Netherlands
organization for
applied scientific
research

IZF

TNO Institute for Perception

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 1989-31

HOW DO EXPERTS SOLVE UNFAMILIAR PROBLEMS: A PRELIMINARY STUDY

J.M.C. Schraagen

12

Classifications:

Number of pages:	34	**Report:**	**unclassified**
		Title:	**unclassified**
		Abstract:	**unclassified**

CONTENTS

Report No.: IZF 1989-31

Title: How do experts solve unfamiliar problems: a preliminary study

Author: Drs. J.M.C. Schraagen

Institute: TNO Institute for Perception

Date: September 1989

No. in Program of Work: 733.1

SUMMARY

A largely unanswered question in the literature on problem solving is how experts solve unfamiliar problems. Do they resort to weak methods of search and analysis similar to those used by novices? Or do experts who have acquired powerful processes of reasoning in one domain apply those processes to solving problems in areas where specific solution methods have not been worked out? An initial attempt at answering this question was made in the present study. The domain chosen was experimental design. Subjects with varying levels of experience with designing experiments were asked to think aloud while they were designing an experiment in the, to them, unfamiliar area of sensory psychology. The results showed that experts quickly translated the unfamiliar problem into more abstract and familiar terms with which they could retrieve an experimental paradigm from memory. In contrast, novices only used a very general idea of what an experiment should look like, and hence could not provide as many details as the experts. The results further showed that experts have acquired powerful strategies for understanding the problem and evaluating designs. They can apply these strategies to unfamiliar problems. Novices apparently lack these strategies.

Rap.nr. IZF 1989-31 Instituut voor Zintuigfysiologie TNO, Soesterberg

Hoe lossen experts onbekende problemen op: een vooronderzoek

J.M.C. Schraagen

SAMENVATTING

Een nog grotendeels onbeantwoorde vraag in de literatuur op het gebied van probleemoplossen is hoe experts onbekende problemen oplossen. Vallen zij terug op zwakke zoek- en analysemethoden zoals beginners die gebruiken? Of passen experts krachtige redeneerprocessen toe in gebieden waar zij specifieke oplosmethoden nog niet hebben verkregen? In het huidige onderzoek werd een eerste poging gedaan deze vraag te beantwoorden. Het gekozen domein was het opzetten van onderzoek. Proefpersonen met verschillende ervaringsniveaus in het opzetten van onderzoek werden gevraagd hardop te denken terwijl zij een experiment moesten opzetten in het, voor hen onbekende, gebied van sensorische psychologie. De resultaten lieten zien dat experts het onbekende probleem snel vertaalden in meer abstracte en bekende termen waarmee zij een experimenteel paradigma konden oproepen uit hun geheugen. De beginners daarentegen maakten slechts gebruik van een algemeen idee van hoe een experiment er uit zou moeten zien, en konden dientengevolge niet zo veel details geven als experts. De resultaten lieten verder zien dat experts krachtige strategieën hebben verworven om het probleem te begrijpen en het design te evalueren. Zij kunnen deze strategieën toepassen op onbekende problemen, terwijl beginners deze strategieën missen.

1 INTRODUCTION

A largely unanswered question in the literature on problem solving is how experts solve unfamiliar problems (for a review, see Greeno & Simon, 1988). Do they resort to weak methods of search and analysis similar to those used by novices? Or do experts who have acquired powerful processes of reasoning in one domain apply those processes to solving problems in areas where specific solution methods have not been worked out and stored in memory? An initial attempt at answering this question was made in the present study. The domain chosen was experimental design.

Designing experiments is a complex cognitive skill requiring various kinds of knowledge, ranging from knowledge of design principles to knowledge of how to control for various irrelevant factors. Knowledge of design principles and concepts is usually acquired by students in an introductory course on experimental design. This knowledge is very general, hence widely applicable. In line with recent proposals (e.g. Anderson, 1983), we will call this knowledge of principles, concepts, and relations between concepts, "declarative knowledge". As students gain more experience with designing experiments, we will assume they will develop domain-specific knowledge. This knowledge is developed by applying general problem solving strategies to the declarative knowledge base (Anderson, 1987). By gaining more experience, students will thus learn when to apply what part of their knowledge. For example, they will form rules about how to operationalize concepts, how to choose and present stimulus material, how to select and instruct subjects, and so on. With enough practice in one or more experimental paradigms, these rules will become highly specific. For example, when working in the area of lexical memory, it is important to control for word frequency. This domain-specific knowledge is often called "procedural knowledge", and the process by which this knowledge develops is called "knowledge compilation" by Anderson (1982, 1987). In this paper, knowledge about how to classify and repair an impasse during problem solving, sometimes called "strategic knowledge", will also be viewed as procedural knowledge.

Put differently, with increasing expertise there is a transition from general and flexible, but "weak" problem solving methods to specific and inflexible, but "strong" problem solving methods (Newell, 1969). According to Newell, "weak methods" do not guarantee good solutions to a problem, whereas "strong methods" do. Weak methods require little knowledge about the task, and can therefore be used in many domains.

As people gain more knowledge about a task, more specialized methods arise, but the weak methods are always available when there is little knowledge of the task (Laird, Rosenbloom, & Newell, 1986). Adding knowledge to a system (human or otherwise) can limit the amount of search necessary for solving a problem. In the limit, search is altogether abandoned and the problem is solved by direct recognition. The ability to solve problems therefore depends on the amount of domain-specific knowledge a system possesses (Lenat & Feigenbaum, 1987).

This account of how cognitive skills are developed raises several interesting questions concerning the area of experimental design:

1) What are the differences, if any, between experts and novices in experimental design? Do experts have more declarative knowledge, more procedural knowledge, or both?

2) Is knowledge of design principles sufficiently general so as to enable experts and novices alike to come up with good experimental designs in areas they are relatively unfamiliar with? In other words, will experts and novices perform alike on a novel design problem?

To answer the above questions, we studied six novices and five experts solving a novel design problem. Since the problem was novel to the subjects, they had to rely on general design knowledge. Before turning to more specific predictions, we will first give an outline of a framework for experimental design.

2 A FRAMEWORK FOR EXPERIMENTAL DESIGN

Experimental design, like architectural design or musical composition, can best be regarded as an ill-structured problem (Simon, 1973). This is because more than one design may be adequate for answering a research question and there is no definite criterion to decide which is the appropriate design for a given question. Another reason the task is ill-structured is because the number of potential extraneous variables that have to be controlled is very large. All this does not imply, as Simon (1973) has argued, that design tasks require qualitatively different mechanisms than the ones already known. The cognitive architecture of a problem solver remains the same, whether it solves ill-structured or well-structured problems. Basically, ill-structured problems are transformed into a number of smaller, well-structured problems that can be solved.

Empirical evidence for the above argument was adduced in the previous section. In various design tasks, ranging from software design to architecture, experts were found to use a divide-and-conquer strategy. This strategy can only be applied effectively when the problem solver has extensive domain-specific knowledge, since it is only then that subproblems can be recognized and solved by knowledge stored in LTM.

We may therefore, on the basis of the available evidence, predict what course the design process may follow (see also Malhotra, Thomas, Carroll, & Miller, 1980). Of course, the empirical evidence may falsify these predictions, but at least they allow us to look more closely at the data. We will first describe what the expert's problem solving process might look like, given the framework sketched above, the empirical evidence available, and a task analysis. After that, we will consider how a novice's problem solving process might differ from the expert's.

As far as the stages of problem solving are concerned, at the highest level we will distinguish between a problem understanding and a problem solution phase. This is more or less a logical requirement, since in order to solve a problem one first has to understand the task requirements, i.e. what exactly has to be solved. Problem understanding is a very shallow process when the problem is well-known. With more difficult problems, the two phases may occur repeatedly, since problem solution will usually deal with well-defined subproblems, and a new phase of problem understanding may occur after one subproblem has been solved and the problem solver goes on to examine the problem description again. In the course of solving the problem, the problem solver gradually comes to understand the problem better and better, i.e. formulates the problem more productively (Duncker, 1945). In this way, more elements are added to the design as the problem gets better understood.

In the problem understanding phase, the problem solver constructs a problem space to solve the problem in. The problem space is constructed by reading the problem statement. From the problem statement the goal is determined, in this case: design an experiment in order to answer a certain question. In order to design an experiment the problem solver first has to disambiguate the problem statement. In other words, the problem solver has to find out exactly what is being asked. In this phase, constraints are generated, and subgoals are set. Fulfilling these subgoals leads to satisfaction of the parent goal: the design.

The subgoals and constraints activate a design schema stored in LTM. The design schema is an abstract plan for how the experiment should be conducted. It may be more or less specified, depending on whether the research question is familiar or not (Friedland, 1979, Friedland & Iwasaki, 1985). The abstract plan is compared with the actual requirements, differences are noted, and operators are proposed to reduce the differences. When the problem statement is ambiguous, or the problem solver can think of more ways of answering the question, more design schemata may be instantiated at once. These design schemata may be successively refined, or the problem solver may alternate between problem understanding and problem solving, instantiating schemata one at a time. Justifications of particular design decisions are made on the basis of general design principles.

After a design schema has been instantiated, the problem solving phase proper is entered. In this phase, the design schema is successively refined: internal and external validity are determined, and the efficiency of the design is checked. Internal validity is determined by generating and controlling for irrelevant variables; by generating, choosing, and evaluating (a) dependent variable(s); and by determining the power of the experiment. Power is determined by number, selection, and assignment of subjects, and by reliability of measurements. External validity is determined by defining a target population and drawing a random sample from it; and by variable and subject generalizability. Efficiency is checked by considering the time it takes to run the experiment, which may be determined by number of subjects required and resources (people, computers, money) available.

The sketchy design is evaluated every now and then to determine whether all the above factors have been specified sufficiently. When the various constraints, such as what variables to control for, have been enumerated, they have to be put in a certain temporal order. If the temporal order of the design's components is mixed up, the design, and therefore the experiment, is faulty, and no solid conclusions can be drawn from it. For example, the order in which measurements should be taken, and the order in which subjects receive each stimulus should be specified. When all slots in the design schema have been instantiated, and put into an appropriate temporal order, and all subgoals have thus been fulfilled, the problem is solved, unless more design schemata have been activated.

The model sketched above is idealized in that not all subjects will conform to all parts of it. This applies to both novices and experts. Still, we may be able to point to a number of differences, based on the available empirical evidence, between the way novices and experts design experiments:

1) Novices will probably spend less time understanding the problem than experts.

2) Novices will not show evidence of using design schemata, since their knowledge of design is either deficient or else they do not know when to apply what schema. Lacking schemata to structure their problem solving, their problem solving will be fragmentary and incomplete.

3) Lacking adequate design knowledge, novices are not able to rephrase the problem statement into abstract design terms. They will therefore follow closely the literal problem instructions. They will also not be able to successively refine the problem.

In order to empirically determine the differences between novices and experts in the area of experimental design, an experiment was carried out. Eleven subjects were asked to design an experiment to determine whether one can identify different brands of cola on the basis of taste alone. The idea for this experiment came from reading Johnson and Solso's (1978) book on experimental design. All subjects could be expected to know various characteristics of cola beverages, such as smell, colour, sweetness, temperature, etc. Thus, we would be able to concentrate on design knowledge alone, since domain-specific knowledge was equated as far as possible across different skill levels. Note that by "domain-specific knowledge" we do not mean "design knowledge", but "knowledge of the domain in which a design has to be set up". In this way, too, we would be able to determine whether general skills, such as successive refinement, would be used by experts in domains they had never experimented in.

3 METHOD

3.1 Subjects

Eleven subjects participated in the experiment. Six of them were novices (N) who had never designed an experiment themselves, although three subjects (N4-6) had been involved in psychological research as research assistants. The other three had the Dutch equivalent to a Master's Degree in computer science (N1), business science (N2), and physics (N3). Five subjects could be called "experts" (E), although their skill level varied. Two subjects held Ph.D. degrees, one of them in psychology (E1), the other in physics (E4). A third subject was close to his Ph.D. degree in psychology (E2). These three subjects had designed experiments for at least ten years. The fourth subject had designed experiments for almost three years (E3). The fifth subject had designed experiments for at least five years (E5). Although their skill levels varied, we decided to place subjects in either a novice or an expert category, based on their experience with designing experiments themselves. Note that the term "expert" in this study does not imply expertness in designing experiments in the area of taste. The experts in this study were good at designing experiments in their respective areas, and none of these areas included "taste".

3.2 Task

Subjects received the following instructions:

> Try to devise an experiment to find out whether one can identify different brands of cola on the basis of taste alone. Take three brands of cola (Pepsi, Coca Cola, and an own brand) and have a group of people taste these. Except for taste of cola you want to eliminate all other factors that can play a role with cola identification. Please indicate as detailed as possible how such an experiment should look like according to you. In doing so, indicate what irrelevant factors can play a role with cola identification, and how you think to eliminate the influence of those factors in your experiment.

The emphasis on controlling irrelevant factors was meant to elucidate the concept of "experiment" for the novice subjects.

3.3 Procedure

Subjects were tested individually. The instructions were given to them on a piece of paper, and they were asked to think aloud while designing the experiment. Their verbal protocols were tape recorded. Since this was a pilot study the experimenter asked subjects some questions after they had finished their task. Subjects were given unlimited time; most of them finished within thirty minutes. Most of the time, subjects finished when they had no more to say.

4 RESULTS

In this section, we will describe both quantitative and qualitative results. The quantitative results are concerned with the number of utterances in a certain category a subject made. Two separate analyses were carried out. The first, described in paragraph 4.1, was based on criteria derived from a handbook on experimental design. The second, described in paragraph 4.2 and 4.3, was based on empirically derived criteria. The qualitative results are concerned with the declarative and procedural knowledge being used by a subject, as inferred from the protocols.

4.1 Quantitative results

Subjects' protocols were transcribed and analyzed according to the following coding scheme. Prior to the experiment, a list of five criteria of research design was drawn up on the basis of a handbook (Kerlinger, 1973, pp. 322-326). These criteria were:

1) Goal of the experiment
2) Internal validity
3) Number of designs
4) External validity
5) Evaluation of the design

Each criterion was subdivided into a number of aspects, e.g. does subject modify problem description; what irrelevant factors are mentioned and/or controlled for; does subject mention power of the experiment; does subject take generalizability into account; does subject evaluate the design in terms of efficiency or answering the research question, etc. A total of seventy-one aspects as applied to the cola-design were put forward.

The protocols were scored as follows: when a subject took one of the seventy-one aspects into account, by mentioning them in his or her protocol explicitly, he or she received one point for that aspect. After this was done, the points received were categorized into one of the five classes mentioned above and were added. Thus, a subject could receive a number of points for "goal of the experiment", "internal validity", etc. The more points in a category, the more aspects a subject had mentioned in that category. In this way, we could determine, in a fairly objective way, whether experts would mention more aspects concerning problem understanding, and whether they would come up with more designs, as predicted. To ensure reliability of coding, six protocols (three of experts, three of novices) were coded blindly by a second rater. Of the 426 aspects, the two raters disagreed on 49; the remaining 377 aspects were agreed upon as being either mentioned (60) or not mentioned (317). On the basis of these scores, the inter-rater-reliability was considered satisfactory.

Table I Number of aspects mentioned by novice (N) and expert (E) subjects on five criteria.

Subject	I	II	III	IV	V	Total
N1	0	4	1	0	0	5
N2	0	6	1	0	0	7
N3	0	9	2	1	0	12
N4	0	3	2	0	0	5
N5	2	2	2	1	2	9
N6	0	9	1	1	1	12
E1	1	7	2	2	0	12
E2	3	12	4	0	2	21
E3	2	9	3	0	1	15
E4	2	10	3	0	0	15
E5	3	15	2	3	1	24

A Mann-Whitney U test was carried out with the five criteria as dependent variable and level of expertise (expert or novice) as grouping variable. Overall, the experts mentioned more items than the novices ($p < .05$). As predicted, the experts mentioned significantly more items from the first and third categories (both p's $< .05$). Experts thus looked more at the problem description and came up with more designs than the novices. Unexpectedly, however, experts also came up with significantly more irrelevant factors and ways of controlling them than novices ($p < .05$). This was unexpected since controlling for irrelevant factors seems to require domain-specific knowledge, and we can assume that both experts and novices have the same amount of knowledge about cola beverages. We will return to this result below. Experts and novices did not differ on categories four and five, that is external validity and evaluation of the design. One would have expected the experts to spend more time evaluating the design, but perhaps there were too few utterances in this category (seven out of fifty-five possible utterances) to yield any meaningful pattern.

To return to the number of irrelevant factors, one possibility why experts mentioned more of these is that they came up with more designs than novices and consequently could come up with a number of irrelevant factors for each design. Against this, one could argue that some (possibly large) number of irrelevant factors is relevant for each design, and that once a subject had mentioned them, he or she would not mention them again with a second design. To test these two possibilities, the number of irrelevant factors was plotted against each design mentioned. Thus, one would get an indication of how many irrelevant factors were mentioned with each design. The results showed that experts mentioned 35 irrelevant factors with the first design they came up with, whereas novices mentioned 30 irrelevant factors with the first design. This difference was not significant (Mann-Whitney $U=22$, $p=0.20$). This means that the finding of experts mentioning more irrelevant factors than novices can indeed be attributed, post hoc, to their coming up with more designs than novices.

There was some difference between experts and novices in the nature of the irrelevant factors mentioned. Novices mostly focused on visible characteristics of cola, e.g. smell and colour, and ways of controlling for these factors. Experts, on the other hand, asked themselves whether smell should be eliminated. They all thought this to be undesirable, because of the close relation between taste and smell. Experts also paid more attention to the interaction between subject and cola, e.g. the satiation that occurs when subjects have to drink large amounts of cola. Finally, the experts came up with typical psychological factors to control for, such as the "experimenter bias effect". In order to control for this effect, some of the experts mentioned they would use a "double blind study", in which both the subject and the experimenter are unaware of the brand of the cola presented on a certain trial.

4.2 Analysis of verbal data

A problem with protocol analysis is determining what constitutes a verbal statement: a proposition, a sentence, a couple of sentences? Another problem is determining what category a particular verbal statement falls into. A third problem concerns the number of categories one should use: fewer categories lead to a more robust, but also a more trivial, model.

As far as the problem of determining what constitutes a verbal statement is concerned, the following approach was taken: all information enabling one to make a coding decision should be contained in one unit. Since subjects often paraphrased the same information in a second or third sentence, this approach resulted in units ranging in sentence length from 1 to 6. Median sentence length was 1.7.

We have tried to solve the second problem by having naive subjects sort verbal statements into categories they had chosen themselves. To this end, six subjects, all familiar with the area of experimental design, were given 58 statements drawn from the protocols of the five expert subjects. Statements were selected such that a wide range of types of statements was included. Protocols of experts rather than novices were chosen since these were thought to contain the widest variety of statements. The statements were typed on cards, and the subjects were asked to sort these cards into as many categories as they thought appropriate. Cards were presented to the subjects in a random order. This is an important precaution, since wrong coding decisions can be made when coders know what statements precede and follow the statement to be coded. A coding decision will then be made based on assumptions of what subjects should have said, instead of what they actually said.

A particular statement was then compared to all other statements: if another statement belonged to the same category, this pair of statements received a score of 1; if another statement was put into another category, the pair received a score of 0. In this way, a matrix of zeros and ones was constructed for each subject, for all pairwise comparisons. These matrices were then averaged across subjects. The reason for averaging was that we were primarily interested in obtaining a robust number of categories, and not in individual differences, however interesting these might be for other purposes.

The matrix thus obtained can be viewed as a similarity matrix: the more often two statements were put into the same category, the more similar they are. The matrix was analyzed by means of a hierarchical clustering procedure, using the "group average method". This method has been shown to give better recovery of the true cluster structure than the single and complete methods (see Milligan & Cooper, 1987, for a review).

Interpretation of the tree structure resulting from the cluster analysis of the verbal statements showed six meaningful, stable clusters:

- Understand
- Operationalize dependent and independent variables
- Controlling for irrelevant factors
- Effects of repeated measurements on the same unit (e.g. drinking one cola affects the taste of the next one)
- Effects of repeated treatment-measurement pairs on the same unit (e.g. carry-over effects)
- Global temporal structure (procedure).

Obviously, not all six category labels will be perfect descriptors of every individual statement. To the experimenter's opinion, some statements could better be placed into a different category: 12% of the statements were thus not described adequately. This could be due to inadequacy of the clustering algorithm, to noisy data, or some other factor.

There is another way of assessing the validity of the results obtained with the cluster analysis. This is via a syntactic and semantic analysis of the verbal statements. Each verbal statement presumably contains some cues on the basis of which subjects categorize this statement. For example, statements containing words such as "randomize" or "counterbalance" would be put into the subcategory "repeated treatment-measurement pairs"; statements containing words such as "identify", "recognize", "preference", and "taste" would be put into the category "understand". Thirteen key words were selected, and each of the 58 statements was scored on presence or absence of these key words. This syntactic analysis was complemented by a semantic analysis. The semantic analysis was used in those cases where statements contained several key words, and thus could not unambiguously be classified into one category. In these cases, the most important key word was determined, as subjects in a sorting task presumably do. Thus, each statement received a score of 1 or 0 on a key word. This matrix was analyzed via a hierarchical cluster analysis, and the resulting tree structure was compared to the tree structure derived from the subjects' sorting of problem statements. Seventeen percent of the statements analyzed syntactically and semantically was misplaced. This means that, on average, 83% of protocol statements can be classified correctly by means of a fairly objective procedure. We have used this procedure to classify the remaining protocols.

4.3 Results from protocol analysis

Appendices A and B show two fully coded protocols: one of a novice and one of an expert.
Table II shows the distribution of verbal statements in all protocols over the six categories mentioned above, as well as the Evaluate category.

Table II Distribution of verbal statements over categories (in brackets: relative frequency).

category	experts	novices
understand	26 (29%)	5 (12%)
operat-vars	15 (16%)	14 (32%)
irrel. fact.	12 (13%)	10 (23%)
rep. measures	12 (13%)	2 (5%)
rep. treatm.meas.	11 (12%)	8 (19%)
global temp.str.	10 (11%)	3 (7%)
evaluate	6 (6%)	1 (2%)
	N = 92	N = 43

A Mann-Whitney U test carried out on the total number of verbal statements showed that experts produced significantly ($p < .05$) more verbal statements than novices. A Mann-Whitney U test was further carried out on the relative frequencies within each category, with level of expertise as grouping variable and the seven categories (including the evaluate category) as dependent measures. The only significant difference ($p < .05$) was in the Understand category: as predicted, more of the experts' than the novices' utterances could be placed in this category. These results resemble those obtained in paragraph 4.1 with a different scoring system.

The experts' problem understanding episodes were not confined to the first part of the protocols. Instead, they were scattered throughout the protocols. The experts reverted to problem understanding in four cases:

- at the beginning, in order to disambiguate features in the problem statement that were ambiguous (e.g. the word "taste").
- during a designing episode to check whether the design obtained so far was sufficiently detailed.
- when they reached an impasse, e.g. when they tried to operationalize a variable and had to read the problem statement in order to make a decision.
- at the end of a designing episode, when the answer was stated and evaluated, and subjects looked at the problem statement to see if their answer matched the problem requirements; if it did, they could either quit, or come up with a different conceptual model.

In contrast, the novices' problem understanding episodes consisted mainly of extracting features from the problem statement, e.g. controlling for irrelevant factors. They then set a subgoal to solve these features. For instance, subjects N2 and N3 both started by saying: "The first thing to know is what are the irrelevant factors".

Interestingly, there were no significant differences between experts and novices in problem solving episodes, apart from the number of designs they came up with: experts came up with 2.8 designs on average, novices with 1.5 designs on average (see section 5.1). This is not to say that the quality of the designs experts and novices came up with are the same. One way of judging this quality is via independent raters who rate the quality of the design blindly. We have not looked into this any further, since a primary aim in this research was a description of the design process, not a judgment of the design product.

Strategic knowledge

Strategic knowledge controls and monitors the execution of a task. This is necessary when knowledge is either incorrect or insufficient and an impasse arises. One would expect novices to be particularly "impasse-driven", since their knowledge is most often incorrect or insufficient. This has been shown to be the case in thermodynamics problem solving (Jansweijer, 1988). It is not clear what to expect for design problems. In contrast with thermodynamics problems, in which the end state is well-defined, design problems have an ill-defined end state. It may well be that novices, not knowing exactly what "an experiment" means, have a simpler task than experts. Experts know what it means to design an experiment and can constantly check what they have achieved so far against some self-imposed goal. Novices may not be "hindered" by such a goal and may consequently encounter fewer obstacles than experts.

The data showed that particularly the moderately experienced experts generated a lot of "monitoring statements". By "monitoring statements" we mean evaluative comments on particular design decisions, noticing of impasses, or checking whether there are more things to do; we also include the resulting actions to resolve impasses, e.g. ignore impasse, read problem statement again, ask experimenter, choose one of several possibilities. The monitoring statements were kept apart from the coding of the rest of the protocols. The experts generated 30% monitoring statements, whereas the novices generated only 5% (these are percentages of the total number of statements coded). The two more experienced experts generated only one monitoring statement in total, while the three moderately experienced subjects generated the remaining 27 statements.

These results show that the novices did not encounter many difficulties as a result of incorrect or incomplete knowledge. This is surprising, since our novices obviously had less knowledge about design than experts. Why, then, could they still solve the problem? The answer could lie in the open-ended nature of design problems. Since there are several designs possible with our problem statement, the main problem becomes one of narrowing down the possibilities until one design can be chosen. This is exactly what distinguished experts from novices in this experiment: the experts spent more time understanding the problem statement than the novices. Once the experts had chosen one conceptual model, their problem solving behaviour could not be distinguished from that of the novices. The novices, on the other hand, just read the problem statement once, and then came up with only one design. They did not switch back and forth between problem understanding and problem solving, as did the experts.

The novices did not lack domain-specific knowledge about cola. They were therefore able to come up with ways of controlling for irrelevant factors. They also possessed some knowledge about randomization and counterbalancing, although they often just used terms such as: "using different orders". What they lacked, however, was a clear conceptual model of how the design should look like, based on a thorough analysis of the problem statement. They could not therefore systematically refine their model, and use it to evaluate their intermediate results. This explains why the novices uttered so few monitoring statements and encountered so few difficulties. The knowledge they possessed was sufficient for solving the problem, but insufficient for providing them with a norm against which to check their intermediate results.

5 GENERAL DISCUSSION

The main question underlying the present experiment was:

- How do experts behave on relatively novel problems?

To answer this question we have confronted some experts in the area of experimental design with a novel design problem. None of the experts had ever designed an experiment in the area of taste. They therefore had to rely both on general design skills, and on general problem solving skills. In order to disentangle the two kinds of skills, we have compared the experts with a group of novices, who presumably did not possess any general design skills.

Our results showed that the main difference between experts and novices on a novel design problem lies in the problem understanding phase: experts spent more time analyzing the problem requirements, they went back and forth between problem understanding and problem solving more often than novices, and they came up with more designs than novices. Thus, the first prediction mentioned in chapter 2 is indeed confirmed: novices spend less time understanding the problem than experts.

A second prediction made was that novices lack design schemata to structure their problem solving, with the result that their problem solving will be fragmentary and incomplete. This prediction was falsified. Novices possessed a rudimentary design schema that was temporally organized: they knew they had to come up with a stimulus, that this stimulus then needed to be presented to a subject, and that the subject then had to give a response. This general schema was sufficient for them to come up with an answer.

The third prediction was that novices will follow closely the literal problem statement, and will not be able to successively refine the problem. This prediction was confirmed. As suggested in the introduc-

tion, a possible reason for this may be that novices are not able to rephrase the problem statement into abstract design terms. However, the results suggest that this explanation is incomplete. An alternative explanation should also take into account the finding that experts evaluate their solutions more extensively than novices. This explanation hinges on two factors:
1. The expert's knowledge base
2. The expert's evaluation strategies.

We postulate that the differences in problem understanding between novices and experts can be explained by their different knowledge structures. We assume that experts possess a (possibly very large) number of design schemata. These schemata are indexed by the type of problem the expert has to solve. A design schema contains slots for (among others):
- independent variable
- dependent variable
- control variable
- procedure
- subjects.

Each slot contains information about:
- range of values
- constraints on those values
- methods to choose values.

Knowledge about the range of possible values and constraints on those values constitutes the declarative knowledge. These are the "facts" that someone knows; for example, one may know that the range of "sex of subjects" (a subslot of "subjects") is: male-female, and that for a particular experiment one only wants females as subjects (constraint). Knowledge of methods to choose values constitutes procedural knowledge. There may be several types of methods available, ranging from highly general (e.g. choose value at random, choose default value) to highly specific methods (e.g. if temperature is not a relevant factor, choose colas with the same temperature). The methods are ordered such that the specific methods are tried first, and the general methods are only tried when the more specific ones fail.

Expert problem solving partially consists of selecting the right kind of design schema and filling in the slots of the schema. During problem understanding the expert, when confronted with a relatively new problem, tries to translate the problem statement into more abstract design terms. These abstract design terms are the cues with which design schemata can be retrieved from long-term memory. The experts' knowledge may be viewed as being organized into default hierarchies (Holland, Holyoak, Nisbett, & Thagard, 1986). Default hierarchies serve two purposes:

- knowledge is organized hierarchically, from highly abstract to highly concrete; abstract knowledge enables people to deal with novel situations;
- at any level, rules generate default expectations that can be overridden by more specific expectations; in this way, even when detailed knowledge is unavailable at all levels, the system can still use all its knowledge in a flexible manner.

Rules are responsible for the phenomenon that experts often use whatever opportunity suggests itself, and do not always work strictly top-down (Hayes-Roth & Hayes-Roth, 1979). For example, after having just read the problem statement, one expert remarked: "Immediately a methodological idea of a double-blind experiment suggests itself". The same expert remarked later on, after finally having found the correct formulation of the problem: "Now a 100,000 other things come up again. There are so many types of designs that you can set up". In the latter case, the final solution of the problem is familiar to the subject, so it need no longer be constructed, but can be reproduced as a whole, as soon as the problem is stated (Duncker, 1945, p.11). Thus, experts come up with more the better they understand the problem. The reason they can come up with more is because of their more extensive knowledge. By reformulating the problem again and again, more and more of the knowledge is accessed. This explains why the experts' protocols contained more statements overall, and why they came up with more designs than the novices.

In this experiment, an example of a default hierarchy (from general to specific) might be:

- experiment
- recognition experiment
- conceptual model
- rough temporal structure.

First, when reading the problem statement, the expert discovers that the task involves "designing an experiment". This will produce certain default expectations, such as: "expect independent variable". These default expectations can be overridden by more specific expectations, such as: "independent variable: cola".

Second, reading that the experiment involves "identifying different brands of cola", may trigger a "recognition schema". In one sense of the word, "identify" means "recognizing something you have seen before". The recognition schema will produce certain expectations, e.g. "first give test stimulus, then target stimulus".

Third, the conceptual model may be viewed as the first, highly general, statement of the design, e.g.: "I want to recognize one out of many, what they do on TV, you have to recognize Pepsi" (Subject E5). Note the use of analogy: the expert uses available knowledge, in this case a commercial on TV, and transfers it to the problem at hand. Use of analogies is characteristic for problem solvers who are confronted with a novel problem (Anderson, 1987).

Fourth, the rough temporal structure may be viewed as the outcome of the design process, which takes as its input the conceptual model. The conceptual model is successively refined and elaborated. This is accomplished by rule groups that generate expectations. For example, when a reference cola is given to a subject first, and a test cola secondly, a rule will fire that will note a possible influence of the first cola on the taste of the second cola. Other rules will suggest ways of dealing with this unwanted influence, e.g. by having subjects drink water in between. In this way, there are several rule groups dealing with irrelevant factors, repeated measurements on the same unit, repeated treatment-measurement pairs on the same unit, and operationalizing variables.

We did not find evidence for a default hierarchy with the novices: they only seemed to work at the lowest level. Eighty percent of their utterances could be classified as belonging to this level, while for the experts this number was fifty-six percent. However, expert problem solving not only consists of selecting and filling in the right kind of schema. In this experiment, experts also more extensively evaluated all the intermediate products they came up with. We consider this a purely strategic factor, not dependent on domain knowledge, since experts did not possess any more cola knowledge than novices. In fact, experts and novices could not be distinguished in terms of the number of irrelevant factors and control variables thev came up with. Of course, an evaluation strategy is knowledge, too, but it is part of the expert's general design knowledge, acquired from experience.

In conclusion, researchers with a lot of experience in designing experiments differ from novices in at least two important respects:
1. Amount and structuring of their knowledge
2. Strategies for understanding the problem and evaluating designs.

These two factors are closely related, since the large knowledge base can only be accessed by properly formulating the problem statement, and designs can only be evaluated properly by the availability of a large knowledge base. Novices lack the proper strategies and the amount of knowledge necessary for coming up with more than the barest essentials.

Future research

Of course, this account of expert-novice differences in the area of experimental design leaves open a number of questions. We will mention a few that merit further investigation:

1) The problem statement in this experiment was rather open and, at some points, ambiguous. This may have caused some of the differences found between experts and novices, and among experts themselves. For example, the problem statement focused heavily on irrelevant factors, and this may have caused the novices to think that this is the only important issue in design. The experts were sometimes bothered by the lack of a clear goal. Some of the experts chose a specific goal, while others did not. Therefore, in a next experiment there should be a clear rationale for why a design should be chosen. Also, subjects should receive instructions to pay attention not only to irrelevant factors but also to the procedure, the dependent variable, the target population, number of subjects, and instructions to subjects. In this way, we may be able to find out whether novices lack knowledge, or simply forget to mention important design elements.

2) It may be interesting to compare domain experts (e.g. experts in sensory research) with general experts, i.e. psychologists with a lot of experience in designing experiments but not in the particular area of sensory research. Both groups of experts share general design knowledge, but the domain experts also possess domain-specific knowledge. Comparing the two groups may yield insight in two questions:
   - what is the nature and extent of domain-specific knowledge?
   - how do general experts solve relatively novel problems, lacking domain-specific knowledge?

The second question was partly answered in this experiment, by comparing general experts with novices. However, it may be interesting to see whether general experts differ from domain experts.

3) Besides looking at the nature of the problem solving process in a descriptive way, one might also evaluate the designs subjects come up with. In this experiment, we have not looked at the quality of the designs subjects came up with.

4) The novices in this experiment did not have any theoretical knowledge about design. They may therefore have had wrong ideas about what constitutes a "good design". This problem may be partly solved by more elaborate instructions to subjects, as mentioned above. Another possibility may be to use subjects who have taken a course in experimental design, but who have very little practical experience with designing experiments.

5) The coding scheme used in this study was not based on a task analysis and did not contain any psychological assumptions. It merely was a scheme for classifying statements into categories that were derived from methodological handbooks. Another coding scheme should be put forward in which the categories have psychological significance (Breuker, Elshout, Van Someren, & Wielinga, 1986).

REFERENCES

Anderson, J.R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J.R. (1983). The architecture of cognition. Cambridge, Mass.: Harvard University Press.

Anderson, J.R. (1987). Skill acquisition: compilation of weak-method problem solutions. Psychological Review, 94, 192-210.

Breuker, J.A., Elshout, J.J., Van Someren, M.W., and Wielinga, B.J. (1986). Hardopdenken en protocolanalyse. Tijdschrift voor Onderwijsresearch, 11, 241-254.

Duncker, K. (1945). On problem solving. Psychological Monographs, 58, Whole No. 270.

Friedland, P. (1979). Knowledge-based experiment design in molecular genetics. In Proceedings of the Sixth International Joint Conference on Artificial Intelligence, Tokyo.

Friedland, P.E., and Iwasaki, Y. (1985). The concept and implementation of skeletal plans. Journal of Automated Reasoning, 1, 161-208.

Greeno, J.G., and Simon, H.A. (1988). Problem solving and reasoning. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey, & R. Duncan Luce (Eds.), Stevens' handbook of experimental psychology (pp. 589-672). New York: John Wiley & Sons.

Hayes-Roth, B, and Hayes-Roth, F. (1979). A cognitive model of planning. Cognitive Science, 3, 275-310.

Holland, J.H., Holyoak, K.J., Nisbett, R.E., and Thagard, P.R. (1986). Induction: processes of inference, learning, and discovery. Cambridge, MA: MIT Press.

Jansweijer, W.N.H. (1988). PDP: An artificial intelligence approach to problem solving and learning-by-doing in a semantically rich domain. Dissertation (in Dutch, with an English summary). Amsterdam: University of Amsterdam.

Johnson, H.H., and Solso, R.L. (1978). An introduction to experimental design in psychology: a case approach. New York: Harper & Row.

Kerlinger, F.N. (1973). Foundations of behavioral research (2nd edition). London: Holt, Rinehart and Winston.

Laird, J., Rosenbloom, P., and Newell, A. (1986). Universal subgoaling and chunking: the automatic generation and learning of goal hierarchies. Boston: Kluwer Academic Publishers.

Lenat, D.B., and Feigenbaum, E.A. (1987). On the thresholds of knowledge. In Proceedings of the Tenth International Joint Conference on Artificial Intelligence, Milan.

Malhotra, A., Thomas, J.C., Carroll, J.M., and Miller, L.A. (1980). Cognitive processes in design. International Journal of Man-Machine Studies, 12, 119-140.

Milligan, G.W., and Cooper, M.C. (1987). *Methodology Review: clustering Methods.* Applied Psychological Measurement, 11, 329-354.

Newell, A. (1969). Heuristic programming: Ill-structured problems. In J. Aronofsky (Ed.), Progress in Operations Research III. New York: Wiley.

Simon, H.A. (1973). The structure of ill-structured problems. Artificial Intelligence, 4, 181-201.

Appendix A: Encoded protocol of subject N1.
(Translated from Dutch).

Index to encoding categories:

Underst.: understand
Op.Vars.: operationalize variables
Contr.NF: control of irrelevant factors
Rep.Meas: repeated measurements on the same unit
Rep.TMP.: repeated treatment-measurement pairs on the same unit
Glob.TS.: global temporal structure
monitor : monitoring statement

Protocol	Encoding
Pour out cola into three glasses and put them on a row. Then put three bottles with the brands of cola next to them. After they have tasted the cola, people should put the glass next to the bottle of which they think the cola is in the glass.	Op.Vars.
Problem is, they aren't allowed to see or smell the cola, of course. So then it doesn't work.	monitor
Blindfold them and clothes-peg on their nose.	Contr.NF
Then ask them: this is glass number 1, glass number 2, glass number 3, and each time they think: this glass is this brand, then you could have them label the glass.	Op.Vars.
But then you should have them taste successively first Pepsi, then Coca Cola, then the own brand. Then you put the glasses into a new order.	Rep.TMP.
And then they can take a glass, taste, and then they have to say: I think this is Coca Cola, can I taste Coca Cola again. And then they take the third glass, own brand, and at a certain moment they will say: yes, glass number 1 is this brand, glass number 2 is that brand, glass number 3 is that brand.	Glob.TS.
To summarize: reference row, of which you may always taste; then fill three glasses with each brand, put them in an arbitrary row, blindfold, clothes-peg (not too important). Each time they have tasted they may say: this is Coca Cola, or they may go back to the reference row and taste	Glob.TS.

Appendix B: Encoded protocol of subject E5.
(Translated from Dutch).

Index to encoding categories:

Underst.:	understand
Op.Vars.:	operationalize variables
Contr.NF:	control of irrelevant factors
Rep.Meas:	repeated measurements on the same unit
Rep.TMP.:	repeated treatment-measurement pairs on the same unit
Glob.TS.:	global temporal structure
monitor :	monitoring statement
Evaluate:	evaluate

Protocol	Encoding
I would first try to define what "taste" means. Whether "taste" really refers to the taste buds, or whether taste is a kind of overall feature that includes factors such as smell, flavour, and colour. When you want to do it as a test for your own cola, then you would like a nice colour. Of course, you should prevent the name "Pepsi" from appearing on the cola. That is obvious. But the colour can be important. When you have decided what your goal is, what you want to achieve with it, do you want to say something about the cola, about the taste, might be important. If you say: just taste, then what you mean is maybe: what you have in your mouth, so no smell.	Underst.
You have to decide upon that in advance. Do you want to include it or not? Let's assume we will include it.	monitor
Let us prevent them from seeing what cola it is. Cola has been poured out and what would you have to do next? What other factors? We have just mentioned those other factors. We have said: taste also includes smell as a factor.	Contr.NF
Uh..., further, first problem of course always with these kinds of sensory experiments is that the first one that you have tasted, when you have come to the third you have already forgotten that taste. So you should, if you want to compare: Pepsi-Coca Cola, Pepsi-own brand, those are the kind of standard comparisons that you have.	Rep.Meas.
I would not know what the effect is if you had just started. You have not had any cola and you start to taste cola, whether that influences your taste a lot. If you have finished two glasses of cola, if you get Pepsi for the third	Rep.Meas

time, whether that Pepsi tastes the same.
So you need to control for a primacy effect.

So then you get a standard thing with a-b-c, c-b-a, or something. Or a-b-c, b-a-c, and c-b-a. At least if you presume there is an effect of order. And I think that effect is very strong with taste. — Rep.TMP.

Further, take wine-tasters as an example. They don't drink it, of course. That is of no use at all, what matters is taste, so that is what you want here, too. So you make sure they spit out everything. Maybe you should even rinse with a little bit of water. Quantities--small glasses. — Rep.Meas
Let's see, what else. Right, irrelevant factors... What would be irrelevant here? — monitor

Yes, you would do it like this: you have two glasses, Pepsi and Coca Cola, and then you have two glasses with those other combinations, Pepsi and own brand, and Coca Cola and own brand. Those are the combinations and you put them in a variable order. You cannot say after tasting once: I like that one the best, so you have them make pairwise comparisons all the time. — Glob.TS.

Then you need a scale to indicate your taste, so you do it very abstract, you can define taste as: good-not good. Or you can get something like: sweet...not sweet. — Op.Vars.

Let's see, you really say: "identify", here, don't you. — monitor

So you really have to recognize. — Underst.

So you really mean: it's a kind of memory experiment. — Underst.

Do you need to recognize one out of three, or do you have to be able to name all three of them? That will be two then, because the third is determined. — Underst.

Let me not ask, since you have not come up with it yourself. — monitor

So I want to know one out of many, what they do on TV, you have to recognize Pepsi. O.k. if I want to recognize one out of many, yes, you should do it like this: one out of many, and with one subject... — Underst.

You would have them taste first, and then taste three other colas. But if the next one

that you taste is the same again, for example, if you get Pepsi, you have Pepsi, then surely there is a high chance of saying: hey, that is the same.	Rep.Meas
Oh yes, you got to take that chance.	monitor
O.k. then I would leave the pairwise comparisons, since, although? Let me leave the pairwise comparisons, and get another type.	Evaluate
Another type would be that you have them taste Pepsi first, and then they would have to remember the taste, and then you get those other three. In an arbitrary order.	Glob.TS.
And then you would like to have each of them in a different order.	monitor
Yes, for example, you first get Pepsi as the target cola, and then you get the order Pepsi-Coca Cola and own brand, and then they have to say which one it is. And then you really should continue, then you should give them a different order.	Glob.TS.
The question is whether you need to taste again. If you don't have it tasted again, there is a chance that the taste will be forgotten. Depends a bit on what you want. Yes, depends a bit on what you want. That's the design for..., then you can take a number of orders.	monitor
And the criterion with all those things is, I would say, that effects of order are undesirable, so you want to exclude those. You do that with all those different orders. Both with the previous as well as this experiment. Should always be done.	Rep.TMP.
Further, if you want to compare something, it is very important that you remember it well. Because otherwise you are doing a memory experiment and that is not what you want. You really want to compare, so if you want to recognize taste, you need to make it vivid all the time. The taste really needs to be clear to you.	Rep.Meas
So that is your choice. Those would be the general principles that play a role in all these kinds of experiments.	monitor
And then the usual things: not letting know what orders there are, perhaps not even to the experimenter, as is often done with drugs	Rep.TMP.

research. The experimenter does not know what drug there is to prevent him influencing the subject.

If you want to judge three brands at the same time...	Underst.
you can first identify Pepsi on the basis of taste, and then the Coca Cola as the target drink, and do those combinations again, and then the own brand and those combinations again, and then see how well you pick out that cola.	Glob.TS.
That is labour intensive, then you need a lot of orders. Then the idea is to remember a taste and pick it out.	Evaluate
You can also say: what is taste really? You can ask yourself that question beforehand. What do they really taste? Is it sweetness, some titillation?	Underst.
You could do it like this, that is a completely different experiment. Nice experiment perhaps, not so standard.	monitor
You could learn the taste of those three in advance, just by saying: this is Pepsi, taste it, this is Coca Cola, tastes like this. You may describe the difference, I don't know whether that is a sensible thing to do, may be, may be instructive, see what criteria they have. And then have them learn it themselves. Right, then they have built up a certain criterion, suppose it works, and I think it will work, there is a difference, and you will do the experiments.	Glob.TS.
Then those order effects are not important any more. Perhaps a very efficient method.	monitor
Then you give them again, but now blindfolded, or without brand name; blindfolded, since there may be differences in colour. No fizzing, there may be a difference in amount of fizzing.	Contr.NF
And then you have them taste, and then you have them say for each drink what it is.	Oper.Vars
If you have Pepsi Cola, then there is the chance of getting one of those three, so you have to equally divide those chances.	monitor
So then you give Pepsi, Coca, and own brand, and then again, do it a couple of times.	Rep.TMP.

Now you don't have that effect of forgetting of taste, becomes less, since we have learned it.	monitor
If you have done a set, three orders or two, then you have to have them learn it again. The criterion may have been changed.	Rep.Meas
It is nice to know whether they change the criterion, because they can set a criterion beforehand by knowing what the names are. But if they don't know the names they will lose the criterion. The criterion is no precise enough to identify a cola in the multitude of tastes. Maybe they can learn a new criterion.	monitor
You could register a shift in criterion, and the nice thing is of course, since we are experimental psychologists after all, to be able to say something about the taste. One would like to go one step further and say: people describe something in a certain way, what do they describe, what aspects of the drink do they attend to?	Oper.Vars
Although, if you ask people beforehand to describe, then it is questionable whether that agrees 100% with what they use while actually judging the colas.	monitor
But if you ask people to be analytic then maybe you will be able to get those effects. Maybe you will have them saying: it's a bit sweeter than the other one, but still...and so on.	Oper.Vars
If those brands differ in degree of sweetness then there is not much to talk about. Then it is very simple. But if degree of sweetness is the same more or less, then it becomes more complicated.	monitor
That final experiment would perhaps be the most interesting, perhaps the most efficient, the most-- I don't know whether it is the most reliable-- perhaps you should do both of them. First experiment-- is not important any more. The second was: taste one and have it identified out of a set of three. So each time: yes-no. Second experiment: learn three tastes, so really learn a criterion...and then say with each drink what it is.	Evaluate

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) TD 89-1059	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER IZF 1989-31
4. PROJECT/TASK/WORK UNIT NO. 733.1	5. CONTRACT NUMBER B89-35	6. REPORT DATE September 12, 1989
7. NUMBER OF PAGES 34	8. NUMBER OF REFERENCES 19	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

How do experts solve unfamiliar problems: a preliminary study

11. AUTHOR(S)

J.M.C. Schraagen

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)

A largely unanswered question in the literature on problem solving is how experts solve unfamiliar problems. Do they resort to weak methods of search and analysis similar to those used by novices? Or do experts who have acquired powerful processes of reasoning in one domain apply those processes to solving problems in areas where specific solution methods have not been worked out? An initial attempt at answering this question was made in the present study. The domain chosen was experimental design. Subjects with varying levels of experience with designing experiments were asked to think aloud while they were designing an experiment in the, to them, unfamiliar area of sensory psychology. The results showed that experts quickly translated the unfamiliar problem into more abstract and familiar terms with which they could retrieve an experimental paradigm from memory. In contrast, novices only used a very general idea of what an experiment should look like, and hence could not provide as many details as the experts. The results further showed that experts have acquired powerful strategies for understanding the problem and evaluating designs. They can apply these strategies to unfamiliar problems. Novices apparently lack these strategies.

16. DESCRIPTORS	IDENTIFIERS
Problem solving Memory Transfer of training	

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) -

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.
