Netherlands
organization for
applied scientific
research

IZF

TNO Institute for Perception

DTIC FILE COPY

AD-A218 056

Netherlands
organization for
applied scientific
research

TNO-report

TNO Institute for Perception

P.O. Box 23
3769 ZG Soesterberg
Kampweg 5
3769 DE Soesterberg, The Netherlands
Fax +31 3463 5 39 77
Phone +31 3463 5 62 11

IZF 1989-29

THE INFLUENCE OF RELIABILITY OF INFORMATION ON DECISION PERFORMANCE

J.H. Kerstholt

12

Number of pages: 25

Classifications:

Report:	unclassified
Title:	unclassified
Abstract:	unclassified

CONTENTS

Report No.: IZF 1989-29

Title: The influence of reliability of information on decision performance

Author: Drs. J.H. Kerstholt

Institute: TNO Institute for Perception

Date: September 1989

No. in Program of Work: 733.2

SUMMARY

Judgement tasks require knowledge of relations between symptoms (cues) and diagnosis. These relations are often probabilistic in nature, and decisions therefore have to be based on more or less unreliable cues. Many so-called Multiple Cue Probability Learning (MCPL) studies have shown that learning such relations is seriously impaired, even by a small degree of uncertainty. On the other hand, there is ample evidence that in natural judgement tasks such as medical diagnosis or market analysis, people are well able to learn from experience. In the present experiment learning under uncertainty is investigated in a relatively natural task that corresponds more closely with typical diagnosis tasks than the MCPL tasks. Subjects were able to learn the underlying model in the probabilistic mode, be it at a lower rate. The delay only concerned the information selection process, not that of information integration.

Rap.nr. IZF 1989-29, Instituut voor Zintuigfysiologie TNO, Soesterberg

De invloed van de betrouwbaarheid van informatie op beslissingsgedrag

J.H. Kerstholt

SAMENVATTING

Beoordelingstaken vereisen dat men kennis heeft van de relaties tussen symptomen en de diagnose. Deze relaties zijn vaak probabilistisch van aard, waardoor beslissingen gebaseerd zijn op onbetrouwbare informatie. Uit de zogenaamde Multiple Cue Probability Learning (MCPL) studies blijkt dat, zelfs als de onbetrouwbaarheid van de informatie gering is, dergelijke relaties zeer slecht worden geleerd. Aan de andere kant blijken mensen in een natuurlijke, complexe en onbetrouwbare omgeving, veel van ervaring te leren. In het huidige onderzoek werd de invloed van onbetrouwbare informatie op het leerproces onderzocht in een relatief natuurlijke taak, die meer overeenkomt met typische diagnosetaken dan de MCPL-taken. Zelfs met onbetrouwbare informatie waren proefpersonen in staat het onderliggend model in deze taak te leren, zij het dat ze meer oefening nodig hadden. De vertraging betrof alleen het leren selecteren van de juiste informatie en niet het leren van de juiste beslissingsstrategie.

1 THEORETICAL BACKGROUND

1.1 A framework to investigate decision making

A typical judgement task, such as diagnosis in medicine or fault finding in a production environment, requires that one knows the relations between variables, e.g. the relations between various symptom configurations and the underlying disease (Brehmer, 1980). An influential framework to investigate how such relationships are learned has been provided by Brunswik (1952, 1956). Originally his model only concerned perceptual abilities, but it has been extended to the study of judgement and decision making (Hammond, McClelland and Mumpower, 1980).

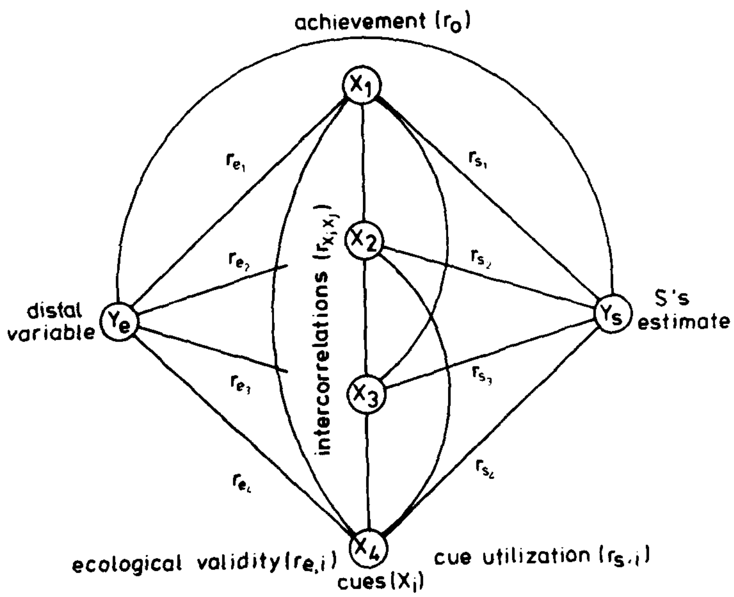

Fig. 1 Brunswik's lens model (adopted from Hursch, Hammond and Hursch, 1971).

Brunswik distinguished three major elements in a judgement situation: a subject, the environment, and cues. The environment, the distal variable, scatters its effects to a number of cues and all the rays are brought together again to a second focus, leading to an estimate of the distal variable by the subject (see figure 1). A distal variable can for example be the size of an object or the disease of a patient. Proximal variables (cues) are defined in terms of observable stimuli. In the case of diagnosis cues will consist of, for example, test results and the patient's complaints. The agreement between distal variable, cues and a subject's estimate is expressed by a correlation coefficient. Ecological validity, depicted on the left side of the figure, is expressed by the correlations between distal variable and cues (re,i). These correlations indicate the relative frequency of the association between a cue with a distal variable. The functional validity of a cue, or cue utilisation, is the relative frequency of its association with the person's response (rs,i). Furthermore, the cues are related to each other, expressed by the correlations between the cues. Non-zero correlations make the information available to the subject redundant. The goal of the subject is to obtain an accurate representation of the environment in order to optimally predict the distal variable. A physician for example has to know how symptoms are related to diseases, in order to judge (predict) the value of the actual state of the system. The extent to which subjects reach this goal, the achievement, is expressed by the correlation between estimates and values of the distal variable.

With this model, Brunswik has stressed two aspects of major relevance to the study of decision making. First, both the environment and the person should be seen as systems, each with properties of their own. This implies that research should equally concern the structure of the subjects' perceptual or information processing system and the structure of the environment. Secondly, it is important to consider the uncertainty that is faced by the person. The relations between distal variable and cues may contain some error, causing a probabilistic criterion-cue relationship. This implies that the information available to the person varies in reliability as predictors for the criterion.

1.2 'Multiple cue probability learning' studies

Multiple Cue Probability Learning (MCPL) studies evolved directly from the ideas of Brunswik and have mainly concerned learning under uncertainty. Although many variants exist, MCPL tasks typically require subjects to predict some criterion (Y), based on the presentation of one or more cues that can take on different values (Hammond, McClelland and Mumpower, 1980). The subject makes an estimate of the criterion value and is given feedback, indicating the true criterion value. The subject's task is to use this feedback about the discrepancy between the predicted and true value to achieve a more accurate translation of cues to criterion. For example, a subject is presented with the information that cue 1= 4 and cue 2= 3. In order to predict the criterion value, the subject has to hypothesise a relation between Y, cue1 and cue2. Suppose this subject assumes that Y=cue1+cue2, he/she will respond with '7'. The subject then receives feedback about the true value of the criterion. This would for instance be '3', if the function was Y=1/4 cue1 + 2/3 cue2. By continuously comparing own estimate with true score, the subject has to infer the relation between criterion score and cues. To make the relationships between criterion and cues probabilistic, some random error is added as well. Different task variables have been manipulated, such as function form relating cues to criterion (e.g. linear or inverted U-form), number of parameters, weights, and amount of error. In this way, it could be investigated how environmental factors influence decision performance.

From a number of experiments within the MCPL-paradigm, Brehmer (1980, 1987) concluded that people are not able to perform optimally in probabilistic tasks, even when only a moderate amount of error is included. Especially when cues and criterion are related in a non-linear way, learning rapidly deteriorates (Klayman, 1988b). Generally, people assume a deterministic linear rule, by which cue values can be transformed to predict the criterion score, and they show limited knowledge of probabilistic relationships and how they should be dealt with. Even when subjects received information about the nature of the probabilistic relations, the optimal rule was not learned (Brehmer and Kuylenstierna, 1978; Johansson and Brehmer, 1979).

An advantage of laboratory tasks such as those used in MCPL studies, is that much experimental control is possible, since the relationships to learn are known to the experimenter. This not only provides a norm with which to compare subjects' performance, but also allows for the investigation of the effects of various task variables, such as function form, on decision performance. The disadvantage of not having a normative model is rather prominent in studies concerning clinical judgements. In personality assessment for example, or diagnosis based on projective tests, experienced clinicians are not more accurate than less experienced ones (Garb, 1989). However, as is shown by Brunswik's model, two relationships may account for this result. Either humans cannot learn these relationships, or there is no clear relationship between cues and true scores.

However, despite their Brunswikian origins, cue learning tasks are only an incomplete realisation of the ideas of representative design (Klayman, 1988b). The basic principle is that decisional abilities should not be evaluated in impoverished environments because these abilities have developed to deal with complex real world situations. This implies that the cognitive processes that are used by subjects in an impoverished environment may differ from the ones that are developed to adapt to the real world. The present experiment is an attempt to investigate learning under uncertainty in a task environment that corresponds more closely with natural judgement tasks.

1.3 Characteristics of natural judgement tasks

Compared to judgement tasks such as diagnosis, MCPL tasks are restricted in various ways. In the following paragraphs we will discuss three general aspects of natural judgement tasks which are absent in MCPL tasks. These aspects are meaningfulness, cue selection, and choice.

1.3.1 Meaningfulness

In most MCPL studies only abstract cues have been used in order to investigate the application of statistical knowledge without possible confounding effects of general world knowledge. In a natural context however, people always have at least some knowledge about the variables on which the decision is based. Prior general world knowledge may interact with learning new relationships. It has indeed been found by Sniezek (1986) that performance increased when the cues were given meaningful labels. She suggested that this effect relates to the reduction of cognitive load with regard to hypothesis generating and testing. With abstract cues many hypotheses can be generated, even with few cues. The presence of meaningful labels however, provides directions and possible relations in the data, which reduces the number of hypotheses dramatically. Furthermore, the learning process is facilitated by a reduced memory load. In order to learn one has to remember hypotheses that are already falsified, and systematically manipulate specific variables. This process is easier with meaningful information than with abstract cues.

1.3.2 Information selection

In MCPL tasks the subject only has to integrate information, since all cues are a priori relevant. However, in order to extrapolate to real world situations it has to be known whether information integration is the only relevant decision process in tasks such as diagnosis.

Of importance here are the results found by 'expert-novice' studies. Rather than studying the gradual experience gained in a laboratory task, these studies have examined people who have learned from experience in applied settings. They have investigated how experts and novices in a particular domain, such as physics or engineering, cope with decision problems commonly encountered there (Johnson, 1981; Woods and Roth, 1988). A consistent finding in this research area is that experts have a much larger and better structured knowledge base than inexperienced subjects. In solving a problem, experts outperform novices not in the 'form of reasoning' but in the 'content of reasoning'. For example, in the medical domain it has been assumed that with experience more disease models and more features are added to memory. This enhances the process of identifying the correspondence between symptoms and diseases (see for a review Schraagen, 1986; or Boshuizen, 1989, for the development of medical expertise). To conclude, one important characteristic of expertise is knowing which information to select. Therefore, if research is supposed to address tasks such as diagnosis, it should investigate not only information integration, but information selection as well.

1.3.3 Choice

In MCPL studies researchers have been mainly concerned with prediction. This means that only the outcome of the decision process is taken into account. Subjects have to make predictions with regard to the criterion value, and when their predictions match actual outcome (expressed by a high correlation), their decision is considered to be accurate.

It might be questioned however, whether the exact prediction of a criterion is a critical ability to real world decision tasks. First, people can often achieve a considerable amount of predictive accuracy even if they treat all cues as being of equal importance (Einhorn and Hogarth, 1975; Dawes and Corrigan, 1974). In the medical domain experts even avoided complex statistical comparisons in reaching a diagnostic conclusion (Connolly and Johnson, 1980).

Secondly, in many tasks hypothesis testing and information use form an interactive process, rather than a strict bottom up combination of information. A major strategy used by clinicians is to generate several diagnostic hypotheses early in the decision process. This permits

the use of a hypothetico-deductive method in which data are collected with a view to their usefulness in choosing between several hypotheses (Elstein, 1978; Connolly and Johnson, 1980). The decision maker has several hypotheses available and by learning the values of several cues, some hypotheses are dropped, while others are confirmed.
Since choice between several options seems to be more compatible with diagnosis tasks than prediction, we have provided our subjects with a choice task. They have to select relevant information sources, which corresponds to assigning a weight of 0, 1, or -1 to a cue. They can then request the actual information values for several options, and decide which option is the optimal one. Choice will lead to a different information integration strategy than prediction, since the values can be compared on several options. In the following section we will therefore discuss typical choice strategies that are distinguished in behavioral decision research.

A commonly made distinction in decision strategies is whether they are compensatory or noncompensatory (Payne, 1982; Billings and Marcus, 1983). A decision strategy is regarded compensatory when trade-offs between high and low values on various dimensions are taken into account. When subjects use a compensatory strategy, they calculate the overall worth of each alternative, and choose the one with the highest worth. On the other hand, subjects may use cut-off points. In this case, the subjects drop an option from the total set, as soon as a value, or a combination of values, does not meet the cut-off point. When trade-offs are not taken into account, subject are said to use a noncompensatory strategy. Examples of noncompensatory strategies are the satisficing strategy and the elimination-by-aspects strategy. Satisficing means that an alternative is selected, scored on several dimensions, until a value does not meet a criterion. At that point the alternative is eliminated from the option set. As soon as an alternative meets all criteria, that alternative is chosen. When an elimination-by-aspects strategy is used, the subject selects the most important dimension, rather than an alternative. Each alternative receives a score on this dimension. The alternatives that do not meet the cut-off point are eliminated. This process continues until only one alternative is left.
In our task the normative model is compensatory. This means that subjects not only have to select the relevant information, but have to integrate the information in a compensatory manner as well.

1.4 Summary and research questions

Typical MCPL studies have led to the conclusion that people cannot perform optimally when they have to deal with unreliable information. However, several aspects of MCPL tasks are not in accordance with diagnosis tasks as found in applied settings. Therefore, we have extended the task in order to answer two questions:

1. Can subjects learn relationships accurately through the use of unreliable information in a more natural task environment?

2. What are the differential effects of reliability of information on information selection and information integration? Does reliability affect the selection of relevant information, the selection of irrelevant information, the integration strategy that is used, or some combination of these?

2 EXPERIMENT

2.1 Method

Stimulus material

Following previous work by Keren and van Doorne (1987), and van Doorne (1987) a computergame was used. The game is called MILSIM (Milkyway Merchant simulation), and deals with a fictitious trading company operating on a market in outer space. The company sells and buys commodities on star centres with the goal of optimising profit.
On each trial subjects are successively shown four screen presentations:

1. The introduction screen, indicating on which star centre in space the subjects are positioned, and in which commodity they will trade.
2. The transaction screen, indicating how much is available of the commodity and the price per unit. A menu indicates three possible courses of action: to buy the commodity, to request information, or to depart. Departure to a specific star centre implies that all units of the commodity that are bought, will be sold on that centre.
3. The information selection screen displays seven aspects. These aspects specify all the information that can be selected, i.e. market position, buying price, experience with centre, demand,

competitiveness, traveling costs, and service. Only three aspects are relevant, i.e. demand, bying price and traveling costs. The values of this information, appear in matrix form on an auxiliary display screen. The columns are formed by the alternatives (star centres) and the rows indicate the information sources. The values of market position, experience, competitiveness and service can take on three values: good, medium and bad. The other variables were of a quantitative nature.

4. The feedback screen gives information concerning obtained profit, maximal profit, and the star centre where this would have been gained. A menu allows for the start of a new trial. On each trial subjects are positioned on a different star centre and and they are required to trade in a different commodity.

Feedback information on actual and maximal profit, is based on true values, and is calculated according to the following formula:

Profit= (amount of the commodity sold * selling price) - (amount of the commodity bought * buying price) - travelling costs.

For the calculation of the maximal profit the variable 'demand' is used instead of the variable 'amount of commodity sold'. The system calculates the profits of all alternatives and selects the one with maximal profit.

Procedure

Subjects were instructed to imagine that they worked for an agency that traded in various commodities. The goal was to gain maximal profit through buying an amount of a given commodity on the current star centre and selling it on another one. There were six star centres: Mira, Spica, Bella, Sun, Alcor, and Vega. The subjects thus had to choose one of the five star centres, where they could deliver the commodity. Before the decision was made the subjects were told that information could be requested about seven aspects for each of the five alternatives, i.e. market position, demand, experience with centre, service on centre, buying price, travelling costs and competitiveness. They were informed that not all information was equally relevant. The subjects were instructed to ask for as much information as they needed to make a good decision. When they had requested enough information and decided where they wanted to go, they bought a certain amount of the commodity and left for their destination. The choice of a star centre implied that they wanted to sell all units that were bought of the commodity at that particular centre.

After arriving at the centre they were given feedback about the profit they had obtained. They were also given information indicating which star centre would have been optimal, and how much profit they could have made on this centre. At this point, the end of the trial, the information was still available on the data screen. They then proceeded with the next trial on another, randomly chosen, star centre with another commodity. The number of trials depended either on reaching the criterion score, defined as obtaining maximal profit in three consecutive trials, or the time limit of four hours.

Subjects

Five female and seven male subjects participated in the experiment. They were all students at the University of Utrecht (mean age: 22,3 years, σ=1,3). They were paid Fl.50.- for participation. For one subject naivety towards the research questions was doubted. Since this could bias the results, her responses were left out of the analyses.

Design

Subjects were randomly assigned to either the 'Reliable Information' (RI) condition, or the 'Unreliable Information' (UI) condition. In the RI condition, the subjects could use information that revealed the true values. In the UI condition, subjects received information to which error was added. The error was normally distributed, with a standard deviation of 0.05 times the value of the true score and with a mean of zero. The implication of this error was that even if a subject in the UI condition selected the correct information, and integrated the values as specified by the normative model, there was still a probability of 25% that a wrong decision would be made. The subjects in this condition were told that the information they received were only estimates of the true values.

To investigate information usage, referred to as information integration hereafter, we used a computerised information board technique (Payne, 1976; Svenson, 1979; Billings and Marcus, 1983). The main purpose of this technique is to identify the strategy that a subject has used to integrate information, by means of the information search pattern. Information boards consist of a matrix in which the columns correspond to the options that can be chosen, and the rows correspond to the attributes, which are the cues available to the subject. The subject can obtain one piece of information (one cell of the matrix) at a time. Decision rules were categorised as either compensatory or noncompensatory. If the subject requested the values for all alternatives, of the variables judged relevant, the strategy was considered compensatory. A strategy was considered noncompensatory when different amounts of information were requested for the various alternatives.

2.2 Results and discussion

Except for one subject all persons were able to learn the underlying model, within the limit of 4 hours. The subject who did not learn the model had been assigned to the RI condition, which means that it was not due to unreliable information. It is known that causal analysis, which means relating outcome feedback to own task behaviour, is an important factor for improving performance (Schweiger, Anderson and Locke, 1985). Through informal verbal protocols it was noticed that this subject, unlike the others, did not show this behaviour and did not try to formulate hypotheses about the underlying model. Therefore, a plausible explanation for not learning the model is the subject's passive attitude towards the goal of the experiment.

The most conspicuous difference between the Reliable Information (RI) condition and the Unreliable Information (UI) condition is the speed with which the model is learned by the subjects. The RI-group needed on average 10.8 trials (σ=7.6) to learn the underlying model, whereas in the UI-group on average 20 trials (σ=4.8) were needed to attain optimal performance ($t(8)=2.4$, $p=.04$).

Our first research question, whether subjects would be able to learn the underlying model if information is unreliable, can be answered affirmatively. Unreliable information slowed down learning but the subjects still learned to select the correct variables, and chose the correct option as specified by the normative model.

Since we are particularly interested in the differential effects of unreliable information on information selection and information integration, we will discuss the results for these two phases seperately. Because these analyses relate to the learning curves of the subjects, the subject who had not learned the model, was left out of the analyses.

Figure 2 depicts the percentage of trials in which all relevant information was used as a function of condition and trial block. One trial block comprises 8 trials. For each trial block a Pearson chi-square was calculated.

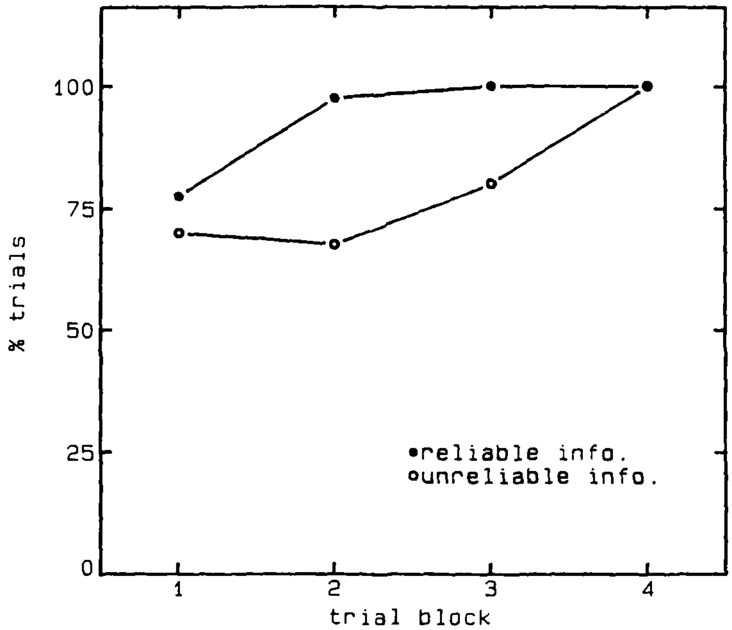

Fig. 2 Percentage of trials in which all relevant information was used as a function of condition and trial block. Each data point comprises 5 (subjects) x 8 trials.

In the first trial block no difference exists between the RI-condition and the UI-condition ($\chi^2(1)=.58$, $p=.45$); In both conditions relevant information is ignored in approximately 25% of the trials. In the second and third trial block however, a significant difference emerges between the RI- and UI-condition ($\chi^2(1)=12.5$, $p=.0001$ and $\chi^2(1)=8.9$, $p=.003$ respectively). In these two trial blocks relevant information is ignored more often in the UI-condition than in the RI-condition. In trial block 4 all relevant information is used in each trial.

Why did subjects in the RI-condition learn to select all relevant information more efficiently than the subjects in the UI-condition? An important factor is the quality of the feedback in both conditions. When subjects in the RI-condition calculated the profits as defined by the normative model, the feedback concerning maximal profit would match their calculations exactly, since they used true values. This way their hypotheses are directly confirmed. In the UI condition there would in practice always be a discrepancy between own calculations and the profit obtained, because an error term was included in the subject's information. Even so, the choice, i.e. the star centre, would be correct in 75% of the trials, if the normative model was used. In 25% of the trials in the UI condition, the subjects would not choose the optimal alternative, even though they used the correct procedure. In these cases the feedback is misleading, because it would disconfirm the normatively correct hypotheses. The fact that no exact confirmation was obtained when the normative model was used and a chance of 25% on misleading feedback, left the subjects in the UI-condition with more uncertainty about the accuracy of their mental representation of the system than subjects in the RI condition.

In Figure 3 the relation between percentage of trials in which all irrelevant information was ignored and trial block is expressed for both conditions. Again each trial block comprises 8 successive trials, and chi-squares were calculated for each trial block.

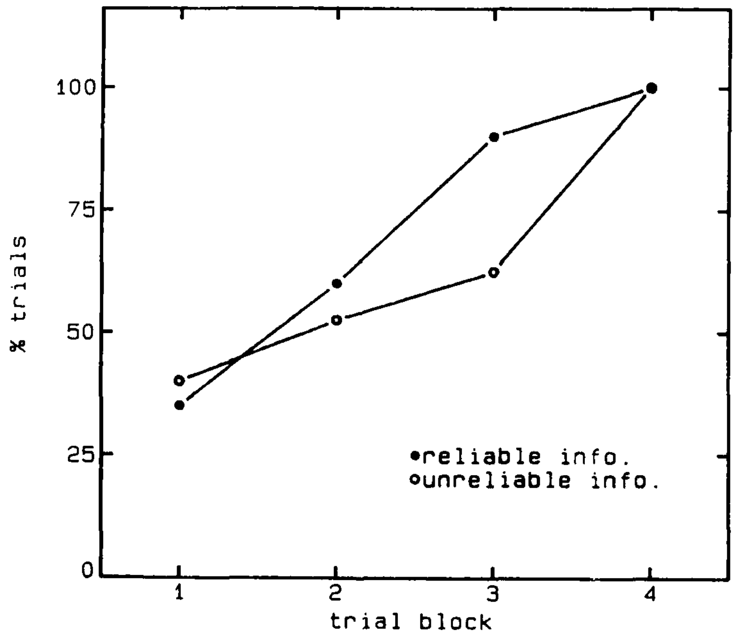

Fig. 3 Percentage of trials in which all irrelevant information was ignored as a function of condition and trial block. Each data point comprises 5 (subjects) x 8 trials.

In trial block 1 and 2 no significant differences between the two conditions were found $\chi^2(1)=.21$, $p=.64$ and $\chi^2(1)=.46$, $p=.50$, respectively). The two groups only differ in the third trial block ($\chi^2(1)=8.4$, $p=.004$). In this phase, irrelevant information was used in an additional 30% of the trials in the UI condition.

The uncertainty caused by unreliable information therefore not only impedes the selection of relevant information but it also causes more subjects to include irrelevant information in their judgement. In trial block four, where the subjects have learned the model, all subjects fully ignore irrelevant information.

In the RI condition, subjects took significantly longer to ignore all irrelevant information than to include all relevant information ($t(4)=-4.8$; $p=.008$). It may be important here that the relevant and irrelevant information differed in measurement level. The relevant variables were all quantitative, whereas the irrelevant variables had qualitative values. From verbal protocols it appeared that subjects often perceived the qualitative variables as providing chance information: a bad market position, for example, would indicate that the chances are high that not all products can be sold, and would thus involve a risk. If information is perceived as chance information, more trials are needed to see its predictive worth.

Figure 4 shows the percentage of trials in which a compensatory strategy was used for each trial block. In the present study a compensatory strategy was defined as normatively correct.

No significant differences in strategy use exist between the two conditions. The chi-squares for the successive trial blocks are $\chi^2(1)=.08$, $p=.79$; $\chi^2(1)=1.3$, $p=.26$ and $\chi^2(1)=1.9$, $p=.17$ respectively. This implies that unreliable information has no overall effect on the strategies used while learning the model. Both groups start with non-compensatory strategies such as satisficing and elimination by aspects, and gradually learn to choose a compensatory strategy. In trial block 4 a compensatory strategy was used in each trial.

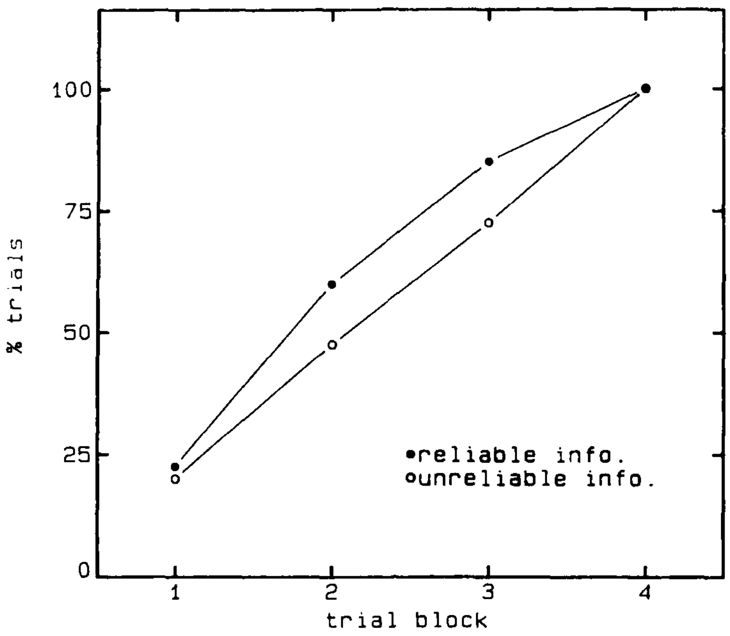

Fig. 4 Percentage of trials in which a compensatory strategy was used as a function of condition and trial block. Each data point comprises 5 (subjects) x 8 trials.

3 GENERAL DISCUSSION

With regard to our research questions the results are unequivocal. They showed that unreliable information delayed, but still allowed subjects to learn the underlying rules of a system model. More in particular, it delayed the learning of which information to select or ignore, and not the learning of how to combine the information once selected. In the present section we will briefly discuss some more general issues related to these results. We will first discuss the implications for extrapolating typical MCPL results to diagnosis tasks. Secondly, we will consider an alternative interpretation of our results concerning strategy choice.

Contrary to the typical MCPL results we found that our subjects could learn from unreliable information. This means that, under the present circumstances, people are able to deal with probabilistic information. This does not suggest conclusions about learning under uncertainty in general however, since our task environment differed from MCPL tasks in a number of respects. A more important implication of this result is that it questions the correctness of extrapolating results from MCPL tasks directly to natural tasks such as medical diagnosis.

The differences between our task environment and a typical MCPL task can be qualified by two more general design issues: the normative model with which performance is compared and the task characteristics. A normative decision model specifies the rules that people should follow in order to reach an optimal solution. MCPL tasks used a statistical rule, regression analysis, as a norm for performance. Many studies using statistical rules as a norm indicate serious biases in decisional behaviour (e.g. Jacob et al., 1986), and this has led to the conclusion that humans are bad intuitive statisticians. However, in order to generalise to specific tasks such as diagnosis, it is of major importance to know to what extent knowledge of these rules is a critical ability. In a complex dynamic environment, with much redundant information, optimal rules may differ from the ones prescribed by statistics. This would argue for increased effort in investigating the necessary components for optimal decision making in natural tasks.
Related to this issue are the characteristics of the task environment. It seems that in order to investigate decision abilities, the experimental environment should enable these abilities to be used. For example, if accurate hypothesis testing would be dependent on meaningful cues, and reasoning processes are intertwined with domain specific knowledge, the task environment should somehow contain these aspects. The general issue is that decision making may well be bound to constraints imposed by the environment in which particular decisional abilities have been learned. One cannot a priori assume that it is a general, formal ability, which can be employed in abstract task environments. Several studies, which addressed the influence of natural task characteristics on the rationality of decision making lend some support to similar conclusions. For example, decision strategies that are considered as biases in static environments may actually be functional in dynamic environments (Hogarth, 1981), and strategies that are considered nonoptimal, may actually be most efficient under time pressure (Payne et al., 1988). Furthermore, rational choice may need a different interpretation when short run or long run expectations are considered (Lopes, 1981). These studies all demonstrate the importance of incorporating critical ingredients of the natural environment in one's task environment.

The second research question addressed the differential effect of unreliable information on selection and integration of information. Our results showed that although unreliable information delayed the selection of correct information, no effect was found for information integration. In both conditions, subjects learned in the same manner how to combine information in the normative way. Each subject started out with a noncompensatory strategy and gradually learned to take trade-offs into account, that is, employ a compensatory strategy. The usual explanation is that a compensatory strategy requires more mental capacity than a noncompensatory strategy (Payne, 1982; Olshavsky, 1979). Thus, our subjects should be gradually able to use a compensatory strategy as a result of a reduced number of dimensions over learning trials.

Results from the present experiment, however, are not fully in line with the information load explanation. Subjects receiving unreliable information started using a compensatory strategy just as soon as those receiving reliable information, even though the former ones considered more information for a longer time period. This means that learning to use a compensatory strategy is not directly related to the selection of only relevant information. A more plausible explanation, then, is that subjects learned the value of integration strategies, independent of the information that is selected. More specifically, they learned that they had to take trade-offs into account, and thereby use a compensatory strategy.

A major difference between these explanations, information load and strategy value, is whether strategy choice is driven by factors external or internal to the decision maker. The information load explanation suggests that decision makers mechanically respond to task complexity because of structural capacity limitations. The strategy value explanation on the other hand, assumes that rational choices are made, based on strategies that the decision maker considers to be optimal given the decision problem at hand. This implies that subjects have to know how to evaluate strategy value, and employ the one most useful given the task environment. Although our results suggest the last explanation, more specific research is needed in order to draw well-based and more robust conclusions.

REFERENCES

Billings, R.S. and Marcus, S.A. (1983). Measures of compensatory and noncompensatory models of decision behavior: process tracing versus policy capturing, Organizational behavior and human performance, 31, 331-352.

Boshuizen, H.P.A. (1989). De ontwikkeling van medische expertise; een cognitief- psychologische benadering, Ongepubliceerd proefschrift: Rijksuniversiteit Limburg.

Brehmer, B. (1980). In one word: not from experience, Acta psychologica, 45, 223-241.

Brehmer, B. (1987). Note on subjects' hypotheses in multiple-cue probability learning, Organizational behavior and human decision processes, 40, 323-329.

Brehmer, B. and Kuylenstierna, J. (1978). Task information and performance in probabilistic inference tasks, Organizational behavior and human performance, 22, 445-464.

Brunswik, E. (1952). The conceptual framework of psychology, In: International encyclopedia of unified sciences, 1, Chicago: University of Chicago Press.

Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.), Berkeley: University of California Press.

Connelly, D.P. and Johnson, P.E. (1980). The medical problem solving process, Human pathology, 11, 412-419.

Dawes, R.M. and Corrigan, B. (1974). Linear models in decision making, Psychological bulletin, 81, 95-106

Doorne, van H. (1987). Milsim: simulatieprogramma voor het onderzoek naar complexe besluitvormingsprocessen, Report IZF 1987 I-5, TNO Institute for perception, Soesterberg, The Netherlands.

Einhorn, H.J. and Hogarth, R.M. (1975). Unit weighting schemes for decision making, Organizational behavior and human performance, 13, 171-192.

Elstein, A.S., Shulman, L.S. and Sprafka, S.A. (1978). Medical problem solving: An analysis of clinical reasoning, Cambridge Mass.: Harvard University Press

Garb, H.N. (1989). Clinical judgment, clinical training, and professional experience, Psychological bulletin, 105, 387-396.

Hammond, K.R., McClelland, G.H. and Mumpower, J. (1980). Human judgment and decision making: theories, methods and procedures, New York: Praeger.

Hogarth, R.M. (1981). Beyond discrete biases: functional and dysfunctional aspects of judgmental heuristics, Psychological bulletin, 90, 2, 197-217.

Hursch, C.J., Hammond, K.R., and Hursch, J.L. (1964). Some methodological considerations in multiple cue probability studies, Psychological bulletin, 71, 42-60.

Jacob, V.S., Gaultney, L.D., and Salvendy, G. (1986). Strategies and biases in human decision making and their implications for expert systems, Behaviour and information technology, 5, 119-140.

Johansson, R. and Brehmer, B. (1979). Inferences from incomplete information -A note, Organizational behavior and human performance, 24, 141-145.

Johnson, P.E., Dur n, A.S., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P.J. and Swanson, D.B. (1981). Expertise and error in diagnostic reasoning, Cognitive science, 5, 235-283.

Keren, G. and Van Doorne, H. (1987). Simulating decision making processes in computer systems: a proposal for an experimental paradigm, Report IZF 1987-37, TNO Institute for perception, Soesterberg, The Netherlands.

Klayman, J. (1988a). Cue discovery in probabilistic environments: uncertainty and experimentation, Journal of experimental psychology: learning, memory and cognition, 14, 2, 317-330.

Klayman, J. (1988b). On the how and why (not) of learning from outcomes, In: B. Brehmer and C.R.B. Joyce, Human judgment: The SJT view, Amsterdam: Elsevier.

Lopes, L.L. (1981). Decision making in the short run: notes comments and new findings, Journal of experimental psychology: human learning and memory, 7, 5, 377-385.

Olshavsky, R.W. (1979). Task complexity and contingent processing in decision making: a replication and extension, Organizational behavior and human performance, 24, 300-316.

Payne, J.W. (1976). Task complexity and contingent processing in decision making: an information search and protocol analysis, Organizational behavior and human performance, 16, 366-387.

Payne, J.W. (1982). Contingent decision behavior, Psychological bulletin, 92, 2, 382-402.

Payne, J.W., Bettman, J.R. and Johnson, E.J. (1988). Adaptive strategy selection in decision making, Journal of experimental psychology: learning, memory and cognition, 14, 534-552.

Schraagen, J.M.C. (1986). Expert-novice differences and their implication for knowledge elicitation techniques, Report IZF 1986-34, TNO Institute for perception, Soesterberg, The Netherlands.

Schweiger, D.M., Anderson, C.R. and Locke, E.A. (1985). Complex decision making: a longitudinal study of process and performance, Organizational behavior and human decision processes, 36, 245-272.

Sniezek, J.A. (1986). The role of variable labels in cue probability learning tasks, Organizational behavior and human decision processes, 38, 141-161.

Svenson, O. (1979). Process descriptions of decision making, Organizational behavior and human performance, 23, 86-112.

Woods, D.D. and Roth, E.M. (1988). Cognitive engineering: human problem solving with tools, Human factors, 30, 415-430.

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) TD 89-1058	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER IZF 1989-29
4. PROJECT/TASK/WORK UNIT NO. 733.2	5. CONTRACT NUMBER B87-38	6. REPORT DATE September 6, 1989
7. NUMBER OF PAGES 25	8. NUMBER OF REFERENCES 32	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

The influence of reliability of information on decision performance

11. AUTHOR(S)

J.H. Kerstholt

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Institute for Perception
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

TNO Division of National Defence Research
Koningin Marialaan 21
2595 GA DEN HAAG

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

Judgement tasks require knowledge of relations between symptoms (cues) and diagnosis. These relations are often probabilistic in nature, and decisions therefore have to be based on more or less unreliable cues. Many so-called Multiple Cue Probability Learning (MCPL) studies have shown that learning such relations is seriously impaired, even by a small degree of uncertainty. On the other hand, there is ample evidence that in natural judgement tasks such as medical diagnosis or market analysis, people are well able to learn from experience. In the present experiment learning under uncertainty is investigated in a relatively natural task that corresponds more closely with typical diagnosis tasks than the MCPL tasks. Subjects were able to learn the underlying model in the probabilistic mode, be it at a lower rate. The delay only concerned the information selection process, not that of information integration.

Foreign [illegible]: Netherlands, Translations, Decision Making (EG)

16. DESCRIPTORS	IDENTIFIERS
Decision making Task complexity Learning	

17a. SECURITY CLASSIFICATION (OF REPORT) -	17b. SECURITY CLASSIFICATION (OF PAGE) -	17c. SECURITY CLASSIFICATION (OF ABSTRACT) -
18. DISTRIBUTION/AVAILABILITY STATEMENT Unlimited availability		17d. SECURITY CLASSIFICATION (OF TITLES) -

Page - 26

VERZENDLIJST

1. Hoofddirecteur van de Hoofdgroep Defensieonderzoek TNO

2. Directie Wetenschappelijk Onderzoek en Ontwikkeling Defensie

3. (Hoofd Wetenschappelijk Onderzoek KL
Plv. Hoofd Wetenschappelijk Onderzoek KL

4,5. Hoofd Wetenschappelijk Onderzoek KLu

6. (Hoofd Wetenschappelijk Onderzoek KM
Plv. Hoofd Wetenschappelijk Onderzoek KM

7. Wnd. Directeur Militair Geneeskundige Diensten
Cdre vliegerarts H.H.M. van den Biggelaar

8. Inspecteur Geneeskundige Dienst KL
Brig.Gen.-arts B.C. Mels

9. Inspecteur Geneeskundige Dienst KLu
Cdre J.Th. Versteeg

10. Inspecteur Geneeskundige Dienst Zeemacht
Cdr-arts A.J. Noordhoek

11, 12, 13. Hoofd van het Wetensch. en Techn. Doc.- en Inform.
Centrum voor de Krijgsmacht

LEDEN WAARNEMINGS CONTACT COMMISSIE

14. Maj.Ir. W.C.M. Bouwmans
15. LTZAR1 F.D.J.R. Feunekes
16. Dr. N. Guns
17. Drs. C.W. Lamberts
18. Ir. P.H. van Overbeek
19. Drs. W. Pelt
20. Maj. dierenarts H.W. Poen
21. Drs. F.H.J.I. Rameckers
22. LTZSD2OC KV Drs. M.B.A.M. Scheffers
23. Prof.Ir. C. van Schooneveld
24. LKol.Drs. H.W. de Swart
25. Kol. vliegerarts B. Voorsluijs

Extra exemplaren van dit rapport kunnen worden aangevraagd door tussenkomst van de HWOs of de DWOO.
