In Rasmussen's model it is still an open issue *what* rules and *what* knowledge are required at the respective levels. We cannot design user interfaces on the base of formally specified mental models about technical processes, but what we can do is to convey plant information to the user in order to support the spontaneous process of building a model.

In general - and this is easier said than done - the user interface should be designed to support normal operations in an easy way. In addition, the interface should provide a picture clear enough of the process to allow learning as well as alternate operations, if this is required. The aspects related to the monitoring and control computer and the task to be performed are treated in the next chapters.

knowledge-based level and the rule-based level respectively in Rasmussen's action model). Conceptual knowledge is more suitable to investigate and to search causes from observation of their effects; it is therefore of immediate importance in failure diagnosis. Operational knowledge is more relevant to select the right cause (if known) that leads to the desired effect. The emphasis on conceptual or operational knowledge in training should therefore be related to the type of task and the range of actions that is required by the process operators.

An additional, and relevant, result is that the transfer of knowledge also depends on its type. The transfer of operational knowledge is much easier than that of conceptual knowledge. This might be one the reasons for the success of graphical user interfaces (Macintosh, MS-Windows, X-Windows). These interfaces provide a common operational mode for all application programs that run under them, so that the users do not need to learn anew how to operate different programs (aside from their specific functions).

2.8 Conclusions

Mental models is a difficult and elusive issue to deal with. The concept of mental models can nevertheless provide useful information about how to shape and configure the interaction between people and complex systems. What we can analyse and study are not mental models as such, but rather information on how people interact with complex devices, get new knowledge and use it in a course of action.

The action model proposed by Rasmussen (Section 2.1) is adequate to describe work actions in general. For the normal operations of a system, rules of the IF-THEN type are usually sufficient. But the less foreseeable are operations, e.g. to control complex and potentially dangerous processes, the more important is the need for autonomous decision-making. Here, conceptual / theoretical knowledge about the technical process is probably more useful than strict operational knowledge. In addition, access to current process information (via the control computer and the user interface) and even explicit organisational support and responsibility delegation are other important factors for success in control operations.

Morris and Rouse (1985), so that probably no direct parallel can be drawn between the different conclusions they draw.

According to Kieras and Bovair, the user should get a "device model" and then make his own assumptions about how to operate the device. But how can the teacher select and define the required device model information to allow the student to make the correct inferences? The teacher acts on the base of a mental model of the trainees' mental models (in Norman's notation, $C(M(T))$). It is not certain that the teacher always identifies the correct information to teach, especially in consideration that technical systems where extended training is needed are usually very complex. On the other hand, if the operator infers the operational rules on his own, he gets both a better conceptual knowledge and the application-oriented instructions needed to carry out a task. Under this light, the approach may be worth further consideration.

Dixon and Gabrys (1991) investigated a different aspect related to training, namely the effect of the *transfer* of type of knowledge - operational or conceptual - on the success in the operation of a device. Their results indicate that prior knowledge of similar devices can aid in learning how to operate a new device. They suggest that this benefit might derive from the operational similarity of the tasks in the experiments (the simulated operations of different complex technical systems) rather than from any knowledge of what the devices do or how they operate.

When subjects learn a task that is *operationally similar* to a previously learned task, they can get an advantage from previously learned information. This holds even when there is little in common between the two tasks in terms of what they are intended to accomplish. *Conceptual information* instead provides a lesser benefit in knowledge transfer. Dixon and Gabrys do not rule out that conceptual information may be useful in the transfer of operational knowledge between similar devices, but they indicate how their experiments did not conclusively support this hypothesis.

Brief considerations on mental models

All the experiments reported here indicate that there is a difference between conceptual and operational knowledge (which would roughly correspond to the

> related to an explicit knowledge of system dynamics. Alternatively, it is conceivable that effective control behaviour may be related to having an understanding of system dynamics, but that this understanding may be in the form of a 'process feel' and may not be obtained via verbal instruction."

The current emphasis on the importance of theoretical knowledge about a system might then be disproportionate to the actual value of such knowledge. Instead, the content of instruction should be more directly related to what the operator may be required to do, both in familiar and unusual situations. However, no procedural training can enable an operator to deal with all possible unusual situations. The more it is expected that unusual situations may arise, or if these actions might have dangerous consequences, the more the operators should be trained as to be able to deal with them. Theoretical system knowledge is here essential.

A different experiment to evaluate the role and the importance of a mental model in learning how to operate a device was performed by Kieras and Bovair (1984). They tested different subjects by using a control panel replicating "Star Wars" hardware and terminology (in this way they intended to offset the effect of any previous school knowledge in physics that some of the test subjects might have).

The results by Kieras and Bovair showed that device model (conceptual) information can have definite and strong facilitative effects in learning how to operate complex devices. The device model supports operations because it makes possible specific inferences about what the operating procedures must be. Consequently, training program should provide such knowledge about the internal workings of the system that allows the user to infer exactly how to operate the device.

Kieras and Bovair observe that the device model information must support inferences about the exact and specific control actions. The relevant how-it-works knowledge can be very superficial and incomplete. Nevertheless, even without a full understanding of the system, the user must be able to derive the procedures for its operation.

The general validity of the results of Kieras and Bovair is questionable. For the first, the system they investigated was much simpler and therefore less representative of real complex systems than the one used in the experiment by type, or at the rule-based level, much more oriented to the practical operations. The issue is then about teaching how to *control* a system (**operational knowledge**) or how the system *functions* (**conceptual knowledge**). Some experiments have been carried out to relate the type of training, conceptual or operational, with success in actual control tasks. These experiments provide workable results and elude the difficulty of explicitly resorting to and characterising mental models.

An experiment to investigate the effect of the type of knowledge about a process on success in its control is reported by Landeweerd (1979). His work focused on the relationship between the *internal representation* (i.e. the mental model) of a process and the control behaviour in relation to diagnosing and correcting faults in a simulated process control situation. His results indicate that in tasks requiring fault correction, where the subjects have to take action, knowledge of cause-effect interactions (operational knowledge) leads to better performance. On the contrary, in tasks requiring fault diagnosis and where the subject has to analyse a cause from the effects, performance seems to be more related to a more abstract, conceptual, knowledge of the process structure.

More recently, also Morris and Rouse (1985) investigated what kind of knowledge, operational or conceptual, the operator of a dynamic system needs to have in order to work effectively. They conducted an experiment with a computer-driven simulation of a dynamic production process consisting of nine tanks, some of which are interconnected by pipes and valves. The plant operator's task was to supervise the flow of fluid through the series of tanks.

Morris and Rouse compared different training programs. At the end, they came to the conclusion that a deep training program, organized in order to provide conceptual understanding of the internal operation of a system, leads to worse results than a procedural-oriented, operational training program. According to Morris and Rouse

> "There is little or no conclusive evidence that providing operators with information about theoretical aspects of system functioning enables them to be better operators. In fact, in research in which subjects were given instruction in the theoretical basis of system functioning there was no apparent advantage to having been given such information [...]. It is quite possible that being able to control the system is not directly

certain number of difficulties from control theory: non-linear and time-varying characteristics, time constants over different orders of magnitude and multiple internal connections. The system can be identified by six properties (or "states" in control language); these properties can be inferred from the values of 11 measurable variables.

This study focused on the subjects' capability to predict the operation of the system on the base of the measured variables. Subjects were asked to judge the predictability of each system state or variable on the base of the others.

An analysis of the experimental data suggested the qualitative conclusion that "novices organise their knowledge in static terms, i.e. the 'surface structure' of the system, whereas the experts conceptualise system relations along more dynamic and integrated 'deep-structure' lines. This result clearly corresponds to the finding in other knowledge domains that experts tend to hold more abstract conceptualisations than do novices." The experiment showed also that experts seemed to make the distinction between system properties and system variables much more sharply than novices.

2.7 Mental Models and Training

The above-described analysis by Hopkins, Campbell and Peterson (1987) indicates that there are differences in conceptual models, but not how these differences can influence control tasks. Different types of experiments are required to indicate to what extent abstract conceptualisations of a system are necessary for successful operations. This information would be useful to select the contents in the training of operators of complex systems. Currently, training programs offer instruction in the theoretical principles upon which the system is based and perhaps some experience with simulators. It is often assumed that such instruction leads to satisfactory performance.

There is still no widespread agreement on the importance of a model in learning how to operate a device, or being able to operate it once it is learned. The basic question addressed by the researchers is whether the model provided by training should be at knowledge-based level, i.e. of a more theoretical and conceptual

Klein and Calderwood are critical about the real benefits of the prevailing paradigms in decision research. As they put it, such paradigms are in fact based "in simplified and highly structured laboratory tasks, and have therefore limited utility in operational domains characterized by high time pressure, uncertainty, ambiguity, continually changing conditions, ill-defined goals, and distributed decision responsibilities."

Classical decision models depend on a number of simplified assumptions. In experimental settings, the goals can be isolated and carried out independent of context. In theoretical models probabilities can be accurately estimated; choices, goals and evidence are carefully defined, and every subgoal is independent of the others.

But in practice - so do Klein and Calderwood report - "Less than 20% of the decisions involved concurrent deliberation, in which more than one course of action was considered and contrasted. The classical concept of a decision event, or moment of choice, is exactly this type of conscious evaluation of several different options, and yet it occurred infrequently in our sample. [...] Although the commanders could clearly recognize and admit to making mistakes, 'workable', 'timely' and 'cost-effective' were much more meaningful criteria."

The simulation by Brehmer and the analysis by Klein and Calderwood indicate that there is little in common between theoretical failure analysis with its probabilistic approach and the way people think in actual emergencies. Assuming a probability-based mental model to cope with unexpected situations is therefore not the correct approach. On the other hand, the indication of how people have dealt with real emergencies can provide clues for the design of technical systems to better support decision management under stress and time pressure.

Mental models of physical systems

Hopkins, Campbell and Peterson (1987) describe an experiment to investigate the nature of mental models of physical systems. In their experiment, the representation of a physiologic (mechanical heart/blood vessel) system has to be inferred on the basis of raw data. Test subjects are a group of students and a group of experts, i.e. people who do not know the system under test, but who possess a basic background in physics and physiology. The test system shows a

- the subjects do not manage to form any truly predictive models of the system
- the subjects do not realise the implications of feedback delay, e.g. in this simulation they do not use the commands which give some freedom of action to the units that were sent in the forest. And as things get more difficult near the end of a problem, when most of the wood has burned down, the test subjects are even less willing to delegate responsibility.

Brehmer's conclusions in relation to system design support the idea to delegate responsibility among the subunits instead of relying on a central coordinator who is unable to cope with system delays:

> "if feedback delays cannot be engineered out of the system, the decision-making powers in the system will have to be distributed throughout the system, and steps must be taken to prevent the central decision maker from assuming the total control that he cannot exercise with any success. Systems should therefore be designed so that they do not require perfect decisions."

The example of fire-fighting is also used by Klein and Calderwood (1991). In this case the behaviour of real firemen in action was analyzed, mostly with help of interviews. The data they collected illustrate aspects of thinking under stress conditions and in life-threatening situations.

When planning is analyzed in theory, it is assumed to be oriented to a "decision tree" with the possible alternatives represented by branches, each having a probability or feasibility value. On the contrary, in real-life emergency situations there is no time to evaluate possible alternatives. Subjects do not evaluate different alternatives according to their relative weights, but rather select a choice as soon as it seems to be feasible, under the unconscious consideration of the other parameters.

According to the study by Klein and Calderwood, firemen do not think in terms of the probability-tree method to evaluate alternatives for action (as implied by many models). Instead, the possible alternative solutions are evaluated one by one. When there is an indication that a particular strategy will not work, it is immediately discarded. As soon as one option for action seems to be successful, it is immediately applied.

In general, this new experience confirms what had been found earlier about the difficulties people have in dealing with complex systems. Subjects tend to act in a "feedback" fashion, approaching each problem on its own and only after this has developed. Often they do not treat systems as such, i.e. as interconnected entities, but just as a collection of independent variables. Subjects tend also to keep the status quo stable and disregard developments and their consequences. In addition, subjects prognosticate the development of variables by means of linear extrapolation; they assume a linear development even when the situation clearly shows that it must be non-linear.

These simulations show an additional result that has direct relevance for process control. Some of the process variables are stable because of the action of feedback loops. The test subjects usually do not recognize this effect of feedback mechanisms and believe instead that the related resources are inexhaustible. This leads often the user to "stretch" the use of these resources up to a point where the system cannot longer cope with it.

In the last experiment, a group of managers was tested against one of students. In general, managers scored better. Both groups had access to exactly the same type of information about the system, but managers seemed to know better how to adapt themselves to new kinds of situations. The managers probably learned the more effective approach from constantly having to make decisions in real-life situations.

Decision-making in emergency situations

In a companion paper to Dörner (1987), Brehmer (1987) investigated the difficulties in decision-making in complex situations. The example Brehmer took is a study on how different test subjects approach a fire-fighting task. People playing the (simulated) role of the fireman dispatcher to deal with a forest fire seem to react well to contingency situations, but:

- the subjects do not learn to differentiate between more efficient and less efficient fire-fighting units
- delay of feedback has truly disastrous effects on the subjects' ability to control the system

- individuals change the topic under consideration relatively quick and often jump from one topic to the next, treating all of them superficially
- subjects enclose themselves in those areas which don't seem to present them any difficulty. These are usually the least problematic and therefore unimportant areas.

In addition, subjects tend to reduce the problem complexity to fewer and fewer causes. In Dörner's own words: "Reductive hypotheses are very attractive for the simple reason that they reduce insecurity with one stroke and encourage the feeling that things are understood". The author also observes that "certain deformations and degeneration of reason are to a certain extent natural effects of general human behaviour patterns".

Dörner's paper makes another further important point, namely that correct decision making is matter of method rather than of experience and knowledge. In other words, a person who thinks rationally will tend to approach all problems in a coherent fashion. A person prone to mess up things will do it in several areas.

A similar experiment with a different scenario substantially confirms these results (Dörner, 1990). After a catastrophe occurred in the Moro region (Bukina Faso, West Africa), a development aid project was organized to create better living conditions for the semi-nomadic people living there. Using deep wells, more water was made available, which led to higher agricultural outputs and increased cattle stocks. Even the tsetse fly had been brought under control. But at the end the cattle exceeded the capacity of the available pasture area; decrease in rainfall led to food shortage and eventually to a famine in the human population.

Fortunately, Bukina Faso exists only in form of computer software and the catastrophe was no more tragic than downward pointing trend curves on output graphs. Again, the aim of the simulation was to investigate how people meet the various demands in a complex situation like this "dynamic decision problem". This system is complex, consists of a large number of variables, and evolves dynamically on its own without external intervention. It is difficult to manage such a system, because the implicit requirement to cope with it is to "look ahead". But what many subjects tend to do is to react only to contingent data, without planning for future actions.

way; in a comprehensive work (Gentner and Stevens, 1983) it is indicated that many people hold incorrect mental models even for simple physical phenomena like elementary dynamics and electrical theory.

A large part of the research on mental models has been dedicated to office systems and in particular to word processors. The aim of such research was to conceptualise the operation of these systems and therefore give indications about how to design them in order to be easier to operate. Because of the intrinsic difference between sequential and real-time computing, not all the results so obtained can be directly extended to real-time systems: a word processor is entirely under the control of the user, while a real-time system is not.

Research about mental models in the real-time interaction with complex system is often carried out by evaluating the performance of test subjects in simulated scenarios. Other researchers conduct field investigations about real-life situations, like e.g. how emergencies have been managed in concrete cases.

A first example of a simulated scenario is described by Dörner (1987). With help of a computer program, the test subjects played the role of mayor in Lohhausen, a small (virtual) city of 3,500 inhabitants. The economy of this city depends mainly on a municipal industrial enterprise, a manufacturing plant for the production of watches. The city has also its administration, shops, schools, etc. On the computer, subjects were able to influence the production of the city factory, to change taxation rates, to create or cancel jobs, to decide about housing construction, and so on. The results of the runs were quite different for the different subjects. While some managed the task very well, others showed a very poor performance and led the city to bankruptcy and ruin.

Dörner, a psychologist, relates the cause of the errors to a number of traits that are very human indeed. According to him:

- most people are not interested in finding out the existent trends and developmental tendencies at first, but are interested instead in the status quo
- all subjects have difficulties with exponential developments
- there is a tendency to think in causal series, not in causal nets
- poor performers show a low assessment of their own ability to act

control of the foreseen operations. A model is also present at the skill-based level; it is for example the "feeling", that the activation of a switch makes something happen in the plant or process. This model cannot be provided by theoretical training and courses but has to be learned from practice.

There is no general agreement about the importance of a formal model to carry out control operations. And there is no indication that people with a strong theoretical background can carry out operations better than other people without this knowledge. In general, actions carried out at the lower levels are faster and much more effective than those that require more intensive thinking. On the other hand, high-level thinking is required to explore new situations, for example to find the reason for a failure. Mental models for the same target system T can also take different forms depending on their purpose. The mental models of a car by a car-repairman is very different from that of the sports driver. The driver probably could not do his job if he constantly thought about everything that might go wrong with the engine.

The industry, much more interested in practical aspects than in theory, has long taken the approach that operators can do their job with comparatively little training and thus does not overestimate the importance of mental models. The personnel of many complex systems receive only little training about the processes they control, so that they have to build their own mental models and simplified operational schemes from experience. This low-level knowledge is periodically expanded and integrated with theoretical courses.

2.6 Experimental Results about Mental Models

The management of complex systems

Different experiments have been carried out in order to get insight into the structure and function of mental models. The approach of these studies is behavioural-oriented rather than cognitive-oriented: what is actually studied is how people deal with complex systems. From these studies insight on the nature of mental models as well as how people manage in different situations may be gained. It is known that even simple facts are often conceptualised in the wrong some physical description for T and $M(T)$ is how the user sees T. It is almost given that $C(T)$ and $M(T)$ are different[1].

In the design of user interfaces we can act - at least in part - on the relation mapping T with $M(T)$. We want to influence the formation of $M(T)$ in such a way to enable the operator to control a process. To do this, what we can manipulate (at least in part) is the interface between T and $M(T)$. In this respect we operate on the base of *our* formal model of the system under control, $C(T)$, and our own ideas about the user's mental model, $C(M(T))$.

The most immediate - and for an engineer almost natural - assumption, is that the user's long-term memory contains a model of the process to control. In formal terms, $M(T)=T$. This view implies that process operators know all handbooks and drawings by heart and can act immediately on an emergency, possibly even with anticipatory commands. It is assumed that, if needed, the operator may select appropriate controls faster and more precisely than the control computer and at a speed compatible with that of the technical process.

Such an assumption is naive and wrong for several reasons. In the first place human thinking is much less structured and detailed than formal information contained in handbooks. In addition, humans can also restructure and adapt their thinking patterns adapting them to contingent situations, while information recorded in books remains the same all the time. And mostly, people in charge of a complex process T do not have a copy of T in long-term memory, to intervene with appropriate functions of T^{-1} when needed.

A mental model is built in course of time and can constantly be modified and extended. At the knowledge-based level (Section 2.1) the model has a theoretical form, and is oriented to explain how a system works. This information can be provided with training and courses and is built in general on basic facts from chemistry, physics, electrical engineering, and so on. At the rule-based level the model consists in rules, possibly in IF-THEN fashion, and that are apt for the

---

[1] Norman's original notation is lower-case t. This may however lead to confusion in notations like $C(t)$, $M(t)$, etc. because of the close similarity with the notation for functions of time (which they are not). For this reason, we will use capital T to denote a target system.

In psychology, there are different approaches to the study of mental models, two of which are of particular interest. In the **cognitive** approach the goal is to build an internal model on the base of experiments. In the **behavioural** approach instead, an internal model is not essential; what matters is the reaction of a person to specific stimuli. According to this view, the mind is a **black-box** that can be studied and defined in terms of stimuli and responses.

What is the importance of a mental model in process control? An interesting indication of the role of models in control tasks is given in the novel "Reason" by Isaac Asimov (1950). A robot called Cutie is highly successful in a complicated control task, the redirection of a powerful electron storm which could destroy the Earth. Cutie (who eventually saves our planet) follows a conceptual model entirely different from the one he was instructed for. Cutie does not believe in the existence of the Earth ("just a dot on a radar screen"), yet he is able to perform the expected tasks by monitoring dials and gauges ("I kept all dials at equilibrium in accordance with the will of the Master"). The astronauts who watch him perform this task affirm then quite appropriately "Then what's the difference what it believes!".

The difficulty of working with mental models does not only lie in the nature of the problem itself. As Wilson and Rutherford (1989) in a comprehensive study indicate,

> "the concept of user mental models appears somewhat confused and incoherent. Indeed, Norman's (1983) observations that mental models are incomplete, unstable, non-exclusive, and unscientific - among other attributes - might also characterise the whole enterprise of applying them, judging from the literature".

Wilson and Rutherford point out how even professionals in related fields (like e.g. psychologists and human factor experts) do not understand the same concepts for mental models. The case is not better with engineers in search of formal descriptions.

Norman (1983) differentiates between the actual target system *T*, the conceptual model of the system seen by the system designer or an external observer, *C(T)*, and the user's mental model of the target system, *M(T)*. *C(T)* is mostly related to potentially competent personnel must balance own judgement, compliance with orders from above, other external pressures (e.g. scheduling requirements) and even the risk of turning out as scapegoats if something goes wrong when they make unusual decision.

Many operators want to have responsibility and actively participate in the configuration and control of the process. As Olsson and Brehmer (1990) and Olsson and Lee (1992) point out, supervisory control should not reduce the operator to passively reading on a screen that the computer is operating correctly, but instead carry out tasks that are more integrated with the work of the company management and with the goals of the organization.

All these issues lie beyond the issue of the human-computer interface and must be solved at other levels. Human-computer interaction is just one of the many relevant components. Yet, if other aspects are left untouched, even the best user interface cannot do much to improve operations quality and workers' satisfaction or reduce the incidence of errors. Working on the user interface alone will not lead to the envisioned results for a system and its goals.

2.5 Mental Models

The issue of **mental models** is central to the whole problem of human-computer interaction and operation of complex systems. Mental models provide the foundation for many operational decisions. Actions at the knowledge-based level in Rasmussen's model (Section 2.1) imply the existence of mental models.

Very little is known about the nature of mental models, just as little is known about the nature of thought in general. In this section there will be no attempt to present a new theory of mental models (or of the thought process), but this problem will neither be eluded by labelling facts instead of explaining them. Scientists and philosophers have long dealt with the problem of thought and intelligence, but we still don't have a general theory. This does not mean that one should not speculate on the nature of mental models or not try to gather information on the subject.

This does not mean that the best process display would show a woman on a high-resolution computer screen more or less undressed as a function of the process parameters in order to draw the operators' attention. It rather indicates that we should not only rely on models like Rasmussen's (Section 2.1), error classification schemes, GOMS, Norman's action model and others as if they gave the whole picture. System engineers and user interface designer must know to what extent these models are applicable and where their limits are. The intrinsic playful nature of people was perfectly understood by the developers of the Macintosh computer, who introduced subtle playing cues in the interface and so helped it to its success.

An additional aspect that should deserve more attention is the role played by traditions, established patterns or pure conservatism, and that can be enormous. For example, we know that normal QWERTY keyboards are not optimized (they were actually designed to slow down typing, so that the keys would not jam in the first mechanical typewriters). The first row above contains all letters needed to type the word "TYPEWRITER", so that the first vendors, who did not possess typing skills, still managed to carry out a basic demonstration of the product. An optimal keyboard layout has been defined (the "Dvorak" design), where the keys are organized in order of their frequency (e.g. the keys "T", "H", "E" are situated close to one another) so that they can be typed with no effort, and assigned to the most agile fingers. Well, very few people use the Dvorak-keyboard, although many would agree on the advantages of its design. All the human factors and user interface experts I met had normal keyboards, although Dvorak keyboards can be purchased. The reason is simple, after having gone through the process of learning how to type (on standard keyboards), nobody wants to retrain to use a different keyboard.

It is known by everybody with direct workplace experience that imponderable factors like degree of responsibility, relation with the supervisors, possibility for career advances together with many others have a definite and strong impact on motivation and thus on performance at work. Reason (1990) cites the example of some companies with very high quality standards, where in high workload situations the conventional, rank-based management structure changes into a work structure based primarily on competence and where the formal ranks lose their importance. Several organisations have not yet understood this fact. Many

TOP 40 NEWSGROUPS IN ORDER BY POPULARITY (FEB. 93)

	(a)	(b)	(c)	(d)	(e)	(f)
1	190000	91%	1	14.2	10.7%	news.announce.newusers
2	150000	87%	11	436.9	8.4%	news.answers
3	150000	85%	1120	2385.6	8.2%	misc.jobs.offered
4	140000	68%	1593	4128.3	8.0%	alt.sex
5	140000	83%	71	136.7	7.8%	rec.humor.funny
6	120000	56%	2511	99428.9	6.8%	alt.binaries. .pictures.erotica
7	120000	83%	1765	2166.3	6.7%	misc.forsale
8	120000	88%	873	1683.4	6.5%	comp.windows.x
9	120000	53%	1150	6188.8	6.5%	alt.sex.stories
10	110000	91%	88	637.2	6.3%	news.announce.newgroups
11	110000	90%	1616	3352.8	6.2%	news.groups
12	110000	82%	2128	5363.5	6.1%	rec.humor
13	94000	69%	459	2442.0	5.2%	alt.activism
14	94000	88%	1145	2168.3	5.2%	comp.lang.c
15	93000	86%	902	2012.9	5.2%	comp.graphics
16	86000	83%	536	963.4	4.8%	misc.jobs.misc
17	84000	72%	88	1086.4	4.7%	rec.arts.erotica
18	84000	64%	1441	4911.1	4.7%	alt.sex.bondage
19	83000	83%	60	2630.5	4.6%	comp.binaries.ibm.pc
20	82000	88%	822	1435.0	4.6%	comp.unix.questions
21	82000	87%	1344	3112.8	4.6%	comp.lang.c++
22	81000	79%	133	3334.8	4.5%	alt.sources
23	79000	88%	244	1550.9	4.4%	news.announce.conf
24	78000	59%	1434	52584.0	4.3%	alt.binaries. .pictures.misc
25	77000	85%	-	-	4.3%	news.announce.important
26	77000	90%	427	543.8	4.3%	news.newusers.questions
27	76000	85%	2766	4255.4	4.2%	comp.sys.ibm.pc. .hardware
28	74000	82%	313	596.5	4.1%	misc.jobs.contract
29	74000	85%	196	324.5	4.1%	comp.misc
30	73000	74%	3118	8658.6	4.1%	soc.culture.indian
31	73000	11%	20	115.0	4.1%	clari.news.briefs
32	72000	91%	43	1046.9	4.0%	news.lists
33	72000	84%	324	944.8	4.0%	comp.ai
34	71000	79%	2410	6052.6	3.9%	rec.arts.movies
35	70000	65%	1989	3255.1	3.9%	alt.personals
36	69000	81%	423	426.6	3.9%	misc.wanted
37	69000	85%	714	1064.3	3.9%	comp.sys.ibm.pc.misc
38	69000	73%	532	867.6	3.9%	alt.bbs
39	69000	86%	103	5049.6	3.8%	comp.sources.misc
40	69000	80%	94	164.4	3.8%	comp.dcom.lans.misc

Table 2.1 Newsgroups in order of popularity (Usenet Log for February, 1993)
(a) Estimated total number of people who read the group, world-wide; (b) Propagation: how many sites receive this group at all; (c) Recent traffic (messages per month); (d) Recent traffic (kilobytes per month); (e) Share: % of newsreaders who read this group.; (f) newsgroup name.

This is quite a subtle point, which unfortunately is overlooked in most of the literature on "human factors". No formal description can capture the role played by natural, human attitudes. The real user jumps from one theme to another, gets bored, oversees important facts, does not recognize the temporal evolution of things, etc. As Douglas Hofstadter in his work on artificial intelligence *Gödel, Escher, Bach: an Eternal Golden Braid* observes, a computer, if it were really intelligent, would show it by getting bored at doing mathematics.

Electronic mail is a good example to illustrate this point. We have believed for several years that new media (like electronic mail) would *improve* the way we communicate. Wrong. This might hold in several cases, but in general the electronic medium has facilitated the transmission of *higher amounts* of data, not *better* data. There is an even subtler side effect: the wider the communication channel at my disposal, the less time and effort I will spend to improve my message right from the beginning in order to economise on channel use. Let those who receive my message do the job instead. The result is a higher general workload made up by the quantity of information transmitted via the medium.

The initial intention with the introduction of networks was good. It was implicitly assumed that they would improve the quality of work and that they would *only* be used to make work more efficient. But to get an idea of what people find most interesting, it is sufficient to check a report of the most popular newsgroups in network use, in this case Usenet in February, 1993 (Table 2.1). One has to scan the log long down to find that "comp.ai" was called by 72000 (4.0%) of the world-wide users and thus takes the 33rd place in order of popularity. "sci.environment" was selected by 33000 (1.8%) users, ranking 287th in popularity. The most favoured user requests are explicit by themselves.

These data are an example of something that is unexpected first and looks obvious afterwards. The newsgroups on Usenet are quite interesting in this respect because their development is spontaneous and therefore provide a much better picture of users' interest than any questionnaire-based research could do. In Usenet, there is no central authority deciding and organising the themes, but rather the activity is controlled by the user community at large in a direct and spontaneous way. The net result is that computer games and lifestyle conferences get a larger share of interest (measured in user messages or binary amount of traffic) than scientific information. They give a very human picture of the user.

detail, so that for each error a specific course of action can be selected and - if applicable - carried out. The design limits of error avoidance schemes should not be hidden but rather be part of an open, known, and documented design decision.

To check the outcome of control actions and for training it would be quite straightforward to install two terminals in process control rooms, one for the real process monitoring and control, the other for process simulation. The outcome of a command could then be tested against the simulated system before being transferred to the real process. To gain more accuracy, the simulation software would have access to all process monitoring data, but without influencing the process itself. Simulation and training software is still used quite sparingly in the process industry, although it would represent a good substitute for the "vicinity to the process" that is getting lost because of automation.

2.4 Effects of Context, Expectations, and Motivation

An additional aspect of importance for how operations are carried out in practice is **motivation**. This factor is not adequately described by research that focuses on too specific aspects of the interaction between humans and computers. Data collected in restricted settings describes very specific phenomena, but may lose its validity when it is extrapolated and referred to in other contexts. The picture of the user that one gets via laboratory experiments is an **ideal** representation that only in part reflects real people in real environments.

People are people and their basic traits and attitudes that have evolved in million of years do not change just because they sit in front of a computer screen. In the development of the *medium* computer and in almost all related literature there is a more or less explicitly stated quest to improve the way people work, with specific expectations about how this is going to happen. What instead often happens when new tools are put at disposal of *normal* people is that a new kind of symbiosis human-machine is established, usually in a different way than expected.

taken by the designers beforehand, i.e. the user interface should support - or at least not hinder - unusual operations.

Error correction means that the operator or the machine recognize that an error has taken place and acts to correct it. The most common solution used in computer systems is the command "Undo". The computer buffers all commands before executing them. For example, a file would not be immediately erased from memory, but only scheduled to be erased. If the operator changes his mind after having deleted a file, he can always get it back as long as this has not been physically removed from the system.

The "Undo" function can obviously only work when no change in the technical process has yet taken place. In real-time process control there is no possibility to "buffer" commands like in the virtual environment of a computer. For this reason, in process control applications errors should be avoided - as far as possible - from the very beginning. The user interface should therefore include data about possible dangerous states or even a simulation routine to foresee the outcome of an action, holding the above-mentioned considerations about error avoidance. The practical handling of this aspect will be treated in Chapter 3.

Paradoxically, in a complex system errors should not be avoided altogether. Errors represent namely a good experience source. It is no coincidence that an important learning method is called *Trial-and-Error*. In a way similar to the way children learn, experimenting on one's own delivers a "feeling" for carrying out actions, most of all at the sensomotoric level. There is no way to replace feelings gained in this way with theoretical training.

There are obviously situations where "experimenting" with the target system is undesirable or impossible (like in nuclear power plants). Simulation routines can here help collect the necessary experience without risk. A pilot who already crashed several times on the simulator will probably be able to keep his aircraft in the air better than the one who does not know the limits of the machine.

In error management, a strategy should be selected in relation to the work task. For each task a decision about error avoidance or error correction must be taken. If errors are considered to be probabilistic, then the system has to be made insensitive to them as far as possible. Causal errors must be analyzed more in

Rasmussen classifies errors on the base of his action model (Section 2.1). Slips take place at the sensomotoric level. Mistakes take place at higher levels, where planning is concerned. A mistake could be the selection of a wrong rule or a wrong complex decision. An example of slip is a typing error on a typewriter or the shift into the wrong gear in a car. A mistake can take place at the rule-level (e.g. wrong spelling of a word) or at the knowledge-level (use of a correctly spelled word that is wrong in the context of a sentence). In car driving a mistake could be the wrong estimation of a slope leading to the *successful* shift into the *wrong* gear.

Focusing on the slips and mistakes of a single person is however also an implicit way of looking for scapegoats, at least in large, hierarchically structured organisations (where scapegoats have a tendency to be found at the lowest levels). More recent approaches focus instead on complex systems as such (Reason, 1990). According to Reason, "[the operators] at the human-machine interface were the inheritors of system defects created by poor design, conflicting goals, defective organization and bad management decisions. Their part, in effect, was simply that of creating the conditions under which these latent failures could reveal themselves". This means that, in order to make a system safer, all its levels must be considered.

Error management is an important part of human-computer interaction. A good user interface must contribute to the reduction of the number and the consequences of errors. Error management can be approached in two different ways: error avoidance and error correction.

Error avoidance requires that the machine in some way recognises the evolution toward an error situation and alarms the operator about it. The machine would then not accept commands that may have dangerous consequences. This method is already employed, but it is itself not immune from drawbacks. In several work contexts (e.g. in civil aviation) operators are unsatisfied with equipment that explicitly does not carry out commands or that "interprets" how to carry them out. The usual reason behind this criticism is that a system can only be as good as it is designed and built and there is no way the designers can foresee all possible situations and outcomes. The final decision about how to react in emergencies must be left to the operators and not implicitly information, if it is needed to accomplish his tasks. Otherwise, an important indication of malfunctioning and problems in a system might be lost.

The limited capacity of sensory storage and of short-term (working) memory sets also the limits to the amount of information that can be perceived and processed concurrently. This aspect will be treated more in detail in Chapter 4.

2.3 Errors and Error Management

Errors are a natural aspect of all human actions and play a very important role also in process control. Human errors are often cited as the cause of accidents or unacceptable performance in technical systems like aircraft, ships and power plants. For many of these systems it has been estimated that 70-90 percent of all failures can be traced to human error.

According to a comprehensive study by Rouse and Rouse (1983), two major approaches can be taken to characterising human error, probabilistic and causal. In the **probabilistic** model errors are - as the name tells - of random nature and can be treated in a manner similar to that used for hardware failures. No effort is made to investigate errors on an individual basis and to minimise their effect; instead, a general reliability metric is established. This metric is then used to verify whether the reliability of a system meets specified levels.

The **causal** approach to error analysis takes an entirely different standpoint. Errors are not considered to be random events but actions whose cause can be exactly traced back, identified, and possibly removed. The difference in the two approaches can also be characterized as whether the accent is put on *what* errors occur or *why* they occur.

Two known classifications of errors were proposed by Norman and Rasmussen (cit. from Rouse and Rouse). Norman's model is simple and straightforward: errors can be either slips or mistakes. **Slips** are intentionally correct actions that are not carried out. **Mistakes** are correctly performed actions that reflect inappropriate intentions.

Information storage in long-term memory takes place with some kind of encoding. What is stored is the **meaning** and not the **form** of a message or of symbols (e.g. the exact wording of a sentence). The transfer of information to long-term memory does not depend only on a voluntary effort. We remember well things that are quite unimportant, but it is usually difficult to memorise rote data as such, for example during the preparation of an exam, and no matter how much we want to do it. Further, there are indications that memorisation is permanent and that there actually is no such thing as "forgetting". It seems to be instead that already stored items cannot be retrieved, possibly because of wrong or missing cues.

Human memory does not work with direct cell addressing as computers do; it works instead on the base of analogies and associations. Information storage is easier if the new data can be put into an existing "frame", i.e. if the data can be related to information already present in long-term memory (for example the association of the above-mentioned phone number with the city of San Francisco). Memorisation of different facts works also better if these are not presented alone but are put in causal relationships. Similarly, recalling is facilitated by "cues" hinting at some aspect of the data to be retrieved.

Perception and interface design

The conclusions about the role of basic perception skills in relation to interface design are quite straightforward. When a type of interface is substituted with another (e.g. a mechanical manipulator with a computer-controlled one, a mechanical valve with an electrical valve and a local regulator) it should be considered whether there is direct information that in this way is lost. From a shift gear one may feel "in the hand" if the clutch does not work properly; the same is not true for an electrical gear, unless explicit feedback information is provided in some form.

If - and that is the case with computer controlled systems - the information provided by tactile and acoustic perception is lost, it must be determined how this information can be replaced. It is difficult to bring on a screen the equivalent information of motor noise and vibrations (which to an experienced operator convey quite a large amount of information). However, the operator should not be required to control the process "blindly", without resort on this type of given instant and provides us with a base for action. Long-term memory has an almost infinite storage potential, but memorising and recalling takes longer. Information in long-term memory makes a person's entire knowledge and includes everything from the use of language to childhood's memories, from multiplication tables to the name of the King of Ruritania. Information in short-term memory has a retention time of seconds, in long term memory it can last a lifetime.

The distinction between short-term and long-term memory has also a physiological explanation. There is no "spatial" separation in the brain between the two storage areas, but both take place in the whole brain. The difference is in the type of activity. Short term memory activity is more of electrical nature (the distribution of electric fields), while long-term memory consists in neuronal interactions and connections of a more permanent, chemical, nature.

In a comprehensive comparative study, Miller (1956) suggests that in short-term memory there is place for about 7+/-2 information items, also called **chunks**. The same limit holds also for the amount of external stimuli that can be properly identified. New information coming in will erase or displace the existing chunks. Items not thought about decay quickly and are lost from consciousness. "Chunks" are not the same as information bits; in fact a chunk can be very rich in its information content. The items in short-term memory are at about the same abstraction level, or show at least some homogeneity.

A key aspect in the efficient use of short-term memory is therefore **coding**, i.e. how much "raw" information is inserted in a chunk. Organization and relation with previous knowledge help to handle new information more easily. Take for example the number sequence 14154569220. It just looks like an arbitrary sequence of eleven figures. Most people would not be able to recall such a sequence without some effort and would probably soon forget it. But regrouping the sequence as 1-415-456-9220 makes it more manageable, and even more if it is identified as a San Francisco phone number. Regrouping has reduced the number of chunks from 11 to 4, an amount that most people can handle without major effort. A similar example is found with chess players: a master can remember the position of 20 pieces on the chessboard, the novice cannot. The reason is probably that the novice sees the pieces as 20 different items, the master as one or two chunks.

than one message can be perceived and understood at the time. When several messages are heard at the same time, they only generate confusion.

The most important sense in general as well as in the realisation of user interfaces is vision. Hearing is important when acoustic devices are used. Tactile perception was important in relation to the use of controls connected to mechanical equipment (clutches, brakes, the yoke in an aircraft). A practical advantage of those controls was the immediate feedback that the task was accomplished, the "feeling in the fingers" that a clutch is connected or that a fluid starts flowing through a valve. With the introduction of computer-controlled servo devices this type of feeling has been lost or replaced by visual information on a screen.

In vision, the perception of colours is one of the most important factors; colours play an important role also in human-computer interaction. The physiology of vision is fairly well known. The human eye is most sensitive to green and least sensitive to the high-end frequencies of the colour spectrum, blue and violet. The eye focuses different colours at different distances, so that for example if red and blue are close to each other, the eye tries to focus them at different distances and loss of clarity is the result. About 8% of the male population and 0.5% of the female population in Europe and North America have some kind of colour blindness and do not recognize some colours or colour contrasts.

In general, unexpected stimuli are perceived with a certain amount of attention, but the attention decreases when the stimuli are repeated. Other factors that increase the likelihood of perception of a stimulus are intensity, size, contrast and movement. In addition, the brain also "knows" what to look for: in scanning a picture, the eyes tend to fixate where the important features are located.

The dual-memory theory

The information from the sensory organs is first collected into the sensory storage and from there it is transferred to short-term (working) memory where we can consciously pay attention to it. From the short-term memory, and in most cases only with a voluntary effort, information is transferred to the long-term memory. Short-term memory is fast to recall (and forget) from, but we can "see" all the information contained in it at the same time and react quickly on its base. Short-term memory is our consciousness, it holds whatever we think about at a mental process of perception, memorisation and information processing by which the individual acquires knowledge, solves problems, and plans for the future. Cognitive processes can be modelled with experiments and without having to resort to neurobiological explanations.

The total amount of information entering the body is estimated to be one billion bit/s, of which only about 100 bit/s are processed consciously. The brain tends to further reduce the amount of information to process. If too much information is presented at the same time, acting capacity is lost and the attention tends to concentrate only on part of the input data.

Psychologists have dealt for a long time with the problem of sensory perception. Modern psychology differentiates between functional cognitive entities: sensory storage, short-term (working) memory and long-term memory (Figure 2.2). The stages in the information processing by the brain are perception, storing in short-term and long-term memory, planning and conversion in control action.

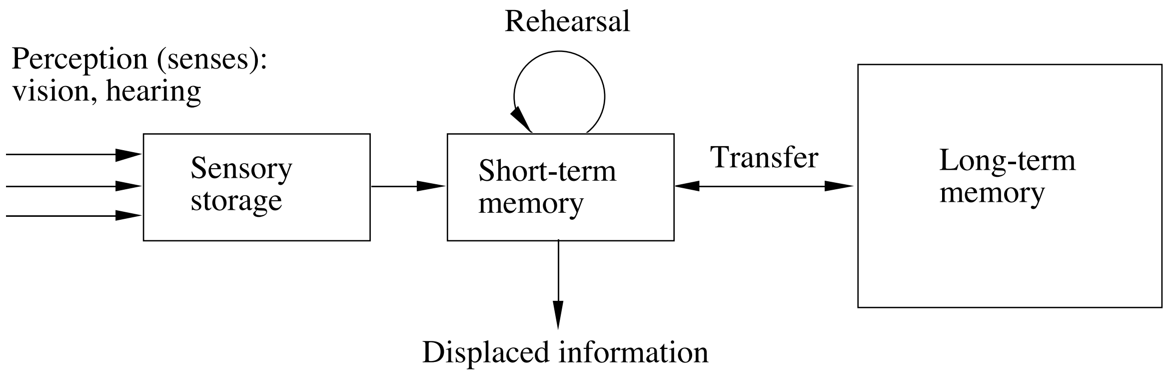

Figure 2.2 The dual-memory model

The senses collect different types of information; information from one sense cannot be directly replaced with information from a different sense. The senses communicate data in different fashions. Vision is parallel and is apt to give clues to the interrelationships among objects. For this reason, vision is the optimal sense not just to *see*, but also to *interpret* the world. Hearing is serial, no more

- executing the action
- perceiving the system state
- interpreting the state
- evaluating the system state with respect to the goals and intentions.

Norman's model can be considered as a refinement of the GOMS method. Both models have however a major drawback because they do not identify the different abstraction levels in planning and evaluating the action course.

In conclusion, Rasmussen's model is adequate for the analysis of process control interfaces where also cognitive activity is involved. The division in skill-based, rule-based and knowledge-based behaviour permits the differentiation among the work tasks and indicates what kind of support must be provided from the machine to the user. At skill-based level this support might consist in immediate feedback about the action, either sensomotoric or visual. For example, in the control of a robot manipulator with a joystick we might want to follow the path by reading the real-time coordinates on displays. Sometimes exact points might be of interest, other times it would be sufficient to check the direction of movement. At the rule-based level, the data and the possible types of action will be shown. A way to do this is to highlight appropriate selections in a menu in relation to the operational context. Similarly, on a control panel, only a subset of the commands might be relevant at a certain time, and these commands could be indicated e.g. with lamps. Finally, at knowledge-based level, the process interface has to be designed in order to allow reasoning about a problem and support decision-making. In some situations this means that large quantities of data might have to be presented to the operators. This aspect will be examined in more detail in later sections.

2.2 Perception and the Dual-Memory Theory

Physiological and sensory perception

The acts of perception and memory processing have been investigated for a long time in psychology. **Perception** means to become aware of something through our senses, i.e. to come to know what is going on around us. **Cognition** is the

At the **rule-based** level, the information is typically treated as **signs**. The signs serve to activate or modify a previously learned action or behaviour. Finally, high-level information must be perceived and communicated in form of **symbols**. In this way, information is useful for reasoning in predicting or explaining unfamiliar behaviour of the environment and, more in general, to support abstract thinking.

The boundaries between the different levels of behaviour are not exactly distinct, so that this model has mainly a qualitative character. It is important to notice how in the different types of actions the efficiency and response speed is highest at the lowest level. The low efficiency of knowledge-based behaviour is compensated by the capacity to adapt the action to new circumstances.

Other models for human behaviour and performance have been proposed. Card, Moran, and Newell in their work *The Psychology of Human-Computer Interaction* (cited in Shneiderman, 1987) describe the **GOMS** (Goals, Operators, Methods and Selection) model. They postulate that users formulate **goals** and subgoals and that each goal or subgoal can be achieved by specific **methods** or procedures. In this model, the **operators** are elementary perceptual, motor, or cognitive acts, whose execution either affects and changes the user's mental state or the external environment. The **selection** rules are used in the choice of one among several methods available for accomplishing a goal.

The GOMS model is comparable with the rule-based level in Rasmussen's model: the operators are related to signals and the selection rules to behaviour. GOMS is however a "keystroke-level" model and does not include explicitly a knowledge level where a new course of action is planned and created anew. GOMS models can be used to represent perceptual, cognitive, and motor activities; they are commonly used to perform detailed keystroke, error, and performance time analysis.

Norman (1986) indicates seven stages of user activity for the process of performing and evaluating an action:

- establishing the goal
- forming the intention
- specifying the action sequence

gear) and finally at knowledge-based level. The experienced car driver can watch the traffic, select strategies, shift the gear and engage in a conversation at the same time. Skilled musicians do not read one note after the other, but recognize immediately more complex patterns like intervals, scales, arpeggios and execute then accordingly.

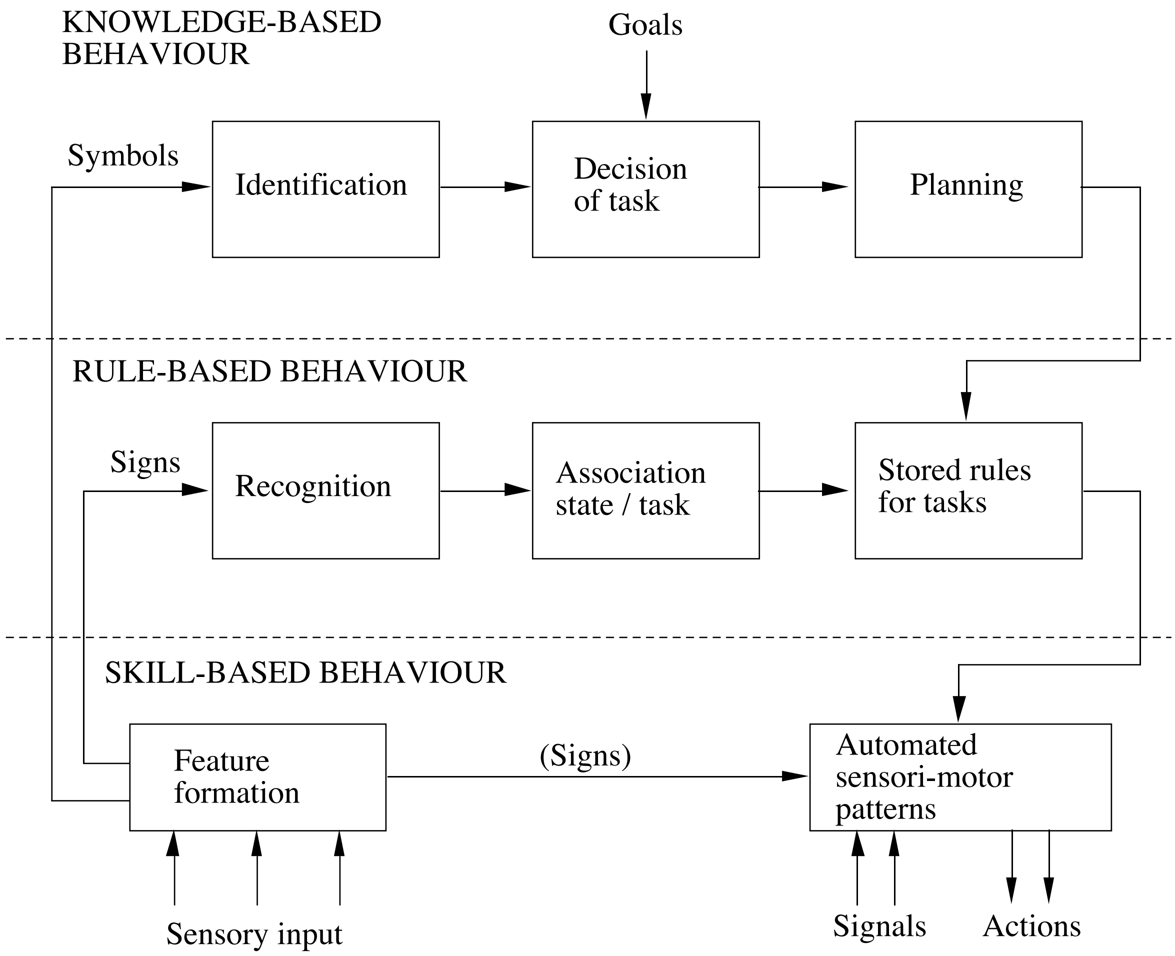

Figure 2.1 Simplified illustration of the three levels of human performance (from Rasmussen, 1983)

Rasmussen differentiates also among the type of communication that takes place at the different levels. At the **skill-based** level the perceptual motor system acts as a multivariable continuous control system. The sensed information is perceived as **signals** that constantly indicate the physical time-space relation to the environment.

In the analysis of human-computer interfaces, the necessity to differentiate among action levels appears early. It is reasonable to assume that typing a few keystrokes is a different action than deciding *what* to type. Almost all behavioural models of the human user show therefore a layered structure, with the physiological functions at the lowest layer and the autonomous thinking and decision-making at the highest. The number of layers varies from one model to another. Models with several layers might not be wrong in principle, but it is difficult to see their utility in practical interface design.

One of the most successful layered models related to the operation of technical systems has been provided by the Dane Jens Rasmussen. According to Rasmussen (1983), there are three layers, for skilled-based behaviour, rule-based behaviour and knowledge-based behaviour (Figure 2.1).

The **skill-based behaviour** represents sensory-motor performance during acts or activities which take place as smooth, automated, and highly integrated patterns of behaviour and without conscious control. At the next level is the **rule-based behaviour**, that takes place in familiar situations and is controlled by **stored rules** or **procedures** which have been learned or derived empirically during previous occasions.

Of course, a person does not face only familiar situations. Under unknown and unfamiliar conditions, when no previous know-how or rules are available, the control of action must take place at a higher level, in which performance is **goal-controlled** and **knowledge-based**. In this situation, the goal is explicitly formulated, based on the analysis of the environment and the overall aims. At this level of functional reasoning, the structure of the environment to act upon is explicitly represented by a mental model on which the selected course of action depends.

To exemplify Rasmussen's model it is sufficient to consider two complex tasks like driving a car or playing a musical instrument. At the beginning, the learner must consciously evaluate every situation, formulate a plan (for shifting the gear or moving the hand on a keyboard) and then carry it out. At this stage, attention is high and efficiency and performance are low. With time and experience, actions are carried out more and more automatically, first at rule-level (e.g. by recognising directly and without a conscious analysis when it is time to shift the

2 User-related Issues in Human-Computer Interaction

This chapter deals with issues related to user psychology. The results reported here are models and experimental findings of interest in the design and evaluation of the user interface for control tasks. In Section 2.1 is introduced Rasmussen's action model; this model is used as reference in the rest of this chapter as well as in other parts of this Thesis. Section 2.2 deals with the basics of perception and human memory. Errors as a fundamental part of human behaviour and error management by the machine are treated in Section 2.3. Section 2.4 deals with the importance of motivation in human performance. The issue of mental models for process control is treated in Section 2.5 and experimental results are presented in Section 2.6. Section 2.7 examines the problem of knowledge and training in relation to mental models.

2.1 A Model for Human Behaviour

In order to better evaluate and analyse the interactions between user, technical process, interface and goals (Section 1.2) it is necessary to have internal models. However, what can be done quite directly with engineering methods for goal, technical process and process computer cannot be done for the user. There have been many attempts to model the human being as if it were a machine, a computer, an "expert system" and the like, but all had to be dropped. The only "physical" modelling of the human being still in use in human-machine interaction contexts is the description of sensomotoric reactions in terms of transfer functions. A model of the human brain and of the thought process is still a long way off, if it ever can be built.

A different approach, that is also much more useful in practice, is to model human behaviour in relation to a work task. In this way, the question of the internals of thought and reasoning can be eluded altogether.