

---

U M E Å P S Y C H O L O G I C A L R E P O R T S

Supplement Series No. 9 1982

DEPARTMENT OF PSYCHOLOGY  
UNIVERSITY OF UMEÅ  
S-901 87 UMEÅ, SWEDEN



LEARNING OF PROBABILISTIC INFERENCE TASKS: EFFECTS OF UNCERTAINTY  
AND FUNCTION FORM

Håkan Alm

---

LEARNING OF PROBABILISTIC INFERENCE TASKS: EFFECTS OF UNCERTAINTY  
AND FUNCTION FORM

Akademisk avhandling

som med tillstånd av rektorsämbetet vid Umeå universitet för vinnande  
av filosofie doktorsexamen framlägges till offentlig granskning vid  
Psykologiska institutionen, Seminarium 2, Rådhusplanaden 2, Umeå  
universitet den 1 juni 1982, kl 10.15.

av

Håkan Alm  
fil kand



## ABSTRACT

Alm, H. 1982. Learning of probabilistic inference tasks: Effects of uncertainty and function form.

This thesis is concerned with the problem of how people learn to use uncertain information for making judgments. The general framework for the thesis is Social Judgment Theory (SJT). First the SJT paradigm, and some research conducted within the paradigm, is briefly described, and a series of four empirical studies is summarized. The studies are concerned with two factors that have been found to have great effect on subjects' achievement in cue probability learning (CPL) tasks: task predictability, and the form of the function relating cue and criterion. The effects of these two factors were studied in experiments employing cue-probability learning tasks. The studies concerned with task predictability addressed the following questions (a) Do subjects understand the probabilistic nature of CPL-tasks? (b) Are subjects able to detect that a random task is, in fact, random, a study undertaken to test an aspect of Seligman's "theory of helplessness". This was also an attempt to bring emotional factors more in focus. (c) Do subjects use data from the task only to test hypotheses, or do they use data also to construct hypotheses?

The results showed that (a) subjects do not seem to be able to cope with probabilistic tasks in an optimal statistical manner. Instead they seem to use a deterministic approach to the tasks, because they do not understand the probabilistic nature of the task. (b) Task predictability affects subjects' mood, but not in the way predicted by Seligman. (c) Subjects seem to use data from the task only to test their hypotheses. The results thus supported the hypotheses sampling model for the learning of CPL-tasks.

As for the factor of function form, the following questions were addressed, (a) What hypotheses about relations between variables do subjects have? (b) Is the difficulties subjects have in learning complex rules in CPL-tasks due to a low availability of hypotheses about complex rules? The results showed that, (a) the hypothesis hierarchy as revealed in the present experiments was in general agreement with earlier results. However, few nonlinear hypotheses were observed, and other rules than functional rules were observed. (b) The difficulties subjects have to learn complex rules in CPL-tasks do not seem to be caused by low availability of rules.

Finally, some suggestions are given for how the SJT-paradigm should be developed. Specifically, it is suggested that the effects of emotional factors should be given more attention, and that the paradigm should be turned into a more general hypothesis testing model.

Key words: Inference, learning, cognition.

The present doctoral dissertation is comprised of this summary and the following studies, referred to by their roman numerals:

- Study I      Alm, H., & Brehmer, B. Effects of task predictability on subjects' performance and mood in probabilistic inference tasks - a test of Seligmans theory. Umeå Psychological Reports, No 149, 1980
- Study II     Alm, H., & Brehmer, B. Hypotheses about cue-criterion relations in linear and random inference tasks. Umeå Psychological Reports, No 164, 1982
- Study III    Alm, H., & Brehmer, B. How subjects explain their failure to learn probabilistic inference tasks. Umeå Psychological Reports, No 158, 1981
- Study IV     Alm, H. Effects of pretraining on the construction of complex rules in probabilistic inference tasks. Umeå Psychological Reports, No 162, 1982

#### ACKNOWLEDGEMENTS

I am indebted to my supervisor, Berndt Brehmer. He has guided, stimulated, advised, supported me through this work. He has provided a friendly and very stimulating atmosphere in which hard work was turned into exciting work (although still hard). He has also travelled quite a lot to do all this. Without - all his help and wisdom, this thesis would probably never have been finished.

I have also received help, critical comments and support from my colleagues, and I thank you all for that. Especially I want to thank Jörgen Garvill for many exciting theoretical discussions and for methodological assistance.

For help in running one experiment I wish to thank Margareta Grensjö. Roger Johansson and Lars-Erik Warg have assisted me in the computational work, Per Asplund has fixed my "computer problems". Mona Wiklund has typed most of the reports, Siv Jungebjörk and Brita Westling have also typed reports, and Siv has drawn all the figures, Stig Lindkvist has put the papers together into

reports. Finally I wish to thank my wife for support and encouragement, and my daughter for showing me how hard it is to make correct judgments.

This dissertation has been supported by grants from the Swedish Council for Research in the Humanities and Social Sciences.

This thesis is concerned with the problem of how people learn to use uncertain information for making judgments. Such judgments are an important part of many decisions (see, e.g., Hammond, 1974), for decision problems is often characterized by the fact that what one needs to know is distal (not manifest) in time or space. Consequently, the decisions have to be based on judgment, or prediction, rather than fact. Examples of such predictions of future events, may be found in weather forecasts or prognoses about economical matters.

In situations of this kind, the person must use the information available to draw conclusions about something not available. This means that s/he must make a guess, or try a hypothesis about the event s/he wants to predict.

Many situations of great importance for man are of this kind. This has also been recognized by many psychologists. Thus, Jerome Bruner (1957) wrote that "to go beyond the information given" is the criterion of knowledge. Tolman and Brunswik (1935) and Hammond (1974) also pointed to the importance of inferences.

A characteristic feature of many decision problems is that the information which is available does not allow perfect predictions of the distal state. One reason for that may be that we do not know all the cues which should be used to predict perfectly. The relation between what one knows and what one wants to know is therefore often uncertain.

### Social Judgment Theory (SJT)

SJT (Hammond, Stewart, Brehmer & Steinmann, 1975) is a general framework, or a "pretheoretical framework" (Brehmer, 1979a), for the study of judgment. It is used to describe the problem area, and provides some guidelines for the analysis of the area. Within this general descriptive approach, models are being developed to explain particular phenomena, such as the hypothesis testing model of learning (Brehmer, 1974, 1979a) and a structural model for interpersonal conflicts (Brehmer, 1976a; Brehmer & Hammond, 1977). Work within the paradigm has consisted both of laboratory studies and applications of laboratory findings (Hammond, et al., 1977) and there is a close interaction between the laboratory and "real life".

SJT has its roots in the ideas of Egon Brunswik (1952, 1956). According to Brunswik, the task for psychology was to study the interaction between organism and environment. Specifically, "both organism and environment will have to be seen as systems, each with properties of its own, yet both hewn from basically the same block. Each has surface and depth, or overt and covert regions" ... Brunswik (1957, p.5). Both systems must be treated with the same respect. It is necessary to investigate the structure of the environment as carefully as the structure of the organism.

To study the interaction between the two systems, one must have compatible concepts for both systems, otherwise the systems cannot be related and compared. This principle is called "the principle of parallel concepts" in SJT. According to Brunswik, and to proponents of SJT, the organism has adapted to the environment; the environment existed before the organism. Therefore environmental concepts should have priority, and the descriptions of the systems should rely on environmental concepts.

It should however be noted that not all proponents of SJT share this view. Brehmer (personal communication) argues that since the environment to a large extent is a creation of man, it would be possible to argue that cognitive concepts should have priority. Brehmer (1979a) also argues that it is impossible to see a cognitive task as something independent of the person dealing with the task. Cognitive tasks are created by the interest or motivation that a subject has and do not have an existence which is independent of a knower.

However, in this thesis, the view that environmental concepts should have priority is adopted.

Brunswik maintained that the relation between cues and events in the distal region is an uncertain relation, that is, no one-to-one relation exists between this two layers. Brunswik meant that all important tasks have a probabilistic nature (an uncertain relation between cues and events), and proponents for SJT are of the same opinion (see Brehmer, 1976b for a discussion of that topic).

This has important implications for the description of the environment and

the organism. Specifically, it implies that it is necessary to have a descriptive system which can handle uncertainty, i.e., statistical concepts must be used. According to the principle of parallel concepts, such concepts must then be used both for the organism and for the environment. These ideas are also expressed in Brunswik's so called "lens model", see Fig. 1.

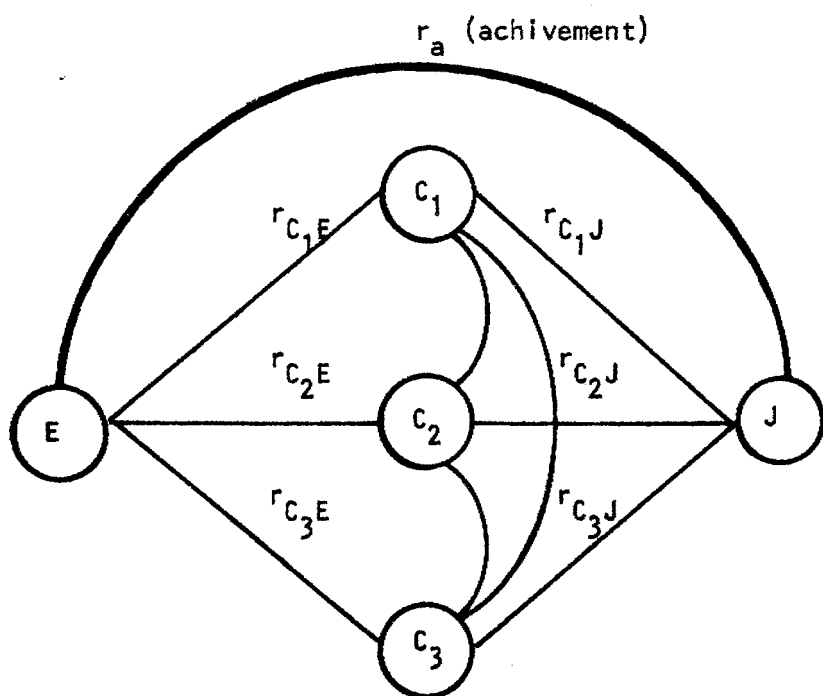


Figure 1. Brunswik's lens model.

To the left in the figure we have the ecological side, defining the cognitive task. A distal event (E) gives rise to three events in the proximal layer, (C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>). The relation between the event (E) and the cues (C<sub>1</sub> - C<sub>3</sub>) representing the event is uncertain, and can be described in terms of correlations, ( $r_{C_1E}$ ;  $r_{C_2E}$ ;  $r_{C_3E}$ ). The multiple correlation,  $R_e$ , between the cues and the event is a description of how much of the variation in E that can be predicted from these cues in a least squares sense. There are also correlations among the cues, the environment is a "causal texture"



(Toman & Brunswik, 1935). An important implication of that is that if one cue is not available, the organism can use another cue. Brunswik called this phenomenon "vicarious functioning", i.e., one goal can be reached in more than one way (equifinality).

The right side of the lens represents the cognitive system. At first we can note that the model is symmetric. The ecology is logically and historically prior to the organism, and the organism must, adapt to the environment. To achieve this adaptation, the organism must have knowledge about the environment. One way of conceptualizing this is to say that the cognitive system must be a model of the environment (see Brehmer, 1979a).

The cognitive system is analyzed in terms of the relation between the cues and the judgments (J) made by the organism. (We are here making the assumption that cues are veridically perceived.)

The multiple correlation,  $R_s$ , between cues and judgments expresses the consistency (predictability) of the organism. Finally, the long arch denotes the correlation between the distal event in the ecology (E), and the distal event in the organism (J), that is, the judgment the organism makes. This relation is called achievement, and is measured in terms of the correlation between E and J, which is a measure of the degree to which the organism has succeeded with the inference of the distal event.

The relation between environment and organism has first got a mathematical expression by Hursch, Hammond & Hursch (1964) and was later modified by Tucker (1964), see equation 1.

$$r_a = G R_e R_s + C \sqrt{1-R_s^2} \sqrt{1-R_e^2} \quad \text{Eq. 1.}$$

where

$r_a$  = achievement, the correlation between E and J.

G = the correlation between the linearly predictable variance in the task system and that in the cognitive system. It expresses the degree to which the organism has detected the linear aspects of the task.

$R_e$  = the linear predictability in the criterion, the distal event in the ecology.

$R_s$  = the linear predictability of the organisms' judgments.

$C$  = the correlation between residual values in the criterion and judgment after linear components have been partialled out.  $C$  shows the extent to which the organism has detected the nonlinear aspects of the task.

In most cases, however, it is feasible to reduce all relations to linear form - and to use a simpler version,

$$r_a = G R_e R_s \quad \text{Eq. 2}$$

In this version,  $G$  is an index of the extent to which the systematic aspects of the cognitive system match the systematic aspects of the task,  $R_e$  gives a measure of the total predictability of the task, and  $R_s$  gives an index of the predictability on consistency of the cognitive system.

From equation 2 it can be seen that task predictability ( $R_e$ ) introduces a limit for a subjects achievement ( $r_a$ ). A subjects achievement can never exceed task predictability. To reach this optimal level of achievement, a subject must find the correct model for the environment ( $G=1$ ) and use that model with perfect consistency ( $R_s = 1$ ).

A conclusion here is that a low degree of achievement does not necessarily mean that a subject is performing very badly. It may simply mean that  $R_e$  is very low, and that a subject may be behaving as well as he possibly could. Low achievement (when the effect or  $R_e$  is considered) may, for instance, mean that (a) the subject has not been able to find the correct model for the environment, or (b) the subject has found the correct model for the environment, but is using the model in an inconsistent manner.

#### Cognitive tasks within SJT

A cognitive task can be described in two different ways, (1) in terms of content, that is, what the task is about, for example, whether it is a weather

forecast or a medical diagnosis and (2) in terms of formal characteristics. Concerning effects of content, see Muchinsky and Dudycha (1975) and Warg and Brehmer (1982a,b).

The formal characteristics of a task is a number of general dimensions which are common for all inference tasks, independently of the content of the task. There are seven formal characteristics, which are divided into two classes.

The first of these is called surface characteristics, and comprises (1) number of cues, (2) measurement level for the cues, nominal or quantitative, (3) intercorrelations among the cues (e.g. intersubstitutability of cues). These three characteristics concerns the proximal layer of the task. The second class is called system characteristics, and consists of (1) the form of the function connecting each cue and the distal variable, (2) relative weight for each cue, that is, the importance for each cue, (3) integration rule for combining the information from the cues into the value of the distal variable, (4) predictability of the system, given the optimal integration rule. System characteristics concern the relations between the cues and the distal variable.

According to the principle of parallel concepts, both the ecological system and the cognitive system must be described in terms of these dimensions, i.e., with the help of statistical concepts. This idea does not imply that the cognitive system works according to probabilistic principles; statistics has complete certainty as a limiting case.

#### Learning tasks.

In SJT, so called cue probability learning tasks (CPL-tasks) are used to analyze how people learn to use probabilistic cues for making inferences. A typical CPL-task requires the subject to learn to use one cue (single-cue probability learning task, SPL-task) or many cues (multiple-cue probability learning task, MPL-task) to predict the state of a criterion variable. Generally both cues and criterion are quantitative variables, although tasks with nonmetric cue and criterion variables have also been studied (e.g., Björkman, 1973).

The tasks are constructed to include the dimensions mentioned above. Most often they lack meaningful content. There are some exceptions, see e.g.,

Muchinsky and Dudycha (1975); Warg and Brehmer (1982 a,b). To investigate the effects of content is now seen as a very important task in SJT. Proponents of SJT choose not to sample actual tasks from the environment despite, their Brunswikian orientation. Such a sampling of actual tasks would only give information valid for tasks of a given kind in a certain cultural age. This information would, in the worst case, say nothing about the new inference tasks which are constantly created (Brehmer, 1979a). Instead of such a substantive representative sampling of tasks, formal representative sampling (Hammond, 1966) is used. This means that one uses the formal characteristics of the task, and samples in such a way that all the formal characteristics of inference tasks will in the end be covered. This would, hopefully, provide information valid for all types of tasks, independently of cultural age. The information found in the laboratory is then often applied in real life settings (see Hammond, 1971, Hammond et al., 1977). This is a way of testing laboratory findings to see if they are valid outside the laboratory, but such studies also provide important information for further laboratory work.

CPL-tasks are designed to simulate probabilistic inference tasks which require people to make repeated inferences of the same kind. An example of such tasks is clinical inference using psychological tests, e.g., MMPI. (For a comparison of results from studies of CPL and clinical inference, see Brehmer, 1976b.)

In the experiments the subject is provided with a number of cue-criterion pairs, and has to learn the relation between the cues and criterion on the basis of outcome feedback, i.e., by first observing the cue information, then making a prediction of the criterion, and finally observing the actual criterion value. The procedure can be varied in a number of ways, e.g., feedback or the prediction can be left out.

### Hypothesis testing models

This year hypothesis testing models celebrate their 50th anniversary in psychology. The first model of this kind was proposed by Krechevsky in 1932, who quite contrary to the "Zeitgeist", developed a hypothesis testing model for discrimination learning in the white rat. Krechevsky had noted that his rats did not behave in a random manner before learning the correct discrimination. Instead, the rats behaved in a systematic manner until they finally

found the correct solution, and this he took to define the rats' hypotheses. There were many critics of that kind of "mentalism", and Krechevsky's results were soon explained away in behavioristic terms. As a consequence, hypothesis testing models left the stage momentarily.

Then, in 1956, Bruner, Goodnow and Austin published a theory for concept learning in humans. In the theory they assumed that subjects learn concepts by testing hypotheses about the concepts. They also assumed that the subjects' hypotheses could be verbalized.

In the sixties, some prominent learning theorists, Estes and Restle, changed their views. Estes, 1960, realized that learning in many tasks proceed in an "all-or-none" fashion, thus supporting a hypothesis view on learning. An article published in 1960 by Estes discussed this problem, and probably had a great influence on the thinking of many psychologists. Restle presented a mathematical theory of discrimination learning, founded on the assumption that subjects used "strategies" to solve the problem.

Important followers of Estes and Restle are, among others, Bower and Trabasso, who refined Restle's model (see, e.g., Levine, 1975). Levine, finally, developed a hypothesis testing model for the learning of concepts. He also modified some of Restle's assumptions, and made some very important research concerning how subjects sample and test hypotheses. Levine et al. also investigated hypotheses in children, and found that children's behavior also can be described in terms of hypotheses. They also found that subjects do not seem to sample hypotheses in an random manner, but rather according to some strategy.

It should be noted that the learning tasks used by all these hypothesis models have a special property. In these tasks, the subjects can get their hypotheses by sampling dimensions from the stimuli. The stimulus material consists of already familiar concepts, which can be used to form hypotheses. This guarantees that the subjects have the hypotheses relevant to the task although he may not think of it. A conclusion of this is that we do not know if these theories really can say something interesting about situations where subjects have to learn new concepts (see Bolton, 1972).

A CPL-task, on the other hand, is a rule learning task, and for such a task no single instance can define the correct rule, and the hypothesis must come from the subject (see Brehmer, 1980a for a discussion).

To explain the effect of function form on subjects' learning of probabilistic inference tasks, Brehmer (1974) developed a hypothesis-testing model. The model is built on the assumption that the subjects have a hierarchy of hypotheses about functional relations between scaled variables, and that they sample hypotheses from the hierarchy according to their strength, when learning a CPL task.

Learning of CPL tasks is seen as a two-stage activity. In the first stage, the subject sample a hypothesis from an internal hierarchy of hypotheses. This hypothesis is then tested against data. When the subject has found the appropriate rule, s/he must learn to use that rule, that is, to find the parameters of the rule. According to Brehmer (1974), the subjects have an internal fixed hierarchy of hypotheses about functional relations. Hypotheses are assumed to be sampled, one at the time, from this hierarchy, according to their sampling probabilities. The rule with the highest sampling probability was the positive linear rule, followed by the negative linear rule, the inverted U-shaped rule, and finally the U-shaped rule. According to this formulation of the model, the subjects can only learn the rules that exist in their hierarchy. Other rules have to be approximated by means of already stored rules, or not be learned at all. The first results (Brehmer, 1974; Brehmer et al., 1974) agreed closely with the predictions from the sampling model, based on the assumption that subjects have a limited number of hypotheses, and that they sample and test hypotheses according to their relative strength. Specifically, the results showed that (a) rate of learning, as well as the frequency and order of occurrence of various hypotheses could be predicted from indices of hypothesis strength, (b) the frequency of various incorrect hypotheses is independent of the functional rule in the task, (c) that the subjects had very few rules in their hierarchy, (d) that new hypotheses occurred only in the beginning of an experiment, and that there was considerable resampling of previously discarded hypotheses.

Later studies showed, however, that subjects are able to learn rules that presumably do not exist in their hierarchy. (Brehmer, 1976c, 1980b; Brehmer &



Kuylenslierna, 1979). These studies employed tasks with J-shaped rules. The results showed that some of the subjects did in fact use approximations to that rule, such as a positive linear or a U-shaped rule, but about half of the subjects did in fact learn the J-shaped rule, thus falsifying the original sampling model. The model was now changed to a "hypotheses-construction" model (Brehmer, 1979a). According to this model, subjects can expand their pool of hypotheses by combining a set of elements into new rules. The subjects still use data only to test their rules, and building blocks for the rules are sampled from stored material. A specific hypothesis consists of a number of such elements, sampled together, and combined in a certain manner.

Brehmer (1980a) argued that hypotheses generated by the subjects are not restricted to functional relations. Hypotheses may differ in generality, and functional rules exists side by side with nonfunctional rules. That is, the subjects have a hierarchical arrangement of rules of different generality, a decision tree starting with the decision "rule - no rule", see Figure 2. The selection of rules is still assumed to be independent of the task, that is, the task is only used in order to test the rule. This is supposed to be true for rules at all levels of generality.

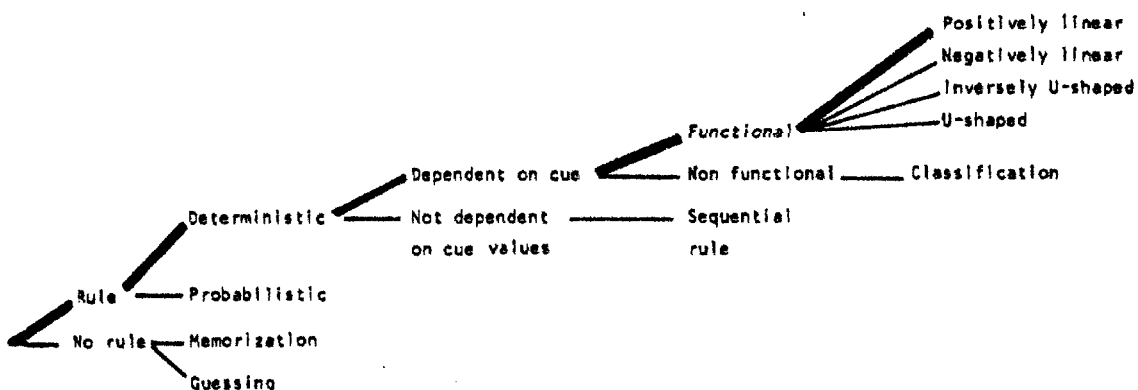


Figure 2. Summary of results form studies on probabilistic inference tasks.

Hagafors (1981) suggested a modification of the hypothesis testing model. Hagafors rejects the idea of a task-independent hypothesis sampling procedure, and proposes that the hypotheses depend on the subjects representation of the task. Situational variables affect the subjects' representation of the task, which in turn, has effects on the subjects approach to the task, as well as on how s/he is able to utilize feedback. A subject is performing three activities to cope with a task.

- a. Hypothesis - selection, which is assumed to be released by the relation between task - representation and previous experiences and existing schemata.
- b. Hypothesis - validation, which also is assumed to be regulated by how the task is represented. The representation is related to earlier experiences and previously acquired schemata for hypothesis validation.
- c. Hypothesis - adaption, the search for appropriate parameters of an accepted hypothesis.

These activities (a-c) have their counterparts in the model by Brehmer (1980a). What Hagafors has added is that it is necessary to consider the subjects representation of the task.

#### A brief outline of the present research program and problemarea

It has been found that two factors have a great effect on subjects achievement ( $r_g$ ) in CPL-tasks. The first factor is task predictability, and the second is the form of the function relating cue and criterion (e.g., Brehmer, 1980a). Within the general framework of SJT, these two problems were addressed in experiments employing CPL-tasks.

#### The effects of task predictability

Under this heading the following questions are addressed:

- (a) Results both from laboratory studies (e.g., Brehmer & Kuylénstierna, 1978) and from applied work (e.g., Brehmer, 1976b) show that subjects fail to adopt an optimal strategy in probabilistic inference tasks. One attempt to explain

subjects inconsistent and suboptimal behavior in CPL-tasks hypothesizes that lack of understanding of the probabilistic nature of the task is an important reason (Brehmer & Kuypenstierna, 1978, 1980). Part of the aim of studies I and III was to test this explanation.

(b) Seligman's theory of "learned helplessness" (Seligman, 1975) assumes that depression is caused by the discovery that something cannot be controlled, i.e. that there is no correlation between one's actions and the outcome. An important question, then, is whether people can discover that a correlation is zero. In study I, this problem was studied as a practical application of the question if subjects understand randomness. This study also constitutes an attempt to bring emotional factors into the SJT-paradigm.

(c) According to Brehmer (1980a) data from the task in a probabilistic inference task are used only to test hypotheses, but not to form hypotheses. This conjecture was tested in study II.

#### Effects of function form

Under this rubric, the following questions are treated:

(a) The nature of the hypothesis hierarchy assumed in Brehmer's (1974) hypothesis testing model, was investigated with a new method in study II.

(b) Study IV, finally, reports a test of the hypothesis that lack of availability of complex rules is an explanation for the fact that subjects have difficulties in learning complex rules in CPL-tasks.

#### The effects of task predictability

Task predictability refers to the level of uncertainty in the cue-criterion system, and it is usually defined in terms of the correlation between the cues and the criterion,  $r_e$ . As explained above, this is an important parameter of all inference tasks because it gives the upper limit of achievement for that task, i.e., it defines the highest correlation a subject can achieve between his judgments and the criterion values, if he uses the cues optimally. The condition of optimality is given by statistical decision theory, and for the present case, statistical decision theory specifies that the cues in the task should be used with perfect consistency, regardless of the predictability of

the task (see Hursch, Hammond & Hursch, 1964). This means that the correlation between the cues and the judgments should be unity, regardless of the correlation between the cues and the criterion (given that  $r_e = 0$ ). The results show, however, that people performing probabilistic inference tasks do not follow the recommendations of statistical decision theory. They are generally inconsistent, i.e., the correlation between cues and judgments is less than unity, and the level of inconsistency varies with the predictability of the task. This has been shown in a number of studies, including SPL-tasks (e.g., Naylor & Clark, 1968), MPL-tasks (e.g., Uhl, 1966), interpersonal learning (e.g., Brehmer, 1973a), policy conflict (e.g., Brehmer, 1973b) as well as in clinical inference by experienced clinicians (Brehmer, 1976b).

Instead of a statistical approach, subjects seem to use an inappropriate deterministic approach, and try to find a rule which will give them the correct answer on every trial instead of a statistical rule which minimizes the errors, (see Brehmer & Kuylensstierna, 1978, 1980; Johansson & Brehmer, 1979; Brehmer, 1980b).

Two hypotheses have been proposed to explain this phenomenon: (a) the subjects' information processing capacity is too limited (Kuylensstierna & Brehmer, 1981), (b) the subjects do not understand the probabilistic nature of the tasks.

A way of testing this hypothesis would be to supply what should be critical information to the subjects by means of instruction. If the explanation for their suboptimality is that they lack information, such instructions should lead to more optimal behavior, and by varying the instructions, one should be able to assess exactly what knowledge the subjects lack for selecting the strategy they should. This approach has been tried in a number of studies (Brehmer & Kuylensstierna, 1978, 1980; Johansson & Brehmer, 1979; Kuylensstierna & Brehmer, 1981), but with little success; instructions do not seem to improve the subjects strategies.

Another approach to this problem is represented by this thesis, where the aim was to investigate how the subjects understand uncertain information.

Specifically, the aim is to assess what information about probabilism subjects

are able to pick up from CPL tasks. This should give useful information about whether subjects lack appropriate schemata to interpret uncertainty or not.

The purpose of studies I and III was to investigate to what extent the subjects interpret and handle randomness. These studies used some methods to assess the subjects' understanding of randomness. Specifically, a variety of verbal reports, including the subjects' explanations for their failure to learn the task, were used.

Study I also aimed at testing Seligmans "theory of helplessness". This may be seen as a more practical application of the question if subjects really understand randomness. It is also an attempt to bring emotional factors into the SJT-paradigm.

Two hypotheses in the "theory of helplessness" are of interest here. The first hypothesis is that the subjects are able to learn when the correlation between their actions and events in the ecology is zero. From a hypothesis testing view of learning, this is possible only if the subjects have a hypothesis about randomness. This question is of general interest because it concerns what kind of hypotheses subjects have.

The second is the hypothesis that there are effects of task predictability on the subjects' mood. This question is of interest because many decisions are probably made under various affective states. If task predictability affects the subjects' mood, the next question concerns if mood can affect, for instance, the subjects' ability to learn functional relations. This second question is, however, outside the scope of the present thesis.

#### Do subjects understand randomness?

If subjects understand randomness, they should be able to learn that an objectively random task is in fact random. As already mentioned, there are some studies within the SJT-paradigm which indicate that subjects do not adopt a statistical approach to CPL-tasks (see above for references). Instead they adopt a deterministic approach. If the subjects adopt a deterministic approach to a CPL-task that is random, then the search for a function relating cue and criterion can go on "for ever". Theoretically, there is an infinite number of mathematical functions, and it is always possible to get

validating evidence for some rule. The subjects can of course try many rules, and if no rule works, s/he can give up and blame the task for the failure. To declare that the task is impossible is, however, not the same as claiming that it is a random task. Rather, the task may be impossible for the subject (but not for other subjects), or it may be impossible for the subject right now (but not in the future). To realize that a task is random, the subject must realize that neither s/he or anyone else can do it better (with the exception for unusual luck), now or in the future, but that it is inherently impossible to make perfect predictions.

To investigate this problem, two experiments were conducted. Experiment 1 in study I contained a random and one nonrandom CPL-task, and some questions aimed at assessing subjects' thinking about the task. Experiment 2 in study I contained two nonrandom CPL-tasks, one low, and one high in task predictability. Experiment 2 also contained more questions which aimed at assessing subjects' thinking about the task.

### Result of experiment 1

The results of experiment 1 showed that most of the subjects did not behave in a random manner, although subjects in the random condition were less orderly than subjects in the nonrandom condition. Verbal reports about randomness revealed that more subjects believed in order in the random condition than in the nonrandom condition. There was no relation between the strategies of the subjects and their judgments about the nature of the task. It seemed that about half of the subjects who believed in randomness had a poor understanding of the concept of randomness. In sum, even though the results of experiment 1 are not totally clear, it is at least doubtful whether the subjects are able to use a statistical approach. A reasonable assumption from the results of experiment 1 is that at least some subjects lack the appropriate schema for interpreting a random task. Experiment 2 in study I was performed to remedy some shortcomings of experiment 1. Many subjects in the nonrandom condition considered the task to be random. This suggested that the manipulation of the independent variable was not powerful enough.

### Results of experiment 2

Experiment 2 showed that all but three subjects in the low validity condition made correct judgments about the nature of the task. It also showed that the



nature of the questions used to assess the subjects belief about randomness is important. Specifically, the results of experiment 2 showed that some subjects answered "no" in a consistent manner to the question about a relation between cue and criterion, but they did this when the task was not random. This finding points to the fact that it is not sufficient to have a hypothesis about randomness. The subject must also be able to test that hypothesis, or be able to refute all other hypotheses. So, even if some subjects in fact have a random hypothesis that does not mean that they will be able to use it, or that they ever will use it.

It seems clear from these two experiments that many subjects have a far from perfect understanding of relations between events. They often make incorrect judgments about a relation between two events, and there is no correlation between how they behave and their judgment about the task. This points to the importance of developing a theory about the connection between subjects cognitive activities and their behavior.

To make a further test of subjects understanding of randomness, study III was performed. One of the aims of study III was to investigate how subjects explain their suboptimal performance in CPL-tasks. The explanation for failure was thought to shed some light on subjects understanding of probabilistic tasks. The idea that subjects explain their performance in achievement oriented situations in terms of the four causal elements ability, effort, task difficulty and luck was used (Weiner et al., 1971). The prediction was that the subjects would be most willing to explain failure in terms of task difficulty, followed by low effort. For the elements low ability and bad luck no definite prediction was made.

The hypothesis was tested using a random and a highly predictable task. The use of a highly predictable task made it possible to investigate if the magnitude of errors would affect the subjects attributions for failure.

### Results of study III

The results of study III showed that the subjects did in fact explain their failure in terms of situational factors. Specifically, they used task difficulty to explain why they did not succeed. The results also showed that there was no difference in attributions between the random and the nonrandom task. These results were interpreted to mean that the subjects did not use a statistical approach to these tasks. The fact that they attributed failure in a random task to task difficulty and not to randomness pointed to that direction. Also the fact that the magnitude of error did not have any effect on subjects' attributions argues for the hypothesis that the subjects use a deterministic approach to these tasks. The fact that the subjects fail, rather than the magnitude of their failure, seemed to be the important fact. To conclude, when faced with a probabilistic task, subjects fail to realize that the task is probabilistic, and go searching for a deterministic rule. When they do not find this rule, they consider the task difficult, rather than random.

### Conclusions from studies I and III

Taken together, the results from studies I and III suggest that subjects take a deterministic rather than statistical approach to CPL-tasks. If that is true, a likely explanation for subjects' inconsistency when performing CPL-tasks is that they are trying to find the correct parameters for a deterministic rule. In a task where the rule to learn is a linear one, the subjects will constantly change the intercept and slope on the basis of outcome feedback in an attempt to find the rule. It should, however, be noted that this is not necessarily the only explanation. In experiment 2 in study I, it was found that some subjects answered "no" in a consistent manner to the question about randomness, but answered so when the task to learn was nonrandom. This could mean that they have a hypothesis about randomness, but they are unable to test it correctly.

It must also be noted that the fact that the magnitude of error did not have any effect on the subjects attribution may be a result of the research design. In a recent study by Roger Hagafors (1982), using a within persons design, it was found that task predictability had effects on the subjects perceived performance and confidence. This shows that subjects are able to discriminate error magnitudes, but does not affect the general conclusion that subjects do not understand the implication of a probabilistic task. It

would, however, be of interest to replicate the experiment in study III, using a within persons design. However, these results, together with other results (e.g., Wagenaar, 1970; Tversky & Kahneman, 1971, 1973) suggest that a very likely conclusion is that subjects really are taking a deterministic approach to these kind of tasks. The subjects seem to lack the appropriate schemata to interpret randomness.

### Effects of task predictability on the subjects mood and performance - a test of Seligmans theory

As mentioned earlier, the testing of Seligmans theory may be seen as a practical application of the question if the subjects really understand randomness, and as an attempt to bring emotional factors into the SJT-framework.

Seligman (1974, 1975) has proposed a theory, the "theory of helplessness" according to which lack of predictability may lead to depression. A basic assumption in that theory, then, is that subjects are able to learn that an objectively unpredictable task is in fact unpredictable. According to a hypothesis testing view of learning this is possible only if the subject has a hypothesis about randomness. As previously mentioned, many subjects do not seem to have such a hypothesis and those who have it adopt it when the task is not random. These results clearly contradict Seligmans statement: "Learning that events are independent of responding has a basic, simple, and indispensable place in the real life of men and animals" (Seligman, 1975, p.18).

The present experiment was, however, concerned with the other part of Seligmans theory; that predictability will have an effect on subjects' mood. Experiment 1 in study I included one random task and one task with low predictability. After the completion of the task, the subjects were asked some questions designed to assess what they thought about the task, and a mood questionnaire.

### Results of experiment 1

There were some differences in mood between the two groups, but the effects were rather weak. When the subjects were divided into two groups according to whether they believed the task to be random or not, practically all mood differences disappeared. That is, there were effects of objective randomness, but not of subjective randomness. In experiment 2 in study I, two

nonrandom tasks were used, one with low predictability, and one with high predictability, as well as a mood questionnaire.

### Results of experiment 2

A comparison between mood ratings in the random task in experiment 1 and the high predictability task in experiment 2 revealed some helplessness effects, but they were rather weak. It also revealed that subjects in the random condition rated themselves as more aggressive. This was interpreted to mean that a random task induces frustration rather than depression.

Some results from study III are also of importance here. The results from that study showed that subjects explained their failure to learn a CPL-task in terms of situational factors, especially in terms of "task difficulty". This finding explains why the subjects in experiment 1 and 2 in study I seemed to be frustrated rather than depressed; if the reason why one does not succeed is in the nature of the task, it makes sense to be aggressive, rather than depressed. These results also show that it is necessary to take subjects' attributions about their failure to perform well on some task into account to understand the effects that the task may have on their mood.

### Discussion of the results from studies I and III

As already mentioned, subjects seem to have difficulties to cope with tasks that contains some measure of random fluctuations. This finding is also in line with those reported by Wagenaar, 1970 and Tversky and Kahneman, 1971, 1973. Concerning emotional reactions, task predictability seemed to have an effect on the subjects' mood. An important problem for future studies is whether emotional reactions affect subjects' performance in CPL-tasks.

Concerning Seligman's theory, it could be possible that a subject not necessarily has to learn that an event is independent of its responses, in order to necessarily get depressed. It may be enough if a subject notes that s/he can not reach important goals or ideals, or can not avoid negative events. If so, the theory might work even if the subjects are unable to detect randomness.

### Effects of task predictability on subjects hypotheses

Brehmer (1974) has distinguished between two kinds of hypotheses testing

models: (a) a hypothesis sampling model, and (b) a hypothesis construction model. According to a hypothesis sampling model, a subject has a preestablished hierarchy of hypotheses. A hypothesis construction model, on the other hand, assumes that the subjects are able to construct new hypotheses when needed. A hypothesis sampling model predicts that the same hypotheses will be tried, regardless of the nature of the task, i.e. data from the task do not affect the selection of hypotheses but are used only to test hypotheses. One of the aims of study II was to assess the role of data from the task. In study II, one random and two nonrandom (one with low, and one with high predictability) CPL-tasks were used. If data are used only to test hypotheses, the same hypotheses should appear in all three conditions. Some results from study I is also of importance here.

### Results of study II

The main results were that the higher the task predictability, the greater the number of hypotheses involving functional rules. There were significant differences between all three conditions. It was also found that guessing strategies were more frequent in the low validity conditions. On the whole, there were no difference between the random and the low validity condition. There was no difference between the conditions in how many different functional rules were reported. Subjects had very few hypotheses, and practically no new hypotheses were introduced after the first block, and there was considerable resampling of old and discarded hypotheses. When the hypotheses in the hierarchy failed, the subjects seemed to switch to a guessing or memorization strategy, rather than to a construction process. All these results are generally consistent with a hypothesis sampling model of learning of probabilistic inference tasks. It was also found that the subjects tried the same hypotheses in all cue validity conditions, suggesting that the selection of hypotheses is indeed independent of the input data from the task.

### Results from study I

In study I it was found that 12 out of 16 subjects in the random condition claimed that there existed a correlation between cue and criterion. Only 6 out of 16 subjects in the low validity condition claimed that there was a correlation between cue and criterion. This finding is of importance, and suggests that the subjects may need some order in a task to consider it random. Similar findings are reported by Wagenaar (1970) for a different kind of

task.

#### Discussion of the results from studies II and I

The results showed that the subjects' guessing strategies were a function of task predictability. The lower the task predictability, the more guessing strategies. This raises two questions, first, why did subjects in the low predictability conditions switch to guessing strategies more often than subjects in the high predictability condition? One reason could be that testing of hypotheses is harder when task predictability is low (see Brehmer, 1979b). The feedback that the subjects receive varies a lot, and this makes interpretation difficult. Sometimes a low criterion value will be followed by a high criterion value, and sometimes by a low value. The first outcome could be a support for a negative linear rule, the second for a positive linear rule. This line of argument presupposes, that subjects may test one rule and note the magnitude and direction of the deviation, and that these two factors affect the choice of the next rule. For instance, suppose that the subject predicted a low criterion value for a low cue value, and received a high number. The subject may then note, not only that the prediction went wrong, but also in what direction it went wrong. As a consequence, the subject may search through his/hers hierarchy for a new rule that could fit the actual feedback values. This search process may take place on the falsifying trial, or on the next trial. Earlier results (e.g., Brehmer, et al., 1974) have shown, however, that the frequency of incorrect hypotheses is independent of the rule to learn in a task, and this argues against the idea that the feedback affects the choice of hypothesis. However, earlier experiments have not manipulated feedback values in a systematic manner. Rather, feedback values have been randomized and therefore it has not been possible to study the relation between hypotheses and data in detail. An important problem for further studies would be to run an experiment in which the feedback values were varied systematically and to make a closer examination of the relation between feedback and hypotheses possible.

The greater ambiguity in the feedback in a low validity condition may have the result that the subjects switch to a guessing strategy. In a high predictability condition the feedback values vary too, but not enough to support both a positive and a negative linear rule. This smaller deviation may have the effect that the subjects change the parameters, that is, the slope and



intercept, of their linear rule rather than the general nature of their rule. They are thus less confused and frustrated, and therefore switch to guessing strategies in a lesser degree. This attempt to explain the higher frequency of guessing strategies in lower cue validity conditions would need experimental tests, however.

As mentioned earlier, task predictability affects the subjects' mood. A very interesting question hinted in the hypothesis above is the possibility that emotional factors have an effect on the subjects choice of hypotheses. This could be tested by comparing subjects who vary with respect to emotional factors, and assessing what hypotheses they produce in different tasks. This is, however, beyond the scope of this thesis. The other question concerns what the subjects are doing when they claim that they are "only" guessing. Earlier studies (Brehmer et al., 1974) have shown that subjects do not behave in a random manner when they say that they are guessing. It is possible that the subjects are trying new rules, or even constructing new rules, when they say that they are only guessing. It that is no, more construction of rules would have occurred in the low predictability conditions. According to the new version of the hypothesis-testing model (Brehmer, 1980a) subjects sample rules until there are no more rules, and after that they start to construct new rules. A low predictability or a random task would, according to this view, invite the subjects to construct a lot of new rules. But, on the other hand, if a low predictability task makes testing of hypotheses difficult, then subjects will only get confused and frustrated, and under these circumstances, construction of new rules is unlikely. To investigate these questions, would be to run an experiment with a random task with verbal hypotheses on every trial. When the subjects claim that they are only guessing, their strategy should be investigated in more detail using a series of blank trials. Such an experiment would shed some light on what the subjects are trying to do.

#### Function form

A number of studies have investigated the effect of the form of the function relating cue and criterion. The results show that (a) tasks with linear rules are learned faster than tasks with nonlinear rules (Brehmer, 1969, 1973a; Carroll, 1963; Deane, Hammond & Summers, 1972; DeKlerk & Oppe, 1972; Earle, 1973; Hammond & Summers, 1965; Sheets & Miller, 1979; Summers & Hammond, 1966; Summers, Summers & Karkau, 1969), and (b) tasks with posi-

tive linear rules are learned faster than tasks with negative linear rules (Björkman, 1965; Brehmer, 1973b; DeKlerk, DeLeeuw & Oppe, 1966; Naylor & Clark, 1968). These results are explained in a general way by the hypothesis testing model (Brehmer, 1974). However, there are still a number of unanswered questions.

First, what kind of hypothesis about relations do subjects have? The kind of hypotheses a subject has will determine what relations s/he is able to learn. Earlier assessments of subjects' hypotheses have some limitations, and therefore a new method was used to assess the hierarchy in study II. Some of the results from study IV are also of interest here.

#### The form of the hypothesis hierarchy

Brehmer (1974) suggested that inference tasks are learned through a hypothesis testing process, and developed a hypothesis sampling model to account for the learning process. According to that model, the subjects have an internal hierarchy of hypotheses about functional relations. Rules are sampled, one at the time, from this hierarchy. The rule with the highest sampling probability was the positive linear rule, followed by the negative linear rule, the inverted U-shaped rule and finally the U-shaped rule. Later, Brehmer (1980a) argued that hypotheses generated by subjects are not restricted to functional relations, but the hypotheses about functional rules are just the lowest level in a hierarchically ordered system of hypotheses.

Earlier studies of subjects hierarchies have some limitations. They come from two sources. The first was a rule production experiment where subjects were asked to produce whatever rule they could think of (Brehmer, 1974). The problem with this experiment is that the subjects' motivation to produce rules in the absence of feedback may be low. Furthermore, the subjects may not employ the same criterion for what constitutes a rule as the experimenter. For example, many subjects considered variations in parameters of a rule to define a new rule.

The other kind of experiments used recording of hypotheses in the learning of tasks with linear or quadratic rules (Brehmer, 1974; Brehmer, et al., 1974). In these experiments there was a rule which was correct, and the subjects also found that rule. This may have led them to omit other, more uncommon

rules.

A better way to assess the subjects hypotheses may be to give them a task where the correlation between cue and criterion is zero. Both of the objections raised above could be avoided by using such a task. Sniezek and Naylor (1978) used such a task to study subjects hypotheses, but they assessed the hypotheses by fitting polynomials to the subjects' responses in blocks of trials in which they received feedback. Because of that, there was no guarantee that the subjects used the same hypothesis during the whole block. Despite that, the experiment yielded results in good agreement with the hypotheses sampling model.

Study II used a random task and two nonrandom tasks plus verbal reports on every trial.

In study IV a pretraining procedure was used to encourage subjects to produce as many rules as they could. The subjects were encouraged to produce many and new rules, new according to the definition of the experimenter.

#### Results of studies II and IV

The result of study II showed that the subjects had a stronger preference for linearity than observed in earlier studies. Very few nonlinear hypotheses were observed. Some subjects claimed that there was no rule in the task. That finding pointed to the existence of rules other than functional and nonfunctional ones, a finding now incorporated in the extended version of the hypothesis hierarchy (Brehmer, 1980a). The results of study IV on the other hand, showed that many subjects (81%) could be encouraged to produce nonlinear rules. It was also possible to encourage subjects to produce more rules than showed by earlier studies. The subjects produced an average number of 5.2 rules, somewhat more than the 2.9 rules reported by Brehmer (1974). Linear rules were produced more often than nonlinear rules, a finding in agreement with earlier findings (Brehmer, 1974). It was also found that the subjects produced an average of 2.40 rules during the learning phase of study IV compared with 5.20 rules during the pretraining. This may be a result of difficulties in testing rules, or due to differences in task demands. It was also found that two subjects had the correct rule during the pretraining, but not during the learning phase. This points to the fact that it is not enough to have the cor-

rect rule, it is also necessary to have the ability to abandon other, less useful rules. Taken together, the results from study II and study IV points to the importance of the situation in which subjects rules are assessed. The subjects may therefore have more hypotheses than shown in experiments so far, and it must be important to vary the task and context in which subjects hypotheses are assessed. The sampling probability, and/or availability of hypotheses may be affected by the situation.

It is also difficult to know whether subjects sample hypotheses from a pre-established hierarchy, or if they construct new hypotheses. For instance, in study IV the subjects were encouraged to produce many and new rules, and in that study it was impossible to know when the subjects sampled old hypotheses or constructed new ones. It may be possible to investigate this problem by means of a "think aloud" method, or by asking the subjects if they are using old knowledge, or combining old knowledge in new ways.

The other problem concerns the fact that subjects learn complex rules slowly and inefficiently, if at al. The reason for this phenomenon is not clear, but two ideas have been proposed to explain this, (a) the subjects have bad access to complex rules, (b) they have difficulties in rejecting simpler rules. It is known that subjects have difficulties in testing rules, but it is not known if bad access to more complex rules also can be a part of the explanation. This is investigated in study IV.

#### Effects of pretraining on the construction of complex rules

In study II it was noted that subjects generally had a very strong preference for linear rules, and especially for the positive linear rule. This preference has been called a "Positive set" (Naylor & Clark, 1968). This "perverse pervasiveness of linearity" (Green, 1968) is interesting in many ways. First, it argues against a view of learning as a copy of the task to be learned. If the subjects just make an internal copy of the learning task, their ability to learn functional rules would have no limits. According to this view, the results of study II would be hard to explain. It also argues against some construction views of learning of probabilistic inference tasks (Carroll, 1963; Björkman, 1965). According to these views, the subjects store cue-criterion pairs, and they then fit polynomials to these pairs.

A likely explanation for the preference for linearity is that the subjects are trying hypotheses about the task. A most interesting question for a hypothesis testing view of learning is why the subjects have such a preference for linearity. Brehmer (1974) has explained this finding by postulating that the positive linear rule is the strongest one.

According to the hypothesis testing model there are at least two reasons why the subjects do not discard the positive linear hypothesis very easy. The first reason is that the probabilistic character of the task makes testing difficult, and if the subjects cannot reject the positive linear rule, there is no reason for sampling or construction of new rules. It is known that the subjects have difficulties in testing their rules (Brehmer, 1979b). The other reason is that the availability of other rules is limited. For instance, in a situation involving stress, the preference for the dominant rule may increase.

It is known that CPL-tasks induces emotional reactions, like frustration, (Alm & Brehmer, 1980) see study I. The lower the task predictability, the more emotional reactions. This finding would lead us to expect a higher degree of construction of functional rules when task predictability is high than when it is low. This agrees quite nicely with the findings of Brehmer (1980b) which indicate that subjects find J-rules to a lower extent when task predictability is low than when it is high.

The purpose of study III was to investigate if lack of availability could be a part of the explanation for the preference for linearity. Availability of the rules was manipulated by a pretraining procedure involving rule production and rule practice. After that followed the test phase where the subjects were required to learn a SPL-task with a complex rule. (A J-shaped rule, with a task predictability of 0.90, given the correct rule). The prediction was that pretraining would facilitate learning of that rule.

#### Results of study IV

There were practically no effects of the pretraining. The results show that having the components necessary for construction of a complex rule is not enough. These results suggest that one of the reason why subjects learn non-linear rules slowly and inefficiently is that they have difficulties in rejecting simpler rules, rather than low availability of the components needed to con-

struct complex rules. One of the reasons for that may be that the subjects adopt a deterministic approach to these tasks instead of a statistical one, i.e. they are trying to find a perfect rule for the task to be learned. If they start with a positive linear rule, they are then probably trying to find the correct parameters for that rule, a search that can go on for a long time. Sometimes the subject will make a perfect prediction by pure chance (more often when task predictability is high) and knowing that subjects often ignore negative instances (Wason, 1960) this will serve as a validation of the positive linear rule.

Another reason for the preference for linearity may be that a linear rule is easier to use than a nonlinear. This hypothesis does not explain the fact that a positive linear rule is easier to learn than a negative linear rule. A third possibility is that in order to learn a task, you must have something to start with, some idea to note deviations from. To note deviations, the best thing to have may be something simple and easy to use. The simplest functional rule is a linear rule. Its regularity makes it easy to detect new patterns. But again, this idea can not explain why there is a difference in learning rate between positive and negative linear rules. The results also showed that it is necessary, but not sufficient, to have the components needed for construction of a complex rule. It is likely that the subjects, in addition to having the correct building blocks, also must have some rule for how to combine these blocks, and the ability to abandon other rules, as well as the motivation to do this. In other words, it must be concluded that the nature of the construction process is little understood so far, and more research in this area is needed. Some important questions for future studies are the following: Will subjects construct more complex rules if they are more skilled in testing rules? If they are given rules for how to combine building blocks, will that affect their ability to construct complex rules? What are the effects of emotional reactions on the subjects ability to construct complex rules? Why do some subject learn complex rules, while others do not? These questions can be tested, for instance by first training subjects to test rules, and/or to combine building blocks, and then assess the effects of this training on their ability to learn complex rules. It seems likely that a subject must have both of these abilities to be able to construct a complex rule. That is, pretraining involving both testing and construction to should be superior to pretraining which involves only one of these abilities.



It may also be possible to select subjects with one or both of these abilities, and predict their performance in tasks involving complex rules. Concerning emotions, it is possible to select people who score high and low on certain emotions, such as anxiety, or to induce stress in a learning situation. It seems likely that subjects scoring high on stress will be less skilled in constructing complex rules than those scoring low on stress. Some studies on problem solving (e.g., Sarason, 1961) have shown that stress can impair performance on a difficult task.

### General conclusions

The present studies seem to warrant the following conclusions:

1. The subjects do not seem to be able to cope with tasks concerning randomness. Instead of using a statistical approach to CPL-tasks, they are using a deterministic approach.
2. Task predictability affect subjects' mood. When task predictability was low, subjects seemed to be frustrated. Their frustration probably was an effect of their attributions of their failure to task difficulty.
3. Task predictability also had an effect on the subjects functional rules. The higher the task predictability, the higher the proportion of functional rules. The subjects guessing strategies were also affected by task predictability. The higher the task predictability, the lesser the proportion of guessing strategies. It was also found that (a) the subjects had few hypotheses, (b) practically no new hypotheses were introduced after the first block, (c) there was a considerable resampling of hypotheses, (d) the hypotheses were independent of the input data from the task. All these results support a hypothesis sampling model of the learning of CPL-tasks.
4. The form of the hypothesis hierarchy was in general agreement with earlier results. However, very few nonlinear rules were observed, and this may mean that earlier studies have overestimated the proportion of nonlinear rules. It was also found that subjects had other rules than functional rules in their hierarchies.

5. It was found that the problem to learn complex rules in CPL-tasks may not be a problem of low availability of rules. It was also found that having the components of a complex rule is a necessary but not sufficient condition for construction of a complex rule. The other conditions needed for the construction of a complex rule seem to be: (a) an ability to abandon dominant hypotheses, (b) an ability to combine simpler ideas into more complex ones.

### SOME SUGGESTIONS FOR IMPROVING THE SJT-PARADIGM

#### The lack of concern for the effect of states other than normal states on human judgment and decision making

Besides some studies on judgment and psychoactive drugs (Hammond & Joyce, 1975) and some studies on judgment and alcohol (e.g., Brehmer & Almqvist, 1977) not much have been done in this field. It should, however, be noted that the interest for "cold cognition" is not unique for the SJT-paradigm, not much have been done elsewhere (within judgment and decision making or within cognitive psychology). The research concerning psychoactive drugs is here of great importance, and should give the SJT-paradigm some advantage compared with other approaches. However, nothing have been done to make a more direct study of the possible effects of various emotional states on human judgment and decision making. As previously mentioned, it seems likely that many decisions are made during stress and other emotional states. Therefore it would be of importance to study the effect of various emotional reactions on judgment and decision making. A step in this direction has been taken in this thesis.

#### Social hypothesis testing theory

As mentioned before, the SJT-paradigm is not intended to be anything but a framework. Within its scope it has (so far) included the following topics: (a) learning under uncertainty, (b) interpersonal learning, (c) interpersonal conflict, (d) judgment and decision making in groups (Hammond et al., 1980).

Common to all these situations is that a subject (or many subjects) must use something available to draw conclusions about something not available. In the case of interpersonal learning, for example, the "pupil" has to learn a hypo-

thesis from the "cues" the "teacher" produces.

Interpersonal conflict is often a result of inconsistent application of judgment policies (Brehmer, 1976a). The persons in conflict are using each others judgments as cues to infer the policy of the other. So, all these situations can at least partly be described as judgmental situations. This means that subjects must test hypotheses about the situations or persons involved. For instance, hypotheses about other peoples motives and intentions (see Hammond & Brehmer, 1973). For the topic learning under uncertainty we have a hypothesis testing model, and what is needed here is some form of a general hypothesis-testing model which could cover all these areas, including not only hypotheses about task relations, but also about other persons.

## References

- Björkman, M. (1965). Learning of linear functions: a comparison between a positive and a negative slope. (Rep. Psychol. Lab., Univer. Stockholm, No. 183.
- Björkman, M. (1973). Inference behavior in non-metric ecologies. In L. Rappoport & Summers (Eds). **Human judgment and social interaction**. New York: Holt, Rinehart and Winston.
- Björkman, M. (1978). *Ekologisk psykologi*. Almqvist & Wiksell Förlag AB, Stockholm.
- Bolton, N. (1972). *The psychology of thinking*. London: Methuen.
- Brehmer, B. (1969). Cognitive dependence on additive and configural cue-criterion relations. **American Journal of Psychology**, 82, 490-503.
- Brehmer, B. (1973a). Effects of cue validity on interpersonal learning of inference tasks with linear and nonlinear cues. **American Journal of Psychology**, 86, 29-48.
- Brehmer, B. (1973b). Policy conflict and policy change as a function of task characteristics. II. The effect of task predictability. **Scandinavian Journal of Psychology**, 14, 220-227.
- Brehmer, B. (1973c). Effects of task predictability and cue validity on interpersonal learning of inference tasks involving both linear and nonlinear relations. **Organizational Behavior and Human Performance**, 10, 24-46.
- Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. **Organizational Behavior and Human Performance**, 11, 1-27.
- Brehmer, B. (1976a). Social judgment theory and the analysis of interpersonal conflict. **Psychological Bulletin**, 83, 985-1003.
- Brehmer, B. (1976b). Note on clinical judgment and the formal characteristics of clinical tasks. **Psychological Bulletin**, 83, 778-782.
- Brehmer, B. (1976c). Learning complex rules in probabilistic inference tasks. **Scandinavian Journal of Psychology**, 17, 309-312.
- Brehmer, B. (1979). Note on hypothesis testing in probabilistic inference tasks. **Scandinavian Journal of Psychology**, 20, 155-158.
- Brehmer, B. (1979a). Preliminaries to an psychology of inference. **Scandinavian Journal of Psychology**, 20, 193-210.

- Brehmer, B. (1979b). Note on hypotheses testing in probabilistic inference tasks. **Scandinavian Journal of Psychology**, 20, 155-158.
- Brehmer, B. (1980a). In one word: not from experience. **Acta Psychologica** 45, 223-241.
- Brehmer, B. (1980b). Effect of cue validity on learning of complex rules in probabilistic inference tasks. **Acta Psychologica** 44, 201-210.
- Brehmer, B., & Almqvist, P. (1977). Effect of a minimal dose of alcohol on the learning of linear and nonlinear inference tasks. **Umeå Psychological Reports**, No. 129.
- Brehmer, B., & Hammond, K. R. (1977). Cognitive factors in interpersonal conflict. In D. Druckman (Ed.), *Negotiations: Social Psychological Perspectives*. Beverly Hill, Calif.: Sage.
- Brehmer, B., Kylenstierna, J., & Liljegen, J. (1974). Effects of function form and cue validity on subjects' hypotheses in probabilistic inference tasks. **Organizational Behavior and Human Performance**, 11, 338-354.
- Brehmer, B., Kylenstierna, J. (1978). Task information and performance in probabilistic inference tasks. **Organizational Behavior and Human Performance**, 22, 445-464.
- Brehmer, B., Kylenstierna, J. (1979a). Length of the postfeedback interval and learning of probabilistic inference tasks. **Scandinavian Journal of Psychology**, 20, 151-154.
- Brehmer, B., & Kylenstierna, J. (1980). Content and consistency in probabilistic inference tasks. **Organizational Behavior and Human Performance**, 26, 54-64.
- Bruner, J. S. (1957). Going beyond the information given. In H. Gruber, J. Jessor & K. Hammond (Eds.), *Cognition: The Colorado Symposium*. Cambridge, Mass.: Harvard Univ. Press.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Brunswik, E. (1952). *Conceptual framework of psychology*. Chicago: Univ. Chicago Press
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

- Brunswik, E. (1957). Scope and aspects of the cognitive problem. In H. Gruber, R. Jessor & K. Hammond (Eds.), *Cognition: The Colorado Symposium*. Cambridge, Mass.: Harvard University Press.
- Carroll, J. D. (1963). Functional learning: The learning of Continuous functional mappings relating stimulus and response continua. *Educational Testing Service, Princeton, Research Bulletin RB-63-26*, 1963.
- Deane, D. H., Hammond, K. R., & Summers, D. A. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology*, 92, 20-26.
- DeKlerk, L. F. W., DeLeeuw, J., & Oppe, S. (1966). Funktioneel leren. *Hypotese*, 11, 10-19.
- DeKlerk, L. F. W., & Oppe, S. (1972). Single-cue learning under different probabilistic feedback conditions. Department of Psychology. University of Leiden, Report No. E 032-72.
- Earle, T. (1973). Interpersonal learning. In L. Rappoport & D. A. Summers (Ed.), *Human Judgment and Social Interaction*. New York: Holt, Rinehart & Winston.
- Estes, W. K. (1960). Learning theory and the new "mental chemistry". *Psychological Review*, 67, 207-223.
- Green, B. F. (1968). Descriptions and explanations: A comment on papers by Hoffman and Edwards. In B. Kleinmuntz, (Ed.), *Formal representation of human judgment*. New York: Wiley.
- Hagafors, R. (1981). Learning of probabilistic inference tasks: the importance of situational variables for the subject task interaction. Doctoral thesis at Uppsala University.
- Hagafors, R. (1982). A note on confidence and performance in judgment. *Uppsala Psychological Reports*, No. 313.
- Hagafors, R., & Brehmer, B. (1980). Effects of informations-presentation mode and task complexity on the learning of probabilistic inference tasks. *Scandinavian Journal of Psychology*, 21, 109-113.
- Hammond, K. R. (1966). Probabilistic functionalism: Egon Brunswik's integration of the history, theory, and method of psychology. In K. R. Hammond (Ed.), *The psychology of Egon Brunswik*. New York: Holt, Rinehart and Winston.
- Hammond, K. R. (1971). Computer graphics as an aid to learning. *Science*, 172, 903-908.

- Hammond, K. R. (1974). Human judgment and social policy. Inst. Behav. Sci., Univ. Colorado: Program on Human Judgment and Social Interaction, Rep. 170.
- Hammond, K. R., & Brehmer, B. (1973). Quasi-rationality and distrust. Implications for international conflict. In L. Rappoport & D. Summers (Eds.), Human judgment and social interaction. New York: Holt, Rinehart and Winston.
- Hammond, K. R., & Joyce, C. R. B. (1975). Psychoactive drugs and social judgment. New York: Wiley.
- Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). Human judgment and decision making. Praeger.
- Hammond, K. R., Rohrbaugh, I., Mumpower, I., & Adelman, L. (1977). Social judgments theory: Applications in policy formation. In M. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes in applied settings. New York: Academic Press.
- Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. Kaplan & S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press.
- Hammond, K. R., & Summers, D. A. (1965). Cognitive dependence on linear and nonlinear cues. **Psychological Review**, 72, 215-224.
- Hogarth, R. (1980). Judgment and choice. John Wiley & Sons.
- Hursch, C. J., Hammond, K. R., & Hursch, J. L. (1964). Some methodological considerations on multiple-cue probability studies. **Psychological Review**, 71, 42-60.
- Johansson, R., & Brehmer, B. (1979). Inferences from incomplete information A note. **Organizational Behavior and Human Performance**, 24, 141-145.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. **Psychological Review**, 80, No. 4.
- Krechersky, I. (1932). "Hypotheses in Rats". **Psychological Review**, 39, 516-532.
- Kuylensstierna, J., & Brehmer, B., (1981). Memory aids in the learning of probabilistic inference tasks. **Organizational Behavior and Human Performance**, 28, 415-424.

- Levine, M. (1975). A cognitive theory of learning. Hillsdale, N.J.: Larence Earlbaum.
- Muchinsky, P. M., & Dudycha, A. L. (1975). Human inference behavior in abstract and meaningful environments. **Organizational Behavior and Human Performance**, 13, 377-391.
- Naylor, J. C., & Clark, R. D. (1968). Intuitive inference strategies in interval learning as a function of validity magnitude and sign. **Organizational Behavior and Human Performance**, 3, 378-399.
- Sarason, I. G. (1961). The effects of anxiety and threat on the solution of a difficult task. **Journal of Abnormal and Social Psychology**, 62, 165-168.
- Seligman, M. E. P. (1975). Helplessness. W. H. Freeman and Company, San Francisco.
- Sheets, C., & Miller, M. J. (1970). The effect of cue-criterion function form on multiple-cue probability learning. Institute of Behavioral Science, University of Colorado, Program on Cognitive Processes Report No. 127.
- Snizek, J., & Naylor, J. C. (1978). Cue measurement scale and functional hypotheses testing in cue probability learning. **Organizational Behavior and Human Performance**, 22, 366-374.
- Summers, D. A., & Hammond, K. R. (1966). Inference behavior in multiple-cue tasks involving both linear and nonlinear relations. **Journal of Experimental Psychology**, 71, 751-757.
- Summers, S. A., Summers, R. C., & Karkau, V. T. (1969). Judgments based on different functional relationships between interacting cues and a criterion. **American Journal of Psychology**, 82, 203-211.
- Tolman, E. C., & Brunswik, E. (1935). The organism and the causal texture of the environment. **Psychological Review**, 42, 43-77.
- Tucker, L. R. (1964). A suggested alternative formulation in the developments by Hirsch, Hammond, and Hirsch, and by Hammond, Hirsch, and Todd. **Psychological Review**, 71, 528-530.
- Tversky, A., & Kahneman, D. (1971). The belief in the law of small numbers. **Psychological Bulletin**, 76, 105-110.



- Uhl, C. N. (1966). Effect of multiple stimulus validity and criterion dispersion on learning of interval concepts. **Journal of Experimental Psychology**, 72, 519-527.
- Wagenaar, W. A. (1970). Appreciation of conditional probabilities in binary sequences. **Acta Psychologica**, 34, 348-356.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. **Quarterly Journal of Experimental Psychology**, 12, 129-140.
- Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, L., & Rosenbaum, R. M. (1971). Perceiving the causes of success and failure. Morristown N.J: General Learning Press.