ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES 56, 190–212 (1993)

Predictor Representation and Prediction Strategies

YOAV GANZACH

School of Business Administration, The Hebrew University, Jerusalem, Israel

An important difference between visual and numerical representation of the predictor is that the former entails a frame of reference against which the predictor can be evaluated, thereby facilitating reliance on the representativeness heuristic. The main hypothesis examined in Experiment 1 is that reliance on representativeness will increase when predictions are based on predictors that are represented visually rather than numerically. In addition, it is hypothesized that factors such as predictor validity, number of predictors, and amount of outcome feedback interact with predictor representation in determining the influence of representativeness on predictions. Experiment 2 examines the hypothesis that reliance on representativeness will increase if subjects are provided with a frame of reference for the predictor by informing them about the predictor's scale. In both experiments, extremity and consistency of predictions are used as indicators for reliance on representativeness. The results of Experiment 1 indicate that: (1) Predictions are more extreme when they are based on visually represented predictors, but this effect depends on the amount of outcome feedback and disappears in high predictor validity, and (2) predictor representation interacts with the number of predictors in determining prediction consistency, but not in determining prediction extremity, but this interaction depends on predictor validity. Experiment 2 replicates Experiment 1 with regards to extremity but not with regard to consistency.

Research in Cue Probability Learning (CPL)—the learning of relationship between predictor and outcome through the reception of outcome feedback—indicates that in making predictions people develop rules and apply strategies, rather than simply learning a stimulus–response relationship between predictor and outcome (Brehmer, 1974).[1] These rules, however, differ from the probabilistic rules necessary for achieving optimal performance in such tasks. Consider for example, the positive linear relationship between predictor and outcome, the easiest rule to learn in CPL tasks. People do not apply appropriate probabilistic strategies even

I thank Paul Andreassen for many helpful discussions. The research was supported by the Kmart Center. Address correspondence to Yoav Ganzach, School of Business Administration, The Hebrew University of Jerusalem, Jerusalem 91905, Israel.

[1] The terms predictor, prediction, and outcome are used rather than the terms cue, response, and criteria, which are more commonly used in the CPL literature, since the emphasis in this paper is on prediction heuristic.

0749-5978/93 $5.00

when they have been given considerable explanation about the probabilistic nature of the relationship or provided with the appropriate probabilistic rule or with aids that illustrate this rule (Brehmer & Kuylenstierna, 1978, 1980, 1981; Brehmer & Johnansson, 1979).

What, then, is the strategy used when the rule relating predictor to outcome is positive-linear? In my view, people rely, to a large extent, on the representativeness heuristic (Kahneman & Tversky, 1973). According to this heuristic, people first learn the distribution of the predictor and outcome variables. Subsequently, they use a matching strategy, whereby the predicted value is chosen so that its extremity (deviation from central tendency) matches the extremity of the predictor. The matching strategy leads to systematic differences between intuitive predictions and normative predictions. Normative predictions are regressive: The position of the predicted value on the distribution of the outcome is less extreme than the position of the predictor on its distribution; the less valid the predictor, the less extreme the prediction. On the other hand, intuitive predictions are—at least when not much learning is involved—nonregressive, and, therefore, more extreme. The position of the predicted value on the distribution of the outcome is roughly equal to the position of the predictor on its distribution, and the validity of the predictor is, to a large extent, ignored. Indeed, there are findings that support the view that subjects use the representativeness heuristic in CPL tasks. For example, Brehmer and Lindberg (1970) observed "over-shooting" in prediction: The slope of the regression line relating subjects' predictions to predictor values is steeper than the slope of the regression line relating outcome values to the predictor values; the lower the correlation between predictor and the outcome, the larger this effect. This is exactly the pattern of predictions one would expect if subjects predict by representativeness.

However, the representativeness heuristic is not the only heuristic used in making predictions. Heuristics that lead to regressive predictions are likely to "compete" (Agnoli & Krantz, 1989) with representativeness. For example, Ganzach and Krantz (1991) showed that in addition to the representativeness heuristic, other heuristics operate in determining intuitive predictions. The operation of these heuristics can lead to regressive predictions which can be quite different from normative regressive predictions. For example Ganzach and Krantz (1991) showed that, unlike normative predictions, intuitive predictions are asymmetrically regressive: They are more regressive when they are made from a low level of the predictor than from a high level (see also Ganzach & Krantz, 1990, for an additional example of a prediction strategy that leads to regressive predictions which are dissimilar from normative predictions).

If the representativeness heuristic is not the only heuristic used in intuitive predictions, factors that facilitate or hinder its use may exist.

One such factor is likely to be the mode of predictor representation. The main hypothesis of Experiment 1 is that when predictions are based on visual representation of the predictors (e.g., bar graphs), people are more prone to use the representativeness heuristic than when the predictors are represented numerically and therefore make more extreme predictions.

The reason for the increased tendency to use the representativeness heuristic when predictors are represented visually is a consequence of the ease by which the extremity of the predictor can be determined in this type of representation. Visual representation provides the prediction maker with a natural frame of reference, or scale, on which the extremity of the predictor can be assessed. Consider, for example, the case in which the predictor is represented as a bar on a computer screen. The extremity of the predictor is easily determined by comparing the length of the bar to the length of the computer screen. This is not the case when the predictor is represented numerically. In this case there is not an obvious frame of reference to which the predictor can be compared.

In this paper I also explore that moderate the influence of the representativeness heuristic on predictions and, as a result, interact with predictor representation. One such factor is experience. While earlier work argued that prediction biases in general (Brehmer, 1980), and biases that result from predicting by representativeness, in particular (Kahneman & Tversky, 1973), are rather unamenable to experience, more recent research has shown that experience can improve probabilistic judgment (e.g., Zukier and Pepitone, 1984; Nisbett, Krantz, Jepson, & Kunda, 1983). These findings were interpreted in terms of a decrease in the tendency to rely on representativeness and an increase in the tendency to adopt strategies that result in more accurate predictions (Nisbett *et al.*, 1983; Ganzach & Krantz, 1990). In the context of CPL, the influence of experience (operationalized as outcome feedback) on the use of the representativeness heuristic has never been investigated directly. However, there are findings—albeit conflicting—that are relevant to this issue. While Brehmer and Lindberg (1970; Fig. 3) and Brehmer (1973; Fig. 2) found that continuous outcome feedback does not influence the extremity of predictions in Single Cue Probability Learning, Ganzach and Krantz (1990; Fig. 3) found that the extremity of prediction in this task does decrease with outcome feedback. This latter finding suggests that the influence of the representativeness heuristic on predictions declines with exposure to continuous outcome feedback.

In Experiment 1 (as well as Experiment 2), this latter finding regarding outcome feedback is replicated. Furthermore, the interaction between outcome feedback and predictor representation is investigated. Such an interaction should be expected since, if the tendency to rely on representativeness is higher in visual than in numerical representation, then in the former there is a larger discrepancy between intuitive predictions and optimal predictions, and, therefore, people have "more" to learn from outcome feedback. As a result, it is expected that the difference in prediction extremity between visual and numerical representation will decrease with experience, primarily as a result of decrease in extremity in the visual representation condition.

Another factor that may moderate the influence of representativeness on predictions is predictor validity. The pattern of the interaction between this factor and the representation factor should also be indicative of the operation of the representativeness heuristic. First, if subjects rely more on representativeness when predictors are represented visually than when they are represented numerically, prediction extremity should be less sensitive to predictor validity in visual representation than in numerical representation. Second, the difference in prediction extremity between visual and numerical representation should be larger in low predictor validity than in high predictor validity. There are two reasons for this latter expectation. First, when predictor validity is high, adherence to prediction strategies that lead to systematic biases from optimal predictions in general, and adherence to the matching strategy, in particular, is low, because feedback provides a fairly clear indication of the bias. The second reason is that the higher the validity, the smaller the discrepancy between optimal predictions and predictions by representativeness.

Yet another factor in Experiment 1 was the number of predictors, whether the explained variance in the outcome is accounted for by one or two predictors. Predictions by representativeness that are based on two predictors imply an averaging strategy, i.e., a strategy in which the extremity of the prediction is matched to the extremity of the *two* predictors. When the stimuli-generating model is an additive model, these predictions are less extreme than optimal predictions (Lichtenstein, Earle, & Slovic, 1975; see also Birenbaum, 1976). *A fortiori,* they are less extreme than predictions by representativeness based on one (equally valid) predictor.[2] Thus, a main effect for the number of predictors factor is ex-

[2] For example, consider two orthogonal predictors with equal (normal) distributions, each having a correlation of .6 with the outcome. Assume that the perceived extremity of the predictors and the outcome can be approximated by their Z-scores. If the values of the predictors in Z-scores are 1 and 2, their average extremity is 1.5 Z-scores, and therefore predictions by representativeness will lead subjects to choose a prediction whose value is 1.5 Z-scores on the outcome scale. On the other hand, if the two predictors are represented by an equally valid one predictor, the extremity of this predictor is $.7071 * 1 + .7071 * 2 = 2.13$ Z-scores (in the current experiment each of the two predictors explain half of the variance in the equally valid one predictor), and therefore predictions by representativeness will lead subjects to choose a prediction whose value is 2.13 Z-scores on the outcome scale. Note that the normative prediction in this case is $.6 * 1 + .6 * 2 = 1.8$ Z-scores on the outcome scale.

pected. In addition, an interaction between number of predictors and prediction representation should be expected, since the effect of the number of predictors on the extremity of the predictions depends on the extent to which subjects rely on representativeness which, in turn, is hypothesized to be dependent on predictor representation.

In addition to prediction extremity, prediction consistency is used as a secondary indicator for the tendency to rely on representativeness. For a linear relationship between predictor and outcome, the representativeness heuristic reduces both acquisition and application difficulties (See Hammond & Summers, 1972, for a discussion of these concepts in CPL task) because it implies a linear relationship between predictor and prediction. Thus, it should be expected that the higher the reliance on representativeness, the higher the consistency. Therefore, the main hypothesis regarding the influence of predictor representation on consistency is that in conditions in which reliance on representativeness is high (i.e., visual representation of the predictor) consistency will be higher than that in conditions in which reliance on representativeness is low (i.e., numerical representation).

It should be noted, however, that while the main effect of predictor representation may be similar for consistency and extremity, the pattern of interactions between predictor representation and the other factors may differ. In particular, while extremity is expected to decrease with experience, consistency tends to increase with experience (e.g., Naylor & Clark, 1968; Brehmer & Lindberg, 1970), since subjects learn the linear relationship between predictor and outcome. Furthermore, since consistency is lower in the numerical representation condition, more learning, and therefore larger change in consistency, is likely to occur in this condition. Note that while differences between numerical and visual conditions are expected to decrease as a result of experience for both consistency and extremity, the way by which these differences decrease is expected to differ for these two measures. Learning is likely to influence (decrease) extremity primarily in the visual representation condition and to influence (increase) consistency primarily in the numerical representation condition.

To summarize, the above analysis suggests that the mode of predictor representation should influence both the extremity of predictions and their consistency. However, this influence depends on other factors such as amount of outcome feedback; predictor validity, and the number of predictors.

EXPERIMENT 1

Method

Subjects. Two hundred and ninety first-year Business Administration students participated in the experiment to fulfill a class requirement. Subjects were assigned randomly to eight conditions. (See Table 3 for the number of subjects in each condition).

Design. Each subject predicted an outcome based on either one or two predictors presented either numerically or visually. The explained variance of the outcome was either .95 (high-validity condition) or .7 (low-validity condition). This resulted in a 2 × 2 × 2 between subject design with respect to these three factors. Note that the R-squares used in this experiment are larger than most R-squares in CPL tasks. The reason for this is that the focus of this experiment is on the strategies people use for a positive-linear rule and not on the process of learning such a rule. Large R-squares are likely to facilitate this learning. The instructions below also reflect this consideration.

Procedure. Subjects participated in the experiment in groups numbering between four and eight. After entering the laboratory, they were seated in front of an IBM XT computer and told to read the initial instructions. Subjects were told that in many areas experts are interested in learning how to make optimal predictions and this is the reason why researchers are interested in how people learn to make predictions and how they improve these predictions as a result of experience. A description of the experiment was given, with an emphasis that "The relationship between the predictor(s) and the outcome is positive, that is, the higher the predictor(s), the higher the outcome," and that, "It is almost impossible to predict the outcome precisely. Your task is therefore to make predictions that are as close as possible to the outcome." Subjects then performed six practice trials. After the practice trials, the experimenter checked that subjects understood how to operate the computer and again emphasized the rule relating predictor(s) to outcome as well as the probabilistic nature of the task. Subsequently, subjects completed the 120 experimental trials at their own pace.

In each trial, the computer first displayed the predictor(s). In the numerical conditions the predictor(s) were number(s) located in the center of the screen. In the visual conditions the predictor(s) was a horizontal bar(s) whose length(s) was proportional to the value(s) of the number(s) in the numerical conditions. After 2 s the computer prompted the subjects to type their prediction. Subjects typed their predictions numerically in all conditions. After typing the prediction, the predictors and the prediction were erased and the computer displayed the outcome for 2 s. Subsequently, the outcome was erased, and a new trial began. In all conditions, the outcome was displayed numerically.

There was not a time limit for typing the predictions. To avoid reading inadvertent mistakes by the computer, the predictions were examined, and if they were completely out of range (numbers of four or more digits or two or less digits), subjects were prompted to the prediction again.

Stimuli. One set of stimuli was used for all subjects. The stimuli were created as follows. Eight blocks of 15 trials were constructed by sampling two predictors and a random error from a trivariate standardized normal distribution with covariances 0. To be included in the experiment, a block was required to fulfill the following conditions: (1) the correlation between each of the three variables would not exceed $\pm.1$; (2) the mean of each of the three variables would not exceed $\pm.05$; and (3) the standard deviation of each of the three variables would not exceed $\pm.05$.

Outcome feedback was generated for the low-validity conditions by the equation $E = .592 * C1 + .592\ C2 + .548 * \epsilon$, where E is the outcome, C1 and C2 the predictors, and ϵ is the error, and for the high-validity conditions by the equation $E = .689 * C1 + .689 * C2 + .224 * \epsilon$. The predictor in the one-predictor conditions was generated by summing the two predictors. Thus, total predictability was equal in the one-predictor and two-predictor conditions.

The stimuli that subjects actually received in the numerical conditions were created by transforming the standardized predictors and feedback values to desired distributions. Each of the predictors in the two predictor conditions was transformed to have a mean of 75 and a standard deviation of 13. The predictor in the one-predictor conditions was transformed to have a mean of 150 and a standard deviation of about 18.4. The outcome was transformed to have a mean of 585 and a standard deviation of 50 (numbers were rounded to the nearest whole number). This resulted in a task slope (regression slope relating outcome to predictor) of about 2.28 in the low-validity conditions and 2.65 in the high-validity conditions.

In the visual conditions, the predictors were horizontal bars whose values were proportional to the numerical values of the predictors in the corresponding trial in the numerical conditions. In the one-predictor conditions, an increase of 1 cm in the bar corresponded to an increase of about 3.5 units in the predictor in the numerical conditions. In the two-predictor conditions, an increase of 1 cm in the bar corresponded to an increase of about 2.5 units in the predictor in the numerical conditions.

RESULTS AND DISCUSSION

Extremity. Extremity was defined in terms of the prediction slope, the regression slope relating subjects' predictions to predictor values. This regression slope was calculated for each subject and each 30-trial block.

(The data were also analyzed for fifteen trial blocks and the results, except for small differences, were similar. These differences are mentioned in the text.) To allow for comparability, in the two predictor conditions, the slope was calculated by regressing subjects' predictions on the sum of the two predictors. In the visual conditions, it was calculated by regressing subjects' predictions on the numerical value corresponding to the predictor(s). The regression slopes were subjected to a $2 \times 2 \times 2 \times 4$ analysis of variance with repeated measures on the fourth factor (block). The results of this analysis are summarized in Table 1 and plotted in Fig. 1, where the means are collapsed over the number of predictor factor (neither the main effect nor the interactions involving this factor are significant).

The main hypothesis of the experiment, that predictions are more extreme when the predictor is represented visually than when it is represented numerically, is supported by a highly significant main effect for predictor representation. Other hypotheses are also supported. First, a significant main effect for block indicates that extremity declines during the course of the experiment. Second, a significant representation × block interaction indicates that overall extremity declines more in the visual conditions than in the numerical conditions.

TABLE 1
SUMMARY TABLE FOR ANOVA ON RESPONSE SLOPE

Source	*df*	*F*	*P*
Between groups	289		
Cue representation	1	13.9	.0003
Cue validity	1	14.9	.0001
Number of cues	1	.8	.38
Representation × validity	1	16.1	.0001
Representation × number of cues	1	.2	.64
Validity × number of cues	1	.1	.84
Representation × validity × number of cues	1	.4	.55
Error	282		
Within subjects	870		
Block	3	16.7	.0001
Block × representation	3	6.8	.0003
Block × validity	3	2.7	.05
Block × number of cues	3	1.1	.34
Block × representation × validity	3	5.0	.003
Block × representation × number of cues	3	1.9	.14
Block × validity × number of cues	3	.7	.57
Block × representation × validity × number of cues	3	.7	.57
Error	846		

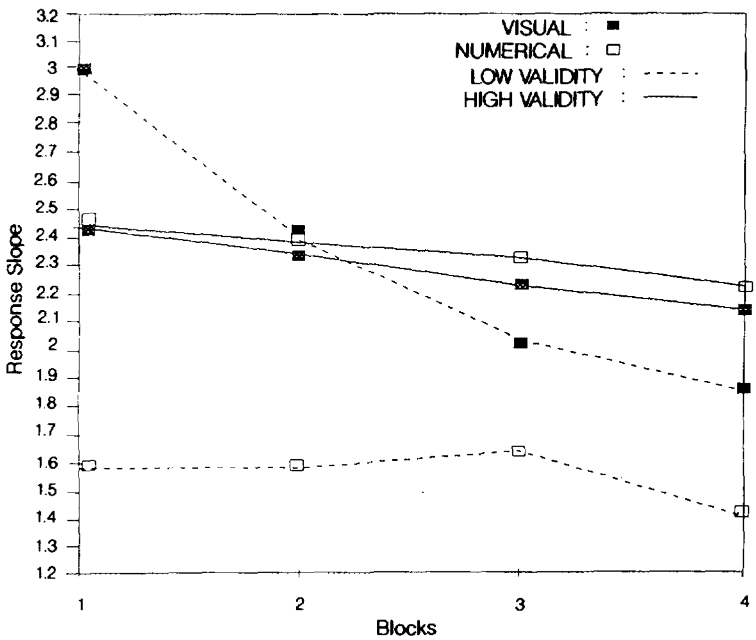

FIG. 1. Mean prediction slope as a function of block, predictor representation, and predictor validity. Data are collapsed over the number of cue conditions. Standard errors are, in an ascending block number, .21, .18, .14, .14 for the visual low-validity conditions; .14, .11, .12, .11 for the visual high-validity conditions; .12, .10, .11, .14 for the numerical low-validity conditions; and .11, .09, .10, .10 for numerical high-validity conditions.

However, these results should be viewed in light of the triple interaction between representation, validity, and block.[3] One feature of this interaction is that in the numerical representation, the slopes in the high-validity conditions are steeper than the slopes in the low-validity conditions during the entire course of the experiment ($p < .0001$ for all four blocks), and the differences between them are approximately constant. On the other hand, in the visual representation, the slope in the first block is steeper in the low-validity conditions (!) [$t(143) = 2.17, p < .03$], while in the rest of the blocks there is no significant difference between the con-

[3] The main effect of predictor validity results from the fact that the predictions in the high-validity conditions are more extreme than the predictions in the low-validity conditions (the mean response slope over blocks and subjects is 2.32 for the former and 1.93 for the latter). This effect is not surprising given that the task slope is indeed steeper in the high-validity conditions, and it is in agreement with previous findings. The other significant effects (the main effect of representation, the interaction between representation and block, and the interaction between representation and predictor validity) are better understood in terms of the triple interaction.

ditions, although there is a tendency for the slope in the low-validity conditions to be more moderate. Thus, for visual representation, prediction extremity is inappropriately sensitive to validity in the first block and not sensitive to validity in the rest of the blocks, while for numerical representation, it is appropriately sensitive to validity in all blocks.

Another feature of this interaction is that in the high-validity conditions there are actually no differences between numerical and visual representation ($p > .5$ for all four blocks). In the low-validity conditions, there are substantial differences between the extremity of predictions in the two representation conditions. These differences, however, decrease during the course of the experiment as a result of the decrease in prediction extremity in the visual representation conditions.

This pattern of results indicates that the subjects in the low-validity visual conditions are the ones most seriously susceptible to biases stemming from representativeness. Unlike the subjects in the low-validity numerical conditions, it is difficult for them to use strategies other than representativeness. Unlike the subjects in the high-validity visual conditions, feedback does not provide them with as clear an indication of possible biases. Also, as previously mentioned, while for these subjects reliance on representativeness can lead to large biases, for subjects in the high-validity visual conditions, reliance on representativeness should not lead to large biases. Subjects in the low-validity visual conditions are also characterized by the continuous moderation of prediction extremity, while subjects in the other conditions appear to have accomplished most of their learning in the first block. The reason is that the subjects in the low-validity visual conditions have the most learning to do, and they have to do it from relatively noisy outcome feedback.

Although moderation of predictions during the course of the experiment is very salient in the low-validity visual condition, it is common to all conditions, as indicated by the main effect of block. While a test for linear trend was highly significant in the low-validity visual conditions [$t(71) = 5.38, p < .0001$], it was also significant in the high-validity visual conditions [$t(72) = 2.72, p < .008$], and it was marginally significant in the high validity numerical conditions [$t(68) = 1.74, p < .09$]. Linear trend in the low-validity numerical conditions was not significant although it was in the expected direction [$t(75) = .86, p < .4$].

The continuous moderation of predictions stands in contradiction to some previous findings (Brehmer, 1973; Brehmer & Lindberg, 1970). The reason for this is related, in my view, to methodological details of the CPL task that, so far, have not received attention. For example, Brehmer used in his experiments rectangular distributions while the distributions in the current experiment are normal. Different forms of distribution will lead to different prediction strategies if one of the strategies available for subjects is averaging two values. One value is an outcome value which is as extreme as the predictor; the other value is a representative value of the outcome distribution, i.e., a central tendency value (Tversky & Kahneman, 1982). In the normal distribution, this central tendency value is much more representative of the distribution and therefore much more likely to influence predictions, since it is also the most frequent value (see Andreassen, 1987, for an example of the influence of central tendency values on prediction). In other words, it could be that what subjects learn from outcome feedback in this experiment is a strategy of "regression to the mode." This hypothesis lends itself to experimental test by examining predictions from the two types of distributions, and it could potentially shed some light on the nature of the heuristics that replace representativeness as a result of experience.

Comparison with normative predictions. In the low-validity visual conditions, the prediction slope is significantly steeper than the task slope in the first block [$t(71) = 3.37, p < .001$]. It is also steeper than the perfect matching slope, the slope that would be obtained from a perfect matching strategy (about 2.72). While this latter effect is not significant, if the first block is divided into two 15-trial blocks, the slope in the first of these two blocks (3.67) is significantly steeper than the perfect matching slope, $t(71) = 3.79, p < .0003$.

In the low-validity numerical conditions, the prediction slope of the first block is more moderate than the task slope [$t(75) = 5.51, p < .0001$], while in the high-validity conditions, it does not differ significantly from the task slope. In the last block, however, the prediction slope is clearly below the task slope for all conditions, including the low-validity visual conditions [$t(71) = 3.06, p < .003$; $t(75) = 6.25, p < .0001$; $t(68) = 4.29$, $p < .0001$; and $t(72) = 4.48$, $p < .0001$ in the low-validity visual, low-validity numerical, high-validity numerical, and high-validity visual conditions, respectively].

This pattern of results departs from a simple model in which subjects rely heavily on representativeness in the early phase of the experiment, subsequently adopting regressive prediction strategies as a result of feedback. In my view, this departure occurs because various heuristics, representativeness being only one of them, determine intuitive predictions. For example, prediction by representativeness depends on the perceived distribution of outcome and predictor rather than the actual distribution. Thus, since extreme values of a distribution are more available in memory (Nisbett & Kunda, 1985), the steep slope in the low-validity visual conditions could be the result of the dispersion of the perceived distributions being larger than the dispersion of the actual distribution (Ganzach, in press).

The overregressiveness observed in the experiment is also likely to stem from the fact that prediction moderation is the result of the operation of various heuristics rather than the learning of a statistical principle. For example, the "regression to the mode" strategy described above could lead to overly regressive predictions if subjects place too much weight on one value—the central tendency value—relative to the other value—the outcome value whose extremity matches the extremity of the predictor.

Finally, why, in contrast to previous CPL experiments, is overregressiveness observed in this experiment? The reason is likely to be the high predictor validities used in the current experiment. When predictor validity is high, the discrepancy between predictions by representativeness and normative predictions is low. Thus, extremism is less likely to be observed in high predictor validities. Indeed, in Experiment 2, where the predictor validity was .5, predictions did not become overregressive even in the fourth block, although moderation of predictions over the course of the experiment was observed. This pattern was also observed in Ganzach and Krantz (1990), where lower validity was used.

Consistency. Consistency was defined in terms of the correlation between predictor and prediction. (For the two predictor conditions, consistency was defined as the multiple correlation between predictors and prediction.) For each subject and each 30-trial block, this correlation was calculated, and subjected to Fisher's Z transformation. The transformed correlations were subjected to a $2 \times 2 \times 2 \times 4$ analysis of variance with repeated measures on the fourth factors. The results of this analysis are summarized in Table 2, and the means and standard deviations appear in Table 3. The main effect for predictor validity results from the fact that predictions in the high-validity conditions are more consistent than predictions in the low-validity conditions (the mean over blocks and subjects is 1.09 for the former and .74 for the latter). The main effect for block results from an increase in consistency during the course of the experiment. Both the main effect for validity and the main effect for block replicate earlier findings (e.g., Brehmer & Lindberg, 1970).

The main effect for representation arises from the fact that the consistency in the visual conditions is greater than the consistency in the numerical conditions (the mean over blocks and subjects is .95 for the visual conditions and .87 for the numerical conditions). This supports the hypothesis that predictions are more consistent when they are based on visual representation. However, the main effect of representation should be viewed in light of its interactions with the other factors. The interaction between predictor representation and block is due to the fact

TABLE 2
SUMMARY TABLE FOR ANOVA ON CONSISTENCY

Source	*df*	*F*	*P*
Between groups	289		
Cue representation	1	3.7	.05
Cue validity	1	64.1	.0001
Number of cues	1	5.9	.02
Representation × validity	1	2.5	.12
Representation × number of cues	1	6.7	.01
Validity × number of cues	1	4.4	.04
Representation × validity × number of cues	1	3.5	.06
Error	282		
Within subjects	870		
Block	3	17.4	.0001
Block × representation	3	2.6	.06
Block × validity	3	1.4	.25
Block × number of cues	3	1.5	.21
Block × representation × validity	3	1.8	.15
Block × representation × number of cues	3	.9	.42
Block × validity × number of cues	3	.8	.48
Block × representation × validity × number of cues	3	.8	.51
Error	846		

Note. The consistency measure is the z transformation of the correlation between cues(s) and response.

that subjects in the numerical conditions increase their consistency during the course of the experiment more than subjects in the visual conditions. In both representations there is a significant linear trend ($p < .0001$) for consistency. However, this trend is stronger in the numerical conditions [the difference between the two trends is not significant for the 30-trial blocks but is significant for the 15-trial blocks, $t(288) = 2.5$, $p < .01$)]. This effect parallels the interaction effect between representation and block in regard to extremity, but in line with our theory, for consistency, these are the subjects in the numerical conditions that have more to learn.

The interaction between representation and number of predictors is due to the fact that in the visual conditions, consistency is not sensitive to the number of predictors, while in the numerical conditions, it is (the results in regard to the visual conditions are in agreement with results reported by Brehmer, 1987). In the numerical conditions, two predictors decrease consistency relative to one predictor. This effect is consistent with the notion that subjects in the visual conditions rely on representativeness, because this heuristic offers them a strategy for integrating the two pre-

TABLE 3

MEAN CONSISTENCY BY CONDITION AND BLOCK

	Low validity				High valid	
	One cue		Two cues		One cue	
Block	Visual	Numerical	Visual	Numerical	Visual	Numerical
n	37	38	35	38	38	34
1	.74	.56	.80	.53	1.00	1.08
	(.39)	(.45)	(.31)	(.31)	(.46)	(.32)
2	.88	.65	.86	.62	1.16	1.43
	(.55)	(.40)	(.40)	(.43)	(.37)	(.43)
3	.75	.78	.87	.71	1.07	1.33
	(.50)	(.49)	(.45)	(.43)	(.46)	(.38)
4	.81	.78	.82	.67	1.13	1.34
	(.53)	(.54)	(.50)	(.52)	(.52)	(.47)
Mean	.80	.69	.83	.63	1.09	1.29
	(.42)	(.39)	(.33)	(.36)	(.39)	(.24)

Note. Entries are the mean Fisher's Z transformation of the correlation between cue(s) and response. Numbers i
deviations.

dictors (i.e., averaging their extremity), which leads to consistent predictions.[4]

It is interesting to compare the results of the analysis of consistency to those of extremity. The main difference is that in regard to extremity, the number of predictors did not make any difference, while in regard to consistency, it did. If the notion that representativeness enhances consistency in predicting from two predictors by reinforcing the use of an averaging strategy is correct, why did it not lead also to less extreme predictions in the two predictors conditions? One answer is that in the CPL task used in this experiment, subjects predicting by representativeness from two predictors are likely not only to average the extremity of the two predictors, but also to evaluate the extremity of the average against the scale of the average rather than against the scale of the individual predictors. Since the average scale is narrower than the predictors scale, this should lead to no differences between the extremity of predictions from two predictors and from one equally valid predictor.[5]

[4] The analysis of variance reveals two additional effects of theoretical interest. First, the interaction between validity and the number of predictors is due to the fact that the difference in consistency between the one-predictor conditions and the two-predictor conditions is larger in high-validity than in low-validity. A possible explanation for this is that low-validity predictors interfere more with the use of an accompanying predictor (Dudycha & Naylor, 1966; see also Brehmer, 1973) and cause more difficulties in using each of the predictors separately, thus "pushing" subjects toward reliance on an averaging strategy, which, in turn, minimizes the differences in consistency between the one-predictor conditions and the two-predictor conditions. This effect depends, however, on predictor representation, leading to triple interaction between representation, validity, and the number of predictors. The effect is rather strong in the numerical representation and nonexistent in the visual representation. One reason for this could be that in the visual representation, the dominance of the representativeness heuristic is such that subjects use an averaging strategy in the two-predictor conditions irrespective of validity. Another feature of this triple interaction is that consistency is greater in each of the visual conditions than in the respective numerical condition except in the case of the high-validity one-predictor conditions, where this relationship reverses. This feature of the triple interaction is likely to be related to the fact that in the one predictor high-validity conditions there are no difficulties associated with the integration of two predictors nor are there serious difficulties associated with the probabilistic nature of the task. Thus, the representativeness heuristic does not "benefit" subjects in these conditions as much as it does in the other conditions. On the other hand, in these conditions the inferior accuracy of the visual representation (the length of the visual scale allowed less accurate perception of changes in predictors in comparison to the numerical scale) could lead to lower consistency.

[5] To use the example in footnote 1, if the standard deviation of each of the two predictors is S, the average of the two predictors is $1 * S + 2 * S = 1.5 * S$. On the other hand, the standard deviation of the average is equal to $\sqrt{.5^2 * S^2 + .5^2 * S^2} = .7071 * S$. Therefore, the extremity of the average on the average scale is $1.5 * S / .7071 * S = 2.12$, which is equal to the extremity of the equally valid one predictor.

This line of reasoning explains why, contrary to our results [as well as Brehmer's (1987)], Lichtenstein *et al.* (1975) found that predicting from two predictors induces reduction in the extremity of predictions. In their experimental paradigm, subjects were first trained in each of the two predictors (which had equal distributions) separately, and only then were they required to make predictions based on the two predictors *without feedback*. In this case subjects are likely to evaluate the extremity of the average on the predictors scale rather than on the average scale, which can lead to reduction in the extremity of predictions.

The differences between the results of consistency and those of extremity also rule out the possibility that the between-condition differences in the extremity measure used in the analysis (prediction slope) are due to differences in noise in subjects' predictions rather than to differences in prediction strategy. First, while the prediction slope decreases continuously, consistency increases, which suggests that the prediction slope does not decrease because of increased noise in subjects' predictions, but rather because of decreased prediction extremity. Second, in regard to prediction slope, the effect of representation depends primarily on predictor validity, while the interaction between representation and number of predictors is nonsignificant. On the other hand, in regard to consistency, the effect of representation depends primarily on the number of predictors, while the interaction between representation and predictor validity is nonsignificant. This discrepancy suggests that in this experiment, prediction slope and consistency are, to a large extent, independent measures of prediction strategy.

EXPERIMENT 2

Visual representation of the predictor is only one way by which a frame of reference for the predictor can be made salient to the prediction maker. Such a frame may become salient even if the predictor is represented numerically, providing the scale of the predictor is known. The main purpose of this experiment is to examine whether awareness of the scale increases reliance on representativeness in the case of nonvisual representation of the predictor.

Method

Overview. Scale awareness was manipulated by presenting the predictor in percentiles in one condition (the percentile representation condition), but not in the other (the standard numerical representation condition). The validity of the predictor was low in both conditions (low-validity conditions were used since these conditions are the critical conditions for testing the influence of predictor representation), and the outcome was predicted on the basis of one predictor (one-predictor conditions were used since no important differences were found in regard to the number of predictors factor in Experiment 1).

Subjects. Forty first-year Business Administration students participated in the experiment to fulfill a class requirement. Subjects were assigned randomly to two conditions, 20 in each condition.

Stimuli. Sampling of stimuli values was conducted separately for each subject. In each trial, a predictor value and a random error were sampled from two rectangular distributions. Outcome feedback was centered around 450, and explained variance was equal to .5. The range of the predictor was 100. In the percentile representation condition the predictor was centered around 50 and in the standard numerical representation condition it was centered around 120. The normative slope in both conditions was 1.5.

Procedures. Subjects participated in the experiment in groups numbering between four and eight. After entering the laboratory, they were seated in front of an IBM XT computer and told to read the initial instructions which explained the task. Subjects were told that they will be asked to predict students' scores in aptitude A from their scores in aptitude B. They were told that the relationship between the two aptitudes is probabilistic.

Following these initial instructions, subjects in the percentile representation condition received a short explanation about percentiles and were told that the scores in aptitude A will be given in percentiles. Subjects in the standard numerical representation condition did not receive this explanation. Subsequently, all subjects performed five practice trials and then completed 120 experimental trials at their own pace. The trials were similar in the two conditions except that in the percentile representation condition subjects were reminded in each trial that the predictor is given in percentiles and that the percentile scale ranges from 1 to 100. Similar to Experiment 1, the predictions were examined, and if they were completely out of range (numbers of four or more digits and two or less digits), subjects were prompted to type their predictions again.

RESULTS AND DISCUSSION

Extremity. Extremity was defined in terms of the prediction slope. This slope was calculated for each subject and each 15-trial block. The slope means by condition and block are plotted in Fig. 2. Standard errors appear in the figure caption. A 2 × 8 ANOVA (representation × block) with repeated measures on the second factor revealed a significant main effect for predictor representation [$F(1,38) = 3.1$, $p < .01$], a significant main effect for block [$F(7,266) = 8.5$, $p < .0001$], and a significant representation × block interaction [$F(1,266) = 5.0$, $p < .001$].

These results replicate the results of Experiment 1. Predictions are

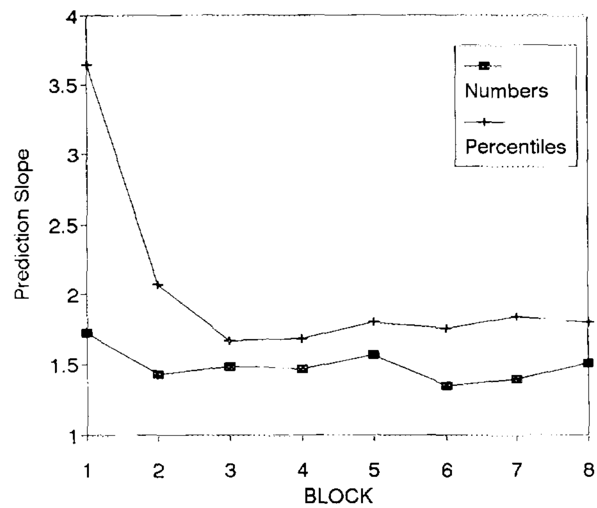

FIG. 2. Mean prediction slope as a function of block and predictor representation. Standard errors are, in an ascending block number, .14, .21, .18, .19, .19, .18, .17 and .15 for the standard numerical representation condition, and .51, .21, .21, .22, .19, .21, .21, and .14 for the percentile representation condition.

more extreme in the percentile representation condition, where predictor representation provides subjects with a frame of reference on which the extremity of the predictor can be easily assessed. Furthermore, there is a learning process in which the extremity of the prediction approaches its optimal value through experience. This learning occurs primarily in the condition in which subjects are susceptible to reliance on representativeness, since these subjects have the most learning to do.

Consistency. Consistency was defined in terms of the correlation between predictor and prediction. For each subject and each 15-trial block, this correlation was calculated and subjected to Fisher's Z transformation. The transformed correlations were subjected to a 2×8 ANOVA with repeated measures on the second factor. The results showed no significant main effect for representation, no significant main effect for block, and no significant interaction. The means of the transformed correlations over the eight blocks were 1.13 (SD = .44) for the percentile representation condition and 1.03 (SD = .53) for the standard numerical representation condition.

Comparison between Experiments 1 and 2. Why do the results of obtained for extremity replicate the results of Experiment 1 while the results obtained for consistency do not? First note that even in Experiment 1, the effect of predictor representation on extremity is much stronger than its effect on consistency, suggesting that extremity is a more sensitive measure of reliance on representativeness. Therefore, the absence of predictor representation effect for consistency in this experiment may be due to lack of power. Second, note that even the strong effect of block on consistency found in Experiment 1 is not replicated here. Subjects reach their maximum level of consistency already in the first block. Furthermore, consistency is higher in this experiment than in the low-validity conditions of Experiment 1, although the predictability of the task is lower. The reason for these differences is most likely that predictor and outcome were labeled in the current experiment, whereas in Experiment 1 they are not. The effect of labeling is likely to reduce the effect of representativeness on consistency. Since labeling leads to high consistency (Adelman, 1981; Miller, 1971; Muchinsky and Dudycha, 1974; Sniezek, 1986, for the influence of labeling on consistency), it is difficult to observe an incremental effect of representativeness on consistency.

The difference between the results obtained for extremity and those obtained for consistency in this experiment rule out the possibility that the between-conditions differences in the extremity measure used in the analysis (prediction slope) are due to differences in noise in subjects' predictions rather than to differences in prediction strategy. Furthermore, as in Experiment 1, this difference suggests that consistency and extremity are, to a large extent, independent measures for prediction strategy.

GENERAL DISCUSSION

It is suggested in this paper that the representativeness heuristic influences predictions in CPL tasks involving a positive linear rule between predictor(s) and outcome. This theoretical framework offers an explanation for phenomena already observed in CPL literature such as overshooting (Brehmer & Lindberg, 1970) or the dependence of prediction extremity on experience with multiple determination (Ganzach & Krantz, 1990), as well as the findings reported here concerning the influence of predictor representation on prediction extremity (and to some extent on prediction consistency).

Another implication of the findings reported here is methodological. In the CPL literature, both numerical and visual representation of the predictor are used. Examples of numerical representation can be found in Dudycha and Naylor (1966), Naylor and Clark (1968), Lichtenstein *et al.* (1975), Sniezek and Reeves (1986), and Sniezek (1986). Examples of visual representation can be found in Hammond and Summers (1972), Deane and Hammond (1972), Brehmer (1973, 1974, 1980, 1987), Brehmer and Kuylenstierna (1978, 1980), and Rothstein (1986). Dual representation is also used (see York, Doherty, & Kamouri, 1987). This paper suggests that more attention should be paid to the method of predictor representation, since the particular heuristic used may depend on the particular representation.

Furthermore, in the applied decision-making literature, a great deal of research has been devoted to the influence of information representation or performance, especially in comparing numerical vs visual representation of the information (e.g., Benbasat & Dexter, 1986; Lusk & Kersnick, 1979; Remus, 1987; Zund, 1978). However, relatively little attention has been devoted to the understanding of the cognitive mechanisms that relate information representation to performance (Benbasat & Dexter, 1985). The current research suggests that such an understanding may be important in guiding the choice of information representation. For example, if predictor consistency is important, visual representation of the predictor may be a better choice than numerical representation, since predictions are more consistent in the former representation than in the latter. However, if minimum absolute error, or lack of systematic bias, is important, numerical representation may be better (especially for low-validity predictors), since predictions by representativeness result in large positive (negative) errors for high (low) values of the predictor.

Although representativeness influences predictions in the current experiments, these experiments suggest also that it is not the only heuristic that influences predictions. Other heuristics "compete" with representativeness. This competition can be conceptualized in terms of the likelihood of the use of representativeness vs the likelihood of the use of the other heuristics. This likelihood, obviously, cannot be observed directly. However, it can be inferred by investigating the influence of factors that facilitate the use of representativeness (e.g., predictor representation, amount of outcome feedback) on prediction parameters that are indicative of its operation (extremity and consistency).

The emphasis of this research was on the competition between the representativeness heuristic and other heuristics in the application of a known rule. In the instructions of Experiment 1, subjects were specifically told about the rule relating predictors to outcome, and the probabilistic nature of the rule was strongly emphasized; in the instructions of Experiment 2, labeling enhanced rule knowledge. Furthermore, the relationship between predictor and outcome in the experiment was positive linear, which is exactly the relationship implied by the use of representativeness. Thus, one research question worth pursuing is the use of representativeness in the application of rules other than the positive-linear rule. Another question worth pursuing is the role of representativeness in the acquisition, rather than the application, of the positive-linear rule. Brehmer (1974) suggested a model similar to the "competing heuristics model" (Agnoli & Krantz, 1989) to explain the acquisition of rules relating predictor and outcome. According to this model, "subjects have a hierarchy of hypotheses [about the possible rule] which differs in strength, but where the actual sampling probability varies with the data presented" (p. 24). Brehmer also found that the most dominant hypothesis in this hierarchy is the hypothesis that the rule is a positive-linear rule. Several explanations were offered for this phenomenon (Brehmer, 1974, 1980). According to the current analysis, the reason for the dominance of this hypothesis is that it accords with representativeness, the most dominant prediction heuristic. Furthermore, Brehmer also suggested that the process by which intuitive predictions are made involves a competition between deterministic and probabilistic prediction strategies and that the former strategies are more dominant, even if the nature of the task is probabilistic (Brehmer, 1980). In this context, the representative heuristic can be viewed as a deterministic prediction strategy, since it is the appropriate strategy if the relationship between predictor and outcome is deterministic (see Ganzach & Krantz, 1990, for a discussion about the relationship between representativeness and deterministic prediction strategies), while the heuristics that lead to regressive predictions can be viewed as probabilistic prediction strategies. Thus, the process of learning from outcome feedback, which is viewed in Brehmer framework (Brehmer, 1980) as competition between various rules and strategies, is viewed in the current framework as competition between various heuristics.

The nature of the heuristics that compete with representativeness is still not clearly understood, and this paper does not address this issue. Nevertheless, the paper suggests that although these heuristics can lead to predictions that are more in line with optimal predictions, they cannot adequately be described by normative statistical rules. This is evident, for example, by the overregressiveness of subjects in all conditions in the last block of the experiment. The understanding of these heuristics is an important task for future research, since they may well explain the cognitive processes that mediate learning from experience.

REFERENCES

Agnoli, F., & Krantz, D. H. (1989). Suppressing natural heuristics by formal instructions: The case of the conjunction fallacy. *Cognitive Psychology*, **21**, 515–550.

Andreassen, P. B. (1987). On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction. *Journal of Personality and Social Psychology*, **53**, 490–496.

Benbasat, I., & Dexter, A. S. (1985). An experimental evaluation of graphical and color-enhanced information presentation. *Management Science*, **31**, 1348–1364.

Benbasat, I., & Dexter, A. S. (1986). An investigation of the effectiveness of color and graphical information presentation under varying time constraints. *MIS Quarterly*, **1**, 59–80.

Birenbaum, M. H. (1976). Intuitive numerical prediction. *American Journal of Psychology*, **89**, 417–429.

Brehmer, B. (1973). Single-cue probability learning as a function of the sign and magnitude of the correlation between cue and criterion. *Organizational Behavior and Human Performance*, **9**, 377–395.

Brehmer, B. (1974). Hypothesis about the relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, **11**, 1–27.

Brehmer, B. (1980). In one word: Not from experience. *Acta Pschologica*, **45**, 223–241.

Brehmer, B. (1987). Note on subjects' hypothesis in multiple cue probability learning. *Organizational Behavior and Human Decision Processes*, **40**, 323–329.

Brehmer, B., & Kuylenstierna, J. (1978). Task information and performance in probabilistic inference tasks. *Organizational Behavior and Human Performance*, **22**, 445–464.

Brehmer, B., & Kuylenstierna, J. (1980). Content and consistency in probabilistic inference tasks. *Organizational Behavior and Human Performance*, **26**, 54–64.

Brehmer, B., & Kuylenstierna, J. (1981). Memory aids in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, **28**, 415–424.

Brehmer, B., & Lindberg, L. (1970). The relation between cue dependency and cue validity in single-cue probability learning with scaled cue and criterion variables. *Organizational Behavior and Human Performance*, **5**, 542–554.

Dudycha, L. W., & Naylor, J. C. (1966). Characteristics of the human inference process in complex choice behavior situations. *Organizational Behavior and Human Performance*, **1**, 110–128.

Ganzach, Y. (in press). Outcome representation and prediction strategies. *Organizational Behavior and Human Decision Processes.*

Ganzach, Y., & Krantz, D. H. (1990). The psychology of moderate prediction. I. Experience with multiple determination. *Organizational Behavior and Human Decision Processes*, **47**, 177–204.

Ganzach, Y., & Krantz, D. H. (1991). The psychology of moderate prediction. II. Leniency and uncertainty. *Organizational Behavior and Human Decision Processes*, **48**, 169–192.

Hammond, K. R., & Summers, D. A. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology*, **92**, 20–26.

Johansson, R., & Brehmer, B. (1979). Inferences from incomplete information—A note. *Organizational Behavior and Human Performance*, **24**, 141–145.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, **80**, 237–251.

Lichtenstein, S., & Earle, T. C., & Slovic, P. (1975). Cue utilization in a numerical prediction task. *Journal of Experimental Psychology: Human Perception and Performance*, **104**, 77–85.

Lindberg, L., & Brehmer, B. (1976). Transfer in single-cue probability learning. *Organizational Behavior and Human Performance*, **16**, 177–192.

Lusk, E. J., & Kersnick, M. (1979). Effects of cognitive style and report format on task performance: The MIS design consequences. *Management Science*, **26**, 982–999.

Naylor, J. C., & Clark, R. D. (1968). Intuitive inference in interval tasks as a function of validity magnitude and sign. *Organizational Behavior and Human Performance*, **3**, 378–399.

Nisbett, R. E., & Kunda, Z. (1985). Perception of social distribution. *Journal of Personality and Social Psychology*, **48**, 297–311.

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.

Remus, W. (1987). A study of graphical and tabular displays and their interaction with environmental complexity. *Management Science*, **33**, 533–541.

Rothstein, R. G. (1986). The effects of time pressure on judgment in multiple probability learning. *Organizational Behavior and Human Performance*, **37**, 83–92.
Sniezek, J. A., & Reeves, A. P. (1986). Feature cues in probability learning: Data base information and judgment. *Organizational Behavior and Human Performance*, **37**, 297–315.
Sniezek, J. A. (1986). The role of variable labels in cue probability learning tasks. *Organizational Behavior and Human Performance*, **37**, 297–315.
Tversky, A., & Kahneman, D. (1982). Judgment of and by representativeness. In D. Kahneman, A. Tversky, & P. Slovic (Eds.), *Judgment under uncertainty: Heuristics and biases*. London/New York: Cambridge University Press.
York, K. M., Doherty, M. E., & Kamouri, J. (1987). The influence of cue unreliability on judgment in a multiple cue probability learning task. *Organizational Behavior and Human Performance*, **39**, 303–317.
Zukier, H., & Pepitone, A. (1984). Social roles and strategies in prediction: Some determinants of the use of base-rate information. *Journal of Personality and Social Psychology*, **47**, 349–360.
Zund, R. W. (1978). An empirical investigation of the dimensionality of the concept of information. *Decision Sciences*, **9**, 187–195.

RECEIVED: October 31, 1990