SOCIAL BEHAVIOR AND PERSONALITY, 2009, **37**(4), 481-490

DOI 10.2224/sbp.2009.37.4.481

MEASUREMENT OF THE CONJUNCTION ERROR IN SOCIAL JUDGMENT: ANSWER CHOICE AND ANSWER JUSTIFICATION

RUSSELL L. KNOTH AND VICTOR A. BENASSI
University of New Hampshire, Durham, NH, USA
GLENN GEHER
State University of New York at New Paltz, NY, USA

People often judge the probability of two events occurring together to be more probable than the less probable of each of these events occurring separately, thereby demonstrating the conjunction error. "Correct" responses are those in which participants rank a single-element statement of low probability as more probable than a conjunction consisting of a low-probability statement and an additional statement. We demonstrated in two studies that task structure was related to the number of people who chose the single-element statement as more probable. However, relatively few participants provided a rationale for their answer choice based on the conjunction rule. In a third study, responses on the answer choice and answer justification measures converged. In addition to serving as a sensitive dependent measure, participants' justifications when answering conjunction problems may provide insight into their reasoning.

Keywords: conjunction error, conjunction fallacy, judgment, answer choice, justification.

A basic principle of probability is the conjunction rule, P (A & B) $\leq$ P (A). The probability of two events, taken together, cannot be more probable than that of either of the individual elements which comprise the conjunction. However, when judging the probability of compound events, people often demonstrate what has been referred to as the *conjunction fallacy* (Tversky &

Russell L. Knoth and Victor A. Benassi, Department of Psychology, University of New Hampshire, Durham, NH, USA; Glenn Geher, Department of Psychology, State University of New York at New Paltz, NY, USA.
Appreciation is due to anonymous reviewers.
Please address correspondence and reprint requests to: Victor A. Benassi, Department of Psychology, Conant Hall, University of New Hampshire, Durham, NH 03824, USA. Phone: +1 603 862 3180; Fax: +1 603 862 1528; Email: **vab@unh.edu**

Kahneman, 1983) or *conjunction error* (Kahneman & Tversky, 1996). People often judge the probability of P (A&B) as being higher than the probability of the less likely component event. This well-established finding has implications for understanding certain types of social judgment. Consider Tversky and Kahneman's (1982, 1983) "Bill problem":

> Bill is 34 years old. He is intelligent, but unimaginative, compulsive and generally lifeless. In school he was strong in mathematics but weak in social studies and humanities.

After reading the paragraph, participants were asked to rank order a number of events based on their statistical likelihood. The events included:

1. Bill plays jazz for a hobby.
2. Bill is an accountant who plays jazz for a hobby.

Because the second statement cannot be more statistically probable than the first statement, participants who rank it as being more probable demonstrate the conjunction error. Tversky and Kahneman (1982, 1983) found that a large majority of people in their studies demonstrated this fallacy on the Bill problem and other similar judgment tasks (typically, 85% or more). They argued that people demonstrate the conjunction error because they apply simple cognitive heuristics when making judgments about an event's probability rather than taking into account the conjunction rule. Representativeness is, perhaps, the most extensively discussed of these heuristics (Kahneman & Tversky, 1972). Participants rank the statement "Bill is an accountant who plays jazz for a hobby" as more probable than the statement "Bill plays jazz for a hobby" because being an accountant is more representative of, or similar to, the paragraph describing Bill. This view has had its critics (e.g., Dulany & Hilton, 1991; Fisk, 2002; Gigerenzer, 1996).

What is to be said of participants who choose a single-element statement as more probable than a conjunction? Some people may rank the single-element statement as more probable than the conjunction statement based on reasoning other than the conjunction rule. The dependent measure of choice in previous studies has been whether participants rate or rank a conjunction or single-element statement as more statistically likely (e.g., Tversky & Kahneman, 1983; Wells, 1985). In this article, we suggest that employing answer choice as the sole measure of the understanding of the conjunction rule may provide an incomplete, and sometimes inaccurate, understanding of participants' reasoning about a variety of events. In addition to answer choice, we used an *answer justification* measure of the conjunction error. We sought to determine whether participants in conjunction problem studies who choose a correct answer do so by applying the conjunction rule or do so based on reasoning other than the conjunction rule.

STUDY 1

Tversky and Kahneman (1983) used a version of the Bill problem in which the components of the conjunction were incompatible:

> The conjunction "Bill is bored by music and plays jazz for a hobby" was judged as less probable (and less representative) than its constituents, although "bored by music" was perceived as a probable (and representative) attribute of Bill. Quite reasonably, the incompatibility of the two attributes reduced the judged probability of their conjunction (p. 305).

Based on an answer choice measure alone, one cannot tell whether the incompatibility between the elements of the conjunction led participants to apply the conjunction rule when judging the statement's probability. In other words, the incompatible-elements version of the Bill problem may have reduced the conjunction error in a manifest manner, but not for the reason that it actually induced the correct statistical reasoning. In Study 1, we used Tversky and Kahneman's (1983) Bill problem with the highly incompatible elements presented in the preceding quote. If the incompatibility of the elements is what led primarily to a reduction of the conjunction error because of an increased reliance on the conjunction rule, participants who chose the single-element answer should write answer justifications based on the conjunction rule.

METHOD

Participants One hundred and fourteen students enrolled in introductory psychology courses at a public university in the northeastern United States participated in the study as part of a course requirement.

Materials and Procedure Fifty-seven students completed a version of the Bill problem in which the elements of the conjunction were highly incompatible (see Appendix). Another 57 students completed the standard version of the Bill problem (Tversky & Kahneman, 1983). The standard version was the same as the incompatible version except that the two element statement was replaced with the statement "Bill is an accountant and plays jazz for a hobby." After the participants chose one of the two statements, they were instructed to describe in writing their rationale for choosing that statement as statistically more probable.

We evaluated answer justifications as demonstrating an understanding of the conjunction rule if the participants viewed the number of elements in a statement as the determinant of its probability. For example, justifications such as "the chances that Bill is an accountant and that he plays jazz for a hobby are less than if he plays jazz" or "the chances are greater that one thing is true than that two things are true" would be coded as responses based on the conjunction rule. We coded other types of justifications as not demonstrating a reliance on the conjunction rule. For example, justifications such as "Bill is an

accountant because he enjoys mathematics" would be coded as not adhering to the conjunction rule.

Results and Discussion

The first author and a research assistant who was not informed about the purpose of the study independently read and coded each answer justification. (We also used this method in Studies 2 and 3.) Initial agreement between the two judges was over 95% in all three studies (cf. Morier & Borgida, 1984, p. 248). The few exceptions were discussed and resolved between the judges.

The results are summarized in Table 1. Based on the answer choice measure, 54% of the participants chose the two-element conjunction statement as more probable on the version of the Bill problem in which the elements of the conjunction were highly incompatible. On the standard version of the Bill problem, 81% of the participants chose the conjunction statement as more probable. A 2 (incompatible versus standard version of the Bill problem) X 2 (correct versus incorrect answer choice) chi-square analysis yielded a significant result, χ^2 (1) = 9.00, $p < .003$, $\Phi = .28$. More participants made the correct answer choice on the incompatible version of the problem. These results are consistent with those reported by Tversky and Kahneman (1983).

In contrast, on both the incompatible and standard versions of the problem, only 1 of the 57 participants in each condition provided a justification based on the conjunction rule. Tversky and Kahneman (1983) indicated that the incompatibility between elements in a compound statement "reduced the judged probability of their conjunction" (p. 305). This statement should not be interpreted to mean that participants necessarily formed their judgments of the incompatible version of the Bill problem based on the conjunction rule. In fact, problems with highly incompatible elements may actually hinder some participants from detecting the relevance of the conjunction rule because, given the description of Bill, they think the two elements do not make sense in combination.

STUDY 2

Wells (1985) reported large effects of variation in task structure on the incidence of the conjunction error. In addition to using the standard procedure of pairing an unrepresentative element and a representative element in a conjunction (UR condition), he also included two other conditions. In one condition, he paired two representative elements (RR); in the other, he paired two unrepresentative elements (UU). Based on previous research, Wells predicted that there would be a high incidence of the conjunction error in the UR condition. He made no prediction about the RR condition. In the UU condition, Wells predicted that the conjunction error should be "minimal or nonexistent, because

the joint outcome is no more representative than are the individual outcomes" (p. 269). Based on a dependent measure derived from participants' probability estimates and comparable to our answer choice measure, 75% of the participants in the UR condition made the conjunction error with the Linda problem (see Appendix). The pairing of two representative elements produced an intermediate effect (25%). The pairing of two unrepresentative elements produced a very low incidence of the conjunction error (8%). We repeated Wells' experiment, with the addition of our answer justification measure.

Method

Participants Seventy-five students enrolled in introductory psychology courses at a public university in the northeastern United States participated in the study on a voluntary basis.

Material and Procedure Students completed the version of the Linda problem used by Wells (1985). They were randomly assigned to one of three conditions (25 students per condition). All participants first read the description of Linda and then made probability estimates for two individual elements. Depending on the condition to which they were assigned, participants gave an estimate based on an unrepresentative element and a representative one, two representative elements, or two unrepresentative elements. They then gave an estimate based on a conjunction consisting of the elements that they had estimated individually. Next, participants wrote a justification for their probability estimate of the conjunction statement.

Results and Discussion

Following the method used by Wells (1985), we coded participant answers as correct if they gave an equal or higher estimate to the single-element statement that received the lower probability estimate than the estimate they provided to the conjunction statement. We coded their answers as incorrect if they gave a higher estimate to the conjunction statement. This was our answer choice measure. We also used the answer justification measure we had used in Study 1.

Based on the answer choice measure, our results replicated closely those of Wells (1985) (Table 1). Eighty percent of the participants in the UR condition demonstrated the conjunction error (compared to 75% in Wells' study). Thirty-six percent of the participants in the RR condition demonstrated the error (compared to 25%). Only 8% of the participants in the UU condition demonstrated the error (compared to 8%). The distribution of answers reliably differed by condition, χ^2 (2) = 27.16, $p < .001$, $\Phi = .60$.

TABLE 1
ANSWER CHOICE AND ANSWER JUSTIFICATION PERCENTAGES AS A FUNCTION OF PROBLEM TYPE

Study 1	Answer Choice[a]	Answer Justification[b]
Incompatible Bill problem	54%	98%
Standard Bill problem	81%	98%
Study 2		
UR[c]	80%	100%
RR[c]	36%	91%
UU[c]	8%	96%
Study 3		
Health Survey	9%	39%
31-year-old Linda	19%	51%

Notes: [a] Incidence of conjunction error based on answer choice.
[b] Percentage of students who wrote answer justifications based on a rationale other than the conjunction rule.
[c] UR = one unrepresentative element and one representative element in the conjunction; RR = two representative elements in the conjunction; UU = two unrepresentative elements in the conjunction.

Our answer justification measure provided different information from the measure derived from participants' probability estimates. One hundred percent, 91%, and 96% respectively, of the participants in the UR, RR, and UU conditions, did not provide answer justifications based on the conjunction rule. These percentages did not differ reliably by condition, χ^2 (2) = 2.26, $p < .32$, $\Phi = .17$. Although 92% of the participants in the two unrepresentative elements condition correctly ordered their probability estimates on Wells' dependent measure, only 4% of the participants wrote answer justifications based on the conjunction rule. This finding is consistent with an implication of Wells' results in his UU condition – namely, that participants judged the elements of the conjunction to be unrepresentative of the information provided in the target paragraph.

STUDY 3

Tversky and Kahneman (1983) described a version of a health-survey problem (see Appendix) that should "induce extensional reasoning" (p. 308) and thus help people avoid making the conjunction error. The problem requires participants to make frequency, rather than probability, judgments. Of 159 participants, Tversky and Kahneman found that only 31% made the conjunction error. Tversky and Kahneman (1983; see also Kahneman & Tversky, 1996) suggested that the frequency version of the problem serves to provide cues that may lead participants to provide a judgment in line with the conjunction rule.

Other researchers have found the frequency versions of conjunction problems may reduce or even eliminate conjunction errors under certain conditions (e.g., Fiedler, 1988; Gigerenzer, 1996; Mellers, Hertwig, & Kahneman, 2001).

Tversky and Kahneman (1983) argued that the conjunction error on the Linda problem occurs because people perceive a connection between the personality description of Linda and the conjunction of bank teller and feminist. When Tversky and Kahneman omitted the personality description and described Linda only as a "31-year-old woman," almost all of their participants "obeyed the conjunction rule" (p. 305). We gave participants either Tversky and Kahneman's (1983) health survey problem or the version of the Linda problem that omitted the personality description. We sought to assess whether or not a high percentage of participants would write answer justifications indicating that they, in fact, obeyed the conjunction rule in making their judgments. If Tversky and Kahneman were correct in suggesting that these problems prompt extensional reasoning, we expected to find low rates of the conjunction error for both the answer choice and answer justification dependent measures.

Method

Participants One hundred and sixty-nine students enrolled in a variety of lower-division psychology courses at a public university in the northeastern United States participated on a voluntary basis.

Procedure One hundred and six and 63 students respectively completed the health survey problem or a modified Linda problem. With the health-survey problem, participants estimated how many of 100 hypothetical respondents in the survey had had one or more heart attacks, how many were aged over 55, and how many were aged over 55 and had had one or more heart attacks. With the Linda problem, participants rank ordered the following three statements in terms of their probability of being true: Linda is a bank teller, Linda is active in the feminist movement, and Linda is a bank teller and is active in the feminist movement. Linda was described only as a 31-year-old woman. After ranking the statements, participants wrote a brief justification for the ranking they gave to the conjunction statement relative to the other two rankings. For the health survey problem, we asked participants whether they gave a lower or higher frequency estimate to the conjunction statement than to either of the first two statements. They then wrote a brief justification for the estimates they provided.

Results and Discussion

With the health survey problem, we coded the answer choice measure as correct if participants gave an equal or higher estimate to the lower of the two single-element statements than they gave to the conjunction; their answers were coded as incorrect if they gave a higher estimate to the conjunction. With the Linda

problem, we coded the answer choice measure as correct if participants gave a higher ranking to the lower-ranked of the single-element statements than they gave to the conjunction statement; their answers were coded as incorrect if they gave a higher estimate to the conjunction. We also used the answer justification measure described in the previous studies.

For the health survey and modified Linda problems, only 9% and 19% of participants, respectively, demonstrated the conjunction error on the answer choice measure (Table 1). These results replicate those reported earlier by Tversky and Kahneman (1983). For the answer justification measure, we found that 39% and 51% of the participants, respectively, demonstrated the conjunction error in the health survey and modified Linda problems. Although the error rates were higher than those shown on the answer choice measure, they were considerably lower than any of the results that we reported for the answer justification measure in the first two studies. The answer justification results for the health survey problem support Tversky and Kahneman's (1983) suggestion that it evokes extensional reasoning, at least for many participants. Further, these results provide qualification to the view of Fiedler (1988) that "it is highly unlikely that the strong effect of replacing 'probability' by 'frequency' is mediated by the cueing or priming of extensional reasoning [i.e., use of the conjunction rule]" (p. 128). About half the participants wrote correct answer justifications to the Linda problem, indicating that they "obeyed the conjunction rule" (Tversky & Kahneman, 1983, p. 305). Still, many of those who ranked the conjunction statement as less likely wrote answer justifications that emphasized a perceived incompatibility between its elements. Perhaps participants' level of general statistical sophistication is an important factor to consider (Benassi & Knoth, 1993).

GENERAL DISCUSSION

We developed an answer justification measure that was useful in assessing the extent of conjunctive reasoning on the same tasks examined by Tversky, Kahneman, and many other researchers. Although several researchers (e.g., Gavanski & Roskos-Ewoldsen, 1991; Morier & Borgida, 1984; Tversky & Kahneman, 1982, 1983; Yates & Carlson, 1986) have examined the reasoning behind participants' answer choices, such justification measures have not been a central focus in many studies of the conjunction error. Kahneman and Tversky (1996) made clear that they "believe that subjective probability judgments should be calibrated" (p. 589) against a normative rule (e.g., the conjunction rule), whereas others have challenged this approach (e.g., Dulany & Hilton, 1991; Gigerenzer, 1996). Whether one agrees with Kahneman and Tversky or with their critics, studies of conjunctive reasoning should include answer justification measures.

One might argue that some participants who chose the single-element statement as more probable did so based on an understanding of the conjunction rule, but they were unable to express their understanding on the answer justification measure. Although this might be the case for some participants, such an argument is inadequate because most students whose answers we judged as incorrect provided justifications that clearly reflected a reliance on a reason *other than* the conjunction rule. They did not struggle for words. Furthermore, our replication of Tversky and Kahneman's (1983) frequency version of the health survey task and the version of the Linda task that did not include a personality description showed that many participants, as expected, articulated the conjunction rule in their written statements. Without a measure of the reasoning behind participants' choices, we would have been left to speculate about how they approached the problem.

REFERENCES

Benassi, V. A., & Knoth, R. L. (1993). The intractable conjunction fallacy: Statistical sophistication, instructional set, and training. *Journal of Social Behavior and Personality*, **8**, 83-96.

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

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

Fisk, J. E. (2002). Judgments under uncertainty: Representativeness or potential surprise? *British Journal of Psychology*, **93**, 431-450.

Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology*, **61**, 181-194.

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

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, **3**, 430-454.

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

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, **12**, 269-275.

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

Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In S. Lichtenstein, D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84-98). Cambridge, England: Cambridge University Press.

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

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

Yates, F. J., & Carlson, B. W. (1986). Conjunction errors: Evidence for multiple judgment procedures, including "signed summation". *Organizational Behavior and Human Decision Processes*, **37**, 230-253.

APPENDIX

CONJUNCTION PROBLEMS USED IN STUDIES 1-3

Bill (incompatible format) (Study 1)

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Concerning the above description, please choose the statement that is more probable:

1. Bill plays jazz for a hobby.
2. Bill is bored by music and plays jazz for a hobby.

Linda (two representative elements condition) (Study 2)

Linda is 31 years old, single, outspoken, and very bright. She majored in Philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. Linda continues to be concerned about issues of discrimination and has organized a group of local women who meet regularly.

Please rate the statistical probability that each of the following statements is true. Rate each statement using a .00 to 1.00 scale, where .00 indicates that the statement has no chance of being true, where 1.00 indicates that the statement is certainly true, and probability estimates between .00 and 1.00 indicate intermediate probabilities that the statement is true.

Please read all three statements before making your probability estimates.

Probability estimates

1. _____ Linda subscribes to a feminist magazine (Ms.).
2. _____ Linda supports the passage of the Equal Rights amendment.
3. _____ Linda subscribes to a feminist magazine (Ms.) and supports the passage of the Equal Rights amendment.

Health Survey (Study 3)

A health survey was conducted in New England with a sample of 100 adult males, of all ages and occupations.

Please give your best estimate of the following values:

1. How many of the 100 participants in the survey have had one or more heart attacks? _____
2. How many of the 100 participants in the survey are over 55 years old? _____
3. How many of the 100 participants in the survey both are over 55 years old and have had one or more heart attacks? _____

Linda (no personality description format) (Study 3)

Here is some general information about a person, Linda.

Linda is a 31-year-old woman.

Please rank in order the following three statements about Linda in terms of their probability (statistical likelihood) of being true.

Give a "1" to the statement that has the highest probability of being true, a "2" to the statement with the next highest probability of being true, and a "3" to the statement with the lowest probability of being true. Please rank each of the statements.

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