A pattern recognition account of decision making

DOMINIC W. MASSARO
University of California , Santa Cruz, California

In the domain of pattern recognition, experiments have shown that perceivers integrate multiple sources of information in an optimal manner. In contrast, other research has been interpreted to mean that decision making is nonoptimal. As an example, Tversky and Kahneman (1983) have shown that subjects commit a conjunction fallacy because they judge it more likely that a fictitious person named Linda is a bank teller and a feminist than just a bank teller. This judgment supposedly violates probability theory, because the probability of two events can never be greater than the probability of either event alone. The present research tests the hypothesis that subjects interpret this judgment task as a pattern recognition task. If this hypothesis is correct, subjects' judgments should be described accurately by the fuzzy logical model of perception (FLMP)—a successful model of pattern recognition. In the first experiment, the Linda task was extended to an expanded factorial design with five vocations and five avocations. The probability ratings were described well by the FLMP and described poorly by a simple probability model. The second experiment included (1) two fictitious people, Linda and Joan, as response alternatives and (2) both ratings and categorization judgments. Although the ratings were accurately described by both the FLMP and an averaging of the sources of information, the categorization judgments were described better by the FLMP. These results reveal important similarities in recognizing patterns and in decision making. Given that the FLMP is an optimal method for combining multiple sources of information, the probability judgments appear to be optimal in the same manner as pattern-recognition judgments.

The present research tests a viable model of pattern recognition as a description of cognitive probability judgments. The model, called a fuzzy logical model of perception (FLMP), formalizes developments in fuzzy logic (Zadeh, 1965), pattern recognition (Selfridge, 1959), and choice theory (Luce, 1959) to provide a systematic account of perceptual judgments. The FLMP has been tested extensively in a wide variety of domains and provides an informative account of the fundamental processes involved in situations in which there are multiple sources of information (Massaro, 1987, 1992; Massaro & Friedman, 1990). Perceivers are able to evaluate multiple sources of information independently of one another, integrate these sources with respect to alternative prototypes in memory, and make a decision on the basis of the relative goodness of match among the viable alternatives. The same operations have been observed in speech perception, written letter and word recognition, object recognition, categorization, memory, and sentence interpretation (Massaro, 1992). The present paper extends the scientific framework, methodological technique, and theoretical approach to judgments about the probability of events. The research tests the hypothesis that subjects often interpret judgment tasks as a pattern recognition. If this hypothesis is correct, the FLMP should describe subjects' judgments well. To begin, a probability judgment task is described.

Probability Judgments about Events

In a provocative series of experiments, Tversky and Kahneman (1983) evaluated whether judgments about the conjunction of events agree with the prescription given by probability theory. As an example, subjects are given a description of a (now familiar) hypothetical person named Linda.

> Linda is 31 years old, single, outspoken, and very bright. She had a double major in philosophy and music. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Subjects are then asked to indicate how probable is the case that (1) Linda is a bank teller, (2) Linda is a feminist, and (3) Linda is a bank teller and a feminist. A majority of subjects claim that the third case is more likely than the first (Stolarz-Fantino & Fantino, 1990; Tversky & Kahneman, 1983; Wolford, Taylor, & Beck, 1990;

Wyer, 1976). This result can be interpreted as a violation of probability theory and set theory because the probability of two events can never be greater than the probability of either event alone. Tversky and Kahneman (1983) call this type of behavior a conjunction fallacy. For reasons developed in the present research, this result will be called a *conjunction effect*.

Decision Making as Pattern Recognition

Tversky and Kahneman (1983) have claimed that subjects do not use probabilities but instead a representativeness heuristic, which usually involves some type of similarity matching. According to this explanation, subjects do *not* follow the instructions of the experimenter in this task. Subjects are instructed to indicate the probability of an alternative given the description of Linda, but they putatively judge the alternative in terms of the degree to which Linda resembles the typical member of that class (bank teller) or (bank teller and feminist). The present research follows in the spirit of Tversky and Kahneman in the sense that subjects are not doing what they are instructed to do. In fact, previous research appears to show that instructions have a relatively small effect on performance. For example, in the Linda task, Tversky and Kahneman found identical results with the Linda problem when subjects were asked to make similarity judgments and when they were instructed to make probability judgments. Macdonald and Gilhooly (1990) found no change in the size of the conjunction effect when believability was substituted for probability in the instructions.

The present hypothesis is that subjects are carrying out pattern recognition (Massaro, 1985) or pattern-based reasoning (Lopes & Oden, 1991). The present inquiry is more optimistic than that of Tversky and Kahneman, who state, "the judged probability (or representativeness) of a conjunction cannot be computed as a function (e.g., product, sum, minimum, weighted average) of the scale values of its constituents" (1983, p. 305). Although this conjunction effect has been interpreted primarily as a violation of probability theory, it can also be interpreted as consistent with the FLMP (Massaro, 1985; 1987, pp. 272–277; 1988). These two interpretations create somewhat of a paradox because the currency and processes of the FLMP are isomorphic to a normative probability account (Massaro & Friedman, 1990).

Central to the FLMP is that two sources of information can be more informative than just one. Consider a situation in speech perception that is analogous to the Linda task. Perceivers listen to an audiotape of a male speaking in a high-pitched voice (falsetto) voice. They are then asked to rate the probability of (1) a male speaking or (2) a male speaking with a falsetto voice. They rate the (2) as more probable than (1). This outcome is not unexpected, but did the subject make a conjunction fallacy? If simple probability theory is taken as the normative model, the conjunction of two properties (male and falsetto voice) should not be judged as more probable than either one alone. In terms of the FLMP, how-

ever, it is assumed that subjects are answering the question of the probability of the message on the audiotape, given (1) a male or (2) a male speaking with a falsetto voice. In this case, it makes good sense to judge the second alternative as more probable than the first. The difference between the FLMP description and the simple probability description centers around the assumption of what subjects are judging (see Bar-Hillel, 1991; Wolford, 1991). According to the FLMP, individuals are likely to judge $P(L|F)$, $P(L|B)$, and $P(L|F$ and $B)$, where L, F, and B correspond to Linda, feminist, and bank teller, respectively. According to the simple probability theory, individuals should be judging $P(F|L)$, $P(B|L)$, and $P(B$ and $F|L)$.

One test between these two possible conclusions is to provide a quantitative test between these two models. Of course, given the previous research, we can expect the simple probability model to give a poor description of the Linda task. However, quantitative tests of this model have not been provided, nor has its prediction been contrasted with other models. If the FLMP gives a good description of the results, we have some evidence for pattern recognition behavior in the Linda task. It would be valuable to contrast this model with a quantitative model of the representativeness heuristic, but one is not available. Tests of other quantitative models are possible, however, and these serve as alternatives to the optimal model.

EXPERIMENT 1
Test of Simple Probability Model and FLMP

Subjects were given the description of Linda and were asked to judge "how likely it was that various statements about Linda applied." Five avocations and five vocations were tested in an expanded factorial design (Massaro & Cohen, 1990). Each possible vocation was paired with each possible avocation, giving 25 unique pairings. In addition, the 10 individual characteristics were presented alone. This design gives $5 \times 5 + 5 + 5 = 35$ experimental questions.

The five vocations ordered from most likely to least likely descriptions of Linda were:

1. social worker for the county
2. public elementary school teacher
3. sales clerk at a book store
4. teller at Bank of America
5. IBM executive

The five avocations were designated as follows:

1. active in a national feminist organization
2. plays violin in an amateur chamber group
3. avid science fiction reader
4. crew member on a local sailing team
5. active in the right-to-life movement

These statements were chosen to span each continuum from being a very probable description of Linda to being very improbable.

Method

Subjects. Thirty subjects were recruited from psychology classes and satisfied part of a course requirement for participating for about 1 h in the experiment.

Procedure. Two subjects were tested at a time, in separate rooms. Each subject was seated at an IBM PC computer. Instructions were presented on the video display monitor, followed by the description of Linda. This description remained in view throughout the experiment. The subjects were instructed to rate the probability of each statement's being true of Linda on a scale from 1 (*completely improbable*) to 9 (*completely probable*). The subjects had as long as they wanted to make each decision and could change the decision before the next trial if they wished. The subjects entered their response by hitting one of the keys 1–9 and then the return key. The next trial was presented after an intertrial interval of 1 sec. The 35 statements were randomized within a block of 35 trials. Two successive trial blocks were presented. The ratings for each subject were linearly translated into values between 0 and 1.

Results

The mean ratings across the 30 subjects are plotted as points as a function of the vocation (left plot), avocation (right plot), and combined vocation–avocation conditions (middle plot) in Figure 1. As can be seen in the figure, the continuum of vocations was more influential than the continuum of avocations. Even so, both variables had statistically significant effects in both the single and the combined conditions (all p values $< .001$). The interaction between the vocation and avocation factors was not quite statistically significant [$F(16,464) = 1.637, p = .056$].

The three circled points in Figure 1 depict the conjunction effect. Bank teller (BT) and feminist (FE) received a higher average rating than did just bank teller.

We now formalize quantitative models to test against these results.

Simple Probability Model

Given the expanded factorial design, all the results can also be used to test a simple probability model (SPM). In all the model tests, it is assumed that the observed rating judgments are on a linear scale of probability. The rating that Linda has an avocation can be defined as simply R(avocation). Similarly, R(vocation) would be the rating given a particular vocation statement. According to this model, the rating given the statement that Linda has an avocation and a vocation is predicted to be equal to the multiplicative combination of these two ratings.

$$R(\text{vocation–avocation} = R(\text{vocation}) \times R(\text{avocation})$$

(1)

In the test of this model, a different scale value is assumed for each level of vocation and for each level of avocation. Given 5 unique avocations and 5 unique vocations, 10 free parameters are necessary to predict the 35 data points.

It can be justly argued that this model is unrealistic because no correlation is assumed between the avocation and vocation in Equation 1. The conjunction effect appears to occur when there is a negative correlation between the two attributes (a feminist is unlikely to be a bank teller). Assuming some correlation between each avocation–vocation pair, however, precludes any test of the simple probability model. Twenty-five free parame-
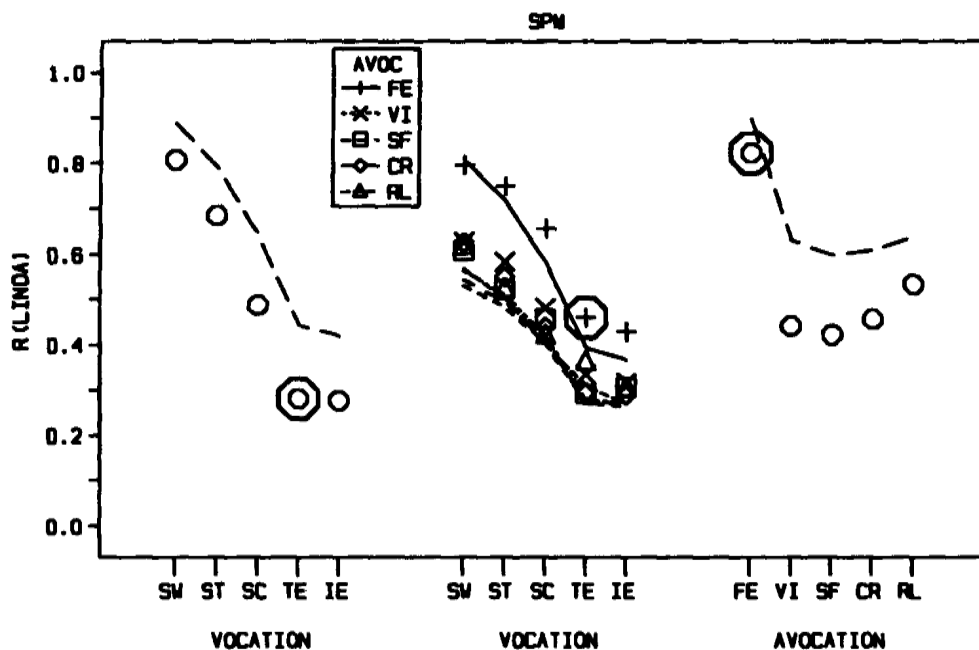

Figure 1. Observed (points) and predicted (lines) rating responses in Experiment 1, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The three circled points from left to right, TE, TE–FE, and FE, replicate the conjunction effect. The predictions are for the simple probability model (SPM).

ters would be needed for the correlation, as well as the 10 parameters required for the scale values for the 5 avocations and 5 vocations. Given the independence assumption in Equation 1 and the conjunction effect, we would expect this probability model to give a poor description of the results. Even so, it is important to test this model because it provides a formal alternative and because independence between the vocation and avocation is assumed in alternative models that are tested.

This SPM and the other models was fit to the results of each subject and to the average results of the 30 subjects, using the program STEPIT (Chandler, 1969). The model is represented to the program in terms of a set of prediction equations and a set of unknown parameters. By iteratively adjusting the parameters of the model, the program minimizes the squared deviations between the observed and predicted points. The outcome of the program STEPIT is a set of parameter values which, when put into the model, come closest to predicting the observed results. Thus, STEPIT maximizes the accuracy of the description of each model. The goodness-of-fit of the model is given by the root mean square deviation (RMSD)—the square root of the average squared deviation between the predicted and observed values.

The lines in Figure 1 give the average predictions of the SPM. The model gave an average RMSD of .1205 in the fits to the individual subjects and an RMSD of .0881 to the average results.

Fuzzy Logical Model of Perception

Testing the FLMP requires knowledge of the candidate set of alternatives that subjects use in the task. Although only Linda is mentioned in the task, it is assumed that subjects generate a contrasting alternative that is the antithesis of Linda. We call this hypothetical person *not(Linda)*. Applying the FLMP to the rating judgments, both the avocation and the vocation are assumed to provide continuous and independent evidence for the Linda and not(Linda). We assume that each subject generates a prototype for Linda and one for not(Linda). Given that we are manipulating the avocation and vocation in the test statements, it is sufficient to assume that these two prototypes differ on these two dimensions. A reasonable prototype for Linda would be

Linda : Intellectual avocation & Altruistic vocation

The prototype for not(Linda) is assumed to be

(not)Linda : (not)Intellectual avocation &
(not)Altruistic vocation.

It should be noted that these descriptors only approximate the subject's actual representation of both Linda and (not)Linda. It is only necessary that the two prototypes differ from one another on the two relevant dimensions.

The FLMP assumes three operations. The evaluation process determines the degree to which each dimension or source of information in the test statement matches each prototype. For each source of information, a fuzzy

truth value between 0 and 1 is assigned for each prototype. This fuzzy truth value represents the degree to which the source of information matches the corresponding entry in the prototype. In fuzzy logic with two alternative prototypes, 0 is no support, .5 is ambiguous support, 1 is complete support, and other values are intermediate degrees of support. If aL_i represents the degree to which the avocation A_i from test statement i supports the alternative Linda, then the outcome of prototype matching for Linda would be

$$\text{Linda} : aL_i,$$

where the subscript i indexes the five levels of the avocation dimension.

In fuzzy logic, the negation of a proposition is one minus its truth value. In this case, it can be assumed that the support from one source of information for not(Linda) is one minus its support for Linda. Thus, the outcome of prototype matching for not(Linda) would be

$$(\text{not})\text{Linda} : (1-aL_i).$$

The integration operation has no material consequences when there is just a single source of information. The decision operation determines the relative merit of Linda and not(Linda) alternatives, leading to the prediction that

$$R(\text{Linda}|A_i) = \frac{aL_i}{\Sigma}, \qquad (2)$$

where $R(\text{Linda}|A_i)$ is the rating of the statement being true of Linda, and Σ is equal to the sum of the merit of the Linda and not(Linda) alternatives. Given that the denominator of Equation 2 is one, the predicted rating is simply aL_i:

$$R(\text{Linda}|A_i) = \frac{aL_i}{aL_i + (1 - aL_i)} = aL_i. \qquad (3)$$

Using a similar logic, the predicted rating for the single vocation conditions can be shown to be vL_j, where the subscript j indexes the vocation dimension.

The same evaluation operation occurs for the test statements containing both sources of information (an avocation and vocation). An important assumption is that the two sources are evaluated independently of one another. Given a prototype's independent specifications for the two sources, the value of one source cannot change the value of the other source. The integration of the two sources is determined by the product of their feature values from the evaluation stage. The outcome of the integration of the two sources of support for Linda would be

$$\text{Linda} : aL_i \, vL_j,$$

where the subscripts i and j index the levels of the avocations and vocations, respectively. Similarly, the outcome of prototype matching for "Not(Linda)" would be

$$(\text{not})\text{Linda} : (1 - aL_i)(1 - vL_j).$$

The decision operation would determine their relative merit leading to the prediction that

$$R(\text{Linda}|A_i \ V_j) = \frac{aL_i \ vL_j}{\Sigma},$$
(4)

where Σ is equal to the sum of the merit of the Linda and not(Linda) alternatives. Ten free parameters, corresponding to the 5 avocations and 5 vocations, are necessary to predict the results.

These predictions of the FLMP were tested against both the results from individual subjects and the average results. The lines in Figure 2 give the average predictions. The RMSD was .0954 for the average fit of the individual subjects and .0465 for the fit of the average results. The latter RMSD is about twice as good as the fit of the SPM. An analysis of variance of the individual RMSDs revealed that the FLMP gave a better fit than SMP [$F(1,29) = 25.596, p < .001$].

Figure 3 gives the parameter values from the FLMP. The scale on the right gives the scale values corresponding to area of the circle. The parameter values for the vocation changed in an orderly manner across the five levels. For the avocation, however, only feminist (FE) gave strong support for Linda and the other four levels were roughly neutral. The avocations were modified in the next experiment to give a better spread across this continuum.

Although the FLMP gave a reasonable description of the results, it should be clear to the reader that the present interpretation of the conjunction effect rests on an important assumption about the prototype alternatives that the subject generates in the task. We assumed that the subject generates a prototype for not(Linda), and the ev-

idence supports this assumption. A better state of affairs would be to give the subject explicit alternatives so that choice alternatives would not have to be generated by the subject or assumed by the theorist. In this case, a direct test of the FLMP would be possible with the assumption that the prototype alternatives being used were equivalent to those given as response alternatives in the task.

EXPERIMENT 2
Modification of the Linda Task

A simple extension of the Linda problem makes the choice alternatives explicit. In this task, subjects are given descriptions of two hypothetical people (Linda and Joan) and have to rate or categorize which of the two people is more likely to have a vocation, an avocation, or both a vocation and an avocation. According to the present framework, this task better represents a pattern-recognition scenario in which the choice alternatives are explicitly given or known in advance. For example, deciding between Linda and Joan is exactly analogous to choosing between /ba/ and /da/ in a speech recognition task, given auditory and visual sources of information (Massaro & Cohen, 1990). To the extent that subjects behave similarly in the task with two explicit choice alternatives and the task with just one, the conjunction effect in the standard task can be described as the engagement of pattern-recognition processes.

Method
In addition to the Linda description, subjects were given a description of Joan.

Joan is 29 years old, married, athletic, and intelligent. She majored in economics, and graduated with honors. She was a writer for the conservative campus newspaper, and participated in intramural sports.
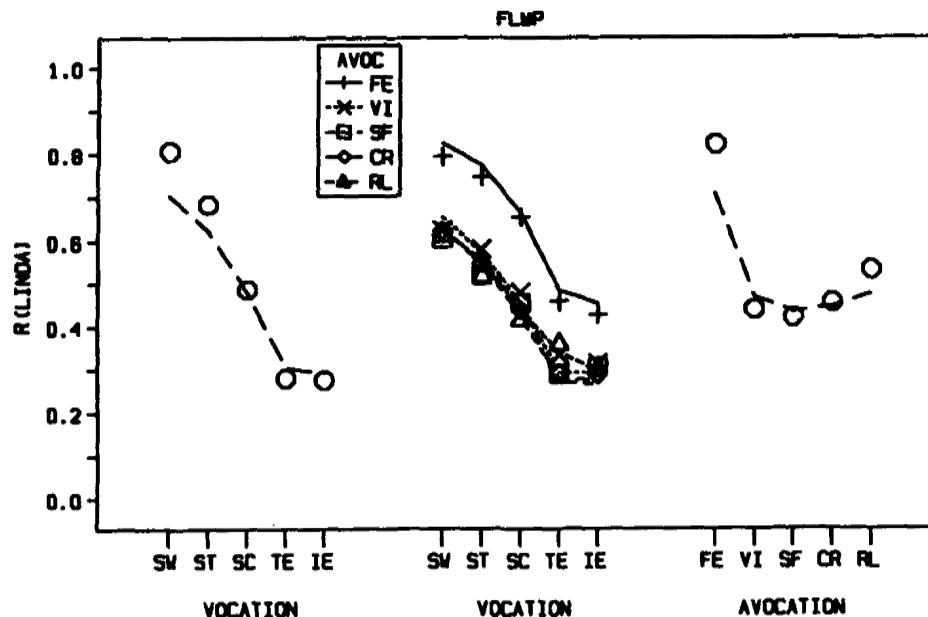

Figure 2. Observed (points) and predicted (lines) rating responses in Experiment 1, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The predictions are for the fuzzy logical model (FLMP).
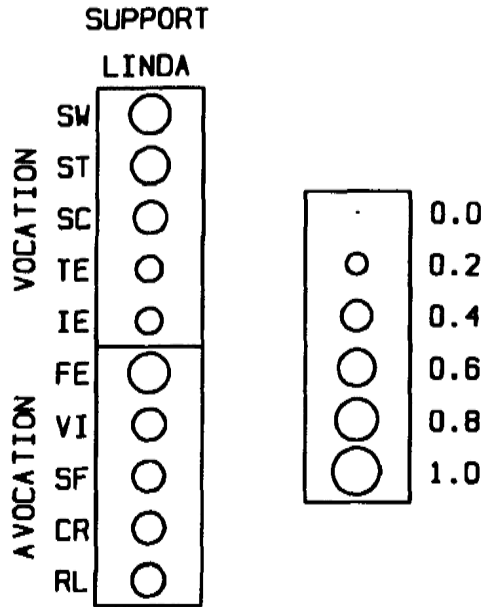
SUPPORT

LINDA

Figure 3. Plot of the fuzzy logical model's parameter values, indicating the degree of support for Linda for the five vocations and five avocations. The scale on the right shows the relationship between area and scale value.

As in the first Linda experiment, five avocations and five vocations were tested in an expanded factorial design. The levels were the same as in the first experiment, except for a few modifications to improve the spread across each continuum. The five vocations ordered from most like Linda to most like Joan were:

1. social worker for the county
2. public elementary school teacher

3. sales clerk at a book store
4. loan manager at Bank of America
5. IBM executive in charge of computer programs

The five avocations were designated thus:

1. active in a national feminist organization
2. plays violin in an amateur chamber group
3. avid science fiction reader
4. competitive Frisbee player
5. active in the Republican party

Subjects. Two groups of 11 subjects each were recruited from psychology classes and satisfied part of a course requirement for participating for about 1 h in the experiment.

Procedure. Two subjects were tested at a time in separate rooms. Each subject was seated at an IBM PC computer. Instructions were presented on the video display monitor, followed by the descriptions of Linda and Joan. The descriptions remained in view throughout the experiment. The subjects were instructed to indicate whether Linda or Joan was more likely to have some avocation, vocation, or avocation and vocation. On each trial, an avocation, a vocation, or an avocation–vocation pair was presented. In the rating group, subjects were instructed "to type a number between 1 and 9 indicating 'whether Linda or Joan is more likely.'" The number 1 would correspond to *definitely Linda* and 9 as *definitely Joan*. The number 5 would correspond to *Linda and Joan equally likely*, and so on for the intermediate numbers. In the categorization group, subjects were instructed "to type a letter 'L' or 'J' corresponding to whether you think Linda or Joan is more likely." These subjects entered their decisions by hitting one of two keys labeled "L" and "J," corresponding to Linda and Joan. The subjects in both groups had as long as they wanted to make each decision and could change their decision before the next trial if they wished. The subjects entered their response by hitting the return key and the next trial was presented after an intertrial interval of 1 sec.

Given an expanded factorial design with two independent variables with five levels of each variable, there were 35 unique trials. The 35 conditions were sampled randomly with replacement in
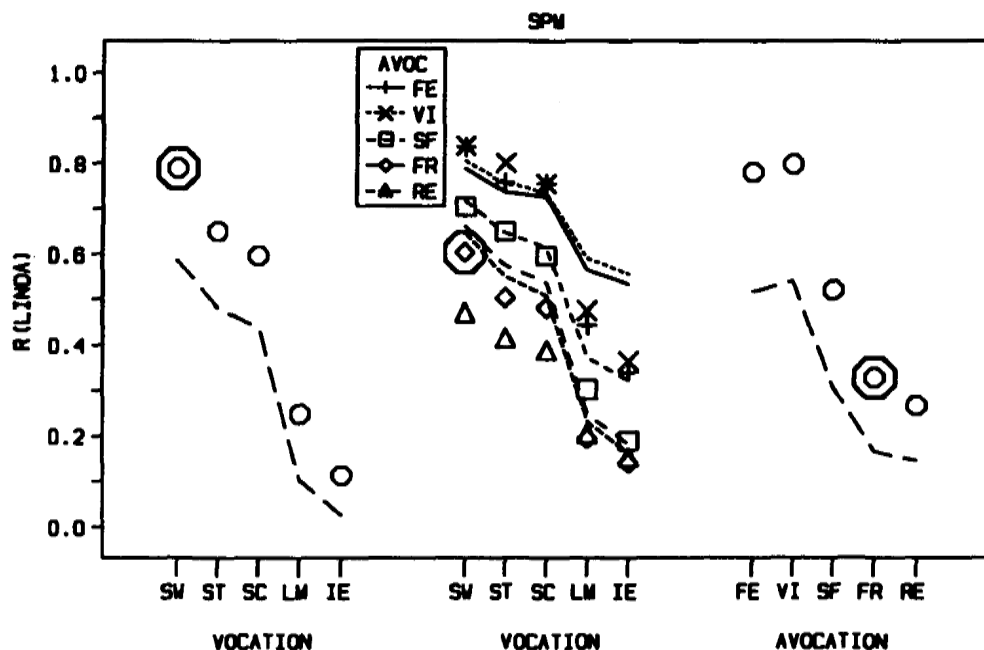

Figure 4. Observed (points) and predicted (lines) rating responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The three circled points from left to right, SW, SW–FR, and FR, replicate the conjunction effect. The predictions are for the simple probability model (SPM).

each block of 35 trials. A given test session had six blocks of 35 trials in each block, and subjects were tested in two sessions. The dependent variables were the average ratings and the proportion of "Linda" judgments for each subject at each of the 35 experimental conditions. Research in other domains indicated that the 12 observations per condition should be sufficient to obtain reliable data for individual-subject analyses and model tests (Massaro & Cohen, 1993).

Results

The average ratings of "Linda" as a function of avocation and vocation are shown as the points in Figure 4. As can be seen in the left and center plots, the rating of "Linda" significantly decreased across the vocation continuum, for both the unidimensional [$F(4,40) = 39.48$, $p < .001$] and the bidimensional [$F(4,40) = 51.14, p < .001$] conditions. Similarly, the right and center plots show that the rating of "Linda" significantly decreased across the avocation continuum, for both the unidimensional [$F(4,40) = 19.99, p < .001$] and the bidimensional [$F(4,40) = 20.39, p < .001$] conditions. There was also a significant avocation $\times$ vocation interaction [$F(16,160) = 4.15, p < .001$] in the bidimensional condition, because each stimulus dimension had a larger effect to the extent that the other was ambiguous.

The mean observed proportion of "Linda" identifications averaged across subjects is shown as the points in Figure 5. As can be seen in the left and center plots, the proportion of "Linda" responses significantly decreased across the avocation continuum, for both the unidimensional [$F(4,40) = 34.70, p < .001$] and the bidimensional [$F(4,40) = 21.31, p < .001$] conditions. Similarly, the right and center plots show that the proportion of "Linda"

responses significantly decreased across the avocation continuum, for both the unidimensional [$F(4,40) = 27.33$, $p < .001$] and the bidimensional [$F(4,40) = 17.54, p < .001$] conditions. There was also a significant avocation $\times$ vocation interaction [$F(16,160) = 3.53, p < .001$] in the bidimensional condition, because each stimulus dimension had a larger effect to the extent that the other was ambiguous.

A necessary first question is whether the conjunction effect was replicated. The three circled points in Figures 4 and 5, respectively, show a conjunction effect. Subjects were more likely to respond that Linda was a social worker (SW) and a Frisbee player (FR) than simply a Frisbee player. Of course, subjects were more likely to respond that Linda was only a social worker than a social worker and a Frisbee player.

Simple probability model. The fit of the SPM was implemented in the same manner as in the Linda task (see Equation 1). The lines in Figures 4 and 5 give the average predictions of the SPM for the ratings and categorization tasks, respectively. The model gives a very poor description of the identifications of both rating and categorization judgments, with average RMSDs of .142 and .256 across the individual subject fits.

Fuzzy logical model of perception. Applying the FLMP to the probability judgments, both the avocation and the vocation are assumed to provide continuous and independent evidence for the alternatives Linda and Joan. Defining the avocation and vocation as the important sources of information, the prototype for Linda would be the same as in the first Linda task:

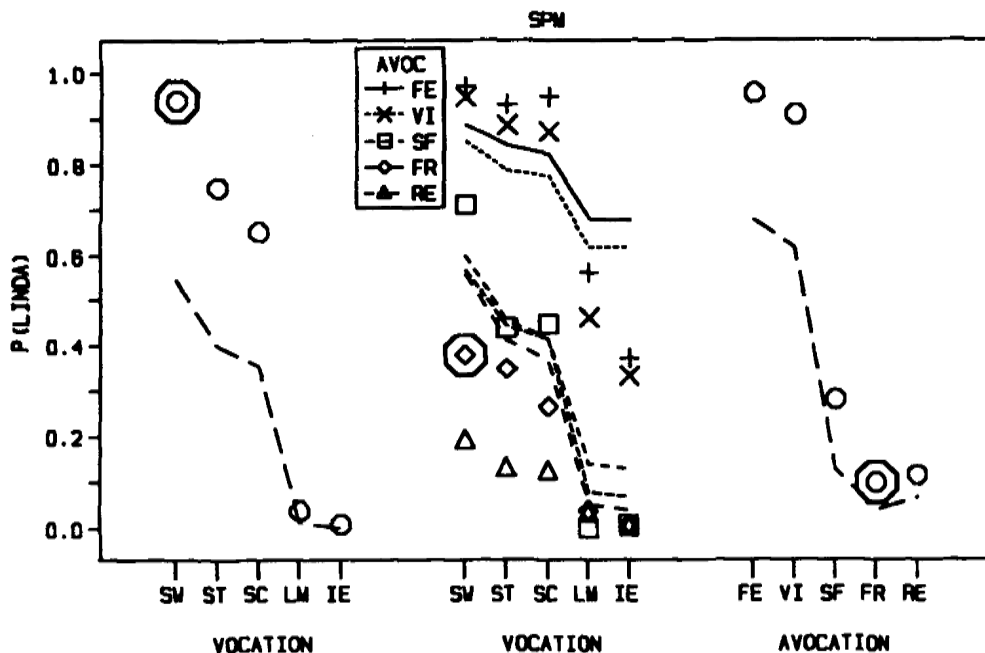Linda : Intellectual avocation & Altruistic vocation.

Figure 5. Observed (points) and predicted (lines) categorization responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The three circled points from left to right, SW, SW–FR, and FR, replicate the conjunction effect. The predictions are for the simple probability model (SPM).

The prototype for Joan would be defined in an analogous fashion:

Joan : Yuppie avocation & Ambitious vocation.

At the evaluation stage, the avocation supports each alternative to some degree and analogously for the vocation. These degrees of support are integrated following the multiplicative rule given by the FLMP. With just two choice alternatives, the predictions do not change when it is also assumed that the degree of support for one alternative is the additive complement of the degree of support for the other (Massaro, 1989). In this case, the prototype for Joan can be defined equivalently to the prototype for not(Linda) in the first experiment. Thus, the avocation support for Joan is given by $aJ_i = 1 - aL_i$, and the value $vJ_i = 1 - vL_i$ gives the vocation support for Joan. The predictions are given by Equations 2–4 and require 10 parameters.

The FLMP was fit to the individual results of each of the 11 subjects in both the rating and the categorization tasks. The lines in Figures 6 and 7 give the average predictions of the FLMP. Figures 8 and 9 give the average best-fitting parameters of the FLMP. The parameter values change in a systematic fashion across the five levels of the avocation and vocation dimensions. The model provides a good description, with an average individual RMSD of .048 for the rating judgments and .040 for the categorization judgments across the individual subject fits. Thus the fit of the FLMP is about three to six times better than the fit of the SPM. Analyses of variance were carried out on the RMSD values given by the fits of the

SPM and FLMP. The FLMP gave a significantly better fit than the SPM for both the rating [$F(1,10) = 104.35$, $p < .001$] and the categorization judgments [$F(1,10) = 105.23, p < .001$].

Weighted averaging model. In pattern-recognition studies, the FLMP is usually compared to several other models. An important contender is a weighted averaging model (WAM). Fantino, Kulik, and Stolarz-Fantino (1993) found some evidence for an averaging model in the Linda task. They claimed that subjects judge conjunctions by averaging the likelihood of their component parts. The present results and tests allow a stronger quantitative test of the averaging hypothesis. In addition, it allows for a weighted averaging, rather than a simple averaging, of the two components (Anderson, 1981). One WAM can be expressed as

$$R(\text{Linda}|A_iV_j) = (p)aL_i + (1-p)vL_j. \qquad (5)$$

The WAM predicts that the rating of "Linda," given two sources of information, is a simple weighted average of the rating of "Linda," given each of the separate sources. In this case, the weight corresponds to relative influence of the avocation and vocation sources of information.

A more appealing averaging model would be to assume that the weight given a particular avocation (vocation) would vary with its actual scale value. A more informative avocation or vocation should receive more weight. Mathematically, this model can be expressed as

$$R(\text{Linda}|A_i \, V_j) = (p_i)aL_i + (p_j)vL_j. \qquad (6)$$

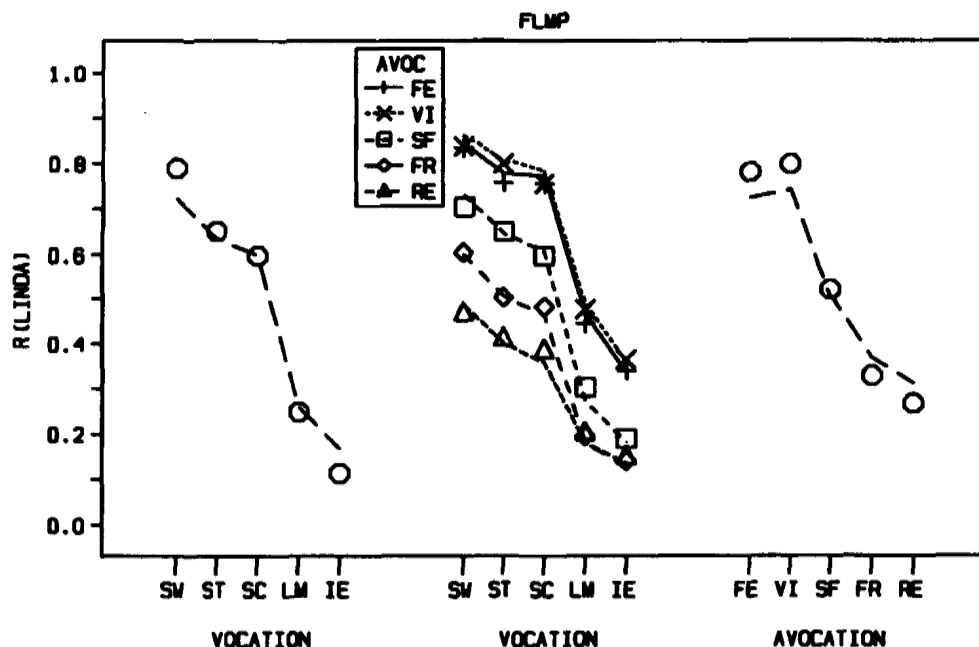In this case, each unique scale value carries a unique

Figure 6. Observed (points) and predicted (lines) rating responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The predictions are for the fuzzy logical model (FLMP).
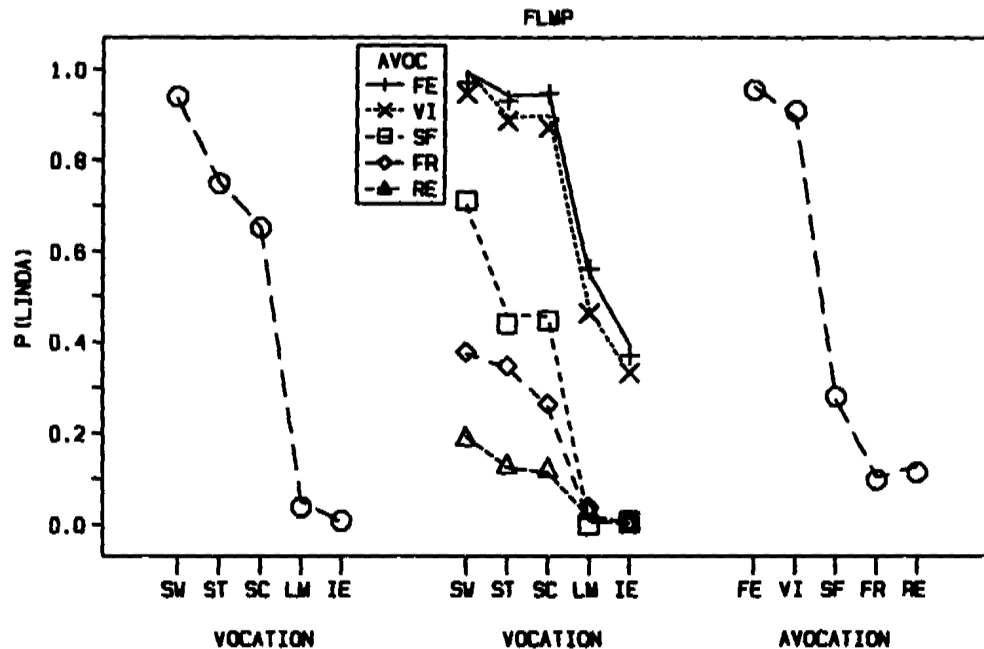
Figure 7. Observed (points) and predicted (lines) categorization responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The predictions are for the fuzzy logical model (FLMP).

weight. The predictions of this model, however, cannot improve on the predictions given by the WAM in Equation 5. The reason is that the weights and scale values change together for each unique avocation or vocation. Thus, the scale value and the weight can be reduced to a single value without changing the predictions, and Equation 6 reduces to

$$R(\text{Linda}|A_iV_j) = aL_i + vL_j. \tag{7}$$

Equation 5 is a more general version of Equation 7, because it allows a weight that is independent of the scale value.

Analogous to the other models, Equation 5 is also used to predict the probability of a "Linda" response in the categorization judgments.

To fit the WAM (given by Equation 5) to the results, each unique level of the avocation requires a unique parameter aL_i, and analogously for vL_j. The modeling of "Linda" responses thus requires 5 avocation parameters plus 5 vocation parameters. The additional p value would be fixed across all conditions, for a total of 11 parameters. Thus, the WAM requires 1 more parameter than the 10 required by the FLMP.

Figures 10 and 11 give the average observed results and the average predicted results of the WAM. As can be seen in the figures, the WAM gave a reasonable description of the rating judgments but a poor description of the categorization results. The average RMSD from the fit of the individual subjects' ratings was .061, only slightly poorer than the fit given by the FLMP [$F(1,10)$ = 1.302, n.s.]. The average RMSD for the categorization judgments was .172, about four times poorer than the fit

of the FLMP. An analysis of variance of the individual RMSDs revealed that the FLMP gave a better fit than the WAM [$F(1,10) = 43.66, p < .001$].

Yates and Carlson (1986) have proposed a modification of the WAM, called a "signed summation" model. In this model, individuals are assumed to categorize events as likely or unlikely, rather than evaluate them on
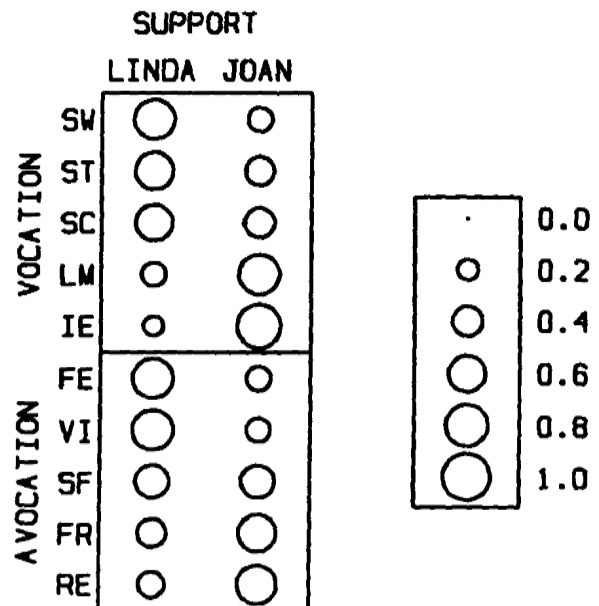

Figure 8. Parameter values of the fuzzy logical model, indicating the degree of support for Linda for the five vocations and five avocations (Experiment 2, rating results).
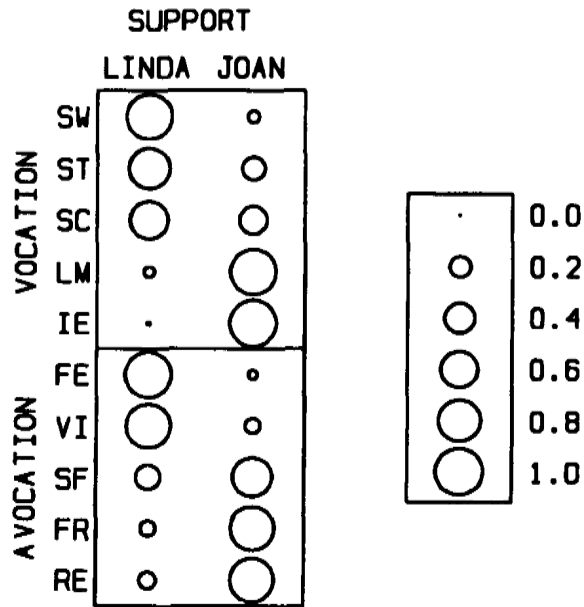
SUPPORT

LINDA JOAN

Figure 9. Parameter values of the fuzzy logical model, indicating the degree of support for Linda for the five vocations and five avocations (Experiment 2, categorization results).

some scale of certainty. This model can predict a conjunction effect, but it can also supposedly predict that the conjunction of two likely events is more likely than either of the events alone. The authors claim that this model is identical to Anderson's weighted averaging model with an initial impression. To test this model

against the present results, 14 free parameters were estimated: the initial impression, the 10 scale values, and 3 weights for the initial impression, avocation, and vocation, respectively. This model gave only a slightly better fit (RMSD = .061 and .170 for the rating and categorizations) than the fit of the WAM. It appears that the model can predict the results only if a different initial impression is assumed for each unique avocation-vocation pair. This assumption would give the model as many free parameters as predicted data points—an unacceptable state of affairs, because this is equivalent to assuming as much as is being predicted.

Discussion

Fine-grained analyses. The present experiment improves on the methodology of previous studies of the Linda task. In previous studies, many subjects are asked just one or a few questions. For example, Fiedler (1988) analyzed seven conjunction problems in terms of the overall proportion of correct answers as a function of the independent variables. No information was given about individual differences. In general, the results are necessarily presented in terms of the proportion of subjects that show a conjunction effect or the group's overall proportion. However, there is necessarily some ambiguity about what a given result of this kind means. When 80% of the subjects show a conjunction effect, we do not know that these subjects will always give the same response with repeated presentations of the same question under the same conditions. Similarly, the 20% who did not give a conjunction effect might have given one had the question been given again at another time.
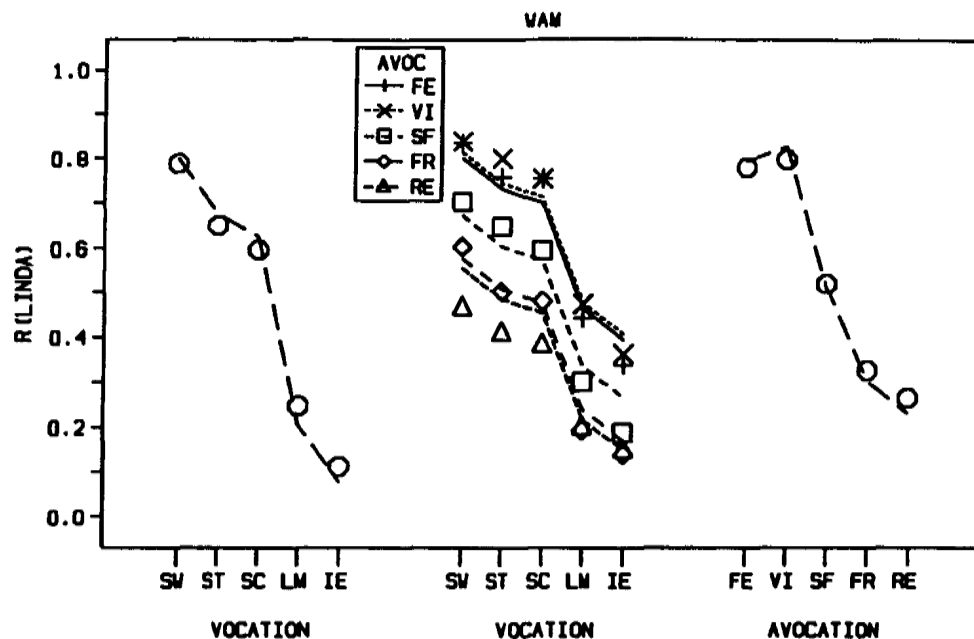

Figure 10. Observed (points) and predicted (lines) rating responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The predictions are for the weighted averaging model (WAM).
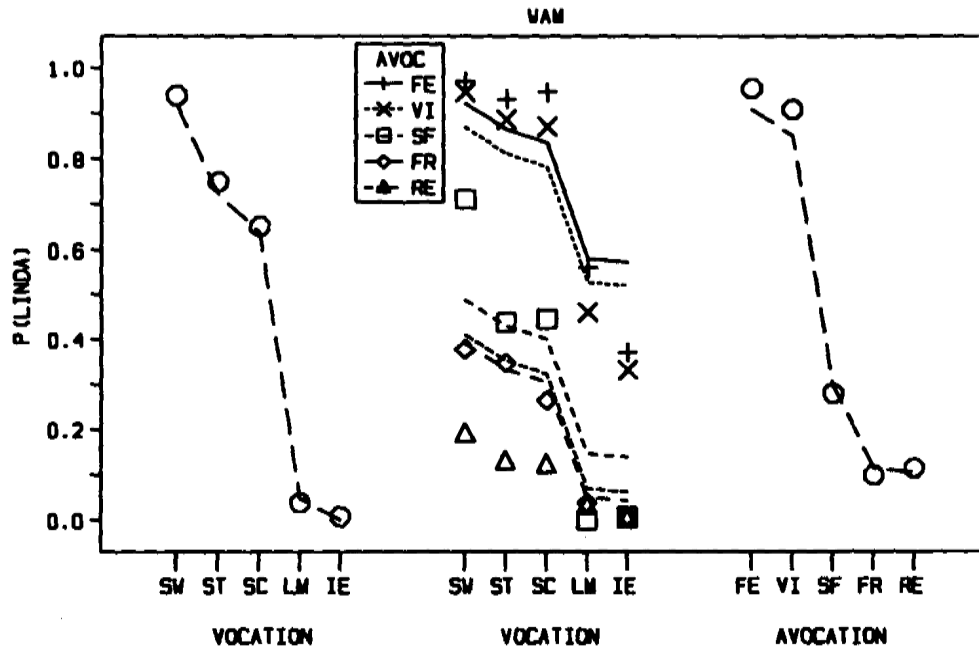
Figure 11. Observed (points) and predicted (lines) categorization responses in Experiment 2, as a function of vocation (left graph), vocation and avocation (center graphs), and avocation (right graph). The predictions are for the weighted averaging model (WAM).

Single-subject analyses are also necessary in tasks of this kind, because not all subjects will order the properties in the same manner. For example, some subjects might decide that it is more probable that Linda plays the violin than that she is an avid science fiction reader. Other subjects might make the opposite decision. Averaging these ratings would dilute the results and would make them inappropriate for tests of mathematical models of integration. When contrasting linear and nonlinear models, averaging results across individuals also tends to favor an additive outcome even though the underlying process might be nonadditive (Massaro & Cohen, 1993).

Averaging versus multiplicative integration. There is a fundamental difference in averaging integration and multiplicative integration within the FLMP (Massaro, 1987, chap. 7). An averaging integration of two sources of information produces an outcome that is less extreme than does either source presented alone. In the FLMP, the operations of integration and decision can lead to a probability or rating judgment of a conjunction that is more true than the judgment of either or both of the single propositions. The present experiments were not completely successful in distinguishing an averaging from a multiplicative integration process. These two forms of integration gave roughly similar descriptions of the rating judgments in both Experiments 1 and 2. However, averaging integration given by the WAM could not accurately describe the categorization judgments. The fit of the WAM was about four times poorer than the fit of the multiplicative integration in the FLMP. It should be noted that it will not be a simple matter to modify some aspect of a WAM to predict categorization judgments.

One might modify the decision stage by allowing some type of nonlinear transformation of the output of integration. However, a scaling transformation on the output of integration cannot override a central feature of averaging integration. Although an averaging integration can explain the conjunction effect in the Linda task, it cannot explain another common finding that two sources of information can be more informative than either source. A loan manager at Bank of America and a competitive Frisbee player is less like Linda (or more like Joan) than either one of these characteristics taken alone.

Optimality of decision making. The present contribution is a normative account of the conjunction effect—that is, subjects are doing exactly what they should be doing when they claim that it is more probable that Linda is a bank teller and a feminist than that Linda is a bank teller. Wolford et al. (1990) arrived at a similar interpretation. They claimed that a Bayesian decision strategy makes sense in some contexts of the conjunction effect paradigm. Specifically, if the subjects interpret the possible outcomes as known, then a Bayesian analysis becomes appropriate. If the possible outcomes are not yet known, probability theory remains appropriate. By varying the context to stress either known or unknown outcomes, the proportion of subjects showing the conjunction effect was modified in the predicted direction. However, over 50% of the subjects still gave a conjunction effect in the unknown context, leading the authors to conclude that subjects may still use the Bayesian model even when it is inappropriate. Wolford et al.'s distinction between known and unknown outcomes does not appear to account for when a conjunction effect will

be found. Within the context of the present perspective, it is claimed that subjects typically follow the FLMP and that a conjunction effect should usually be found.

The present analysis differs from other explanations that have stressed the ambiguity of the scenario established by the Linda problem. Margolis (1987) has argued that the instructions should stress the gambling sense of the term *probability*. In this case, a conjunction fallacy should be avoided. In contrast to his claim, the perspective of the FLMP leads us to expect a conjunction effect even if given his revised wording. Supporting the FLMP prediction, Macdonald and Gilhooly (1990) found little or no effect of formulating the Linda problem in terms of probability rather than belief.

Given the mathematical equivalence between the FLMP and Bayes's theorem (Massaro, 1987, pp. 196-198; 1989), the good FLMP description of the Linda–Joan task demonstrates that decision making can be optimal. The appropriate normative theory for the two-alternative Linda–Joan task is Bayes's theorem—which assumes probability theory. That is, multiplying probabilities is also assumed within Bayes's theorem. When subjects are given a task requiring pattern-based reasoning that parallels pattern recognition, they behave appropriately. When subjects are given a task involving conjunctive events, however, they do not appear to follow the appropriate normative model. Our interpretation, like the original Tversky and Kahneman (1983) interpretation, is that subjects are not doing what they were instructed to do. Macdonald (1986), for example, views performance in the Linda task as uncertain because of ambiguities in the concept of probability, the sample space, and the intention of the experimenter's communication. The present research, however, goes beyond the representativeness heuristic, because the present interpretation claims that subjects are interpreting the task as pattern-based reasoning and performing this type of reasoning optimally. Subjects are carrying out pattern recognition in the standard Linda task even though they are instructed otherwise.

REFERENCES

BAR-HILLEL, M. (1991). Commentary on Wolford, Taylor, and Beck: The conjunction fallacy? *Memory & Cognition, 19*, 412-414.

CHANDLER, J. P. (1969). Subroutine STEPIT—finds local minima of a smooth function of several parameters. *Behavioral Science, 14*, 81-82.

FANTINO, E., KULIK, J., & STOLARZ-FANTINO, S. (1993). *The conjunction fallacy: A case of representiveness or averaging?* Unpublished manuscript.

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

LOPES, L. L., & ODEN, G. C. (1991). The rationality of intelligence. In E. Eells & T. Maruszewski (Eds.), *Probability and rationality: Studies on L. Jonathan Cohen's philosophy of science* (pp. 199-223). Rodopi.

LUCE, R. D. (1959). *Individual choice behavior.* New York: Wiley.

MACDONALD, R. R. (1986). Credible conceptions and implausible probabilities. *British Journal of Mathematical & Statistical Psychology, 39*, 15-27.

MACDONALD, R. R., & GILHOOLY, K. J. (1990). More about Linda *or* conjunctions in context. *European Journal of Cognitive Psychology, 2*, 57-70.

MARGOLIS, H. (1987). *Patterns, thinking, and cognition: A theory of judgment.* Chicago: University of Chicago Press.

MASSARO, D. W. (1985, November). *Conjoining multiple sources of information in pattern perception.* Paper read at the meeting of the Psychonomic Society, Boston.

MASSARO, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry.* Hillsdale, NJ: Erlbaum.

MASSARO, D. W. (1988, August). *A pattern recognition account of decision making.* Paper read at the XXIV International Congress of Psychology, Sydney, Australia.

MASSARO, D. W. (1989). Multiple book review of *Speech perception by ear and eye: A paradigm for psychological inquiry. Behavioral & Brain Sciences, 12*, 741-794.

MASSARO, D. W. (1992). Broadening the domain of the fuzzy logical model of perception. In H. L. Pick, Jr., P. van den Broek, & D. C. Knill (Eds.), *Cognition: Conceptual and methodological issues* (pp. 51-84). Washington, DC: American Psychological Association.

MASSARO, D. W., & COHEN, M. M. (1990). Perception of synthesized audible and visible speech. *Psychological Science, 1*, 55-63.

MASSARO, D. W., & COHEN, M. M. (1993). The paradigm and the fuzzy logical model of perception are alive and well. *Journal of Experimental Psychology: General, 122*, 115-124.

MASSARO, D. W., & FRIEDMAN, D. (1990). Models of integration given multiple sources of information. *Psychological Review, 97*, 225-252.

SELFRIDGE, O. G. (1959). Pandemonium: A paradigm for learning. In *Mechanization of thought processes* (pp. 511-526). London: Her Majesty's Stationery Office.

STOLARZ-FANTINO, S., & FANTINO, E. (1990). Cognition and behavior analysis: A review of Rachlin's *Judgment, decision, and choice. Journal of the Experimental Analysis of Behavior, 54*, 317-322.

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

WOLFORD, G. (1991). The conjunction fallacy? A reply to Bar-Hillel. *Memory & Cognition, 19*, 415-417.

WOLFORD, G., TAYLOR, H. A., & BECK, J. R. (1990). The conjunction fallacy? *Memory & Cognition, 18*, 47-53.

WYER, R. S. (1976). An investigation of the relations among probability estimates. *Organizational Behavior & Human Performance, 15*, 1-18.

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

ZADEH, L. A. (1965). Fuzzy sets. *Information & Control, 8*, 338-353.