus, even if some normative rule is being violated, such behavior is not maladaptive because typically no negative outcome follows (in fact, usually, no outcome follows, because the conditions for the behavior (almost) never occur). But this line of thought rests on little more than speculation. In order to explore the consequences for this notion of rationality, we believe research is needed that would examine behavior in the real world, where real consequence exist for committing such errors. Such research is hard to come by. There may be hints of the generality of the conjunction error to such consequentially important behaviors, but we know of no systematic demonstration of this.

Thus, depending on the notion of rationality one holds, our studies might be more or less relevant for exploring the issue of human rationality. But even if no conclusion can be drawn about human rationality, these studies do help to establish an important result: The conjunction fallacy is not due to a misinterpretation of the problem by participants. We tested different versions of the misinterpretation hypothesis and found evidence to reject them all. So, the advocates of human rationality cannot use this possibility to defend their claim.

Finally, let us explore what can be the practical consequences of our study. Its relevance lies in uncovering the boundary conditions surrounding the conjunction fallacy phenomenon. The body of research on conjunction errors points to the dangers of intuitively choosing courses of action that require the integration of probabilities as potentially leading to grave miscalculations when outcomes are consequential. Our research and that of others further suggest that the tendency for such missteps will be greatly reduced if the individual can be encouraged to evaluate probability or frequency-based estimates of options before undertaking a given course of action. Thus when applied to environments with important consequences attached to decisions (as in the fields of medicine, engineering, forecasting, etc.) these studies imply that using systematic rules for estimation should be the basis for choice.

Appendix A

This Appendix presents the different problems used in Experiments 1 and 2. Each problem could be presented in one of eight versions, based on the experimental

Table A
Conjunction problems used in Experiments 1 and 2

Name/type/ Experiment 1	Problem	Unlikely conjunct	Likely conjunct
Dice problem Game of chance	Consider a game in which there are two dice. One has *five red* sides and *one blue* side. The other has *five yellow* sides and *one green* side. For any given round, both dice are rolled. For the next round to be played, please indicate which of the following outcomes is most probable?	Roll a blue. Roll a blue and roll a yellow. Roll a blue and roll a green.	Roll a red. Roll a red and roll a yellow. Roll a red and roll a green.
Unlikely conjunct in Experiment 1			
Football problem	American football players are earning more and more money but saving less and less. A survey revealed that on average an American football player earns $800,000 a year. Paradoxically, on average an American football player saves no more than $50,000 a year. Suppose we choose at random an American football player.	The individual earns under $400,000 a year. The individual earns under $400,000 a year and saves less than $30,000. The individual earns under $400,000 a year and saves $30,000 or more.	The individual earns over $600,000 a year. The individual earns over $600,000 a year and saves less than $80,000. The individual earns over $600,000 a year and saves $80,000 or more.
Sampling continuous category			
Unlikely conjunct in Experiment 1			
Urn problem Game of chance	Consider a game in which there are two urns. One contains *90 black* balls and *10 white* balls. The other contains *80 orange* balls and *20 brown* balls. For any given round, a ball from each urn is drawn at random. After the results are recorded, the balls are replaced, and the urns are properly shaken. For the next round to be played, please indicate which of the following outcomes is most probable?	Draw a white ball Draw a white ball and draw an orange ball Draw a white ball and draw a brown ball.	Draw a black ball. Draw a black ball and draw an orange ball. Draw a black ball and draw a brown ball.
Unlikely conjunct in Experiment 1			

Table A (*continued*)

Name/type/ Experiment 1	Problem	Unlikely conjunct	Likely conjunct
Movie problem Sampling continuous category Unlikely conjunct in Experiment 1	Hollywood movies are becoming longer and more expensive to produce. The average duration of movies released in the last 5 years is 110 min. The cost of making these movies has also gone up, averaging $21 million per movie during this same period. Suppose we choose at random a movie of the last 5 years. Which event do you think is most probable?	The movie is less than 90 min. The movie is less than 90 min and cost more than $10 million. The movie is less than 90 min and did not cost more than $10 million.	The movie is more than 95 min. The movie is more than 95 min and it costs more than $10 million. The movie is more than 95 min and it does not cost more than $10 million.
Height and weight problem Sampling continuous category Likely conjunct in Experiment 1	Americans are getting bigger and heavier. The average height of a 30-year-old male is 5′10″ and the average weight is 155 lbs. Let's say we randomly pick a 30-year-old male from the population. Which of the following outcomes is most probable?	The individual weighs under 145 lbs. The individual weighs under 145 lbs and is over 5′5″. The individual weighs under 145 lbs and is not over 5′5″.	The individual weighs over 145 lbs. The individual weighs over 145 lbs and is over 5′5″. The individual weighs over 145 lbs and is not over 5′5″.
Scandinavian problem Sampling discrete category Likely conjunct used in Experiment 1	The Scandinavian peninsula is the European area with the greatest percentage of people with blond hair and blue eyes. This is the case even though every possible combination of hair color and eye color occurs in those countries. Suppose we choose at random an individual from the Scandinavian population.	The individual has green eyes. The individual has green eyes and blond hair. The individual has green eyes and does not have blond hair.	The individual has blue eyes. The individual has blue eyes and blond hair. The individual has blue eyes and does not have blond hair.
NBA problem Sampling continuous category Likely conjunct used in Experiment 1	To play in the NBA (National Basketball Association) one needs to be very talented, very athletic, very tall and very lucky. The average height of an NBA player is 6′6″ and the average age is 28 years old. Let's say we randomly pick a player out of the NBA.	The individual is under 6′1″. The individual is under 6′1″ and is more than 23 years old. The individual is under 6′1″ and is not more than 23 years old.	The individual is over 6′4″. The individual is over 6′4″ and is more than 23 years old. The individual is over 6′4″ and is not more than 23 years old.

Table A (*continued*)

Name/type/ Experiment 1	Problem	Unlikely conjunct	Likely conjunct
Hockey problem	Professional hockey has become one of the most violent sports to play. This fact, in turn, makes professional hockey players a risk group for certain accidents and health problems. Suppose that we randomly choose a professional hockey player and review the health reports throughout his career.	The player has diabetes.	The player has experienced a concussion.
Sampling continuous category		The player has diabetes and has experienced some concussion.	The player has experienced a concussion and has suffered some memory loss.
Did not appear in Experiment 1		The player has diabetes but has not experienced any concussion.	The player has experience a concussion but has not suffered any memory loss.
India Problem	India is the Asian country with the greatest percentage of bilingual people (i.e., people who fluently speaks two languages). India also rates the highest in percentage of adults working for foreign companies. Suppose we choose at random an adult from India. Which event do you think is most probable?	N/A	The individual fluently speaks two languages. The individual fluently speaks two languages and works for a foreign company. The individual fluently speaks two languages and does not work for a foreign company.

design. In Experiment 1, only four versions were generated for each problem, with each problem nested within type of problem (likely or unlikely conjunct). In Experiment 2, each problem was additionally represented as each type of conjunct. Table A presents the base problem, the likely conjunct options, and the unlikely-conjunct options under probability choice condition only. The table also designates which form the problem occurred in for Experiment 1. Note that the India problem occurred only in Experiment 1. Because this problem did not generate many conjunction errors, it was replaced in Experiment 2.

References

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology, 65*(6), 1119–1131.

Bonini, N., Tentori, K., & Osherson, D. (2004). A different conjunction fallacy. *Mind and Language, 19*(2), 199–210.

Cherniak, C. (1986). *Minimal rationality*. Cambridge: MIT Press.

Dulany, D. E., & Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition, 9*, 85–110.

Fantino, E., Kulik, J., & Stolarz-Fantino, S. (1997). The conjunction fallacy: A test of averaging hypotheses. *Psychonomic Bulletin and Review, 1*, 96–101.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*, 123–129.

Gigerenzer, G. (1994). Why the distinction between single-event probability and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York: Wiley.

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

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 513–525.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. Cambridge: MIT Press.

Hertwig, R., & Chase, V. M. (1998). Many reasons or just one: How response mode affects reasoning in the conjunction problem. *Thinking and Reasoning, 4*, 319–352.

Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275–305.

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

Messer, W., & Griggs, R. (1993). Another look at Linda. *Bulletin of the Psychonomic Society, 31*, 193–196.

Morier, D. M., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin, 10*, 243–252.

Moro, R. (2007). *Linda is back: Controversies around the conjunction fallacy in the psychology of reasoning*. Unpublished Doctoral Dissertation, University of South Carolina.

Politzer, G., & Noveck, I. A. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of Psycholinguistic Research, 20*, 83–103.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13*, 60–66.

Sides, A. (2000). The conjunction fallacy under probability and betting instructions. MA Thesis, Rice University, Bell and Howell Information and Learning Company.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91*, 296–309.

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

Tversky, A. (1977). Features of similarity. *Psychological Review, 24*, 327–352.

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

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review, 95*, 371–384.

Wedell, D. H., & Senter, S. M. (1997). Looking and weighting in judgment and choice. *Organizational Behavior and Human Decision Processes, 70*, 41–64.

Wells, G. L. (1985). The conjunction error and the representativeness heuristic. *Social Cognition, 3*, 266–279.