

# The conjunction fallacy: Judgmental heuristic or faulty extensional reasoning?

IRWIN D. NAHINSKY, DANIEL ASH, and BRENT COHEN  
*University of Louisville, Louisville, Kentucky*

Twenty subjects trained in graduate-level statistics and 12 subjects employed in a business firm were asked to estimate probabilities of randomly sampling members of specified political parties distributed in known proportions between two pools. There was a strong tendency to commit the conjunction fallacy [e.g.,  $p(\text{republican} \cap \text{Club I member}) > p(\text{republican})$ ] on problems involving sampling the club followed by sampling from the club. This tendency was accompanied by tendencies to confuse conditional with joint probabilities and make such estimates independently of the component event. Such results were less pronounced in situations involving direct sampling from the pools. Results suggest that errors in probabilistic reasoning stemmed from problems in information processing rather than from the use of judgmental heuristics.

Tversky and Kahneman (1983) have attributed fallacies in judgments of probabilities to the use of certain judgmental heuristics that are substituted for appropriate applications of the laws of probability. These authors define a judgmental heuristic as a strategy that relies on natural assessment to produce an estimate or prediction. An example is the representativeness heuristic, which involves an assessment of correspondence between a sample and a population. Such an assessment may produce an estimate incompatible with the laws of probability.

A fundamental relationship in probability stems from the fact that, since  $(A \cap B) \subseteq B$ , then  $p(A \cap B) \leq p(B)$ . Tversky and Kahneman (1983) presented research indicating that individuals often violate this extensional relationship. For example, physicians who were surveyed estimated that the probability of the joint occurrence of two symptoms is greater than that of one of these symptoms in a situation in which the two symptoms would be quite characteristic of a given type of patient. Results of this kind, labeled the conjunction fallacy, were associated with salient representations of schemata. Thus, a symptom pair is representative of a given patient type, and this salience gives rise to a probability estimate that is larger than the less salient representation of the larger set. Such biases are pervasive, as is demonstrated by the fact that individuals trained in probability exhibit the conjunction fallacy with significant frequency.

Although heuristic biases no doubt account for much fallacious reasoning in the realm of probability, it is possible that errors of the type referred to may also be associated with difficulties in information processing and incorrect concepts of probability. Nisbett, Krantz, Jepson, and Kunda (1983) found that individuals think in terms of a relative frequency interpretation when a situa-

tion clearly calls for it. Tversky and Kahneman (1983) noted that the conjunction fallacy occurred, albeit with reduced frequency, even in situations where relative frequencies of events were emphasized to facilitate appropriate extensional reasoning.

We presented a series of probability problems, in which relative frequencies of events were explicit and preexisting biases should have been eliminated in determining probability estimates. Individuals were asked to estimate probabilities of joint events and of component events in a situation involving two dichotomized variables. Estimates were made for problems in which elements were sampled in a simple random fashion or in two stages, with initial random selection of one of the two marginal characteristics. Data were examined for consistent errors that could be attributed to faulty information processing or incorrect reasoning.

## METHOD

### Instrument and Design

Each subject was administered a questionnaire consisting of two sets of four problems each. One set, the two-stage set, consisted of problems exemplified by the following:

Two clubs are available as pools of experimental subjects. Club I contains 24 republicans and 12 democrats, and Club II contains 12 republicans and 24 democrats. An experimenter draws subjects from the pools by first randomly drawing one club and then selecting subjects from the chosen club randomly. The experimenter needs subjects from the same club for a group experiment, but does not want to know which club is drawn in order to avoid bias in assessment of results. Thus, the club that is the source of any sample of subjects is unknown to the experimenter.

Another set, the complete-random set, consisted of problems exemplified by the following:

Four subgroups of subjects are available as an experimental pool. They are composed as follows: 80 republicans over 35 years of age, 20 democrats over 35 years of age, 80 democrats under 35 years of age, and 20 republicans under 35 years of age. Subjects are to be drawn randomly from this pool.

The third author is now at the University of North Carolina, Chapel Hill. Requests for reprints should be sent to: I. D. Nahinsky, Department of Psychology, University of Louisville, Louisville, KY 40292.

Problems varied with respect to club (or age group) sizes (small or large), ratio of one party to another (2:1 or 3:1), and relative club (age group) size (same or different). One group of subjects received four two-stage problems first, followed by four complete-random problems; the problem order was reversed for the second group. The problem set for each subject contained all club (age group) size-party ratio-relative club (age group) size combinations, with assignment of these combinations to problem type balanced between groups. Each problem type contained both levels of each of the variables described for each group. This procedure resulted in a factorial experiment with the three parameters, problem type, and political party sampled as within-subjects factors, groups as a between-subjects factor, and subjects as a nested-random factor. Table 1 shows the eight club (age group) size-party ratio-relative club (age group) size combinations.

Each problem was followed by two questions. Subjects were asked to find the probability that the first subject drawn is a republican. There were two choices allowed for each two-stage problem with equal club size, exemplified by the following (Choice 1 correct):

1. The probability is .50, because each club is of the same size, and the total number of republicans equals the total number of democrats. Therefore, each individual has the same chance of being drawn as any other.
2. This cannot be determined without knowing which club was drawn, because each club is greatly different from the other in composition. Therefore, the probability of drawing a republican is dependent on the club drawn.

When club size was unequal, two choices of the following type (Choice 2 correct) were substituted for the first choice shown above:

1. The probability is .35, because democrats outnumber republicans 13 to 7 overall and each individual should have the same chance of being drawn as any other.
2. The probability is .50, because each club has the same chance of being drawn as the other, and republicans outnumber democrats in Club I in the same ratio as democrats outnumber republicans in Club II.

All complete-random problems allowed two choices for this question, exemplified by the following (Choice 1 correct):

1. The probability is .50, because there are just as many republicans as democrats in the pool.
2. This cannot be determined without knowing whether the individual is under 35 or over 35, because these two age groups are very different with regard to proportion of republicans.

The second question required estimates of the probability for drawing a republican, a democrat, and an individual in each of the four club-party (age-party) combinations. Each estimate was made on a graphic scale of 30 mm extending from .00 (impossible) to 1.00 (certain), marked in four equal intervals.

**Subjects**

Two groups of subjects were sampled. One group consisted of 12 employees of a local firm. All had at least some college background and 7 had some coursework in statistics or probability. A second group consisted of 6 faculty members and 14 graduate students in the psychology doctoral program at the University of Louisville. All had training in graduate-level statistics. Each group was divided equally between problem-presentation orders.

**RESULTS**

Rules of probability require that  $p(\text{party} \cap \text{group } 1) + p(\text{party} \cap \text{group } 2) = p(\text{party})$ , where group *i* refers to club- or age-group *i*, and the *p* values refer to random sampling probabilities for designated categories for the two political parties. Consequently,  $p(\text{party} \cap \text{group } 1) + p(\text{party} \cap \text{group } 2) - p(\text{party})$  was calculated for each subject-party-problem combination from the graphic

**Table 1**  
**Club (Age Group) Size-Ratio-Relative Size Combinations**

	Small		Large	
	Same Size	Different Size	Same Size	Different Size
2:1	12:24, 24:12*	8:16, 32:16	70:35, 35:70	24:48, 48:16
3:1	7:28, 28:7	4:16, 48:12	20:80, 80:20	12:48, 144:36

\*Republicans:Democrats, Democrats:Republicans for numbers in Club I (or over 35 years of age) and Club II (or under 35 years of age), respectively.

scales as a measure of the tendency to overestimate joint probabilities in relationship to the component event, sampling a given party member. This dependent measure was subjected to an ANOVA for each sample.

The ANOVA for the nonuniversity sample produced a main effect only for political party [ $F(1,11) = 4.99, p < .05, \bar{X}(\text{republican}) = 3.83, \text{ and } \bar{X}(\text{democrat}) = 1.95$ ]. A marginal problem-type  $\times$  party interaction [ $F(1,11) = 3.86, .05 < p < .10$ ] showed a difference between parties for the two-stage condition, but not for the complete-random condition. Because problem frequencies favor no party, results have no apparent explanation save for preexperimental biases. A test of significance for the grand mean produced marginal significance [ $F(1,11) = 4.67, .05 < p < .10, \bar{X} = 2.89$ ].

The university sample produced significance only for problem type [ $F(1,19) = 4.37, p < .05, \bar{X}(\text{complete random}) = .64 \text{ and } \bar{X}(\text{two-stage}) = 4.31$ ] and for club-age group size [ $F(1,19) = 4.86, \bar{X}(\text{small size}) = 2.07 \text{ and } \bar{X}(\text{large size}) = 2.92$ ]. The latter result may reflect the possibility that increased group size accentuated computational errors. A test of significance of the grand mean produced marginal significance [ $F(1,19) = 3.87, .05 < p < .10, \bar{X} = 2.47$ ]. The significant problem-type effect warranted separate tests for each problem type, with significance for the two-stage condition [ $F(1,19) = 4.41, p < .05$ ] and nonsignificance for the complete-random condition. (It should be noted that a similar trend obtained for the nonuniversity sample.)

The tendency to overestimate group-party probabilities in relation to party probabilities warranted detailed analyses. A tabulation of conjunction fallacies [ $p(\text{group} \cap \text{party}) > p(\text{party})$ ] revealed 40 for the university sample and 23 for the nonuniversity sample over the 512 subject-party-problem combinations. Furthermore, the two-stage condition produced 33 and 13 such fallacies for the university and nonuniversity samples, respectively, which comprised .18 of the estimates involved.

In order to assess the possibility that confusion of conditional probability with joint probability contributed to conjunction fallacies, the average value of  $p(\text{democrat} \cap \text{group } i) + p(\text{republican} \cap \text{group } i)$  for each *i* was computed for each of the 26 university subject-problems and 21 nonuniversity subject-problems involved in such fallacies. The mean values [ $\bar{X}(\text{university}) = .95 \text{ and } \bar{X}(\text{nonuniversity}) = .85$ ] suggested such a possibility. In

Table 2  
Relationship Between Joint Probability Estimates and Component  
Probability Estimates for Unequal Group-Size Condition

$p(\text{democrat})$ vs. $p(\text{republican})$	$p(\text{democrat and group 1}) + p(\text{democrat and group 2})$ vs. $p(\text{republican and group 1}) + p(\text{republican and group 2})$					
	$p(\text{dem.}) < p(\text{repub.})$		$p(\text{dem.}) < p(\text{repub.})$		$p(\text{dem.}) > p(\text{repub.})$	
$p(\text{dem.}) < p(\text{repub.})$	0*	0†	0	1	0	1
	1‡	0§	0	0	0	0
$p(\text{dem.}) = p(\text{repub.})$	0	4	0	1	1	8
	0	0	0	1	1	2
$p(\text{dem.}) > p(\text{repub.})$	2	3	0	2	23	6
	0	4	0	3	13	9

\*Complete-random, university sample frequencies. †Two-stage, university sample frequencies. ‡Complete-random, nonuniversity sample frequencies. §Two-stage, nonuniversity sample frequencies.

.59 of the university cases and .50 of the nonuniversity cases, values were between .95 and 1.05. Since  $p(\text{democrat}/\text{group } i) + p(\text{republican}/\text{group } i) = 1.00$ , results are compatible with confusion of joint and conditional probabilities.

This confusion suggests that subjects make joint probability estimates independently of estimates of probabilities of component events. In this case,  $p(\text{party} \cap \text{group } 1) + p(\text{party} \cap \text{group } 2)$  would be independent of the corresponding  $p(\text{party})$  estimate. The condition involving disparate group sizes provides a range of  $p(\text{party})$  estimates. A tabulation of instances in this condition in which  $p(\text{party} \cap \text{group } 1) + p(\text{party} \cap \text{group } 2) - p(\text{party}) \neq 0$  for at least one party in a subject-problem produced 52 and 34 such outcomes in the university and nonuniversity groups, respectively. Each of the two kinds of estimates for both parties can produce three possibilities:  $p(\text{democrat}) < p(\text{republican})$ ,  $p(\text{republican}) = p(\text{democrat})$ , and  $p(\text{democrat}) > p(\text{republican})$ . The possibilities derived from each source should be independent of each other if the two estimates are made independently.

Table 2 presents contingency tables for the results for each sample and each problem type. The complete-random condition produced a strong agreement between the two estimates, with 36 of 40 possible correct agreements overall that  $p(\text{democrat}) > p(\text{republican})$ . In contrast, the two-stage condition produced 28 disagreements for the 45 paired estimates. Since both groups exhibited similar patterns, data were combined with a test for independence, producing  $\chi^2(2) = .143$ ,  $p > .90$ . The result supports the hypothesis that subjects make the estimates independently of one another.

Questions preceding the graphic estimates provide indications of the nature of the difficulties in probabilistic inference. Of the 64 two-stage subject-problems involv-

ing unequal club sizes, 40 subjects answered incorrectly that the probability of drawing a republican was equal to that party's overall relative frequency taken over both clubs. This answer indicates a failure to see the contingency between club sampling and subsequent party-member sampling, a failure that may be associated with independence of the two estimates demonstrated above. Only 16 errors of other types occurred on preliminary questions.

## DISCUSSION

Results suggest that conjunction fallacies result from information processing difficulties in attempts to reason with relative frequencies. These difficulties were common for individuals familiar with the rules of probability. The fact that the two-stage condition produced substantially more reasoning errors than did the complete-random condition indicates that difficulties emanated from information processing requirements of the task rather than from judgmental heuristics associated with population characteristics. The two-stage task required understanding of the contingency between club sampling and subsequent selection of its members. Data from preliminary questions indicated a widespread failure to make this linkage. Furthermore, the estimates of joint probabilities suggested a strong tendency to confuse joint probabilities with conditional probabilities. Possibly, the information processing requirements of the task are sufficient to make it difficult to maintain the contingent nature of club sampling in working memory. Hence, subjects may take sampling from a specified club as given. The conjunction fallacy may then appear as a by-product of this error.

## REFERENCES

- NISBETT, R. E., KRANTZ, D. H., JEPSON, C., & KUNDA, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, *90*, 339-363.
- TVERSKY, A., & KAHNEMAN, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review*, *90*, 293-315.

(Manuscript received for publication November 22, 1985.)