DECISION MAKING UNDER UNCERTAINTY
R.W. Scholz (editor)

VERBAL VS. NUMERICAL PROCESSING OF SUBJECTIVE PROBABILITIES

ALF C. ZIMMER

Universität Münster
Psychologisches Institut
F.R. Germany

The problem of deciding between different alternatives with equally positive outcomes has an apparently trivial solution: take the one where the positive outcome is most probable. That this solution is not as trivial as it seems to be can be seen in the fact that it was not until 1660 for the concept of probability to be treated mathematically (HACKING, 1975), thereby providing solutions for how to decide in the case of uncertain events.

As early as 1713 JACQUES BERNOULLI has made the distinction between probabilities - later defined by LAPLACE as the relative frequencies of successes - and subjective"degrees of confidence". By means of this distinction it became possible to account for inconsistencies or irrationalities in human decisions. The course of action taken was assumed to depend on these subjective "degrees of confidence" instead of the objective probabilities as derived from long-run experience. Practical as well as philosophical considerations then led to various attempts to define subjective probability formally (e.g. DE FINETTI, 1937) and to prescribe procedures of determining whether numerical estimates of probabilities given by human raters fulfil the axioms of subjective probabilities (The system of axioms most often investigated is the one due to SAVAGE, 1954). A thorough critique of this approach can be found in SLOVIC & TVERSKY (1974) and SUPPES (1974). LINDLEY (1974) adds a further discussion from the point of view of Bayesian statistics. The core of both SUPPES' and TVERSKY's critique is that human decision makers and probability raters markedly and systematically derivate from what is prescribed by SAVAGE's theory , especially concerning the axiom of indifference.

From the point of view of human information processing the framework for investigating this problem is given in table 1.

Table 1

Human information processing depends on:

1. Modes of representation

(i) propositional
- verbal
- numerical
- generally symbolic

(ii) analogue
- imaginal
- automatic frequency monitoring

these modes differ in: - the transformations which can be applied. The constraints on these representations can be "strict" or "elastic".

- the mental work they impose on the information processing capacity

2. Available procedures

(i) rules (e.g. grammars, arithmetics, "Gestalt laws" in perception, etc.) these interact with the above mentioned applicable transformations and are therefore context specific.

Less context specific are

(ii) heuristics, which can be regarded as tools for narrowing down the number of possible candidates among the transformations in the case of insufficient information

Least context specific and therefore applicable in situations of information overload are

(iii) rules of thumb, which can be applied almost without any constraints. They provide "quick and dirty" procedures for the reduction of information so that the processing overload is relieved. (e.g."Take whatever comes to your mind first".)

Most investigations of subjective probabilities have been much more concerned with the underlying procedures (see bottom half of table 1) than with the modes of representation (see upper half of table 1). These investigators have implicitly assumed that information is stored propositionally and that after retrieval it can be stated numerically. Any inconsistency between information intake (the objective side) and the numerically expressed subjective probability (the subjective side) is then ascribed to the procedures. That is, the subject is presumed to have chosen an inappropriate algorithm. Examples for such proposed procedural mistakes are overconfidence (see Figure 1), conservatism (i.e. sticking to an initial appraisal of a situation in spite of new information available for revision), and negligence of the regression effect (i.e. the implicit assumption of a perfect correlation between the predictor variable and the criterion).

The approach taken here begins by asking how uncertain events might be represented internally. That is, are they represented in (i) a verbal propositional mode, (ii) a numerical propositional mode, or (iii) in an analogue mode of automatic frequency monitoring (see the next chapter). Thus, what is questioned here is not the adequacy of human judgments concerning uncertainty, but rather the individuals' ability to express numerically what is internally represented.

Coping with uncertainty is ubiquitous in humankind (WRIGHT and PHILLIPS, 1980) and expressions for different degrees of certainty can be found in most languages. Yet, as reported above it took until the 17th century to develop these concepts mathematically. Even then it took place only in the European culture where it was motivated by practical problems, such as gambling and insurance. It seems unlikely that the mathematically appropriate procedures with numerical estimates of uncertainty have become automatized since then. It is more likely that people handle uncertainty by customary verbal expressions and the implicit and explicit rules of conversation connected with them. In the research reported here I therefore start from the analysis of meaning of common verbal expressions for uncertain events. These expressions are interpreted as possibility functions (ZADEH, 1978) and the procedures applicable to them (rules as well as heuristics and biases) are modelled in the framework of possibility theory. Since this theory allows

The Automatic Monitoring of Frequencies

Before developing this theory in detail for the purposes of our study it seems to be worthwhile to ask how people gather the knowledge which is reflected in the verbal expressions for the probabilities of uncertain events. One can assume that for repetetive events e.g. the outcomes of ballgames or the daily weather, they monitor the frequency of outcomes automatically and that they revise their knowledge accordingly. HINTZMAN, NOZAWA, and IRMSCHER (1982) have studied this automatic monitoring of frequencies and they report that it is not interfered with by numerical processing. From their result they conclude that this information is stored in a non-numerical analogue mode (see item 1 (ii) in Table 1). Such automatic monitoring of frequencies seems to be a plausible candidate for the mode of representation underlying the generation of judgments concerning subjective probabilities.

In an experiment I have tried to determine whether the frequencies of more than one unattended stimulus attribute can be monitored in order to determine the degree of automaticity and mutual interference of multiple frequency judgments. In the experiment 150 color slides depicting either landscapes or buildings were presented to 80 Introductory Psychology students in a probability learning task, in which they had to predict the content - building vs. landscape - of each upcoming slide. Afterwards they were asked for the relative frequencies of two presumably attended features of the slides, orientation and form, which were uncorrelated with the content of the slides. The question format was either a numerical scale, on which subjects had to mark their frequency estimate, or a sequence of verbal expressions for relative frequencies, out of which they had to pick the best fitting one. For the exact experimental design and the results (mean squared errors) see Table 2.

The results show (i) the unattended features differed in saliency (orientation was easier to monitor than form), (ii) the verbal judgments were slightly better than the numerical ones, (iii) in all cases the second frequency judgments were worse than the first ones, but more important, the impairment was less in all cases where one of the judgments was verbal and it was least when both were verbal. From these results it seems plausible to conclude, first, that more than one unattended variable can be automatically monitored even if they are closely related, but the mode of probing this

Table 2

The verbal judgments have been transformed into numerical values by assuming that the labels partition the range (0% - 100%) equally.

mean squared errors		mode of judgment			
		verbal		numerical	
		other position		other position	
		verbal judgm.	numerical j.	verbal judgm.	numerical j.
Form	position 1	10.1	9.8	12.3	13.1
Form	position 2	15.5	16.2	17.4	22.0
Orientation	position 1	5.7	7.1	8.1	7.9
Orientation	position 2	9.2	8.7	13.8	19.5

knowledge influences the precision of judgments. Second, if more than one judgment has to be made there is inference between them, but the amount of inference depends on the modes of judgments. This apparent superiority of the verbal mode leads to the tentative interpretation that the verbal mode for representing knowledge is able to process information more effectively than the numerical mode, because the rules applicable to processing verbal propositions have been acquired earlier and are therefore more automatized. This would explain why they require less mental processing capacity and consequently are less prone to interference from simultaneous processes. If the overload of the mental processing capacity necessitates the application of heuristics, and if in turn biases in human judgment can be traced back to mistaken applications of heuristics (for an overview see NISBETT and ROSS, 1980), the following conjecture seems to be plausible: Any mode of judgment imposing less mental work load should be more valid than one requiring more mental processing capacity. This conjecture motivated the following experiments, in which I have looked for possible ways of making the information processing easier in tasks which imply the analysis of uncertain events. In order to analyze the efficiency of different modes of representation in information processing it is necessary to devise an interpretation for the meaning of verbal expressions which can be translated into numerical values. Fuzzy set theory provides a tool for such interpretation by means of numerically stated elastic constraints for the applicability of verbal expressions.

A Fuzzy Set Theoretical Approach to Subjective Probabilities

According to fuzzy set theory the meaning of concept "$\underset{\sim}{x}$" in a universe of discourse "U" can be modelled by the possibility function for "$\underset{\sim}{x}$" in "U", which indicates for which states in "U" the concept "$\underset{\sim}{x}$" fits, for which it can be possibly applied, and for which it does not fit at all. Figure 2 depicts the possibility functions for two concepts as well as the basic two binary operations:

(i) the intersection $\underset{\sim}{A} \cap \underset{\sim}{B} = \underset{U}{\mathrm{MIN}} \, (f_{\underset{\sim}{A}} \; ; \; f_{\underset{\sim}{B}})$

(ii) the disjunction $\underset{\sim}{A} \cup \underset{\sim}{B} = \underset{U}{\mathrm{MAX}} \, (f_{\underset{\sim}{A}} \; ; \; f_{\underset{\sim}{B}})$

The version of fuzzy set theory applied here, especially concerning the form of possibility functions, deviates from approaches where the exact form of these functions is determined either theoretically (for an overview see KAUFMANN, 1975, or DUBOIS and PRADE, 1980) or empirically (e.g. HERSH and CARAMAZZA, 1976; ZIMMER, 1980, 1982 a). The appropriateness of these specific membership functions as parts of the mental representation of vague concepts

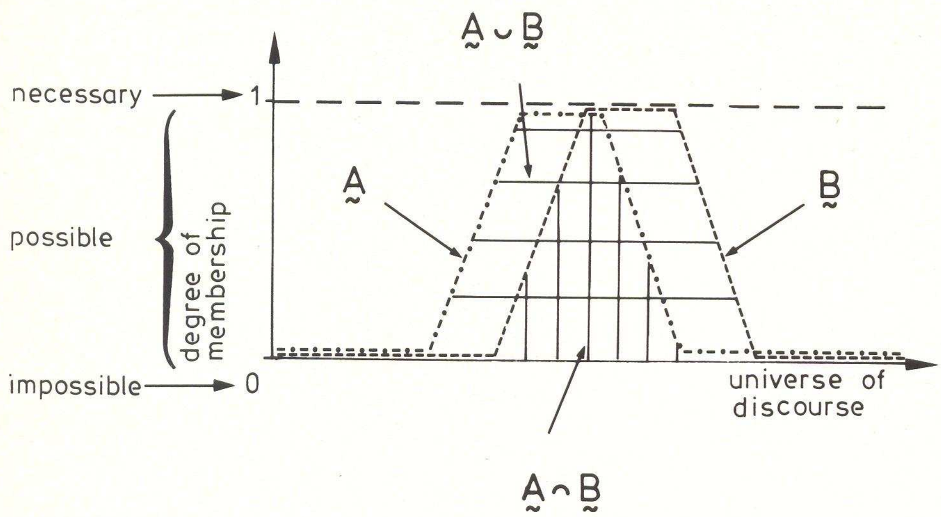

Figure 2. Possibility distributions and binary operations upon them

is at least doubtful. FREKSA (1981, 1982) has shown that for practical purposes such specification is unnecessary. By restricting the possibility to the values: "necessary", "possible", and "impossible", one can avoid the debated question of how exact the information about one's own knowledge can be (cf. NISBETT and WILSON 1977) without losing the amount of specificity necessary to apply fuzzy set theory to verbal concepts. The meanings of these concepts is then given by the elastic constraints imposed upon them by the possibility functions.

Modelling the meanings of verbal labels for relative frequencies is straightforward in this framework. The universe of discourse is the unit interval and the regions of applicability, possible applicability, and inapplicability can then be determined empirically. The resulting possibility functions for a given set of verbal expressions is depicted in Figure 3. The spacing of the possibility functions in the unit interval is reminiscent of WITTE's (1960) treatment of verbal judgments stabilized in memory over time. I (ZIMMER, 1980) have shown that such equidistant and equally shaped categories are optimal from the point of view of language pragmatics, but it has to be kept

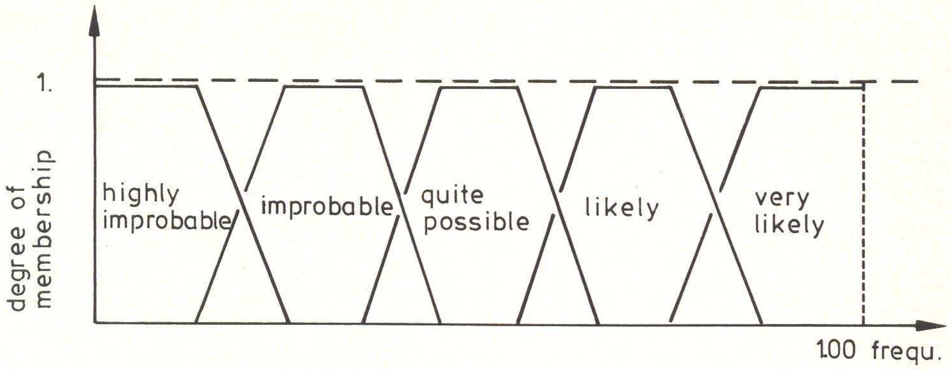

Figure 3. Possibility distributions for the meaning of verbal expressions for uncertainty for a subject using 5 labels

in mind that this communicability constraint is a simplifying assumption, which is not a prerequisite for this kind of modelling. WITTE's approach, however, implies two questionable assumptions: that the verbal labels are uniquely ordered for all subjects and that all subjects are equally familiar with the labels provided by the experimenter.

Empirical Determination of the Fuzzy Meanings of Verbal Expressions

In the preliminary step for my experiment I observed the subject's usage of expressions for uncertain events and asked them afterwards to order the used labels according to frequency. It turned out that the subjects differed markedly in verbal labels used and moreover in the serial order of the same labels used by different subjects. For instance some subjects reported the increasing serial order "possible, likely, probable" whereas others interchanged "possible" and "likely" or "likely" and "probable" (WALLSTEN, note 1, reports similar results). In the following it will be assumed that subjects used these expressions in the way they reported. As it will turn out later this assumption is in line with the empirical results.

The elicitation of the verbal labels that subjects were most familiar with and used most frequently, revealed that most subjects used merely five or six different expressions (see Table 3). A total of twelve different labels was

Table 3

number of labels for uncertain events	4	5	6	7	8
number of subjects	2	87	55	5	1

found. The empirical definitions of the meanings had to be done individually because of the interindividual differences. From the point of view of language pragmatics these individual differences in using the same expressions are a puzzling aspect because they contradict the assumed communicability constraint and cast into doubt the pragmatic value of such verbal judgments. From the interviews with the subjects in the debriefing sessions I learned that these marked individual differences were partially due to the compositions of our sample of subjects: 30 students in psychology, 30 students in business and economy, and 90 drafted soldiers of the German Bundeswehr. Whereas the first two groups had finished courses in statistics and probability theory, in the last group only a small proportion had studied probability theory in high-school. When asked what they would do if they realized that somebody else might misunderstand their verbal frequency expressions, most subjects reported that in such cases they would use the reference to generally known standard situations as a means to resolve this misunderstanding with an analogy. For instance they might point out that a certain event is as probable as winning in the German Lottery.

The empirical determination of the meanings for the individually used verbal expressions is done in the following steps:

(i) survey of recurrently used labels for relative frequencies

(ii) test of equivalence for labels (Do some of the words mean the same for you?)

(iii) determination of the ranges for the labels, s(k), assuming that the labels partition the overall range equally

(iv) sequential estimation of the empirical meanings for the labels used by the subjects under examination:

The first step in the experimental determination of the meanings for the verbal expressions used by the subjects consisted in determining the starting point for the estimation procedure. This was done by taking the midpoint of the range in case of an uneven number of the verbal labels, that is, 50%, or in the case of an even number of categories a starting position was taken slightly moved to the lower or upper part of the range, that is, either 33% or 67%. In order to represent objective relative frequencies for the calibration of the individually used verbal expressions, I showed a window on a computer-driven tv-screen which containted up to 6144 dots (100%). The number of the white dots spaced randomly in the window in relation to the size of the screen represented the objective relative frequencies. After the starting configuration had been shown the subject had to give a verbal expression for the chance that any point inside of the window would be covered by a white dot, when the same frequency of dots were presented again. After the subjects had given their answers the next relative frequency of white dots was determined according to a modified ROBBINS-MONROE procedure with an ever decreasing step size as long as the subjects gave the same verbal label. When the subject changed the label the direction of change for the relative frequencies was changed, too. After more than 4 switches in direction, a new starting position was taken in another part of the range according to the number of verbal expressions used by the subject in question. This recursive procedure was repeated until the boundaries for each verbal label had been determined. The boundaries, that is, the areas where subjects switched from one label to another, were interpreted as the elastic contraints characterizing the meaning of the verbal expression.

A typical result of this experimental procedure is shown in Figure 3.

The results (e.g. Figure 3, label "likely") can be interpreted in the following way: for the subject under consideration the label "likely" fits for events with the relative frequency of occurence between .65 and .75. It might be applied too in the intervals .55 - .65 and .75 - .85 (possible applicability), and it is not a permissible expression for events outside of this interval.

The Modelling of the Availability Bias in this Model

I have argued elsewhere (ZIMMER, 1982 b) that the context specific meanings of quantifiers in colloquial English can be decomposed into knowledge about the occurence of quantified statements in these contexts and into context free meanings of quantifiers. This separability of contextual and general meaning might explain why my subjects are apparently able to distinguish between at least four quantifiers, whereas BEGG (1982), who does not take into account the context, concludes from his data that normal language relies on only three quantifiers. This separability can also be used to model the availability bias in the model developed in this paper. It has been shown (for an overview see TVERSKY and KAHNEMAN, 1973) that judging the relative frequency of a class of events, such as the probability of dying due to an aircraft accident, subjects take into account not only the occurence of this event but also the difficulty or ease they have in coming up with a typical example for this class of events; the more available information about the event there is, the more the probability of occurence is overestimated. This availability bias can be accounted for in our model by the conjunction of the fuzzy meaning of each verbal expression under consideration and a scope function representing the influence of availability. In Figures 4, 5, and 6 scope functions for high, moderate, and low availability are shown. For these scope functions only monotonicity is assumed; the exact shape is to be determined empirically, e.g. by measuring the association speed or a similar variable. The conjunction of these scope functions with the possibility functions for the meanings of the verbal labels leads, after renormalization, to possibility functions of the biased judgments (see figures 7, 8, and 9). The comparison of the biased with the context free possibility judgments reveals that the model predicts an upward shift of the high uncertainty labels in the case of high availability and a downward shift of the low uncertainty labels in the case of low availability, whereas the labels on the other end of the range remain virtually unchanged. This effect is in line with experimental results (e.g. those of SLOVIC, FISCHHOFF, and LICHTENSTEIN, 1979). If the postulated separability of the fuzzy meanings and the availability bias proves to be empirically valid, this model can have fruitful applications in methods for estimating subjective probabilities, because it provides a procedure for debiasing (for different approaches to debiasing see FISCHHOFF, 1982).

Scope functions

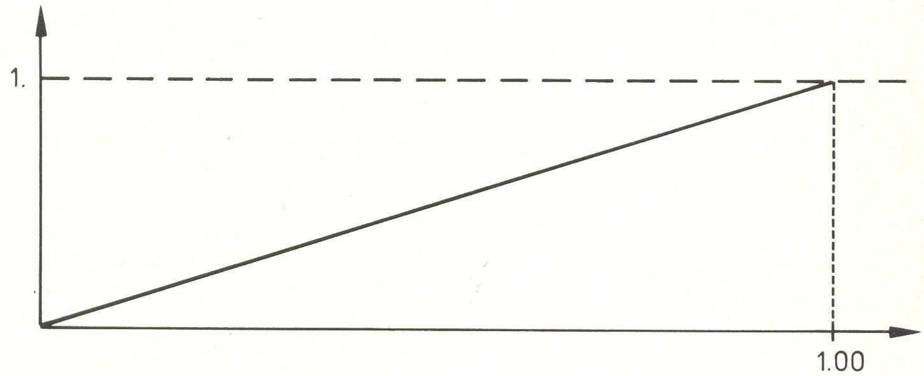

Figure 4. Scope function for a "high availability bias"

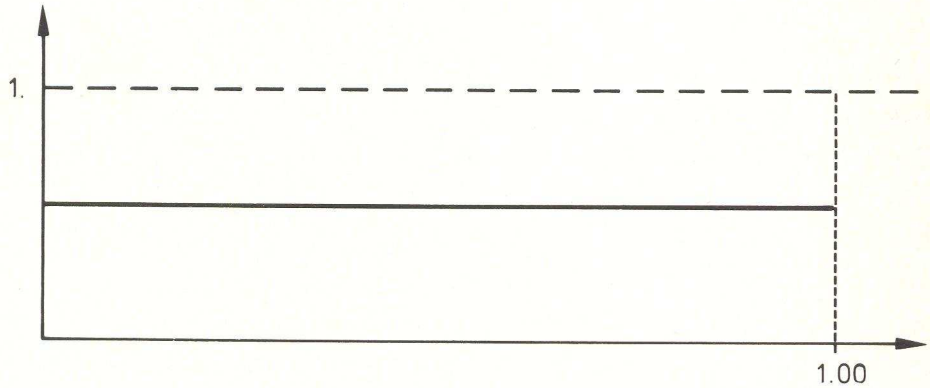

Figure 5. Scope function for a "moderate availability bias"

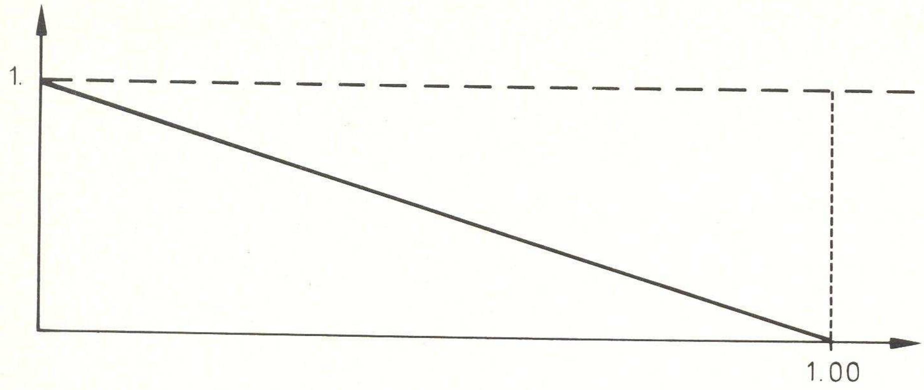

Figure 6. Scope function for a "low availability bias"

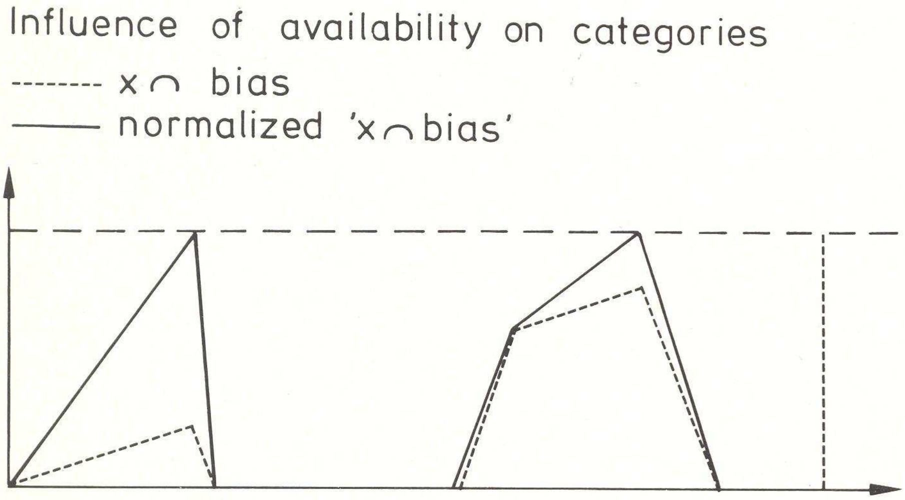

Figure 7. Influence of high availability on the meaning of "highly improbable" and "likely" in Figure 3

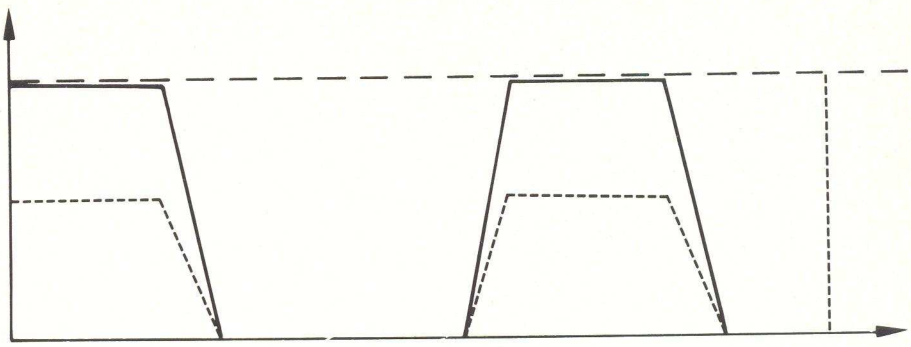

Figure 8. Influence of moderate availability on the meaning of "highly improbable" and "likely" in Figure 3

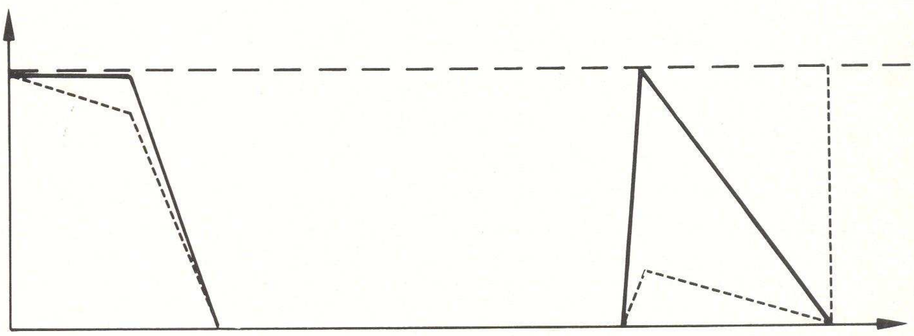

Figure 9. Influence of low availability on the meaning of "highly improbable" and "very likely" in Figure 3

Up to this point I have implicitly assumed that the possibility functions for the verbal labels are well calibrated, which is tautologically true for the relative frequencies of white dots on a tv screen. It can also be shown for tasks usually undertaken to test calibration, where subjects answer questions about almanac knowledge and estimate immediately afterwards the subjective probability of their answer being correct (for an overview see LICHTENSTEIN, FISCHHOFF, and PHILLIPS, 1982).

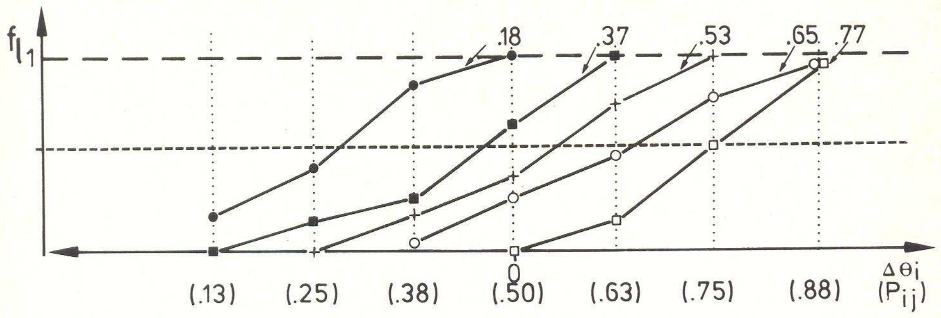

Figure 10. The cumulative frequencies for the verbal labels of one subject. The curves are drawn for items from seven categories of difficulty for this subject, which determine the individual probabilities of successes according to the Rasch model. The curves are labelled with the mean frequencies of the second experiment (see Figure 3). The subject would be perfectly calibrated if the labeling of the curves and the cutting points with the dashed line were the same. This subject is slightly underconfident for hard items but well calibrated for the easy ones.

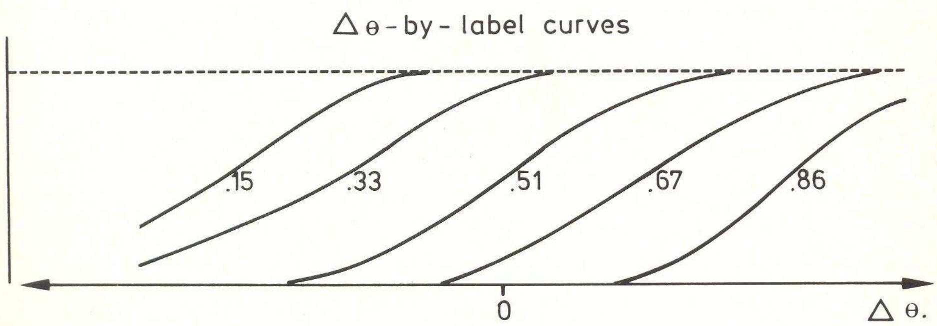

Figure 11. The average cumulative difficulty by label curves for all subjects applying five expressions for their subjective probabilities of successes

Calibration of Verbal Expressions for Probabilities

In order to separate the variables "difficulty of the item" and "of the subject", both of which influence the probabilities of correct answers, the conjoint measurement approach of RASCH (1966) is applied, which provides independant estimates of the ability of a subject, ξ (i), and the difficulty of an item j, δ (j). These estimates in turn allow us to assess the probability that a subject i solves item j, according to the following formula:

$$P(+|\xi_i;\delta_j) = \frac{e^{\xi_i - \delta_j}}{1 + e^{\xi_i - \delta_j}}$$

MAY (note 2) has suggested this approach to calibration which takes into account individual differences. The specific difficulty of a given item j for a subject i is the difference between the difficulty of the item and the ability of the subject. On this scale the relative frequencies of the labels which the subjects use to characterize their probability of success, can be compared with the estimated objective probabilities.

A test consisting of 150 political knowledge items was administered to 90 drafted soldiers of the German Bundeswehr. The items were in an open response format with subjects first answering the item and immediately afterwards giving their verbal expressions for the probability of a correct answer. Figure 10 represents the results of a single subject, who applied five different expressions for his subjective probabilities. Figure 11 gives the difficulty-by-label curves for all subjects using five labels. The curves are fairly parallel and the median values of the possibility functions (ranges of applicability for the expressions) are in good agreement with the probabilities estimated from the model.

In order to make these results more comparable to the ones depicted in Figure 10 I have computed the probabilities of successes for given ability levels and given difficulty levels conditioned on the labels applied. These "objective" probabilities are plotted against the median values of the possibility functions representing the meaning of the verbal expressions in Figure 12.

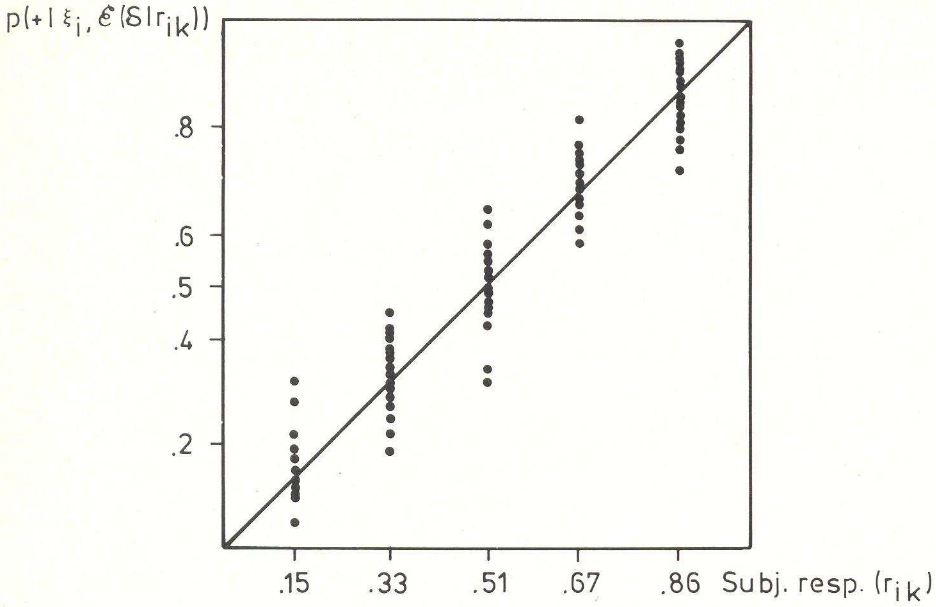

Figure 12. The probabilities of successes for given ability and difficulty levels conditioned on the applied verbal expressions plotted against the median values of the possibility functions for these verbal expressions.

Although there are marked variances, the results indicate that on the whole this group of subjects is well calibrated.[1]

---

[1] SCHUETT (1981) reports that after sufficient training and optimal support by a computerized information processing system his subjects were able to estimate subjective probabilities, which did not violate SAVAGE's (1954) axioms. This result does not contradict the conclusions I have drawn from my experiments; that is, humans usually express their confidence verbally and they are able to perform by means of verbal arguments a task which is as complex as SAVAGE's system of axioms. Furthermore the results reported here indicate that by using this kind of information processing people are able to give veridical judgments without the amount of external support provided in SCHUETT's (1981) experiment.

Further Applications of this Model

As pointed out in the beginning I assume that any mode of processing information which imposes higher mental work load is bound to render any consecutive rating more biased than a mode imposing a lower mental work load. Examples for the negative effect of a high mental work load can be found in the studies done on conservatism in information processing and on the negligence of the regression effect in numerical prediction tasks. In order to find out if the application of verbal expressions in these tasks relieves the biasing effect, I have applied my model to two prototypical experiments from this area.

Conservatism

A typical example for the suboptimal information processing has been demonstrated in experiments on conservatism, where subjects tend to stick to their initial assumptions concerning the probabilities of events despite the fact that in the light of new information they should revise these assumptions. The optimal revision strategy for assumptions in this task is the application of Bayes'theorem, which therefore can be used as a normative standard for the subjects' performance. In a series of experiments PHILLIPS and EDWARDS (1966) have investigated this phenomenon with the result that in all cases conservatism occured. This effect could be slightly reduced by relieving the memory load of the subjects and in one experiment by permitting them to answer verbally. I have replicated part of their experiment III with 30 subjects (undergraduates in economy). Subjects were shown two identical bins pictured side by side on a tv screen. They were told (A) contained 70% bricks and 30% balls and the other (B) 30% bricks and 70% balls. On each trial subjects pushed a button and thereby drew a sample of one - either a brick or a ball - from each of the bins. Subjects were asked to describe verbally the chance of getting a brick or a ball the next time from the bin. On a second screen a tally was kept of the numbers of bricks and balls drawn from each bin, together with the last verbal expression the subject had given for the chances. The results plus the ones from PHILLIPS and EDWARDS (1966, experiment III) are shown in Figure 13. These results clearly indicate that verbal responses together with a visible record induces near optimal, or, Bayesian performance in the subjects. In the light of this results one can assume that the so

Conservatism

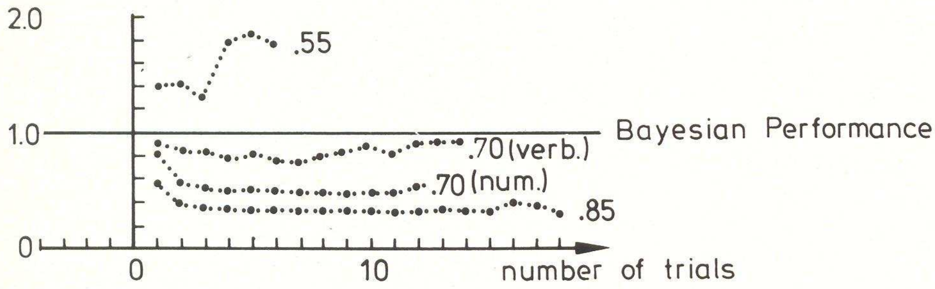

Figure 13. Comparison of the results from PHILLIPS and EDWARD's experiment III with the ones from our experiment with verbal expressions for chances (labelled ".70 (verb.)"

called conservatism in information processing is mainly due to a conservative handling of numbers and computation and is not necessarily due to a biased processing procedure.

Prediction

In their 1973-study KAHNEMANN and TVERSKY have tried to determine which kinds of heuristics and biases influence subjects' performance in predicting future events. They studied categorical as well as numerical predictions and found in all cases that subjects tended to neglect the base rates and exhibited mistaken intuitions about regression. For instance, they acted as if in all cases the correlation between the predictor variable and the criterion were perfect and as if the predictor variable were measured with perfect reliability. In order to determine if these mistaken intuitions are at least partially caused by the apparent difficulties subjects have in processing numerical variables, I have undertaken the following experiment: subjects (30 undergraduate students in psychology who had not attended courses in statistics) were asked to predict the success of students in the university from their performance at highschool. Subjects were given a sample of 75 cases which includes the grades the students received at highschool and in the graduating exam at the university. The results were very similar to the ones obtained by KAHNEMAN and TVERSKY (1973) but when the subjects were asked how certain they were about the correctness of their predictions, they gave quite low subjective degrees of confidence. When probed further about the direction of their probable error of prediction 27 subjects indicated correctly that the true value would probably be closer to the mean performance than they had predicted. Only 3 subjects answered that they could not tell.

This result as well as those reported above seem to indicate that subjects are better able to take into account complex dependencies by means of verbal processing than they are if they are forced to process the same amount of information numerically. As GREGORY (1982) pointed out, numbers and computation form a more recent tool of mind than language and therefore the numerical information processing is less automatic.

A further aspect of the verbal superiority is illustrated by the outcome of an experiment from a different study. In this experiment[2] subjects (24 bank clerks responsible for foreign exchange) were asked to predict what the exchange rate between the US Dollar and the Deutschmark would be four weeks later. 12 Subjects had to give the predictions "in their own words as they would talk to a client", whereas the other twelve were asked to give numerical estimates in percentage of change. Both groups were asked to verbalize the steps they took in order to come up with the prediction. After the predicitions of the first group had been calibrated with a technique similar to the one of our second experiment, they were compared to the one made by the numerical forecasting group. It turned out that the first group was more

---

2 In a different context I have reported other aspects of the results of this experiment.

2. personal communication
   (ZIMMER, note 3)

3. ZIMMER, A. A model for the interpretation of verbal predictions, manuscript, Stanford, 1983, (submitted for publication)

References

BEGG, I. (1982) On the interpretation of syllogisms. Journal of Verbal Learning & Verbal Behavior, 21, 595-620

DE FINETTI, B. (1937) La prévision: des logiques, ses sources subjective. Annales de l'Institute Henri Poincaré, 7, 1-68

DUBOIS, D. & PRADE, H. (1980). Fuzzy Sets and Systems: Theory and Application. New York: Academic Press

FERRELL, W.R. & MC GOEY, D.J. (1980). A model of calibration for subjective probabilities. Organisational Behavior and Human Performance, 26, 32-53

FREKSA, C. (1981) Linguistic Pattern Recognition.Ph. D. thesis, UC BERKELEY

FREKSA, C. (1982) Linguistic description of human judgments in expert systems and in "soft" sciences. In: L.TRONCALE (ed.) Systems Science and Science. Louisville, Ky.:Society for General Systems Research

GREGORY, R.L. (1981) Mind in Science. Cambridge: Cambridge University Press

HACKING, I. (1975) The Emergence of Probability. Cambridge: Cambridge University Press

HERSH, H.M. & CARAMAZZA, A. (1976) A fuzzy set approach to modifiers and vagueness in natural language. Journal of Experimental Psychology, General,105, 254-276

HINTZMAN, D.L., NOZAWA, G. & IRMSCHER, M. (1982) Frequency as a nonpropositional attribute of memory. Journal of Verbal Learning and Verbal Behavior, 21, 127-141

KAHNEMAN, D. & TVERSKY, A. (1973) On the psychology of prediction. Psychological Review, 80, 237-251

KAUFMANN, A. (1975) Theory of Fuzzy Subsets. Vol. 1. New York: Academic Press

LICHTENSTEIN, S., FISCHHOFF, B. & PHILLIPS, L.D. (1982) Calibration of probabilities: The state of the art to 1980. In: KAHNEMAN, D., SLOVIC, P., TVERSKY, A. (eds.) Judgment under uncertainty. Heuristics and Biases. Cambridge: Cambridge University Press

correct and more internally consistent. While this is interesting in itself, another point is more important: The slight difference in the instructions given caused marked differences in the way the subjects performed their task as revealed by verbalizing. The verbal prediction group used quantitative variables (e.g. the GNP increase in percent) as well as qualitative variables (e.g. the stability of the German government) for deriving their predictions, whereas the other group merely took into account those variables which were usually expressed numerically. From this it seems plausible to assume that the reason for the superiority in the verbal forecasting condition is the fact that the knowledge base on which these subjects relied was broader and allowed for more elaboration.

An alternative interpretation of those results is possible too: It might be that the thinking-aloud interfered specificly with the numerical reasoning.

General Conclusion

Although the studies reported here have originated from a perspective of applied psychology, aiming at better ways to implement vague knowledge into decision processes, there may be consequences of this kind of research for cognitive psychology in general. When analyzed by formal means it turns out that decision making and forecasting on the one hand and reasoning and problem solving on the other hand are structurally identical. Therefore it can be assumed that the dependency of cognitive processes on the underlying mental representations might be a quite general problem.

The experimental results presented here indicate that: first, the available procedures for human information processing are strongly dependent on the modes of mental representations, in which they are applied; and second, human subjects are more effective in reasoning with verbal expressions than with numerical expressions, even if the tasks performed rely on frequency information.

Reference Notes

1. personal communication

LINDLEY (1974) Discussion on the papers by Professor Tversky and Professor Suppes. Journal of the Royal Statistical Society, Series B. 36,181-182

NISBETT, R.E. & ROSS, L. (1980) Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, N.J.: Prentice Hall

NISBETT, R.E. & WILSON, T.D. (1977) Telling more than we know: Verbal reports on mental processes. Psychological Review, 84, 231-259

PHILLIPS, L. & EDWARDS, W. (1966) Conservatism in a simple probability inference task. Journal of Experimental Psychology, 72, 346-354

RASCH, G. (1966) An item analysis which takes individual differences into account. British Journal of Mathematical and Statistical Psychology, 19, 49-57

SAVAGE, L.J. (1954) The Foundations of Statistics. New York: Wiley

SCHÜTT, K. P. (1981) Wahrscheinlichkeitsschätzungen im Computer-Dialog. Stuttgart: Poeschel Verlag

SLOVIC, P., FISCHHOFF, B. & LICHTENSTEIN, S. (1979) Images of desaster: Perception and acceptance of risks from nuclear power. In: G.T. GOODMAN & W.D. ROYCE (eds.) Energy Risk Management. London: Academic Press

SLOVIC, P. & TVERSKY, A. (1974) Who accepts Savage's axiom? Behavioral Science, 19, 368-373

SUPPES, P. (1974) The measurement of belief. Journal of the Royal Statistical Society, Series B, 36, 160-175

TVERSKY, A. (1974) Assessing uncertainty. Journal of the Royal Statistical Society, Series B, 36, 148-159

WITTE, W. (1960) Struktur, Dynamik und Genese von Bezugssystemen. Psychologische Beiträge, 4, 218-252

WITTGENSTEIN, L. (1969) Philosophische Untersuchungen. Frankfurt: Suhrkamp

WRIGHT, G.N. & PHILLIPS, L.D. (1980) Cultural variation in probabilistic thinking: alternative ways of dealing with uncertainty. International Journal of Psychology, 15, 239-257

ZADEH, L.A. (1978) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-38

ZIMMER, A. (1980) Eine Formalisierung mnestisch stabilisierter Bezugssysteme auf der Grundlage von Toleranzmengen. In: A. THOMAS & R. BRACKHANE (eds.) Wahrnehmen - Urteilen - Handeln. Bern: Huber

ZIMMER, A. (1982 a) What really is turquoise? A note on the evolution of color terms. Psychological Research, 44, 213-230

ZIMMER, A. (1982 b) Some experiments concerning the fuzzy meaning of logical quantifiers. In:L. TRONCALE(ed.) Systems Science and Science. Louisville, K .: Society for General Systems Science