Journal of Behavioral Decision Making, Vol. 9, 125–139 (1996)

The Natural Selection of Prediction Heuristics: Anchoring and Adjustment versus Representativeness

BENJAMIN CZACZKES
The Hebrew University of Jerusalem, Israel

YOAV GANZACH
Tel-Aviv University, Israel

ABSTRACT

There are several heuristics which people use in making numerical predictions and these heuristics compete for the determination of prediction output. Some of them (e.g. representativeness) lead to excessively extreme predictions while others (e.g. anchoring and adjustment) lead to regressive (and even over-regressive) predictions. In this paper we study the competition between these two heuristics by varying the representation of predictor and outcome. The results indicate that factors which facilitate reliance on representativeness (e.g. compatibility between predictor and outcome) indeed lead to an increase in extremity, while factors that facilitate reliance on anchoring and adjustment (e.g. increased salience of a potential anchor) lead to a decrease in extremity.

KEY WORDS prediction; heuristics; representativeness; probability learning

Many phenomena in judgment and decision making can be explained by the operation of heuristics, or rules of thumb, that simplify the judgment and decision process. While most of the research in this area investigated the operation of single heuristics in isolation (for a summary see Kahneman, Slovic, and Tversky, 1982), recently some researchers have begun to study the interplay between heuristics (Agnoli and Krantz, 1989; Ganzach and Krantz, 1990). In particular, Agnoli and Krantz (1989) suggested a 'competing heuristic' model for the process by which heuristics are selected to determine judgment output (and also Brehmer, 1974, for a similar idea in the context of functional rule learning). According to this model, various heuristics compete for the determination of this output. Some of these heuristics are 'natural', while others are acquired through learning or experience. However, Agnoli and Krantz's work did not involve the study of competition among heuristics in the natural environment, since they examined only how rules acquired through formal training compete with natural heuristics. In this paper we investigate the process by which heuristics become dominant in the natural environment, i.e. an environment in which subjects do not receive any formal training. In particular, we show that in numerical predictions with feedback, the heuristic selected depends on how predictor and outcome are represented.

A heuristic that received considerable attention in previous research, and probably the most ubiquitous heuristic in intuitive predictions, is representativeness (Kahneman and Tversky, 1973).

CCC 0894–3257/96/020125–15

Received 10 March 1994
Accepted 25 October 1995

According to this heuristic, people acquire information about the distribution of the predictor and outcome, and on the basis of this knowledge choose a predicted value whose *standardized value* (a value akin to *Z*-score or percentile) matches the standardized value of the predictor. The representativeness heuristic 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. On the other hand, predictions by representativeness are non-regressive and, therefore, more extreme.

Reliance on representativeness can explain the extremism frequently observed in intuitive numerical predictions (e.g. Brehmer and Lindberg, 1970; Kahneman and Tversky, 1973). However, some findings cannot be explained by a model which assumes that representativeness is the only heuristic influencing prediction. First, this model cannot readily explain how people learn to regress from experience (Ganzach and Krantz, 1990; Ganzach, 1991). Second, this model cannot explain why predictions are sometimes excessively regressive (Yates and Jagacinski, 1979; Ganzach, 1993; and in particular see the results of Study 2 below).

In this paper we suggest that in order to account for these findings it is necessary to assume the operation of various heuristics that compete for the determination of prediction output. While one of these heuristics — representativeness — may lead to excessively extreme predictions, other heuristics may lead to more moderate predictions. One example for a heuristic that may compete with representativeness and lead to moderate predictions is the anchoring and adjustment heuristic (Slovic and Lichtenstein, 1971; Tversky and Kahneman, 1974). In relying on this heuristic in numerical prediction, people use as an anchor a salient value of the outcome, such as a central tendency value, and adjust based on the extremity (deviation from central tendency) of the predictor. Since adjustment is often insufficient, this strategy may lead to regressive (and even over-regressive) predictions.

The three studies reported here manipulate the interplay between representativeness and anchoring and adjustment by varying the representation of predictor and outcome. The main hypothesis of these studies is that a representation that supply people with a natural anchor (e.g. a previous prediction) or a representation that makes a potential anchor salient (e.g. the mean of the outcome distribution) enhances reliance on anchoring and adjustment and decreases reliance or representativeness.

In all three studies predictions are elicited in a single-cue probability learning (SCPL) task, a prediction task in which subjects are asked to use a predictor variable to predict an outcome variable (see Dudycha and Naylor, 1966, for an early example of such tasks, and Klayman, 1988, for a recent review). Extremity, operationalized as *prediction slope*, the regression slope relating predictions to predictor values, is used as an indicator for the heuristics used by the subject. In particular, we assume that if the representativeness heuristic is dominant, the prediction slope should be excessively steep, while if anchoring and adjustment is dominant, the prediction slope should be relatively moderate.

STUDY 1: THE REPRESENTATION OF PRICE CHANGES

In this study, subjects predicted the impact of a series of *changes in stock earnings* on prices of the stock. In one condition they predicted the new prices that followed the changes in earnings, while in the other they predicted the price changes that followed the same changes in earnings. Anchoring and adjustment is more likely to operate in the 'price' condition than in the 'price-change' condition, since the previous price is a natural anchor on which predictions can be based (note that since price reflects earnings, changes in earnings influence changes in price). Therefore, the central hypothesis of this study is that predictions are less extreme in the price condition than in the price-change condition.

Method

Subjects

Eighty-nine Business Administration students (45 in the price-change condition and 44 in the price condition) participated in the experiment to fulfill a class requirement. Subjects were assigned randomly to two conditions and participated in the experiment either individually or in small groups numbering two to three.

Procedure

Subjects in the price condition received written instructions informing them that:

> In the experiment you will be asked to predict the price of a share at the end of the quarter on the basis of the change in earnings during this quarter. The experiment will be conducted in the following way. At the beginning of each trial you will receive the change in earnings. After receiving this information, you will write your prediction. Subsequently, you will receive a feedback about the true price.

Subjects in the price-change condition received the same instructions, but the word 'price' was replaced by 'change in price'.

The instructions explained that there is a positive relationship between the change in earnings and the change in price, but that this relationship is probabilistic. A monetary price of 100 New Israeli shekels (about $40) was offered to the student who achieved the most accurate results. Subjects were told that the criterion for accuracy was minimum prediction error.

Subjects received six practice trials followed by 30 experimental trials. Each subject received two sheets of paper, one for the practice trials and one for the experimental trials, on which subjects were given the changes in earnings and were asked to write their predictions. The feedback of the 'true' price/price-change was supplied to the subjects by the experimenter after they wrote their prediction.

The two conditions were identical in the predictor, but they differed in the feedback. The feedback in the price condition was derived from the feedback in the price-change condition by adding the price-change to the previous feedback (i.e. the previous 'true' price). The first price, 98.3, was written on the top of the page containing the experimental trials.

Stimuli

Thirty values of earning change and price change were sampled from a bivariate normal distribution. The actual correlation between the two variables in the sample was 0.71, their means were about zero, and their standard deviations were 18.1 and 4.9, respectively, which resulted in a regression slope of 0.19. The order of the stimuli was reversed for half of the subjects.[1]

Results and discussion

To compare the predictions in the two conditions we transformed the predictions in the price condition to predictions of price change by subtracting each price prediction (i.e. each response) from

[1] The experiment also involved a 'trend' manipulation where in two conditions a positive constant was added to the price changes (this trend was presented to the subjects as resulting from inflation). Subjects in the trend/price condition saw a positive trend in the prices of the stock, while subjects in the trend/price-change condition saw only positive price changes. This manipulation did not have any effect and therefore was ignored in the analysis.

the previous price (i.e. the previous feedback). Subsequently, the prediction slope was calculated for each subject in each condition by regressing the predictions of price changes on the changes in earnings. The results indicated that predictions in the price-change condition were more extreme than predictions in the price condition ($t(87) = 2.1$, $p < 0.05$). The mean regression slope in the price-change condition is 0.224 (SD = 0.102), while in the price condition it is 0.182 (SD = 0.084). These results are consistent with the hypothesis that anchoring and adjustment moderate prediction in the price condition.

Note also that the difference in extremity between the two conditions does not result from the fact that the predictions in the price condition are more difficult, and therefore associated with more 'noise'. There is no significant difference in the mean unexplained variance of the two conditions, $t(87) = 1.0$, $p > 0.3$.

One additional factor that may lead to differences in extremity between the price and the price-change condition is that it is easier to use representativeness in the latter condition than in the former. The reason is that reliance on representativeness requires the knowledge of an outcome's distribution against which the standardized value of the predictor can be matched. Such distribution is easy to learn in the price-change condition, but is difficult to learn in the price condition. In the latter condition, an additional mental operation is involved — translation of the price feedback into a price change — before it is possible to learn the appropriate distribution. This issue is further explored in the next study.

STUDY 2: THE REPRESENTATION OF DISPERSION AND CENTRAL TENDENCY

The study reported here manipulates the interplay between representativeness and anchoring and adjustment by varying the representation of the dispersions and central tendencies of the predictor and outcome in a standard SCPL task. The predictor was centered around 0, 150 or 585, and the outcome was centered around 0 or 585. The standard deviation of the predictor was either 13 or 50, while the standard deviation of the outcome was always 50. The dispersion and central tendency of the predictor and outcome are varied in five conditions as presented in Exhibit 1. This design involves two sub-designs. In one (conditions D and E), the effect of compatibility in the predictor and outcome scales is examined. In the other (conditions A, B, C and D), the effect of compatibility in centering the predictor and outcome around zero and the effect of saliency of the central tendency of the outcome is examined. Since these two sub-designs are conceptually independent, we divided the discussion of this study into two sections.

Exhibit 1. The conditions of Study 2

		Outcome scale: M = 0, SD = 50	Outcome scale: M = 585, SD = 50
Predictor scale	M = 0, SD = 13	A	B
	M = 150, SD = 13	C	D
	M = 585, SD = 50		E

Method

Subjects

One hundred and twenty-one first-year Business Administration students participated in the experiment to fulfill and class requirement. Subjects were assigned randomly to the five conditions. They participated in the experiment in groups numbering four to eight.

Procedure

After entering the laboratory, subjects were seated in front of an IBM AT computer and told to read the initial instructions which explained the task. The instructions emphasized that 'The relationship between the predictor and the outcome is positive, that is, the higher the predictor, the higher the outcome'. They also emphasized 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'. The subjects then received 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 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, a number located in the center of the screen. Having displayed the predictor, the computer then prompted the subjects to type their prediction. After the prediction was typed, the outcome was displayed. Subsequently, the screen was erased, and a new trial began. No time limit was set for typing the predictions. To avoid reading inadvertent mistakes by the computer, the predictions were examined, and if they were completely out of range (more than eight standard deviations from the mean of the outcome), the subject was prompted to type his or her prediction again.

Stimuli

Four blocks of 30 trials were constructed by sampling a standardized predictor and a standardized random error from a bivariate standardized normal distribution with covariance 0. To be included in the experiment, a block was required to fulfill the following conditions: (1) The correlation between each of the two variables would not exceed ± 0.05; (2) The mean of each of the two variables could not exceed ± 0.05; and (3) The standard deviation of each of the two variables could not exceed 1 ± 0.05.

Standardized outcome feedback was generated by the equation

$$Y = 0.7071^{*}X + 0.7071^{*}\varepsilon$$

where Y is the outcome, X the predictor and ε is the error. The stimuli that subjects actually received were created by transforming the standardized predictors and outcome values to desired distributions (see Exhibit 1).

STUDY 2A

While in Study 1 the emphasis was on factors that influence reliance on anchoring and adjustment (i.e. the existence of a salient anchor), in this part of Study 2 the emphasis is on factors that influence reliance on representativeness. Reliance on this heuristic requires a number of mental operations. First, the distribution of the predictor and outcome has to be learned. Next, in each prediction, the value of the predictor has to be 'translated' to a standardized value. Finally, since prediction is required in terms of a raw score on the outcome scale, this standardized value has to be translated back into a value on the outcome scale.

All these operations are considerably facilitated when there is some compatability (Slovic, 1974; Slovic, Griffin, and Tversky, 1988; Tversky, Sattath, and Slovic, 1988) in the representation of the predictor and the representation of the outcome. One way by which compatibility can be achieved is when both predictor and outcome are represented on the same scale (i.e. same mean and standard deviation). We will label this compatibility *scale compatibility*. First, learning the distributions of predictor and outcome is easier, since there is only one distribution to learn.[2] Second, standardized values can be represented by raw scores, since a particular raw score represents the standardized value for both predictor and outcome. This may obviate the need for translating from predictor raw score to standardized value and the reverse translation from standardized value to outcome raw score.

Below we compare conditions D and E. In condition E, predictor and outcome are represented on the same scale, while in condition D they are represented on different scales. We hypothesize that the representation of both predictor and outcome on the same scale enhances reliance on representativeness, and obstructs the learning of regressive prediction strategies. (Slovic, 1974, used a similar experimental manipulation. However, in his experiment the compatible and incompatible conditions were associated with predictor and outcome having similar or different signs, respectively. In our experiment, the sign of the predictor and outcome is similar in both conditions.)

Results and discussion

The prediction slope was calculated for each subject and each 30-trial block. For the purpose of comparability between the two conditions (as well as comparability with the results of the second part of the study), the predictor values in condition E were standardized to have a mean of 150 and a standard deviation of 13 (that is, to have the same mean and standard deviation as condition D). The mean regression slope by condition and block are plotted in Exhibit 2.

The results of a 2×4 analysis of variance with repeated measures on the second factor (block) revealed neither main effect for condition ($F(1,46) = 0.80$, $p > 0.40$) nor main effect for block ($F(3,138) = 0.86$, $p > 0.50$). However, the interaction between these two factors was significant ($F(3,138) = 2.86$, $p < 0.05$). The interaction stems from a learning process in condition D, which is reflected in a decrease in the prediction slope during the course of the experiment (testing a null hypothesis of no linear trend revealed $t(24) = 2.3$, $p < 0.03$). While in the first two blocks the slopes exceed 2.72, the values of the optimal slope, in the last two blocks they are roughly equal to the optimal slope. On the other hand, in condition E the slopes exceed the optimal slopes in all four blocks, and there is no learning ($t(22) = 0.6$, $p > 0.5$ in testing for linear trend). This pattern of results is consistent with the notion that in condition E feedback does not lead to abandoning the representativeness heuristics, while in condition D there is a learning process in which reliance on representativeness decreases and regressive strategies are adopted.

There is also a difference between the two conditions in regard to the slopes' standard deviations ($F(24,22) = 8.80$, $p < 0.0001$, $F(24,22) = 2.90$, $p < 0.02$, and $F(24,22) = 7.20$, $p < 0.0001$ for the second, third, and fourth blocks, respectively. In the first block the difference is rather small ($F(24,22) = 1.4$, $p > 0.40$). This difference is most likely the result of differences in the adoption of regressive strategies.

[2] It still could be argued that even when predictor and outcome are represented on the same scale, subjects have to learn that the two distributions are the same. However, it is reasonable to assume that it is easier to learn that the two distributions are the same than to learn two different contributions (i.e. in Brehmer's (1974) the hypothesis that the two distributions are the same is stronger than the hypothesis that the two are different).

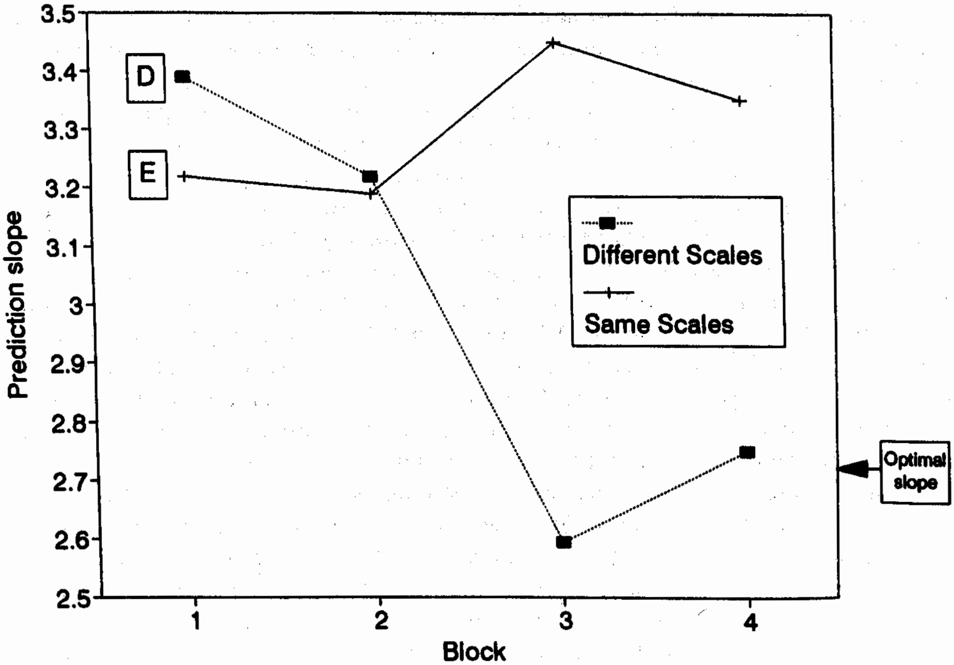


Exhibit 2. Mean prediction slope as a function of condition and block in Study 2A. Standard deviations for condition E are 1.56, 0.7, 0.91, 0.55, for blocks 1 to 4, respectively. Standard deviations for condition D are 1.87, 2.0, 1.55, 1.50

Since the adoption of regressive strategies varies across subjects in condition D, the prediction slope varies more in this condition than in condition E.

STUDY 2B

The first part of this Study demonstrates that under conditions that facilitate the use of representativeness, people do not learn to moderate their predictions. It is not clear from this part of the study what are the regressive strategies that underlie moderation in prediction, when such moderation indeed occurs. Anchoring and adjustment may be one of these strategies. People may learn to anchor their predictions to the outcome mean, which is a salient or representativeness value of the outcome distribution (Tversky and Kahneman, 1982). However, other heuristics that facilitate regressive predictions may also operate here. The design of the first part does not allow for conclusions about the operation of anchoring and adjustment, since the salience of the central tendency of the outcome is not manipulated. On the other hand, the design of the second part allows for a direct examination of the effect of the saliency of the central tendency of the outcome, and therefore an examination of the operation of anchoring and adjustment.

In this part of the study, subjects made predictions based on a predictor which was either centered or non-centered around zero and received an outcome (i.e. feedback) which was centered or non-centered around zero. This resulted in a 2 × 2 between-subject design with respect to these two factors. These factors are labelled as the Predictor Centering (PC) factor and the Outcome Centering (OC) factor.

Anchor saliency is associated in this study with the OC factor, since the central tendency of an outcome which is centered around zero is more salient than the central tendency of an outcome which is not centered around zero. The hypothesis is that when the outcome distribution is centered around zero, predictions are less extreme, since the saliency of the outcome's central tendency facilitates its use as an anchor in an anchoring and adjustment strategy. This hypothesis will be called *the outcome entering hypothesis.*

In addition, the design of the study allows for the examination of the effect of compatibility. Compatability is associated with whether the predictor and outcome are similarly centered. In two conditions (in the condition in which both are centered around zero, and in the condition in which they are both not centered around zero) there is such similarity, while in the other two conditions there is not. We call this *centering compatibility*. We expect that this compatibility will increase reliance on representativeness, since it facilitates matching by making the mental representation necessary for matching easier. While Study 3 directly focuses on the mental representation of predictor and outcome, here it is sufficient to say that in the conditions in which there is centering compatibility, reliance on representativeness is relatively easy. The reason for this is that when predictor and outcome are both centered around zero (and the central tendency is salient), they are likely to be mentally represented for the purpose of matching as deviations from zero (e.g. a Z-score-like representation), while when they are both *not* centered around zero (and the central tendency is not salient) they are likely to be represented by their size, or how 'big' they are (e.g. a percentile-like representation. Conversely, when there is no centering compatibility, matching is more difficult, and therefore less reliance on representativeness should be expected. This reasoning suggests an interaction between the PC and OC factors. We label this interaction hypothesis *the compatibility hypothesis*.

In summary, in this part of Study 2 we manipulate both reliance on representativeness, which is associated with compatibility, and reliance on anchoring and adjustment, which is associated with outcome centering.

Results and discussion

The mean prediction slope by condition (collapsing over blocks) are plotted in Exhibit 3, and the means by condition and block are plotted in Exhibit 4.[3] The between-subjects results of a three-way mixed ANOVA (Block × OC × PC) with repeated measures on block revealed a significant main effect for the OC factor ($F(1, 94) = 16.0$, $p < 0.0001$) indicating that, in line with the outcome-centering hypothesis, predictions are less extreme when the outcome is centered. (The main effect for PC was not significant ($F(1, 94) = 0.26$, $p > 0.9$.) The between-subjects results also revealed a significant interaction between OC and PC ($F1, 94) = 4.9$, $p < 0.03$), indicating that, in line with the compatibility hypothesis, for each level of the OC factor, compatibility with respect to centering tends to increase extremity. Note that the between-subjects effects are stronger in the beginning of the experiment, and weaken towards the end. The variance explained by the between-subjects factors are 0.24, 0.17, 0.07 and 0.10 in the first to the fourth block, respectively. This pattern is due to a learning process that weakens biases due to information representation.

The learning process is apparent from the within-subjects results of the three-way ANOVA. These results revealed significant interaction between block and OC ($F(3, 282 = 7.46$, $p < 0.0001$). While in the first block prediction extremity is very low when outcome is centered, and much higher when it is not, there is not much difference between the conditions in the last block. As can be seen in Exhibit 4,

[3] Note that the regression slopes of condition E of Study 2A can be compared to the regression slopes in Exhibit 4, since the regression slopes of condition D appear in both Exhibit 2 and Exhibit 4.

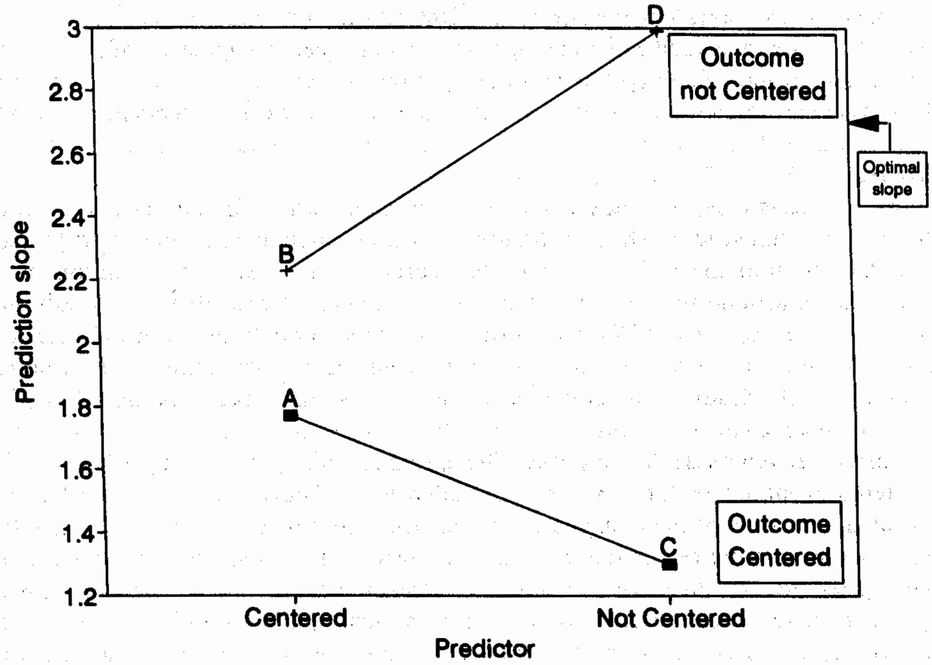


Exhibit 3. Mean prediction slope as a function of condition in Study 2B. Data are collapsed over blocks

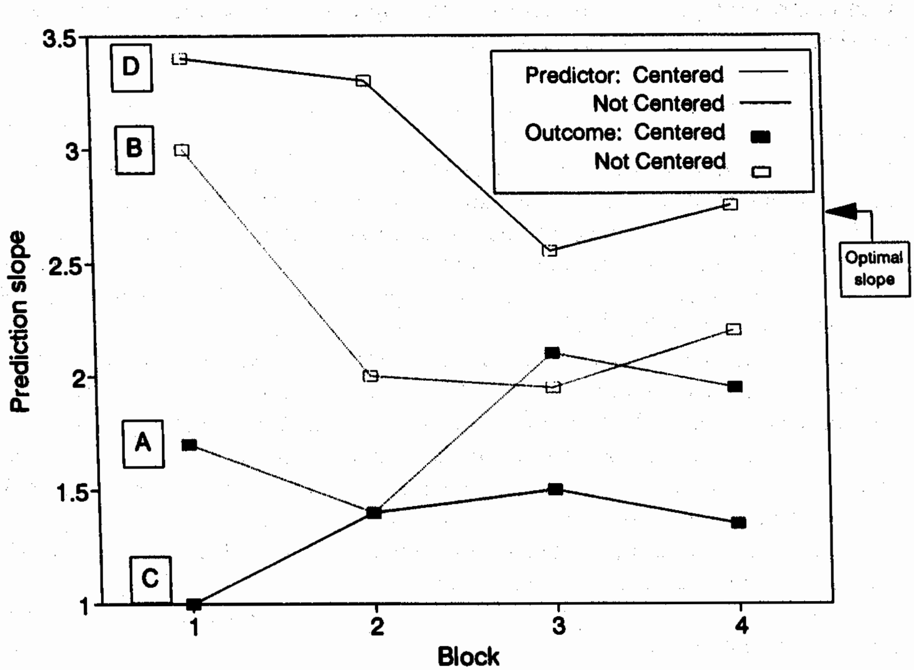


Exhibit 4. Mean prediction slope as a function of condition and block in Study 2B. Standard deviations for condition A are 1.11, 1.42, 1.05, 1.04 for blocks 1 to 4, respectively. Standard deviations for condition B are 2.60, 1.89, 1.90, 2.09. Standard deviations for condition C are 0.77, 1.16, 1.16, 1.07. Standard deviations for condition D are 1.87, 2.08, 1.56, 1.51

an increase in extremity occurs when the outcome is centered (tests for linear trend revealed $t(23) = 1.9$, $p<0.07$; $t(23) = 2.0$, $p<0.06$ for conditions A and C, respectively). On the other hand, extremity decreases when the outcome is not centered ($t(24) = 1.5$, $p<0.15$; $t(24) = 2.3$, $p<0.03$ for conditions B and D, respectively). Both excess extremism and excess conservatism tend to disappear as a result of exposure to outcome feedback.[4]

A comparison of the prediction slope of the first block in the conditions in which outcome is centered to the optimal slope reveals considerable over-regressiveness. The mean slopes for conditions A and C are 1.7 and 1.0, while the optimal slope is 2.72. The null hypothesis of no difference between these two slopes and the normative slope was rejected, $t(23) = 4.4$, $p<0.0002$ and $t(23) = 11.0$, $p<0.0001$ for conditions A and C, respectively. While over-regressiveness in predictions was observed before (Yates and Jagacinski, 1979; Ganzach, 1993), the extent of over-regressiveness in this study far exceeds these previous observations. This over-regressiveness is consistent with an anchoring and adjustment process: Since adjustment is often insufficient (Tversky and Kahneman, 1974), over-regressive predictions may occur.

It is important to note that it is not likely that the differences in extremity observed in the experiment are due to differences in prediction noise. First, the between-condition differences in extremity cannot be explained by between condition differences in prediction noise. A three-way mixed ANOVA on the unexplained variance in subjects' predictions did not reveal a main effect, neither for the PC factor ($p>0.6$), nor for the interaction ($p>0.3$). It revealed a main effect for the OC factor, $F(1,94) = 11.0$, $p<0.001$, but this effect is contrary to what is to be expected from a 'noisy prediction explanation', since it is associated with greater noise when the outcome is not centered than when it is centered (i.e. it is associated with a positive relationship between prediction slope and prediction noise). The ANOVA on the unexplained variance revealed also a significant main effect for block, $p<0.0001$. This effect is associated with a decrease in noise throughout the course of the experiment, which results from the learning of the linear relationship of the task. This effects replicate earlier findings (see, for example, Dudycha and Naylor, 1966).

Second, by examining the unexplained variance in subjects' predictions it is also possible to rule out the artifactual explanation that the over-regressiveness observed in the experiment stems from prediction noise. For example, in the first block, not only is the slope of condition B greater than the slope in condition A but so also is its unexplained variance; $t(1,47) = 2.2$, $p<0.05$ and $t(1,47) = 3.3$, $p<0.002$, respectively. Thus, contrary to a 'noisy-predictions' explanation, predictions are excessively regressive when there is a little noise and excessively extreme when there is a lot of noise.

STUDY 3: ON THE MENTAL REPRESENTATION OF STANDARDIZED VALUE

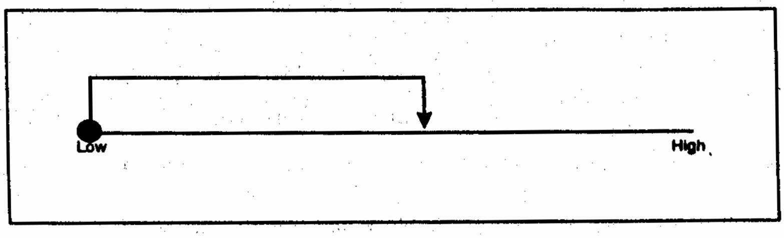
The results of Study 2B suggest that there are two ways by which standardized values can be mentally represented. First, they may be represented by their size. A value can be perceived as being 'small', 'large', 'somewhat larger', 'moderate', etc; that is, the mental counterpart of a percentile score. We will label this representation *size representation*. Second, a standardized value may be represented by its extremity, or deviation from a central tendency value; that is, the mental counterpart of a Z-score. We will label this representation *deviation representation*.

[4] It appears from Exhibit 5 that there is an asymptotic prediction slope which is common to all conditions. This slope appears to be lower than the optimal (least-squares) slope (2.7). However, comparison with the optimal slope is questionable, because noise in the data (e.g. subjects who did not understand the instructions, lack of motivation, key punching errors and many other sources) is likely to have a uni-directional influence on the asymptotic slope, that is, it is likely to decrease (but not increase) the slope relative to the optimal slope.

An increased tendency towards size representation is likely to enhance reliance on representativeness, because it makes matching easy. On the other hand, an increased tendency towards deviation representation may enhance both reliance on representativeness and reliance on anchoring and adjustment. It may enhance reliance on representativeness because it makes matching easier, but it may also increase reliance on anchoring and adjustment because the central tendency of the outcome — being highly salient — is more likely to serve as an anchor. Therefore, other things being equal, we expect higher reliance on representativeness when the mental representation of standardized values is size representation and more reliance on anchoring and adjustment when the mental representation of standardized values is deviation representation.

In the current experiment, we manipulate the mental representation of standardized value by manipulating the *physical* representation of the predictor. The value of the predictor is shown to the subject by moving an indicator on a scale. In the *size representation condition* the natural reference point to evaluate the value of the predictor is the origin of the scale, while in the *deviation representation condition* this reference point is the middle of the scale. Exhibit 5 shows the way the predictor was presented in these two conditions, respectively. We hypothesize that the physical representation of the predictor would influence the mental representations of the predictor, the mental representations of the outcome, and as a result, the tendency to rely on representativeness versus anchoring and adjustment. A size (extremity) representation of the predictor would increase the tendency towards size (extremity) representation of the outcome, which, in turn, would increase reliance on representativeness (anchoring and adjustment).

A



B

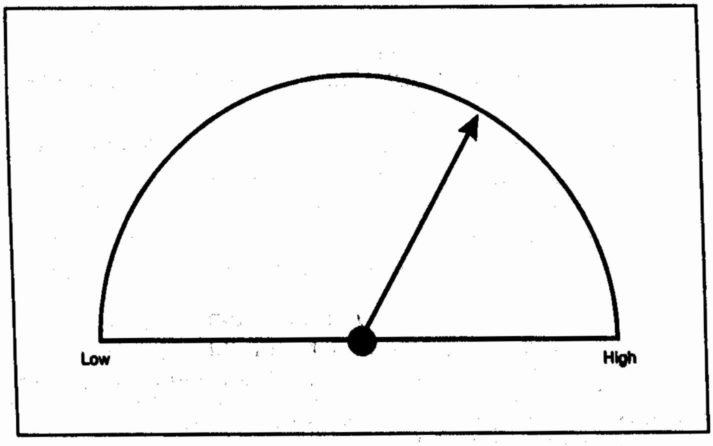


Exhibit 5. The predictor scale in the size representation condition (A) and in the deviation representation condition (B)

Method

Subjects

Forty-three Business Administration students (21 in the size representation condition and 22 in the deviation representation condition) participated in the experiment to fulfill a class requirement. Subjects were assigned randomly to two conditions and participated in the experiment either individually or in small groups numbering two to three.

Procedure

In the experiment, subjects predicted the stopping distance of a car based on its velocity. The velocity was presented on the 'speedometers' of Exhibit 5. The length of the line in Exhibit 5(A) and the length of the arch in Exhibit 5(B) were both 23.9 centimeters. The experiment was run on a computer, and included 95 trials, five practice trials and 90 experimental trials. At the beginning of each trial, the indicator moved from its resting point in the middle of the arch (in the deviation representation condition) or the beginning of the line (in the size representation condition) and stopped at a point that represented the value of the velocity. Subsequently, the computer prompted the subjects to type their predictions. After they typed their predictions, the feedback appeared on the screen. If the prediction was completely out of range, the subject was prompted to type the prediction again.

The feedback was generated by the equation $Y = -82 + 20.18^{*}(X + \varepsilon)$ where Y is the feedback, X is the velocity, measured by the distance (in centimeters) of the indicator from the origin of the speedometer (that is, distance from the 'low' anchor) and ε is an error term. X and ε were orthogonal, and each of them explained half of the variance in Y. The values of X and ε were sampled from a normal distribution with a mean of 11.95 and standard deviation of 4.6 (only values of X and ε which were within 2.6 standard deviations from the mean were allowed). As a result, the mean of Y was 400 and its standard deviation was 130. Only small deviations from these population parameters were allowed in each of the blocks (see Study 2).

Exhibit 6. Mean regression slope by block and condition in Study 3

Block	Condition	
	Extremity representation	Size representation
1	21.04	29.9
	(8.8)	(7.5)
2	20.7	25.3
	(6.3)	(5.7)
3	21.3	23.4
	(5.1)	(6.1)

Note: Numbers in parentheses are standard errors.

Results and discussion

The mean regression slope by 30 trial block and condition are given in Exhibit 6. A 2×3 mixed ANOVA with repeated measures on the second factor (block) revealed a strong main effect for condition, $F(1, 41) = 9.7$, $p < 0.005$. This effect is consistent with the hypothesis that there is more

[5] The difference in extremity between the two conditions does not result from differences in prediction noise, since the two conditions did not differ in their unexplained variance ($p > 0.5$). In addition, the 2×3 mixed ANOVA on the unexplained variance did not reveal significant block effect ($p > 0.4$) or significant condition × block interaction ($p > 0.6$).

reliance on representativeness in the size representation condition and more reliance on anchoring and adjustment in the deviation representation condition. The analysis also revealed a main effect for block, $F(2, 82) = 5.2$, $p < 0.01$, but this main effect should be understood in light of the significant interaction between block and condition, $F(2, 82) = 4.4$, $p < 0.02$. This interaction is associated with a learning process in the size representation condition, but not in the deviation representation condition. In the former condition, predictions are excessively extreme in the first and second blocks of the experiment $t(21) = 6.1$, $p < 0.0001$ and $t(21) = 4.3$, $p < 0.0005$, respectively, while in the last block they approach the optimal slope (20.0).[5]

Comparison with Study 2B

It could be argued that the two conditions of this experiment are similar to the two conditions of Study 2A in which the outcome was not centered, and that the size/deviation representation manipulation is similar to a compatibility manipulation. Thus, compatibility may be an alternative explanation for the difference between the two conditions of the current experiment and size versus deviation representation may be an alternative explanation for the difference between conditions B and D of Study 2B. In our view, however, different processes underlie the results of the two experiments, because the size versus deviation representation explanation cannot explain the differences between conditions A and C in Study 2B. If anything, this explanation would suggest that predictions would be more extreme in the predictor-not-centered condition than in the predictor-centered condition, which is the opposite of what is observed in the experiment.

GENERAL DISCUSSION

There are various natural heuristics that may determine prediction output. The selection of the heuristic that will actually determine prediction depends, to a large extent, on the representations of the predictor and outcome. Such representations favor one heuristic over others if they make the implementation of this heuristic easier. Thus, saliency of a potential anchor increases reliance on anchoring and adjustment, while compatibility between predictor and outcome increases reliance on representativeness. Similarly, deviation representation increases anchoring and adjustment, while size representation increases representativeness.

The results of Studies 2 and 3 also indicate that except for predictor and outcome representation, experience with outcome feedback is also involved in the selection of prediction heuristics. So far, the discussion of the process of learning from experience in SCPL was concerned mainly with the learning of the functional form of the rule relating predictor to outcome (Brehmer, 1974; Klayman, 1988). In this paper, we are more interested in the prediction strategies people use to implement an appropriate (linear) rule in a SCPL task, and in the way these strategies are changed with experience (i.e., how people improve their predictions as a result of feedback). In this sense, the distinction between the learning process described in the current paper and the learning process described in the rule learning literature, is similar to the distinction between application and acquisition (Hammond and Summers, 1972) in multiple-cue probability learning. The SCPL counterpart of acquisition is the process of functional rule learning, and the SCPL counterpart of application is a process by which inaccurate heuristics are replaced by accurate heuristics, or alternatively, a process by which heuristics are changed to become more accurate.

Many facets of functional rule learning can be accounted for by a model, according to which people do not extract the functional rule from outcome feedback, but rather come to the experiment with a limited amount of 'natural' hypotheses about the rule, and test these hypotheses against the feedback

(Brehmer, 1974). In this model, the rules that are tested are ordered hierarchically. Some are 'stronger' than others, i.e. they are considered earlier and more often than others (e.g. a positive linear rule is stronger than a curvilinear one). Learning from experience is viewed as a change in rule strength. The strength of the rule that generates appropriate predictions increases, while the strength of other rules decreases. For example, a single peaked rule gains strength during the course of the experiment when the functional relationship between predictor and outcome is a single peaked relationship (Brehmer, 1974).

Similarly, the results of Studies 1 and 2 can be accounted for by a model in which people do not derive prediction strategies from outcome feedback, but rather come to the experiment with a limited number of natural prediction heuristics. These heuristics compete for the determination of prediction output. The competition is characterized by differential heuristics' strength. Some heuristics have a higher probability of being used than others. This probability depends on two factors. One factor is the representation of predictor and outcome. Thus, in Study 2B (at least in the earlier phase of the experiment) outcome representation influences the competition between representativeness and anchoring and adjustment, while compatibility influences the competition between representativeness and other heuristics that lead to inconsistent predictions. The other factor is experience. Outcome feedback increases the strength of heuristics that produce more accurate predictions, and decreases the strength of heuristics that lead to inaccurate predictions. Thus, in Studies 2 and 3, heuristics that lead to moderate predictions gain in strength with experience when initial prediction is excessively extreme and lose strength when initial prediction is excessively regressive. The dynamics of the competition between heuristics could be viewed, therefore, as an evolutionary process in which there is competition between various heuristics, and those that fit the environment (outcome feedback) are continuously selected.

ACKNOWLEDGEMENTS

We thank Offer Erez and Arye Zvetler for their help in Study 1. Correspondence should be sent to Yoav Ganzach.

REFERENCES

Anderson, N. H. *Foundations of Information Integration Theory*, New York: Academic Press, 1981.

Agnoli, F. and Krantz, D. H. 'Suppressing natural heuristics by formal instructions: The case of the conjunction fallacy', *Cognitive Psychology*, **21** (1989), 515–50.

Brehmer, B. 'Hypothesis about the relations between scaled variables in the learning of probabilistic inference tasks', *Organizational Behavior and Human Performance*, **5** (1974), 1–27.

Brehmer, B. and Lindberg, L. '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** (1970), 542–54.

Dudycha, L. W. and Naylor, J. C. 'Characteristics of the human inference process in complex choice behavior situation', *Organizational Behavior and Human Decision Processes*, **1** (1966), 110–28.

Fitts, P. M. and Seeger, C. M. 'S-R compatibility: Spatial characteristics of stimulus and response codes', *Journal of Experimental Psychology*, **46** (1953), 199–210.

Ganzach, Y. 'Predictor representation and prediction strategies', *Organizational Behavior and Human Decision Processes*, **56** (1993), 190–212.

Ganzach, Y. 'Feedback representation and prediction strategies', *Organizational Behavior and Human Decision Processes*, **59** (1994), 391–49.

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

Ganzach, Y. and Krantz, D. H. 'The psychology of moderate prediction: II. Leniency and uncertainty', *Organizational Behavior and Human Decision Processes*, **48** (1991), 169–92.

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

Kahneman, D. and Tversky, A. 'On the psychology of prediction', *Psychological Review*, **80** (1973) 237–51.

Klayman, J. 'On the how and why (not) of learning from outcomes', Brehmer, B. and Joyce, C. R. B. (eds), *Human Judgment: The Social Judgment Theory View*, pp. 115–62, Amsterdam: Elsevier Science Publishers.

Kornnblaum, S., Hosbrac, T. and Osman, A. 'Dimensional overlap: cognitive basis for stimulus–response compatibility — A model and taxonomy', *Psychological Review*, **97** (1990), 253–70.

Nisbett, R. E., Krantz, D. H., Jepson, C. and Kunda, Z. 'The use of statistical heuristics in everyday inductive reasoning', *Psychological Review*, **90** (1983), 339–63.

Slovic, P. 'Hypothesis testing in learning of positive and negative linear functions', *Organizational Behavior and Human Decision Processes*, **11** (1974), 368–76.

Slovic, P., Griffin, D. and Tversky, A. 'Compatibility effects in judgment and choice', in Hogarth, R. M. (ed.), *Insights in Decision Making: Theory and Applications*, Chicago: University of Chicago Press.

Slovic, P. and Lichtenstein, S. 'Comparison between Bayesian and Regression approaches to the study of information processing in judgment', *Organizational Behavior and Human Decision Processes*, **6** (1971), 649–744.

Tversky, A. and Kahneman, D. 'Judgment under uncertainty: Heuristics and biases', *Science*, **185** (1974), 1124–31.

Tversky, A. and Kahneman, D. 'Judgment of and by representativeness', in Kahneman, D., Tversky, A. and Slovic, P. (eds), *Judgment under Uncertainty: Heuristics and biases*, Cambridge, MA: Cambridge University Press.

Tversky, A., Sattath, S. and Slovic, P. 'Contingent weighting in judgment and choice', *Psychological Review*, **95** (1988), 371–84.

Tversky, A., Slovic, P. and Kahneman, D. 'The causes of preference reversal', *The American Economic Review*, **80** (1990), 204–17.

Yates, J. F. and Jagacinski, C. M. 'Nonregressiveness of subjective forecasts', in *Proceedings of the 11th Annual Meeting of the American Institute of Sciences*, New Orleans, 1979.

Authors' biographies:

Yoav Ganzach is a senior lecturer at the Faculty of Management at Tel-Aviv University. He received a Ph.D. from Columbia University in 1988. His area of interest is behavioral decision making and its application to organizational behavior and consumer behavior.

Benjamin Czaczkes is currently finishing his Ph.D. at the Hebrew University. He received an MBA from the School of Business Administration of the Hebrew in 1991. His area of interest is behavioral decision making, management information systems and human-computer interaction.