



# Exploration strategies, performance, and error consequences when learning a complex computer task

DIMITRI VAN DER LINDEN, SABINE SONNENTAG, MICHAEL FRESE and CATHY VAN DYCK

Work and Organizational Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands; e-mail: ao\_linden@macmail.psy.uva.nl

**Abstract.** When trying to learn a complex task, people can use different strategies. They can use systematic exploration in which they take on an active approach to discover the computer functions and make use of problem solving steps such as planning, evaluation of feedback, and control of emotion and motivation. Alternatively, they can use non-systematic strategies like trial-and-error, rigid exploration, and encapsulation in information seeking. This study examined whether the exploration strategies were related to error consequences and performance when people learned a new computer program. Strategies were assessed by means of coding. Analysis showed strong correlations between strategies, error consequences, and task performance. These results can have implications for training design and human reliability in dealing with complex devices.

## 1. Introduction

Due to the progress in technology, the nature of work is changing. The large amount of automation in industry has made work more complex. Often, employees are expected to work in complex environments in which they have to search for information and make the right decisions. Performance on such tasks depends on the quality of the exploration strategies that are used. Therefore, there is a growing interest in complex problem solving research and in the modelling of strategies during problem solving or learning (van der Schaaf 1993). Exploration strategies during complex problems can be considered either optimal or sub-optimal. Sub-optimal strategies are, for instance, associated with lack of planning and feedback evaluation. Studies of Dörner and co-workers (e.g. Dörner and

Schaub 1994, Dörner and Wearing 1995) showed how sub-optimal strategies led to poor performance in different computer simulations: Participants who used sub-optimal strategies had less insight into the system or were less able to control the system. The simulations in these studies were all complex dynamic tasks. However, the quality of exploration strategies is also important in learning computer applications or other complex devices. For instance, Trudel and Payne (1996) found that people who used reflective exploration strategies learned more of the functions of a digital stopwatch compared to people who used less reflective strategies.

This paper examines the different strategies people use when learning a complex computer task. In the following sections we will first describe some of the exploration strategies used during complex problem solving. Then we will illustrate how such strategies might be related to performance and error consequences.

### 1.1. Problem solving strategies

Studies in the area of complex problem solving and human reliability have concentrated on describing different strategies people use when solving complex problems. For instance, Dörner and Schaub (1994) differentiated between optimal strategies such as systematic exploration, and sub-optimal strategies such as trial-and-error, rigid exploration (which they called dogmatic entrenchment), and encapsulation in information seeking.

Systematic exploration is an optimal pattern of behaviour that leads to the highest probability of successful task completion (Hollnagel 1993a, Dörner and Schaub 1994). The term 'systematic' refers to the process in which a person forms goals and intentions based on hypotheses about how the task should be done. These hypotheses are tested in a logical manner and their outcomes are evaluated in order to plan further behaviour.

Systematic exploration does not only go hand in hand with the adequate co-ordination of problem solving steps but also with a high level of motivational and emotional control (Dörner and Wearing 1995). The processes responsible for this are referred to as self-regulatory processes (Muraven *et al.* 1998). When solving complex tasks, there is a constant risk of making errors and of information overload. This can lead to high levels of arousal and to a subsequent decrease in performance. Motivational and emotional self-regulation has to deal with this arousal and has to make sure that most cognitive resources stay invested in the task instead of in the intrusive negative cognitions resulting from errors (Frese 1995).

The trial-and-error strategy is assumed to be either less effective or less efficient (Hollnagel 1993a, Dörner and Wearing 1995). Crucial for this strategy is a lack of hypotheses to direct behaviour. Trial-and-error resembles the concept of the scrambled mode described by Hollnagel (1993a: 168) in which 'the choice of the next action is completely unpredictable or random'.

Several possible causes can lead to the use of trial-and-error approaches. Lack of knowledge about a particular problem domain is one of these causes. Without a minimum amount of knowledge, it is very difficult to plan or to get an idea where to search for answers. In such a situation, trial-and-error is the only thing one can do. Another possible cause is perceived loss of control over the task. A high level of perceived loss of control, without a high level of self-regulation can easily lead to the negative effects of cognitive intrusions and negative emotions we referred to earlier. The result will be that less planning and feedback evaluation will be done and trial-and-error behaviour will be displayed.

One type of behaviour that often can be observed during problem solving is rigid exploration (Reason 1990, Hollnagel 1993a, Dörner and Schaub 1994). This type of behaviour is characterized by the repetitive nature of action sequences. People working in a rigid way hold on to certain hypotheses for a long time, even if evidence is accumulating that the hypotheses are wrong. Rigid exploration is closely related to confirmation bias (Nisbett and Ross 1980). If the problem space is too large for complete

exploration or if time is limited, then actions will be guided by the hypothesis that has the highest perceived probability of leading to a solution. Once effort is invested in testing a particular hypothesis, the tendency to search for confirming evidence increases. The processing of disconfirming evidence, on the other hand, decreases (Dörner and Wearing 1995). Rigid exploration might be a direct result of declined self-monitoring during problem solving. Self-monitoring can be considered a meta-cognitive process by which people reflect on their own strategies. If self-monitoring is low, people will not notice their tendencies to repeatedly perform the same action sequence without success.

Another type of behaviour one can show when trying to solve a complex problem is encapsulation in information seeking. If one is not able to solve the problem, it can be useful to search for more information. There is, however, a possibility that the newly acquired knowledge will lead to more questions and more uncertainty. This will result in more information seeking. In the end, the gathering of information becomes the main goal. Consequences of excessive information seeking will be that decisions and actions will either be postponed or will not be taken at all. Thus, task performance will suffer. In literature, the negative impact of encapsulation in information seeking on performance has mainly been described in relation to complex dynamic tasks (e.g. emergency situations) (e.g. Funke 1991, Dörner and Schaub 1994, Dörner and Wearing 1995). However, it is likely that any other task, in which decisions have to be made within a limited amount of time, will also suffer from encapsulation (although the authors leave open the possibility that encapsulation might enhance performance on transfer tasks because the people using 'encapsulation' strategies should have required more declarative knowledge in initial tasks).

Until now, we described several exploration strategies people can use during complex tasks. We also argued that the choice of a particular exploration strategy sometimes depends on cognitions about one's inability to accomplish the task (loss of control). Such cognitions might be considered as negative self-evaluations that are accompanied with lower levels of self-efficacy (Bandura 1997). Negative self-evaluations might also be related to low performance and sub-optimal strategies because when people are giving attentional resources to negative cognitions then less of those resources would be available for working on the main task (Kanfer and Ackerman 1989, Mikulincer 1989). Therefore, we also looked at to which extent the participants of our study got involved in negative evaluations about their own performance and capability.

### 1.2. Consequences of strategy use

Whenever people use non-systematic strategies, they will often make many errors (Hollnagel 1993a, Dörner and Wearing 1995). In safe environments such as training sessions, making many errors is not necessarily dysfunctional as long as one is able to learn from them. However, the problem is that using non-systematic strategies leads to a lower degree of learning (Shrager and Klahr 1986). To learn from an error it is necessary to process the information that the error can provide. Furthermore, an error can lead to negative cognitions and emotions that have to be dealt with in order to prevent a breakdown of skills during the rest of the task (Davis 1948). From our descriptions of systematic versus non-systematic exploration strategies it follows that both the processing of information and the regulation of cognition and emotion are parts of systematic exploration and are almost lacking in non-systematic strategies. Thus, we can argue that non-systematic strategies reduce the amount of learning.

It is expected that the different strategies will be directly related to task performance: In our study, performance is the number of sub-tasks participants learn to solve with a computer program. We expect that systematic exploration will be positively related to performance (which would be displayed by a positive relationship between systematic exploration and number of tasks solved). For non-systematic strategies the relationship with performance is expected to be negative (Hypothesis 1).

### 1.3. Error consequences

Strategies might not only be related to differences in learning performance but also to differences in the number and type of errors. In this section, we concentrate on this relationship between strategies, errors, and error consequences in complex tasks. The literature on human error distinguishes between the error itself and the consequence of the error (Hollnagel 1993b, Frese 1995). It is not the error per se that has a large impact on further behaviour. What matters most are the error consequences. This is a major point in the concept of error management. Frese (1995) suggested that errors are omnipresent in human behaviour. Therefore, one should not only concentrate on the prevention of errors but also on the prevention of negative error consequences (Frese and Zapf 1991).

Error consequences might play a major role in the relationship between strategy-use and performance. Under certain conditions, errors could cause shifts in strategy-use, for example, in a situation in which an

error is made that has a serious negative impact on goal attainment. Such error has the potential to cause frustration, which might rise to such a level that a person might switch from systematic strategies (if that was the strategy that that person was initially using) to non-systematic strategies. On the other hand, it is also possible that someone makes an error (e.g. takes a wrong turn, presses a wrong button) that accidentally leads closer to the intended goal. Such an event would have a direct positive effect on performance as well as a positive influence on mood and possibly also on feelings of self-esteem, which are important for complex problem solving (Dörner and Wearing 1995).

If many things go wrong and people feel that they are losing control over the task it might also happen that feelings of panic and time pressure emerge. According to Davis (1948: 22) this happens when people 'should fail to reach the standard of performance which represents success to them'. Such events cause disorganization in subsequent behaviour resulting in more errors. Again, it is assumed that this disorganization is caused by impaired self-regulation—or stated differently, impaired resistance against negative intrusive thoughts. Therefore, we expect a strong relationship between errors during a task and the number of negative self-evaluations (Hypothesis 2).

Loss of control during a task often goes together with bursts of activism. During this state of activism one will act often and rapidly but in a way that is not very useful for goal attainment (non-systematic strategies). This behavioural syndrome resembles the concept of what Dörner and Wearing (1995) called the *emergency reaction*. A possible explanation for this kind of reaction is that it is a demonstration (to oneself or to the 'outer world') that one is still in control and that one is doing something. This gives at least the illusion of competence (although this behaviour might actually worsen the situation).

Based on the separation between the error itself and the error consequence it is useful to differentiate between three types of error consequences. First, negative consequences for direct goal attainment (e.g. errors that block further effective behaviour or errors that destroy earlier constructive work). Second, non-effective actions: erroneous actions that do not have any direct effect at all (e.g. pushing non-active buttons or other worthless attempts). Finally, actions which are in essence wrong (do not lead to the intended goal) but might lead to insight in solving the task one is working on (positive error consequence).

With regard to strategies one might expect that a trial-and-error strategy in which things are tried out in a disorganized way is related to more negative error consequences and more non-effective actions (Hypothesis 2).

esis 3). Strategies such as rigid exploration and encapsulation in information seeking can be associated with low levels of risk taking. Keeping stuck to pre-formed ideas or excessively gathering information without acting a lot will not lead to much progress during the task and it also prevents one from trying out the unknown area within the problem space. Thus we hypothesize that rigid exploration and encapsulation in information seeking will not be correlated with amount of negative error consequences (Hypothesis 4).

To make the hypotheses from the introduction more explicit we present the expected relationships between the strategy variables on the one hand and performance and error variables on the other hand in table 1. In this table a '+' denotes that we expected a positive relationship, '-' denotes an expected negative relationship, and '0' means we expected no relationship.

## 2. Method

### 2.1. Participants

Nineteen first-year psychology students participated in this study. They received extra study credits for participation. Fourteen participants were female and five were male. Average age was  $M = 21.6$  years ( $SD = 5.5$ ). Study participants differed in the amount of pre-knowledge on computers (see section 2.4 on computer knowledge). We were aware of the relatively small sample size. The analyses we did, were however quite detailed in nature and concentrated mainly on the processes during the task. Furthermore, other studies have shown that analysing such a sample can provide insight into the process of exploratory learning (Schrager and Klahr 1986, Chi *et al.* 1989, Trudel and Payne 1996).

### 2.2. Task

We used a task with the spreadsheet computer program Excel (version 4.0 for Macintosh). This task

consisted of changing the format of a table, which included moving text, adding rows and changing colours and borders. Participants worked on this task within an error-training framework. Error training is a training method in which participants have to work on a task with a minimum amount of information provided and in which they are encouraged to learn from errors (Frese 1995). During such task, one can observe almost all of the systematic and non-systematic strategies.

Participants had to discover most of the features of the program themselves by means of exploration. Because none of the participants had any experience with Excel it was inevitable that they made errors. The task turned out to be very difficult for most of the participants.

### 2.3. Procedure

Before the participants worked on the task they had to fill out the questionnaire on the control variable (computer experience). Then they worked on a short introductory Excel task for five minutes. With this introductory task they could get used to the experimental setting and the thinking aloud method. After the introductory task, participants worked on the formatting task. They were given the table formatting assignment, but no instructions on how to do it. They were told to work on the task for half an hour. Participants got the instruction that they did not necessarily have to finish the entire task within the time but they at least had to try to come 'as far as possible'.

During the task, the experimenter sat out of sight of the participants and observed them. Participants were told to keep on talking aloud when they stayed quiet for longer than five seconds, the experimenter used the standard sentence 'keep talking please', to avoid prompting (Ericsson and Simon 1993). During the task, the computer screen the participants worked on was directly connected to a Video recorder (VCR). Every action on the screen was recorded. A small microphone above the screen, which was also connected to the same VCR, recorded the verbalizations.

Table 1. Expected relationships between strategies, performance and error consequences.

	Performance	Positive error consequences	Negative error consequences	Non-effective actions
Systematic exploration	+	+	—	—
Trial-and-error	—	—	+	+
Rigid exploration	—	—	0	+
Encapsulation	—	—	0	0
Negative self-evaluations	—	—	+	+

## 2.4. Measures

2.4.1. *Performance and strategies*: We categorized participants' strategies by coding their behaviour based on the video recordings. Such a method is often used in error analysis and problem solving research and has shown to be valid and reliable (Hollnagel *et al.* 1981, Frese and Zapf 1991).

2.4.2. *Performance*: Within the Excel task, we differentiated between several sub-tasks, such as border thickening and colour changing. The whole Excel task consisted of eight sub-tasks. Performance was operationalized as the number of sub-tasks solved.

2.4.3. *Strategy-use*: As coding units for strategy assessment, we used actions such as pushing a button in the toolbar or handling a menu. These units were the smallest meaningful actions in our experimental task. Thus, every time a participant pressed a button in the toolbar or opened a menu-option in the menu bar, the raters assessed whether this was done in a systematic or non-systematic way. Furthermore, every time such handling of a button or menu-option led to an error, the raters assessed its consequence. Additional coding was done for statements of negative self-evaluation (frequency of occurrence).

It turned out that most of the participants used several strategies during the task. For example, it could be observed that someone worked in a trial-and-error manner for most of the time, but occasionally used systematic exploration. The frequency with which the participants used a certain strategy was what differentiated between them. For the statistical analysis of the strategies and error consequences we used proportions. These proportions were calculated by dividing—for each participant—the number of actions in a category by the total number of actions of that same participant. Thus, for example if a participant performed 30 actions during the task of which 10 led to negative consequences, this participant would have a proportion of 0.33 for this category.

A second independent rater coded eleven of the nineteen videotapes. For every strategy- and error-consequence category, we computed Pearson correlations between the two raters. Descriptions and some examples of the strategies and error consequences that were coded, as well as interrater correlations are given in table 2. Note that strategies such as systematic exploration cannot be operationalised by standard forms of behaviour such as depth-first or breadth-first strategies (Norman 1991). A menu or button-handling event of a participant was rated systematic if there were signs that

this event handling was part of an intentional plan or was evaluated. Events such as saving intermediate results or using the undo-button to undo previous errors were for instance rated as systematic. However, due to the complexity of the task and many different types of behaviour that can occur on such task it would go too far to list all specific events that were associated with systematic or unsystematic strategies.

2.4.4. *Computer knowledge and general intelligence*: We expected that individuals differ on general computer experience and that this variable might have an impact on performance. For instance, knowledge about features that are shared by a wide range of different computer programs might lead to clear advantages on the task. Although experience might influence the way of working of the participants, in this study we were interested in the theoretical relationships between exploration strategies and performance and error consequences. For this reason we included measurements of general computer experience in our study and used it as a control variable. It could also be possible that general intelligence is related to task performance although we expected that within the restricted range of intelligence that is found in our participants (all were university students), effects of intelligence differences might be overruled by computer experience. Nevertheless, we also include intelligence measures in our study as an additional control variable.

2.4.5. *Computer knowledge*: Computer knowledge was assessed with seven questions (5-point Likert scales) concerning participants' experience with operating systems, and the use of other applications (Reliability  $\alpha = 0.75$ ). Response categories ranged from 'never' (= 1) to 'very often' (= 5). This measure showed that 14 participants had never or seldom worked with a Macintosh computer (score of 1 or 2). Four participants had worked often with a Macintosh computer (score 4 or 5). Most participants were familiar with the Windows Operating System (approximately 60% had a score of 4 or 5).

2.4.6. *Intelligence*: Intelligence was measured with a test that was based on the intelligence model of Guilford (1967). The test consists of six sub-test such as reasoning, perceptual speed, verbal and numerical aptitude. All sub-test are based on established aptitude tests and have high reliability ( $M = 0.83$ , range: 0.80 to 0.85) (see, for example, Elshout and Veenman 1992). We used the scores on the sub-tests to construct a single score that represented general intelligence. Intelligence was measured before our study during a mandatory period in which all first year students had to participate.

Table 2. Category descriptions, examples and interrater correlations.

Category	Description	Examples	r
Systematic exploration	Exploring the system in a reflective manner: testing hypotheses, planning, and evaluation of action outcomes	- Considering a menu-option due to its name (e.g. border). - Analysing a menu-option accompanied with statements that suggest the processing of information available in that option.	0.74
Trial-and-error	Exploring the system without displaying signs of hypotheses testing, planning, or evaluation.	- Randomly pushing buttons or menu-options without sense of direction.	0.87
Rigid exploration	Repetitions of ineffective actions (more than twice), or holding on to faulty ideas longer than 'objectively' necessary.	- Keep on trying the option colour palette (more than twice) to colour a section without (any) effect.	0.97
Encapsulation in information seeking	Excessive search for information	- Staying in the help-menu for a long time and reading every bit of information carefully, irrespective of its use.	0.93
Negative self-evaluations	Statements that reflect perceived loss of control of the task	- 'I really cannot do this task'. - 'I have absolutely no idea what to do anymore'.	0.77
Positive error consequences	Erroneous actions that did not lead to the intended goal but that provide obvious useful information about the task.	- Accidentally changing the colours of a proportion of the table when trying to thicken the border.	0.93
Negative error consequences	Erroneous actions that led to consequences that had a direct negative effect on goal attainment.	- Deleting previous work.	0.75
Non-effective actions	Actions that did not have any consequences.	- Pressing an inactive button	0.83

### 3. Results

Before testing the hypotheses, we analysed whether intelligence and computer knowledge were related to performance. As expected, intelligence was not related to performance on the computer tasks. However, computer knowledge was substantially related to task performance ( $r = 0.40, p = 0.01$ ). To examine the effects of strategies on performance beyond the effects of computer knowledge, we controlled for computer knowledge in all further analyses.

Table 3 presents the average proportions of strategy-use plus the range. Thus the proportion of 0.51 in this table implies that on average about 50% of the actions made were rated as being a part of systematic exploration. Although there were sometimes substantial individual differences, all the participants had a score greater than zero for systematic exploration, trial-and-error, and rigid exploration. Encapsulation was the least frequent kind of behaviour that occurred. Approximately 50% of the participants never showed this behaviour at all. For the error consequence categories, 'positive consequences' was the least occurring event. Approximately 20% of the participants did not have any errors with positive consequences. An average of 33% of all actions did not lead to errors. The intended goal was reached in these actions (e.g. making a text bold).

#### 3.1. Strategies and performance

Analysis showed that systematic exploration was positively related to task performance (see table 4) and that rigid exploration and encapsulation were negatively related to performance. Trial-and-error strategy and negative self-evaluations were not significantly related to performance although there was a negative trend.

#### 3.2. Error consequences and performance

Analysis further showed that error consequences were related to performance. Errors with positive consequences were positively related with performance. Non-effective actions and errors with negative consequences were negatively related to performance. The two correlations did not reach significance although they were in the expected direction.

#### 3.3. Strategies and error consequences

It turned out that the quality of strategy-use was strongly related to the types of error consequence (table 4). Systematic exploration was negatively correlated to negative error consequences and non-effective actions.

Table 3. Participants' mean proportions, standard deviations, percentiles, and highest and lowest values for strategy-use and error consequences ( $n = 19$ ).

	M	SD	25th percentile	75th percentile	Lowest value	Highest value
<b>Strategies</b>						
Systematic exploration	0.51	0.18	0.36	0.70	0.19	0.85
Trial-and-error	0.19	0.17	0.05	0.24	0.03	0.47
Rigid exploration	0.17	0.11	0.09	0.22	0.03	0.43
Encapsulation	0.03	0.05	0.00	0.07	0.00	0.25
<b>Error consequences</b>						
Positive error consequences	0.05	0.08	0.01	0.11	0.00	0.21
Negative error consequences	0.18	0.05	0.09	0.25	0.05	0.40
Non-effective actions	0.44	0.11	0.38	0.51	0.31	0.67
Actions not leading to errors	0.33	0.14	0.21	0.41	0.00	0.48

Note: Proportions for strategies do not add up to 100% because not all actions that were coded could be indexed into one of the categories (rest category for strategies was 13%, there were no correlations with other study variables).

Table 4. Intercorrelation matrix strategies, error consequences, and task performance ( $n = 19$ ).

	M <sup>a</sup>	SD	1	2	3	4	5	6	7	8	9	10
1. Systematic	21.84	7.5	–	–0.62**	–0.60**	–0.57*	–0.49*	0.73**	–0.51*	–0.71**	0.59**	–0.45
2. Trial-and-error	8.11	7.3	–0.63**	–	–0.09	0.05	0.37	–0.39	0.62*	0.56*	–0.23	0.41
3. Rigid exploration	7.26	5.0	–0.65**	–0.02	–	0.61*	0.09	–0.56*	–0.13	0.29	–0.52*	0.02
4. Encapsulation	1.42	2.2	–0.60*	0.11	0.60*	–	0.02	–0.47*	–0.19	0.14	–0.60**	–0.15
5. Negative evaluation	2.72	3.2	–0.59*	0.52*	0.03	–0.01	–	–0.57*	0.41	0.80**	–0.29	0.70**
6. Positive consequence error	2.17	1.9	0.77**	–0.37	–0.61*	–0.51*	–0.56*	–	–0.42	–0.81**	0.65**	0.60*
7. Negative consequence error	7.44	4.6	–0.51*	0.61*	0.17	–0.18	0.43**	–0.37	–	0.51*	–0.29	0.52*
8. Non effective action	18.83	6.5	–0.73**	0.56*	0.32	0.16	0.85**	0.79**	0.48	–	–0.40	0.85**
9. Task performance <sup>b</sup>	3.50	2.6	0.68**	–0.09	–0.68**	–0.77**	–0.24	0.60*	–0.33	–0.22	–	–0.02
10. Number of actions	42.47	12.9	–0.45	0.44	0.05	–0.14	0.93**	0.75**	0.95**	0.60*	–0.13	–

Note: Correlations above the diagonal are zero-order correlations. Correlations below the diagonal are the partial correlations (controlled for general computer knowledge). <sup>a</sup>Means denote the averages of the participants. <sup>b</sup>Range 0–8. \* $p < 0.05$ , \*\* $p < 0.01$ .

Systematic exploration was positively correlated with positive error consequences, which can be considered as events that often lead the participant to a solution. In contrast, the trial-and-error strategy was positively correlated with negative error consequences and non-effective actions and negatively correlated with positive error consequences (although not significantly). Rigid exploration and encapsulation in information seeking both were negatively correlated with positive error consequences. There were no significant relationships between these non-systematic strategies and negative error consequences.

Negative self-evaluations correlated negatively with positive error consequences and positively with negative error consequences and non-effective actions. Notable is the very high correlation we found between the total number of actions and negative self-evaluations ( $r = 0.70$ ,  $p < 0.01$  for the correlations corrected for computer knowledge and  $r = 0.93$ ,  $p < 0.01$  for the

simple correlations). This can be an indication of an emergency reaction (Dörner and Wearing 1995) from the participants who performed poorly.

#### 4. Discussion

Our study confirmed our hypothesis that the use of systematic exploration was related to better task performance (Hypothesis 1). Participants who mainly used systematic exploration formed hypotheses about the system. This indicated deliberate efforts in trying to understand the system they were working on. They also tested these hypotheses and evaluated the outcomes of such tests in order to direct further behaviour, which suggests a higher level of planning. Moreover, they reflected more on their own actions and changed their ways of working if necessary. By using this kind of exploration they learned more about the functions of the

program and were thus better equipped to deal with the complex problem.

The pattern of error consequences related to using systematic exploration (more positive errors, less negative errors) confirmed our expectation that, in terms of human reliability, systematic exploration is the most desired kind of behaviour to show when facing complex problems. Furthermore, the results showed that it is useful to focus on specific error consequences in error analysis instead of strictly rating errors as not reaching the goal that preceded the act.

In contrast to our expectation (the second part of Hypothesis 1), trial-and-error did not have a significant negative impact on task performance. Apparently, by trying out different things with trial-and-error one can probably still detect many of the computer functions needed to solve a task. However, trial and error behaviour was associated with more errors with negative consequences and non-effective actions (Hypothesis 3). Thus, there are costs of using a trial-and-error strategy.

Rigid exploration and encapsulation in information seeking both showed similar patterns with regard to performance and error consequences. Use of these strategies was associated with low performance (Hypothesis 1) and a low amount of errors with positive consequences. More rigidity during exploration or encapsulation in information seeking did however not go together with more errors with negative consequences or ineffective actions. This was in accordance with our hypothesis (Hypothesis 4).

Observations during the study showed that participants who got engaged in rigid exploration seemed to be driven by salient information that was available in the program (e.g. menu names or pictograms on buttons). Furthermore, these observations showed that even after several ineffective attempts they held on to this salient information and were reluctant to explore new possibilities, maybe due to uncertainty or to lack of further ideas. These observations have implications for system design because it suggests that under certain circumstances, inadequate menu naming (or button pictograms) can yield rigidity. For instance the use of the terms of background and foreground under the format-option: in attempts to change the colour of the Excel-cells, many participants tried out the option background, which initially looked like the right option. Rigid participants either took very long switching to the counter-intuitive option foreground (which actually was the right option), or did not try this option at all (and subsequently failed to solve that particular sub-task).

It is obvious how such rigidity can lead to low performance: if one does not try out different possibilities fewer options will be discovered. Furthermore, one can imagine that limited exploration also does not lead

to many errors because fewer difficulties will be encountered.

Similar arguments can be used for the relationships between encapsulation on the one hand and errors and performance on the other hand. During encapsulation, participants spent a lot of time gathering all kinds of information without acting a lot. According to the literature, this behaviour occurs if people are overwhelmed by the complexity of the task (Dörner and Scholkopf 1991, Dörner and Wearing 1995). People who become encapsulated in information seeking do not wish to take much risk, hence they will not make many errors. The downside of this behaviour is that encapsulation is a relatively passive state while exploration requires an active pursuit for solutions. Therefore, one will not discover much of the program's functions and subsequently show low performance.

Low performance, many errors with negative consequences, and many non-effective actions also went together with negative self-evaluations (Hypothesis 2). Due to the cross-sectional nature of this study we cannot draw any causal conclusion. Nevertheless, observation during the study suggested that negative self-evaluations often occurred after experiencing negative error consequences or after relatively long series of non-effective actions. It seemed as if negative self-evaluations were a reaction to things that went wrong during the task. The high correlation between negative self-evaluations and the total number of actions is noteworthy (table 4). We believe that this relationship between number of negative self-evaluations and number of actions reflects the emergency reaction in which one starts getting extremely activated and tries out many things which might not be very effective or useful (Dörner and Wearing 1995). However, this behaviour gives the person the sense of doing something, thus protecting the person's self-esteem that is essential for solving complex problems.

If lack of success is indeed an important factor contributing to the emergency reaction, then one can imagine that achieving some intermediary success should make it possible to escape from the emergency reaction. This is in accordance with findings of Hollnagel (1993a: 188) who reported that 'a continued series of failures is likely to reduce subjectively available time and make the situation more tense while a continuous series of successes is likely to improve the situation in every way'.

#### 4.1. *Implications for further research*

The strategy-classifications used in this study were mainly based on theories on complex problem solving



(CPS) from Dörner and co-workers (e.g. Dörner and Schaub 1994, Dörner and Wearing 1995). Most of the research on this has been done with computer simulations of complex tasks such as fire fighting or controlling a city (Funke 1991). In the present study we applied this particular theory of complex problem solving to human-computer interaction with an existing software program (Excel) and showed that the impact of strategies derived from CPS-theory can be generalised to this area.

Further studies might be done to find out how strategies and performance change during the period of working on the task. For instance, it might be of interest to find out whether negative error consequences and success can indeed lead to shifts in strategy-use. Besides this it might be illuminating to more specifically look at the individual steps of the action process of which we assumed to underlie the strategies we studied. This might show that differences in strategy-use mainly reflect self-regulatory processes. Such processes are assumed to be of major importance for regulating self-motivation and negative emotions. Furthermore, several authors have proposed that self-regulation is the mechanism with which a person allocates attentional resources and co-ordinates different cognitive steps in the problem solving process (such as planning) (Kanfer and Ackerman 1989, Farr *et al.* 1993, Frese and Zapf 1994). The idea of a single construct underlying strategies is supported by the high inter-correlations between the strategies (see table 4). Further studies might be done that put more emphasis on the relationship between self-regulation and problem-solving strategies.

#### 4.2. Consequences for practice

Because of the increase of technical interactive devices around us and the increase in complexity of our environment, it is important to know how people try to get insight into complex situations by means of exploration. We showed how people can use different strategies during such exploration and how these strategies relate to error consequences and performance. For training purposes this kind of information can be used to create specific instructions that will guide a trainee to adopt more systematic, and fewer non-systematic strategies. For instance, by providing trainees with abstract information about steps that have to be taken during problem solving. Such instructions should stress the importance of action sequences in which one tries to come up with hypothesis, make plans to test these hypotheses, and evaluate their outcomes. Furthermore, such instructions should also stress the importance of active control of emotions and motivation by

keeping attention focused on the task instead of on intrusive negative thoughts.

Information about exploration strategies might not only be useful for training purposes but also for the analysis of human reliability. The use of trial-and-error during critical situations in high-risk industries can lead to disastrous outcomes. Rigidity can lead to operators spending much precious time on faulty options before they become aware that they are on the wrong track. For developing procedures to enhance the safety of human behaviour it is necessary that future research identifies the different types of strategies during complex tasks including their causes and their consequences.

#### Acknowledgements

This study was supported by a grant from the national Dutch NWO-project 'Mental fatigue at Work' (project number 580-02.103)

#### References

- BANDURA A. 1997, *Self-efficacy: The exercise of control* (New York: Freeman).
- CHI, M. T. H., BASSOK, M. LEWIS, M. W., REIMANN, P. and GLASER, R. 1989, Self-explanations: how students study and use examples in learning to solve problems. *Cognitive Science*, **13**, 145–182.
- DAVIS, D. R. 1948, Increase in strength of a secondary drive as a cause of disorganization. *Quarterly Journal of Experimental Psychology*, **1**, 22–28.
- DÖRNER, D. and SCHAUB, H. 1994, Errors in planning and decision-making and the nature of human information processing. Special Issue: Errors, error detection, and error recovery. *Applied Psychology: An International Review*, **43**, 433–453.
- DÖRNER, D. and SCHOLKOPF, J. 1991, Controlling complex systems: or expertise as 'grandmother's know-how'. In K. A. Ericsson and J. Smith (eds) *Toward a General Theory of Expertise: Prospects and Limits* (New York: Cambridge University Press), pp. 218–239.
- DÖRNER, D. and WEARING, A. J. 1995, Complex problem solving: toward a (computersimulated) theory. In P. A. Frensch and J. Funke (eds) *Complex Problem Solving: the European Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 65–99.
- ELSHOUT, J. J. and VEENMAN, M. V. 1992, Relation between intellectual ability and working method as predictors of learning. *Journal of Educational Research*, **85**, 134–143.
- ERICSSON, K. A. and SIMON, H. A. 1993, *Protocol Analysis: Verbal Reports as Data*, revised edn (Massachusetts: The MIT Press).
- FARR, J. L., HOFMANN, D. A. and RINGENBACH, K. L. 1993, Goal orientation and action regulation theory: Implication for Industrial and Organizational psychology. *International Review of Industrial and Organizational Psychology*, **8**, 193–232.

- FRESE, M. 1995, Error management in training: conceptual and empirical results. In C. Zuccheromaglio, S. Bagnara and S. U. Stuckey (eds) *Organizational Learning and Technological Change* (Berlin: Springer), pp. 112–124.
- FRESE, M. and ALTMAN, A. 1989, The treatment of errors in learning and training. In L. Bainbridge and S. A. R. Quintanilla (eds) *Developing Skills with Information Technology* (New York: Wiley), pp. 65–86.
- FRESE, M. and ZAPF, D. 1991, *Fehler bei der Arbeit mit dem Computer: Ergebnisse von Beobachtungen und Befragungen im Bürobereich* (Bern: Hans Huber).
- FRESE, M. and ZAPF, D. 1994, Action as the core of work psychology: a German approach. In H. C. Triandis, M. D. Dunnette and J. M. Hough (eds) *Handbook of Industrial and Organizational Psychology*, 2nd edn (Palo Alto: Consulting Psychology Press), pp. 271–340.
- FUNKE, J. 1991, Solving complex problems: exploration and control of complex systems. In R. J. Sternberg and P. A. Frensch (eds) *Complex Problem Solving: Principles and Mechanisms* (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 185–222.
- GUILFORD, J. P. 1967, *The Nature of Human Intelligence* (New York, McGraw-Hill).
- HOLLNAGEL, E. 1993a, *Human Reliability Analysis: Context and Control* (London: Academic Press).
- HOLLNAGEL, E. 1993b, The phenotype of erroneous actions. *International Journal of Man Machine Studies*, **39**, 1–32.
- HOLLNAGEL, E., PEDERSEN, O. M. and RASMUSSEN, J. 1981, Notes on human performance analysis (Risø-M-2285), Electronics Department, Risø National Laboratory, Roskilde, Denmark.
- IVANCIC, K. 1997, Errors as means of promoting transfer of training. Unpublished Dissertation, University of New South Wales.
- KANFER, R. and ACKERMAN, P. L. 1989, Motivation and cognitive abilities: an integrative/aptitude-treatment interaction approach. *Journal of Applied Psychology*, **74**, 657–690.
- MIKULINER, M. 1989, Cognitive interference and learned helplessness: the effects of off-task cognitions on performance following unsolvable problems. *Journal of Personality and Social Psychology*, **57**, 129–135.
- MURAVEN, M., TICE, D. M. and BAUMEISTER, R. F. 1998, Self-control as limited resource: regulatory depletion patterns. *Journal of Personality and Social Psychology*, **74**, 774–789.
- NISBETT, R. and ROSS, L. 1980, *Human Inference: Strategies and Shortcomings of Social Judgement* (Englewood Cliffs, NJ: Prentice Hall).
- NORMAN, K. L. 1991, *The Psychology of Menu-Selection: Designing for the Cognitive Control of the Human/Computer Interface* (Norwood, NJ: Ablex).
- QUINONES, M. A. and EHRENSTEIN, A. 1996, *Training for a Rapidly Changing Workplace: Applications of Psychological Research* (Washington, DC: American Psychologist Association).
- REASON, J. T. 1990, *Human Error* (Cambridge: Cambridge University Press).
- SHRAGER, J. and KLAHR, D. 1986, Instructionless learning about a complex device: the paradigm and observations. *International Journal of Man Machine Studies*, **25**, 153–189.
- TRUDEL, C. I. and PAYNE, S. J. 1996, Self-monitoring during exploration of an interactive device. *International Journal of Human Computer Studies*, **45**, 723–747.
- VAN DER SCHAAF, T. W. 1993, Developing and using cognitive task typologies. *Ergonomics*, **36**, 1439–1444.