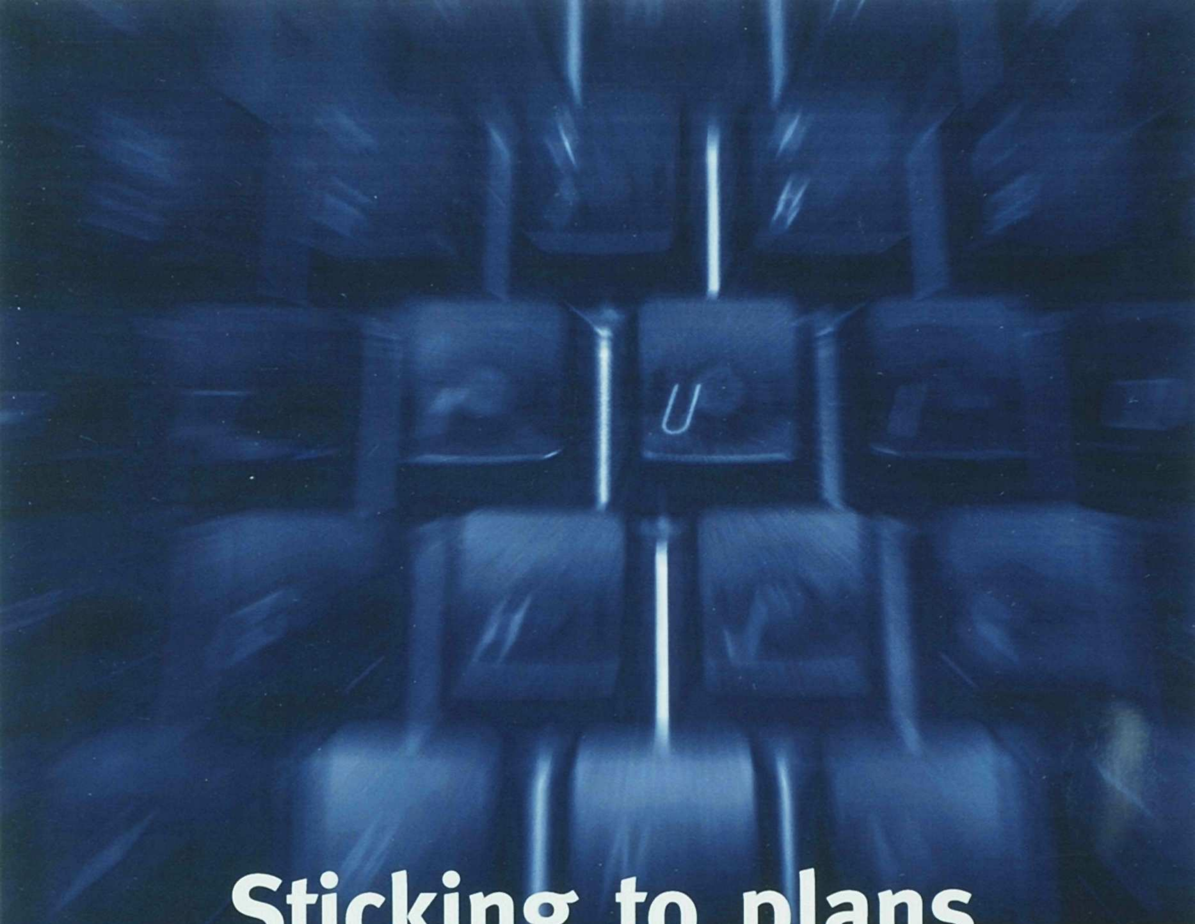


Sticking to plans

Capacity limitation or decision-making bias?

Gerard Meij

STICKING TO PLANS:
CAPACITY LIMITATION OR DECISION-MAKING BIAS?

The research described in this thesis was performed at:
TNO Human Factors Research Institute, Soesterberg

STICKING TO PLANS:
CAPACITY LIMITATION OR DECISION-MAKING BIAS?

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der universiteit
op dinsdag 27 januari 2004, te 10.00 uur

door

Gerardus Meij

geboren te Monnickendam

Promotiecommissie

Promotor:
Prof. dr. J.G.W. Raaijmakers

Co-promotor:
Dr. J.H. Kerstholt

Prof. dr. C.K.W. de Dreu
Dr. P. Koele
Prof. dr. J. van der Pligt
Dr. P.H.M.P. Roelofsma
Prof. dr. M. Zeelenberg

Faculteit der Maatschappij- en
Gedragswetenschappen

Voor mijn vader en moeder

Contents

1. A theoretical approach to cognitive lockup

The first chapter is an attempt to identify possible explanations for cognitive lockup: the tendency to focus on a subpart of a system while ignoring the rest of it. In order to exemplify this phenomenon we first describe the case of flight 401, an example of an airplane crash which was caused by cognitive lockup. Next, we briefly discuss two studies on cognitive lockup that demonstrated the phenomenon in an experimental setting. Since no theoretical explanation for cognitive lockup has been provided to date, we explore three research paradigm where phenomena similar to cognitive lockup have been found: planning, task-switching and decision-making. For each paradigm we discuss the main results with respect to these similar phenomena and we try to identify the explanations that have been provided. These explanations are summed up at the end of this chapter.

Flight 401 of Eastern Air Lines

In 1972 a dramatic plane crash took place. During the landing, the pilot of flight 401 of Eastern Air Lines is warned about a problem with the landing gear. To win time, the pilot cancels the landing, sets the plane in the autopilot mode and starts solving the problem with the landing gear. This problem fully occupies the pilot and multiple warnings about a decreasing altitude (a low-altitude alarm, a remark of the air-traffic controller) are ignored. As a consequence, the plane crashes, resulting in the death of most people aboard.

The National Safety Transport Board concluded that the accident was due to a "preoccupation with a technical malfunction". However, it is not clear what caused this preoccupation. Why did the pilot hold on to solving the problem with the landing gear in spite of several warnings? Or, to reframe this question in more psychological terms, which mechanisms caused the pilot to be captured in a problem-solving task?

A fashionable concept in this respect is that of Situation Awareness. Situation Awareness (SA) is a stage in information processing immediately preceding the stage of decision making. Endsley (1988, 1995) used the following definition of SA: "SA is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988, p. 97). And, because control actions are taken automatically, operators are less aware of the actual situation, which becomes problematic when critical situations occur. In such non-normal, and mostly time-critical, events human intervention is required. But because operators lack knowledge of the details of the system, they are not able to cope with these critical situations.

The concept of SA is fairly broad however. It doesn't tell us which element of the SA – perception, comprehension or projection – contributes to people being captured in a task. So, it remains a question which psychological mechanism caused the pilot of flight 401 to stick to the landing gear problem.

One possible reason why the pilot did not attend to the descending altitude is that he just did not notice the warnings. Maybe he was so caught up in the problem with the landing gear that signals like alarms simply did not reach his attention.

Another reason for sticking to the relative less urgent problem with the landing gear is that the pilot did notice the alarms but was incapable to provide attention to it. The task of solving the landing gear problem may have required so much of his attention that there was nothing left for (switching to) the problem with the altitude.

A final reason for the pilot to stick to the problem with the landing gear could be that he actually made a decision to do so. He decided to continue working on the malfunctioning of the landing gear and to postpone solving the problem with the altitude.

The example of the air crash of flight 401 is not an isolated case; there are multiple examples, especially in aviation and shipping, where accidents happened because operators solely concentrated on a particular subtask, thereby ignoring the rest of the system. Though the consequences of focussing on a subpart of the system can be disastrous, experimental research on this phenomenon has been confined to mere demonstrations of the existence of it.

In the next section we will briefly discuss two experimental studies where being locked up was manifested in the data.

Experimental studies of cognitive lockup

Moray and Rotenberg (1989)

Moray and Rotenberg (1989) were the first to use the term "cognitive lockup" for operators' tendency to deal with disturbances sequentially. As this term reflects the problem adequately, we will use this term throughout this thesis.

Moray and Rotenberg (1998) asked participants to supervise a simulated thermal hydraulic system that consisted of four subsystems. There were two critical conditions: in one condition a single fault appeared in a subsystem and in the other condition a fault in one subsystem was followed by a fault in another subsystem. Eye-movement recordings showed that for the first (and in half of the cases only) fault, attention was given to the corresponding subsystem within 30 seconds, but for the second fault, attention started to shift only 45 seconds after occurrence. According to the authors a first fault absorbs the operators' attention to such a degree that they cannot spend time on the rest of the system.

Kerstholt, Passenier, Houttuin and Schuffel (1996)

Another demonstration of cognitive lockup comes from a study by Kerstholt, Passenier, Houttuin and Schuffel (1996). Participants in their study were required to supervise four dynamic subsystems (navigation, electricity, propulsion and cargo) and to deal with disturbances whenever they occurred. The system included an option to stabilise a subsystem in which an additional fault occurred. Some participants used this option, thereby acknowledging their understanding of the development of a disturbance over time. Most participants, however, handled the disturbances sequentially: full attention was given to the first disturbance(s) and subsequent disturbances were ignored.

These experimental studies are clear demonstrations of operators being locked up in a subpart of the system. However, as we mentioned earlier, none of these studies have produced a theoretical explanation for this phenomenon. Why do operators hold on to solving a minor problem in a subpart of the system while a major problem is evolving somewhere else in system?

In generating possible explanations for cognitive lockup it is useful to review some related phenomena. Cognitive lockup seems a relatively new phenomenon coinciding with the rise of automation but the tendency to stick to a course of action is far from new. It has been identified in a number of adjacent fields, namely planning, task-switching and decision making. In the following sections we will discuss results from these lines of research.

Planning

Planning is a complex process that has been defined in different ways by different investigators. Nevertheless most definitions comprise the following two features: (1) scheduling a series of actions in order to (2) achieve a certain goal. Read (1987), for example, defined planning as the selection and organization of actions to attain certain goals. In the same line, McDermott (1978) conceives planning as the identification and organization of subtasks to execute a problem solution.

These definitions do not recognize the need to revise a plan during execution when environmental changes occur. More recent articles on planning use definitions that do take into account the notion of a changing environment. In their article, Mumford, Schultz, Van Doorn and Judy (2001) provide an elaborate review of the scarce literature on planning. They argue that the early conception of planning provides little room for the adaptive flexibility that seems to characterize most high-level performance. For that reason they define planning as "the mental simulations of actions in a dynamic environment" (p. 214).

In their route planning experiment, Hayes-Roth and Hayes-Roth (1979) acknowledged the importance of simulating the execution of a plan mentally and indicated how simulations can be used to guide subsequent planning. In fact, most recent definitions of planning follow the Hayes-Roth and Hayes-Roth (1979) idea, that planning involves a mental simulation of future action sequences, intended to direct action and optimise the attainment of certain goals (e.g. Berger, Karol and Jordan, 1989; Patalano and Seifert, 1997). Plans are generated and adjusted in interaction with the environment, though the environment in the Hayes-Roth and Hayes-Roth study (a map) remains static throughout the task.

The notion of a dynamic environment in the process of planning is taken into account only recently (Bainbridge, 1997; Brichin and Rachadzo, 1995, and Hine and Gifford, 1997). Hine and Gifford (1997) for example inspected how participants' strategy changed over the course of a simulated common dilemma involving harvest decisions. During the simulation, harvesters had to make a choice at several points in time. After each choice they were provided with feedback of other, fictitious harvesters participating in the simulation. As a result of this repeating interaction with fellow harvesters, the environment is under constant change.

How can cognitive lockup be related to research in the planning paradigm? First of all, people possibly make their initial planning in too much detail. Secondly, people may not be sufficiently susceptible to emerging environmental changes. And, third, people may go through the process of mental simulation too restrictively, leaving individuals unprepared for contingencies.

In the next section we will elaborate on how these three planning errors can result in cognitive lockup.

Detailed planning

Simon and Galotti (1992) suggested that successful planners not only organize their activities but also maintain flexibility in their activity organizations. They appeared to be more sensitive to goal priority than poor planners. Goldin and Hayes-Roth (1980) provided further empirical support for the importance of flexibility. They found that successful planners tended to avoid early commitments to detailed action plans and specific goals. In contrast to unsuccessful planners they first explored the task environment. Also Berg, Strough, Calderone, Sansone and Weir (1998) attribute successful planning to a flexible and adaptive use of models.

Kleinmutz and Thomas (1987) and Kerstholt (1994, 1995) also showed that people do not always exercise flexibility in the selection of strategy. They found that in general participants use a judgement-oriented strategy - a strategy to reduce uncertainty by requesting information - even in situations where an action-oriented strategy - a strategy where people apply actions and react on the observed effects - would have resulted in better performance.

Mumford et al. (2001) argued in this respect for the use of *midrange models*. Midrange models "provide some general direction, albeit direction appropriate to the situation at hand" (p.233). Plan models should neither be highly abstract or overly detailed. Successful planners refrain from committing themselves to detailed action planning and specific goals. They rather explore the context of the task and set up a global initial plan. During its evolvement, the plan can be filled up and adjusted on the basis of environmental cues. This type of planning is also known as "opportunistic planning".

In the planning task of setting out a route to run a set of errands, Hayes-Roth and Hayes-Roth (1979) found that participants, while planning a route according to an initial strategy, noticed opportunities to achieve other goals. For example, a participant planning to proceed toward the bank to satisfy a high- priority goal noticed that her planned route passed by the dry cleaning shop where she needed to pick up a cleaning order. Her initial strategy to satisfy high-priority goals first, was altered into satisfying proximate goals. People, they conclude, modify current planning to take advantage of unforeseen opportunities.

Patalano and Seifert (1997) elaborated on the recognition of opportunities. In their experimental task, participants were presented with a set of goals. They had to imagine being left alone in a friend's dormitory room for a short period of time, during which these goals should be reached. An example of one of the goals is the following:

> *You go to put your hair in a ponytail with Chris's favourite elastic ponytail holder. The ponytail holder snaps out of your hand and flies across the room. It lands atop her bookshelf, too high for you to reach. You cannot stand on the furniture to reach it because dorm furniture is very unstable. But you need to retrieve the elastic band before Chris returns.*

The presented goals were accompanied by a plan. For the example above, this plan was to use a set of encyclopaedias. In a later stage of the experiment, participants were presented with cue objects available in the dormitory room. Participants were asked to record next to each cue any of the goals from the earlier phase that came to mind.

Cues that were identical to objects studied in the goal study phase (e.g. a set of encyclopaedias) resulted in a greater number of remindings than cues that were based on the same abstract plan as the cues in the goal study phase (e.g. a trash can). Apparently, the authors conclude, "participants did not always encode plans at an abstract level" (p.20), which is suboptimal since the retrieval of goals did not always occur when later opportunities to achieve them arose. However, instructions to encode a plan at a more abstract level resulted in recognition of a wider range of opportunities.

To conclude, studies on planning and flexibility show that people are rather inflexible in selecting an accurate strategy or model. In a very early stage they select a model that is too concrete, which hinders them in recognizing opportunities in the environment.

Neglecting to monitor the environment

Good planning requires a general rather than a detailed approach, leaving open the opportunity to adjust or fill out the original plan. This implies that good planners must have the ability for scan the environment on emerging opportunities. Eisenhardt (1989) did an exploratory study on how executives in a rapidly changing environment (the computer industry in the eighties) made decisions. She found that effective decision makers made more use of real-time information (e.g. a sudden profit drop) than less effective decision makers who, on the contrary, based their decisions on long term information (forecasts and trends).

Another observation in Eisenhardt's study was that decision making was more effective when more alternative plans were considered. That there may be an interaction between the availability of plans and monitoring behavior may be inferred from the work by Oswald, Mossholder and Harris (1997). They examined how the availability of plans influenced managers' perception and reaction to their environment and found that the disposal of multiple plans made them more aware of their relative strengths and weaknesses. This way the environment was monitored more analytically and individuals were more prone to identify relevant information. In the same vein, Thomas and McDaniel (1990) found that plan availability influenced the range and relevance of the environmental information being considered.

Xiao, Milgram and Doyle (1997) also stressed the importance of monitoring the environment. They found that an active and direct scanning of the environment stimulated the development of contingency plans. Data from a field study with anaesthesiologists revealed that options changed during examination of a patient's anatomy. For example, the placement of a particular transducer during open-heart surgery places anaesthetists for different potential problems and solutions than during brain surgery. In open-heart surgery the transducer can be placed from inside the patient's chest and reviewing potential problems in this particular case engendered the option of asking the surgeon for assistance. This option does not exist in the case of brain surgery.

In sum, there seems to be an interaction between monitoring the environment and the availability of plans. Direct monitoring of the environment can draw one's attention to emerging opportunities and stimulates the development of contingency plans. The availability of multiple plans, on the other hand, seems to affect the way the environment is scanned.

Lack of mental simulation

As noted earlier, Hayes-Roth and Hayes-Roth (1979) were the first ones who acknowledged the importance of mental simulation. They argued that people mentally simulate the execution of a plan and use the results of this simulation to guide subsequent planning. Mental simulation can be either time-driven or event-driven. In other words, one can simulate walking through a sequence of time units, or one can simulate a plan by mentally moving from one situation to another.

Support for the use of mental simulation in planning was provided by, amongst others, Xiao, Milgram and Doyle (1997). They asked anaesthetists to think aloud during surgical problems. It was found that anaesthetists anticipated potential problems by preparing both physically (assembling and arranging materials needed) and mentally ("if blood pressure rises quickly then..."). In their study, McLenan and Omodei (1996) also found evidence for the use of mental simulation and how this process could be beneficial for planning performance. Fire officers were asked to report all kinds of things that occurred between receipt of a fire call and the beginning of the fire fighting operation. Analyses of their reports revealed that information prior to an operation (e.g. type of structure involved) is used to mentally simulate possible situations and actions. For example, a fire officer who has attended a fire at a particular site will review this experience while travelling to the next similar fire. During this process of mental simulation potential pre-existing situations are activated so that when an actual incident is encountered those preprimed situations are inspected first.

During mental simulation, one generally imagines doing a sequence of actions in a period of time. A finding in the study by Hayes-Roth (1981) and Hayes-Roth and Hayes-Roth (1979) is that participants typically overestimate how much they can accomplish in a given time period. Kahneman and Tversky (1979) termed this finding the "planning fallacy" referring to the tendency to hold a confident belief that one's own project will proceed as planned, even while knowing that the vast majority of similar projects have run late.

Cognitive lockup may be regarded as a result of the underestimation of predicted time. People may structurally underestimate the time needed to solve problems. At the time a second disturbance occurs, people may think they have more time to solve the subsequent disturbance than they actually have, even in the worst case when there is only limited time for handling the second disturbance. They may think they have enough time to complete the first disturbance first and then to solve the second one.

Following the line of Hayes-Roth and Hayes-Roth (1979), Buehler, Griffin and Ross (1994) drew the conclusion that people anticipate that they will finish their own project earlier than they actually do. In their study, university students were asked to estimate the time needed for academic tasks (e.g. completing a thesis) and non-academic tasks (e.g. cleaning one's department). The overall finding was that when they actually performed these tasks, a majority of students did not finish the task within the predicted time.

What psychological mechanism underlies this bias? In the aforementioned study by Buehler et al. (1994), it was stated that the planning fallacy results from an internal focus on predictions of success. When asked to predict time necessary for the completion of a task, people are focussed on the future rather than the past. This future orientation may prevent them from looking backward in time. When they do consider the past, the authors argued, they are usually focussed on previous occasions that justify their optimism. They rarely mention past experiences involving difficulties or delays.

An 'internal' approach on the predictions implies that forecasters apply data from the specific data at hand rather than the distributions of outcomes in similar cases (Kahneman and Tversky, 1979). By neglecting distributional information, people might underestimate the time to complete tasks, even when they have considerable experience.

The planning fallacy appears to be a persistent phenomenon. Several techniques to debias the planning fallacy failed (Byram, 1997; Newby-Clark, Ross, Koehler, Buehler and Griffin, 2000). In a series of experiments, Byram (1997) asked participants to predict how long they thought it would take to assemble a computer stand. He tried to break through people's tendency to underestimate completion times by testing several debiasing techniques. One such technique is decomposition. An explanation for time underestimation is that people tend to look at the task holistically, thereby ignoring task components. By subdividing a task and predicting the time to complete each component, the net prediction should be based on more information than a single prediction. Decomposing the computer stand into three components (a computer table, a key-board tray and a monitor stand) to predict the time for each component did not result in a more accurate net prediction of completion time.

Both Byram (1997) and Newby-Clark (2000) tested another debiasing technique: multiple scenarios. This technique takes advantage of the outcome of the study by Buehler et al. (1994) that when considering past experiences, people focused on previous occasions that justify their optimistic view of completing a task within time. In the multiple scenarios technique participants were asked to elicit predictions for alternative scenarios (optimistic, best guess, pessimistic). Each scenario contained different information and forced consideration of events that might otherwise be ignored. However, neither of the studies found that generation of alternative scenarios reduced the bias to underestimate completion times.

According to Newby-Clark et al. (2000) participants are motivated to disregard pessimistic scenarios of the future. They very much want to believe that they will successfully achieve their goals. Evidence for this notion was provided by an experiment in which this motivational aspect was excluded by letting participants predict someone else's completion times. In that case participants were able to take into account pessimistic scenarios.

Byram (1997) provided additional support for a motivational explanation. In his series of studies on debiasing techniques he found evidence for only one of them: financial incentives. A group of participants who were given explicit incentives for speed before making their prediction- the faster they finished, the more money they would receive- gave shorter predictions for an origami folding task than a group of participants that were not given those incentives. The actual performance time was the same for both groups.

To conclude, mental simulation seems to be a beneficial process for planning. However, people appear to be biased in their simulations of future events; they generally underestimate the time they need to fulfil a sequence of tasks.

In sum, from a planning point of view, cognitive lockup can be explained in three different ways: (1) people commit themselves in an early stage to a detailed plan, which prevents them from recognizing opportunities in the environment, (2) people refrain from a direct and active monitoring of the environment which hinders them in contingency planning, and (3) during the process of mental simulation people generate scenarios that are too optimistic with respect to completion time.

Task Switching

Hitherto, cognitive lockup has been formulated in terms of holding on to a task or sticking to a problem. In terms of the task-switching paradigm, cognitive lockup can be considered as a reluctance to switch to an alternative task or problem. In this respect, a review of task-switching studies can provide insight into why people refrain from switching.

Most studies on task switching are simple reaction-time experiments where the level of control is low. Typical tasks in this branch of studies require participants to respond as quickly as possible - mostly in a manual way - to a simple (visual) stimulus, such as a letter or a symbol. Studies that describe task switching on a higher level of control are quite scarce (Gillie and Broadbent, 1989; Schiffman and Greist-Bousquest, 1992 and Zijlstra, Roe, Leonara and Krediet, 1999). Tasks in these studies are more complex and instead of task-switching people have to deal with interruptions.

For simple reaction time studies, Rubenstein, Meyer and Evans (2001) distinguished two kinds of task switching: task switching in successive tasks and task switching in concurrent tasks. In successive tasks a response to the current task stimulus is typically made before the stimulus for the next task is presented. In concurrent tasks the stimulus for the next task is presented before a response to the current task stimulus is made. So, in concurrent task studies there is an overlap between the tasks.

Successive task switching studies

Jersild (1927) was the first one to report results on task-switching experiments. In his experiments there were always two conditions: an "alternating" condition in which participants had to alternate between performing two tasks (ABABAB...) and a "pure" condition in which participants had to perform just one task (AAAAAA..., or BBBBBB...). In one experiment Jersild displayed a column of 25 two-digit numbers. In the Alternating condition, participants had to subtract 3 from the first number, add 6 to the second, subtract 3 from the third and so on. In the Pure condition participants had to subtract 3 from every number for the first half of the list and add 6 to every number for the second half of the list. The average time for the Pure lists was subtracted from the average time for the Alternating lists, resulting in switch costs of 1.2 sec per item.

Spector and Biederman (1976) used the same task procedure as Jersild and replicated this result. In an additional experiment each stimulus was accompanied by task-relevant cues (i.e. 34 + 3, 56 - 3, 12 + 3). Switch costs were reduced, supporting the notion that the principal determinant of switch costs is the extent to which the stimulus provides a readily discriminable retrieval cue for the currently appropriate mental operation when the participant switches between the two tasks in the alternating blocks. Experiments by Allport, Styles and Hsieh (1994) support this notion. Two stimulus task sets were used: one task set was an ensemble of incongruent Stroop colour words (e.g. RED printed in green colour). Participants could be asked to name either the colour or the word itself. The other set was an ensemble of displays each containing between 1 and 9 tokens of the same digit. Participants were asked to name either the number of digits or the value of the digit itself. Switch costs were higher when participants had to switch within the same task set (e.g. between colour naming and word naming) than when they had to switch between task sets (e.g. between colour naming and value naming). In line with Spector and Biederman (1976) the stimulus in the different task set condition discriminated better which of the tasks had to be performed than the stimulus in the same task set condition.

Rogers and Monsell (1995) mentioned two disadvantages of the way Jersild estimated switch costs. First, in the alternating conditions there is an extra demand besides the activity of switching, namely two task-sets rather than one has to be kept available. As a consequence, the source of switch costs is not clear. Second, the alternating blocks condition is very likely to be perceived as more difficult than the pure blocks condition. Therefore, Rogers and Monsell (1995) devised an alternative paradigm: the alternating runs paradigm. Instead of comparing alternating (ABABAB) and pure (AAAA, BBBB) blocks of trials they had participants alternating between runs of trials on the two tasks (AABBAABB). To help the participant keep track, the stimulus was accompanied by a cue indicating its position in the current run. Switch costs were estimated by comparing performance on trials in which participants had to switch (AB, BA) with performance on trials in which no switch was required (BB, AA).

In this paradigm switch costs were substantial as well. The authors explained these switch costs in terms of task-set reconfiguration (TSR): there is an executive controller that initiates the appropriate task-set. Initiating the appropriate task-set takes time and when the interval between tasks is short, it is reasonable to suppose that switch costs occur only because there is not enough time to complete the TSR process before the next stimulus arrives. So processing the next stimulus has to be postponed until the TSR process is completed. However, when the interval between tasks was set so large that the TSR process could be completed, switch costs were still present. Rogers and Monsell (1995) suggested that besides an endogenous component there is an exogenous component: "stimuli can of themselves activate or evoke in a person a tendency to perform actions (or tasks) habitually associated with them, irrespective of prior intention, and sometimes in conflict with prior intention" (p. 208).

Evidence that the residual switch costs are exogenous comes from a series of experiments by Rogers and Monsell (1995) who used two tasks: classifying a letter as a consonant or a vowel and classifying a digit as odd or even. An irrelevant character accompanied the stimulus (letter or digit). In one condition the irrelevant character was always drawn from a "neutral" non-alphanumeric set (e.g. an @ or a #). This character was neither related to a response in the 'letter' task, nor to a response in the 'digit' task. This condition forms the No-Crosstalk condition. In the Crosstalk condition, the irrelevant character was, in two-third of the trials, a digit in letter-classification trials, and a letter in digit-classification trials – and, in the remaining third of the trials, a neutral character.

Switch costs in the Crosstalk condition were much greater than in the No-Crosstalk condition. Moreover, in the Crosstalk condition the response associated with the irrelevant character could be congruent with the response associated with the stimulus (e.g. the stimulus letter is a vowel for which the right index finger has to be used and the irrelevant character is an odd digit which requires a response with the left index finger) incongruent with the response associated to the stimulus (e.g. the stimulus letter is a vowel for which the right index finger has to be used and the irrelevant character is an even digit which would have required a response with the left index finger) or neutral. The switch costs for congruent trials were slightly smaller than for incongruent trials. The more striking effect however, was that switching was much faster with a neutral irrelevant letter, than with either a congruent or incongruent one.

Apparently, the irrelevant character evokes a concurrent task from which participants are supposed to be switching away. The interfering effect of this task makes it more difficult to switch. The authors argue that this effect, and the overall effect of less switch costs for the No-Crosstalk condition, are symptoms of exogenous control. Or, as Monsell (1996) framed it: "the presence of stimulus attributes associated with a task tends to evoke that task-set and makes it harder to suppress when another task is required" (p.139).

An alternative explanation for switch costs is provided by Allport et al. (1994). They assumed that time costs associated with task switching stem from task-set inertia (TSI). By analogy with memory research they attribute switch costs to proactive interference from competing S-R mappings with the same stimuli, persisting from the instruction on previous trials. Switching from task X to task Y is harder if task X comprises the same type of stimuli as task Y. In one experiment, participants switched between the task of 'vocally reading colour words printed in various ink colours' and the task of 'vocally reporting the numerical value of digits for which multiple copies were displayed in spatial arrays'. In the first phase of the experiment these tasks were performed in alternating-task and repetitive-task blocks. After a few blocks of each kind no switch costs remained. In the second phase of the experiment participants were presented with the same stimuli but with different tasks: 'naming the ink colours of colour words' and 'vocally reporting the numerosity of digits in spatial arrays'. The responses to the stimuli of the second phase of the experiment conflicted with responses to exactly the same stimuli of the first phase, resulting in larger switch costs in the second phase. It is concluded that proactive interference from prior S-R mappings makes task switching during the second phase more difficult.

Concurrent task switching

A paradigm that is very often employed in overlapping tasks is that of the psychological refractory period (PRP). In this paradigm two stimuli are presented in rapid succession. Responses to these stimuli are discrete and usually simple (e.g. pressing a key). The second stimulus (S2) is presented after the first stimulus is presented but before a response (R1) is given. The typical finding in these studies is that the response to the second stimulus (R2) is considerably slowed as compared to when S2 is presented in isolation. De Jong (1995) states that the PRP-effect "appears to reflect a fundamental limitation in people's ability to engage in the performance of more than one task at a time" (p. 2). He assumes that there is a central channel that can deal with only one task or stimulus at a time. He investigated whether there is a control mechanism that allocates the central channel to the tasks and, if there is such a mechanism, how it would work.

In typical PRP-tasks the order of tasks is held constant. De Jong altered the standard paradigm by using a variable order of tasks. The task was either an auditory task or a visual task. Participants were cued about the actual order in every trial. Cues could be either informative or neutral. When informative, cues could be either valid - for example the cue predicted that the auditory stimulus was presented first followed by the visual stimulus which happened accordingly - or invalid - for example the cue predicted that the auditory stimulus was presented first followed by the visual stimulus while in fact the presented order was reversed. Results showed that when the cue was invalid and the time between S1 and S2 was short (100 msec), participants had the tendency to respond to the second stimulus first and then to the first stimulus. This tendency was much weaker when the time between the two stimuli was longer (350 msec). When the cue was valid the tendency to respond to the second stimulus first was also very weak. So, it seems that participants possess a controller that initially allocates the central channel to the task to be performed first and re-allocates it to the other task after the first one is finished. These results contradict the theory that holds that the central channel is not explicitly allocated but is instead simply recruited by the order in which stimuli arrive at the channel. For, according to this 'recruitment theory', participants should always process two stimuli in the order in which they are presented while De Jong's data show that participants in some cases are inclined to process the second task first.

Furthermore, De Jong found that participants not only prepared for the first task but also for the subsequent switch to the second task. In an additional experiment a fixed-order condition - that is a condition in which the order of the auditory and visual task was fixed - was compared with an alternating order condition - that is a condition in which task order regularly alternated between trials. This implies that in the fixed order condition the switch to the second task was the same throughout the block (for example there was always a switch from an auditory to a visual task) whereas in the alternating-order condition the switch alternated (for example, a switch from an auditory to a visual task was followed by a switch from a visual to an auditory task). Since it was assumed that preparation requires time it was predicted that for short intervals between trials, responses to the second stimulus would be substantially delayed in the alternating order condition. In the fixed-order condition this preparation time would not be necessary because one can use the same control structure for each switch. The outcome of the experiment was that responses to the second stimulus in the alternation-order condition were indeed much slower than in the fixed-order condition, especially for short interval between trials. These findings support the idea that participants not only prepare for the first task but also for the subsequent switch to the second task.

Recent studies on multiple task performance lend support to the existence of a controller for simple reaction time tasks (Meyer and Kieras, 1997; Meyer, Kieras, Lauber, Schumacher, Glass, Zurbriggen, Gmeindl and Apfelblat, 1995 and Schumacher, Lauber, Glass, Zurbriggen, Gmeindl, Kieras and Meyer, 1999). All these studies adhere to a model that assumes that people have flexible control over processing a second task. According to this model a so-called lockout point is set for the second task. Processing the second task is suspended temporarily until the first task is judged to be completed. For the first task there is an unlocking event. When this stage is reached, the first task is judged to be completed and the executive process permits processing of the second task. So at that point executive processes unlock the second task so that processing that task continues from the point at which it was previously suspended. The specific location of lockout points and unlocking events are presumed to be optional and are contingent on factors like relative task priority, participants' strategic biases and the amount of practice.

The main observation from the traditional task-switching studies is that switching between tasks always results in switch costs in terms of time. Concurrent task-switching studies demonstrated that extra time is needed for the second task because people are unable to deal with more than one task at a time. Processing the second task is postponed until the first task is completed. Also in studies where simple reaction time tasks follow in close succession there appear to be switch costs which are either explained in terms of task-set reconfiguration or in terms of task-inertia.

As mentioned earlier, tasks in the task-switching paradigm are simple reaction time tasks. Switching between tasks at a higher level of complexity is found in studies on interruptions. Therefore, in the next paragraph we will look more closely at the literature on interruption tasks.

Interruption studies

Already in the 1920s Zeignarik (1927) found that interrupted tasks that could not be completed were better recalled than tasks that were not interrupted. She administered a series of 20 brief simple tasks (e.g. making words from letters or writing names of cities beginning with the letter L). Half of the tasks could be finished whereas the other half were interrupted and could not be finished. Immediately following the completion of the series, the participants were required to recall as many of the tasks as possible. The percentage of interrupted tasks recalled appeared to be significantly higher than the percentage of completed tasks (68% versus 43%). Oviankina (1928) demonstrated that, when interrupted, participants have a tendency to complete the interrupted task. Later studies confirmed this effect (Brown, 1948; Krech, Crutchfield and Livson, 1967).

In a recent study, Zijlstra, Roe and Krediet (1999) tried to identify which cognitive processes underlie the Zeignarik effect. Two groups of secretaries - one from a Dutch university and one from a Russian university - worked on standardized text editing tasks for two days. During the main task the secretaries were interrupted by telephone calls from the experimenter.

The interrupting task could differ in complexity (a 'simple' interruption consisted of a request for some irrelevant information, a 'complex' interruption consisted of a more elaborate task with a greater similarity to the main task). Besides complexity, interruptions could vary in frequency: in the Dutch part of the experiment participants were not interrupted at all, or one or three times during a session, and in the Russian part participants could not be interrupted or two times during a session.

Several dependent variables were measured: performance measures (e.g. the time required for the interruptions and the number of errors in the text editing task), psychological state indicators (scales that indicated participants' well-being) and psycho-physiological state indicators (the cff technique - which stands for Critical Flicker/Fusion Frequency - which reflects the level of activation of the central nervous system). Contrary to expectation, interruptions did not have a detrimental effect on performance of the main task. Participants even spent less time on the main task. Dutch and Russian participants differed in the way they handled interruptions. Russian secretaries executed the interrupting task immediately, whereas Dutch secretaries, being more used to interruptions during work, postponed execution of the interrupting event. Interesting in the performance data is the finding of an after-effect of interruptions: after a complex interruption participants needed more time to disengage from the interrupting task and to reorient themselves on the main task than after a simple interruption.

There was a difference in the emotional and psychological state between Dutch and Russian secretaries as complexity of interruptions increased. While Dutch participants experienced a decrease in negative emotional feelings, Russian participants showed a reduced subjective well being. This divergence was also explained by the difference in professional background: Dutch secretaries may be accustomed to frequent and demanding types of interruptions thereby perceiving complex interruptions as welcome distractions whereas secretaries in the Russian study may be less used to interruptions and consequently perceived these interruptions as stressors.

Gillie and Broadbent (1989) wondered why some interruptions are more disruptive than others. During a computer-based adventure game where participants needed to issue commands to the computer in order to achieve certain goals, they were interrupted by mental arithmetic tasks. The interrupting task varied in length, similarity to the main task and complexity of processing. Results suggested that similarity to the main task and complexity of interruptions were determinants of the disruptive effects of interruptions. The length of the interruption did not seem to be a critical factor.

Like simple reaction time tasks, switching between complex (interrupting) tasks is accompanied by costs. The studies discussed above show that interruptions have disruptive effects on the main task. These costs are generally explained in terms of task-set inertia.

In sum, task-switching studies provide two reasons for cognitive lockup. (1) People may refrain from switching and stick to their initial plan of action because of the costs that accompany task switching. (2) People may not be able to deal with two tasks simultaneously. This second reason for not switching to the second problem is derived from PRP (Psychological Refractory Period) studies. The general explanation for the slower response to a second stimulus presented during processing of the first stimulus than to the same stimulus in isolation is a bottleneck in people's information processing capacities. One simply lacks executing capacity to deal with a second task before the first task is completed.

Decision Making

People have the tendency to stick to an initial plan, even in situations in which it is rational to switch to an alternative plan. In the literature on behavioral decision making this tendency is reflected in a number of biases, four of which - the sunk-cost effect, task completion, the omission bias and the endowment effect - will be discussed shortly.

Sunk cost effect

The sunk-cost effect is the finding that an endeavour is continued once an investment in money, effort or time has been made. An example of the sunk-cost effect comes from a study by Arkes and Blumer (1985). They offered one group of participants the following scenario:

> *As the president of an airline company, you have invested 10 million dollars of the company's money into a research project. The purpose was to build a plane that would not be detected by conventional radar, in other words, a radar blank plane. When the project is 90% completed, another firm begins marketing a plane that can not be detected by radar. Also, it is apparent that their plane is much faster and far more economical than the plane your company is building. The question is: should you invest the last 10% of the research funds to finish your radar-blank plane?*

Another group of participants received a the following scenario:

> *As the president of an airline company, you have received a suggestion from one of your employees. The suggestion is to use the last 1 million dollars of your research funds to develop a plane that would not be detected by conventional radar, in other words, a radar blank plane. However, another firm has just begun marketing a plane that can not be detected by radar. Also, it is apparent that their plane is much faster and far more economical than the plane your company could build. The question is: should you invest the last million of your research funds to build the radar-blank plane proposed by your employee?*

On the question whether they would invest the last one million dollars to finish their radar-blank plane a majority of the participants in the first group answered "yes", whereas a majority of the participants in the second group answered negatively. The difference between the scenarios is that for the first group 9 million dollars have already been invested while for the second group nothing has been invested yet. The fact that so much money has been spent on the research project motivates participants in the first group to keep investing money in it.

The example described above is a clear example of the so-called sunk-cost effect. Arkes and Blumer (1985) explain this effect in terms of wastefulness. Discontinuing an endeavour for which money (or time) has been spent may give the impression that one is spending money like water. And one does not want to appear wasteful.

However, there are also alternative explanations for the sunk-cost effect (Brockner, 1992; Garland, 1990; Staw, 1981, Thaler, 1980). Staw (1981) for instance, supports the idea that the sunk-cost bias results from a process of self-justification. People have a strong desire to be correct and accurate and also to prove this to themselves and to others. A continuation of investments can be regarded as a justification for prior investments.

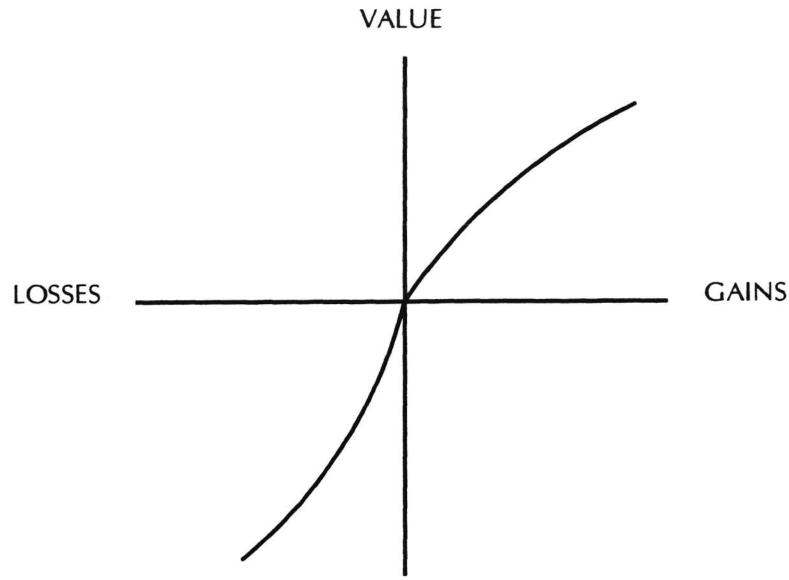


Figure 1.1: *The value function of gains and losses according to prospect theory.*

Thaler (1980) tried to explain the sunk-cost effect in terms of prospect theory (Kahneman and Tversky, 1979). A crucial feature of prospect theory is that choices are not evaluated in terms of final assets but in relation to a reference point (see figure 1.1). An outcome is considered a gain when it is above the reference point and it is considered a loss when it is below the reference point. The value function that is depicted in figure 1 shows that the function is concave for gains and convex for losses. This implies that people are generally risk averse in gain situations and risk seeking in loss situations. Another characteristic of the value function is that it is steeper for losses than for gains, implying that losses loom larger than gains. This means that the pleasure associated with a gain of $100 is less intense than the pain associated with a loss of $100.

According to a prospect theory account, prior investments are not totally discounted. When evaluating subsequent prospects, prior investments are regarded as losses. Still being at the loss side of the value function, which is concave, losses do not cause large decreases in value. However, gains do cause large increases in value. From this point of view, risky behavior – like investing even more – is more likely than total withdrawal that would imply a sure loss.

To conclude, the sunk-cost effect is a robust phenomenon in decision-making literature that has been explained from several theoretical perspectives.

Task completion

Recently, a growing body of literature emphasizes the importance of completion effects (Boehne and Paese, 2000; Conlon and Garland, 1993; Garland and Conlon, 1998 and Moon, 2001). Garland and Conlon (1998) pointed out that many sunk-cost explanations may be confounded with "project completion", the degree to which the project at hand is near completion. In all studies on sunk-cost effects, large sunk costs go together with high project completion and small sunk costs go together with low project completion.

Consider for example the scenario of the airplane company discussed above (Arkes and Blumer, 1985). The overall finding was that participants who were told that already 9 million dollars had been invested show a greater tendency to keep investing in the project than participants who lacked this information. It is true that the degree of investment is higher in the first version of the airplane company than in the second version, but the degree of completion varies as well. In the first version only 10% of the company's own project has to be completed whereas in the second version the project has not been initiated yet. In other words, the full 100% of the project has to be completed. Investment and completion in this example are clearly confounded.

A number of studies have been conducted that manipulated sunk costs and project completion independently. An example of such a scenario is "The Bank Manager"(Garland and Conlon, 1998, study 1). The scenario reads as follows:

You are a loan officer at a large commercial bank. Custom Molds Inc., a manufacturer of plastic injection molds for high tech and precision parts, is one of your clients of long and good standing.

About 1 year ago, the CEO of Custom Molds approached you with a request for funds in order to revamp his manufacturing capabilities in a manner that would allow the firm to gear up for new competition. After long discussion and detailed scrutiny of the project plans, you recommended that the bank approve a $ 10 million loan for this project. The bank did approve up to $10 million for the project, with an agreed schedule of disbursement. The covenants provide for bank monitoring of project progress.

To date, $2 ($8) million has been disbursed to the company.

Over the past few months, industry data and market information have suggested that the firm's competitive position has been negatively affected by the new entrants into this increasingly global market. In fact, just last week, a principal client of Custom Molds dropped the company from its approved vendor list.

The CEO of Custom Molds has now asked you to authorize the next instalment of $2 in order to continue with his revamping project. In his letter to you, he indicated that the revamping project is about 20% (80%) completed.

Failure to authorize the requested funds would place Custom Molds in a very precarious position, with a high probability of default on their outstanding loan.

Sunk costs were manipulated by the amount of money that had been disbursed. Sunk costs were considered low when the amount was $2 million and sunk costs were considered high when the amount was set on $ 8 million. Project completion was manipulated by a percentage provided by the CEO. This percentage was either 20% (low project completion) or 80% (high project completion).

Participants were asked to indicate the probability that they would authorize the expenditure of the project. Results showed only an effect of project completion: when 80% of the project was completed, the willingness to allocate the next $2 million was stronger than when only 20% of the project was completed.

Boehne and Paese (2000) subjected the project completion hypothesis to a stronger test. Garland and Conlon (1998) stated that "as progress moves forward on a project, completion of the project itself takes increasing precedence over other goals (e.g. economic profit) that may have been salient at the time the decision was made to begin the project" (p.2042). However, as Boehne and Paese argued, Garland et al. (1998) provided no evidence for this explanation. Goals such as economic profit could not even have played a role in their studies because the information how much the project would generate once it would be terminated, was missing.

In their study, Boehne and Paese (2000) included this piece of information in the scenario of a real estate development project. Besides an independent manipulation of sunk costs and task completion, they varied the sales price of the real estate. The sales price information allowed participants to calculate the economic value. According to the authors the information may induce participants to engage in an economically rational decision process.

Contrary to their expectations the authors still found a very strong effect of task completion. In fact, when the project was close to being finished, they often recommended completing the project even when it was economically unwise to do so. The sunk-cost manipulation virtually had no effect.

In a recent study, Moon (2001) found evidence for both sunk costs and task completion. Most interesting, he also found an interaction between the two effects: the sunk-cost effect appeared to be present under conditions of high completion, but appeared to be absent under conditions of low completion. The author explains this interaction in terms of psychological pressure. This pressure is much higher in case of high completion and much lower in case of low completion. This explanation agrees with entrapment studies (Brocker, Rubin and Lang, 1981; Teger, 1980) in which it is assumed that the decision-maker must be psychologically triggered to invest further. In the study of Moon, sunk-cost effects were found when participants were psychologically compelled by a nearly finished project.

The task completion effect is a phenomenon that has recently been raised as an alternative explanation for the sunk-cost effect.

The endowment effect

Another example of people's tendency to stick to old plans is the endowment effect. The endowment effect is the phenomenon that people are unwilling to give up an item, with which they have been randomly endowed, for an alternative item. A well-known experiment is that of the coffee mugs (Kahneman, Knetsch and Thaler, 1990). Half of a group of participants were given a coffee mug while the remaining participants received a large Toblerone chocolate bar. After that, participants were given the opportunity to exchange items. Since items were assigned randomly it was expected that half of the participants would be willing to trade their item. However, only a small part was actually willing to do so.

This effect is related to the finding that the lowest selling price for an endowed item is considerably larger than the highest price for which one wants to buy the same item. In a study by Kahneman, Knetsch and Thaler (1990) a group of students were given a coffee mug from their university bookstore and they were asked to indicate for which price they would be willing to sell the mug. Another group of students indicated for which price they would be willing to buy the same mug. The median price was $7.12 for the "sellers" and $2.87 for the "buyers".

The effect mentioned above could be explained by prospect theory: people are loss averse and they have the tendency to weigh losses more heavily than corresponding gains. So, when participants are endowed with a coffee mug and are asked to exchange the mug for a chocolate bar, the loss of the mug looms larger than the gain of the chocolate bar. Once given an item, that item gets a surplus value. As a consequence, people are reluctant to give up what they have.

The status quo bias

The tendency to cling to an initial course of action is also reflected in the so called status-quo bias, reported by Samuelson and Zeckhauser (1988). Participants tended to choose the current state of affairs, although it was no more attractive than other available alternatives. They presented participants with either one of two different versions of a funds investment decision task. In one version - the neutral version - participants are told to picture they had inherited a large sum of money from an uncle. Furthermore, they were told that they had to imagine considering different portfolios. Participants could choose between four different options.

In the other version – the status quo version – participants were told to imagine they had inherited a portfolio of cash and securities from their great uncle. As in the neutral condition, participants had to choose between four different portfolios, one of which was the actual portfolio inherited. So, the status quo was equal to one of the alternatives. The general finding in this experiment was that participants in the status quo version demonstrated a much higher preference for the alternative that corresponded to the status quo than participants in the neutral version.

As the endowment effect, the status quo bias is generally explained in terms of loss aversion. The status quo is a reference point. A switch to an alternative course of action entails expected losses on one dimension and expected gains on the other dimension. One gives more weight to potential losses from switching than to potential gains. Hence, people are unlikely to prefer alternatives for which the expected gains are only slightly higher than the expected losses.

According to some researchers (Ritov and Baron, 1992) the exaggerated preference for the status quo is actually the same phenomenon as the omission bias, people's tendency to prefer omissions over acts. The omission bias is often demonstrated by the following example:

Paul owned shares in Company A. During the last year he considered switching to stock in Company B, but he decided against it. He now finds that he would have been better off by $1,200 if he had switched to stock of Company B.

George owned shares in Company B. During the past year he switched to stock in Company A. He now finds that he would have been better off by $1,200 if he had kept his stock in Company B.

Most participants imagined that George would feel more regret than Paul, even though both are faced with the same final outcome. The only difference is that in George's case the outcome resulted from an action whereas in Paul's case it resulted from inaction. This effect reflects the phenomenon that negative consequences from acts are evaluated as more negative than the same consequences that result from omissions.

Ritov and Baron (1992) argued that sticking to the status quo is actually confounded with a preference for omissions over acts. Presented with scenarios in which the effects of status quo and omission to act were not confounded, evidence was in fact only obtained for an omission bias. However, in a later report, Schweitzer (1994) demonstrated both a status quo and an omission bias.

Several explanations have been proposed for the omission bias. Apart from loss aversion, the omission bias is explained in terms of norm theory and reluctance to choose. According to norm theory (Kahneman and Tversky, 1982) acting is considered as more abnormal because "it is usually easier to imagine abstaining from actions that one has carried out than carrying out actions that were not in fact performed" (p.145). Omission is taken as a reference point and, since acts are considered as more abnormal than non-acts, emotional reactions (e.g. regret) are enhanced.

Another explanation for the occurrence of the omission bias is a reluctance to choose (Ritov and Baron, 1992). Omissions may be perceived as non-acts rather than a deliberate choice "not to act". When confronted with an awkward dilemma one may find it difficult to make a decision and a way to deal with this difficulty is by making no decision at all. Or, as Janis and Mann (1977) put it: "a decision maker under pressure to make a vital decision will typically find it painful to commit himself, because there are some expected costs and risks no matter which course of action he chooses.

One way of coping with such a painful dilemma is to avoid making a decision" (p. 6).

So, loss aversion seems to be the general psychological mechanism underlying both the endowment effect and the status quo bias. The omission bias implies that people prefer the option of non-acting over options of acting. We think that the status quo bias – referring to a preference for the status quo over alternative state of the world – is better applicable to lockup phenomena than the omission bias. The actual problem with being locked up is that people prefer to continue the ongoing course of action rather than any alternative course of action. Doing nothing is no option.

In sum, a review of decision-making literature renders three different explanations for cognitive lockup: (1) loss aversion, (2) sunk costs and (3) task completion.

Summary

Earlier, we stated that a review of research areas related to cognitive lockup could provide entries to investigate this phenomenon. We elaborated on three adjacent fields, planning, task switching and decision making, each generating a number of hypotheses for cognitive lockup. Summarized, the following explanations were discussed.

Planning

1. People commit themselves too early to a detailed plan;
2. People refrain from monitoring the environment;
3. People generate future scenarios that are too optimistic;

Task-switching

4. People lack spare attention to switch to a second disturbance;
5. The costs of switching are perceived too high;

Decision-making

6. People are loss averse: they weigh losses larger than gains;
7. Sunk costs: prior investments are taken into account;
8. Task completion: people have a desire to fulfil a task.

2. Cognitive lockup and task characteristics

In the second chapter we try to identify the main characteristics of the three research paradigms (planning, task switching and decision making). In order to examine the explanations that were provided by each paradigm we implement these task characteristics in the present task setting. The task setting is a simulation of a fire control task that consists of two modes of control: monitoring the environment and fault diagnosis. The system is in a steady state until a fire breaks out. At that moment participants have to detect the fire and start diagnosing the cause it in order to select the appropriate treatment. When there are two fires at the same time, the situation has to be reassessed in order to find out which is the most urgent and needs to be dealt with first. An experiment is conducted with two main conditions: a sequential condition which includes the characteristics (an environmental change, the start of a second problem while being involved in the first one and the presence of prior investments) and a simultaneous condition where these characteristics are absent. The results of the experiment show that cognitive lockup is stronger in sequential scenarios. We therefore conclude that the present experimental task paradigm can be adequately used to study psychological explanations for cognitive lockup. We end this chapter with an overview of where in the thesis different explanations are investigated.

Introduction

We started the introductory chapter with the air crash of flight 401 of Eastern Air Lines. This accident is an example of a supervisory control task where cognitive lockup resulted in a dramatic outcome. The purpose of this thesis is to explore how operators like the pilot of flight 401 get caught by cognitive lockup

In the previous chapter we reviewed research from three different paradigms in which similar phenomena like cognitive lockup occurred. These paradigms provided possible explanations for cognitive lockup. However, the findings from each paradigm are dependent on the characteristics of the task environments that were used and conclusions can therefore not simply be extrapolated to supervisory control tasks. The goal of the present chapter is to relate the explanations that have been put forward in each paradigm to the main characteristics of the tasks that were used and to assess whether these characteristics are relevant to supervisory control. The result provides necessary task characteristics for an experimental task that is useful to investigate psychological mechanisms for behavioral entrapment in supervisory control. In addition, an experiment will be described. The most important goal of this experiment is to show that cognitive lockup occurs. In subsequent chapters, research will be described that aimed at finding the most likely explanation.

Planning

Recent definitions of planning have stressed the importance of an adaptive change of plan to possible changes in the environment. Effective planners need to revise their plans when the environment changes. Research has therefore mainly focussed on how people react to such changes. So, an important feature of tasks that have been used in the planning paradigm is a change in the environment that requires a revision of plan.

This feature is also important in supervisory control tasks, where operators may be confronted with environmental changes while they are doing other tasks. In the case of flight 401, for example, the aircraft started to descend while the crew was involved in dealing with the problem of the landing gear. The altitude problem that occurred was far more urgent, and would require a revision of the initial course of action, that is, to start working on the altitude problem.

To conclude, one of the main characteristics that has been investigated in planning research is the reaction to environmental changes, and this feature is highly relevant to supervisory control tasks as well.

Task-switching

In studies on concurrent task-switching and interruptions people have to deal with multiple tasks simultaneously. They are involved in processing a first stimulus or task at the moment a second stimulus or task is introduced. The central question is how participants deal with a second stimulus or task. Do they proceed with the first stimulus or task or do they switch to the second stimulus or task?

The task characteristic of multiple tasks is also present in critical situations in supervisory control. Subparts of the system are often interrelated and faults may easily propagate through the system. As a consequence the operator has to deal with multiple faults at the same time. In case of flight 401, the pilot also had to deal with several problems at the same time: the malfunctioning landing gear and the decreasing altitude. Even though the second one occurred later in time, it had a higher priority and should have been dealt with first.

Decision Making

Decision making research on phenomena related to cognitive lockup has shown the importance of initial investments. At the moment an alternative task or project is introduced people have already invested in the first task or project and additional investments have to be made to complete it. In deciding whether to continue on the current task or project or to switch to the alternative task or project, these prior and future investments may play a role.

In supervisory control tasks, these phenomena may occur as well at the moment a second disturbance starts. On the one hand investments are made in the first disturbance (time and effort) and on the other hand, because of these prior investments, the completion of the disturbance is likely to have become closer. For the example of flight 401, investments were made in solving the problem with the landing gear. At the time the problem with the descending altitude occurred, the pilot had made initial investments that implied that the problem was also closer to completion.

In sum, the following task characteristics can be identified that may all explain behavioral entrapment: a change in the environment that requires an adaptation of plans, multiple tasks that need to be handled in order of priority and not in order of presentation and the presence of initial investments.

In the next section we will describe how we implemented these task characteristics in the present task environment.

Task characteristics in the present study

For the present thesis we designed a simulation of a shipping control task. Globally, there are two modes of control: monitoring the system and fault diagnosis. The system is in a steady state until a fire breaks out. At that moment, the operator has to detect the fire and start diagnosing the cause of the fire in order to select the appropriate treatment.

When there are two fires simultaneously, the situation has to be (re)assessed in order to find out which fire is most urgent and has to be dealt with first.

Environmental change

The environmental change in the fire control task was induced by starting a second fire when the participant was already involved in fire fighting. This environmental change meant that participants had to reassess the situation, as either fire 1 or fire 2 could have the highest priority. This situation involved the presentation of two fires sequentially.

In order to examine a person's reaction to an environmental change, the sequential condition is compared with a situation in which two fires start at the same time. In this situation participants have the opportunity to assess the priority of both fires and to define a plan that - in contrast with the sequential condition - does not need revision halfway through. This way we could examine the question to what extent participants would interrupt ongoing activities in order to reassess the situation.

Multiple tasks

The task characteristic of multiple tasks is realised in the present task environment by the occurrence of a second fire while one is still involved in the diagnosis process of the first fire. Of main concern with this characteristic is to assess the effect of task complexity, in terms of attentional resources, on cognitive lockup. To that extent we designed the task in such a way that information had to be requested to a) assess priorities and b) select the correct course of action. In the present task environment it was possible to manipulate the way in which information had to requested. By this, we could manipulate the complexity of the task of assessing priorities and the task of selecting the correct treatment.

The possibility to manipulate the complexity of the task and its claim on the available resources enabled us to examine explanations of cognitive lockup in terms of lack of attention.

Costs

When several disturbances occur at the same time a decision has to be made as to which disturbance is handled first. The cost structure for the fire control task can be defined by time. First, a fire has to be dealt with within a predefined time span. Second, requesting information takes time. The answers to questions are not provided immediately but after a delay. In all, this cost structure means that when fires occur sequentially, time investments have been made the moment a second fire starts, and some time period remains until task completion.

Experiment 1

The main goal of this experiment is to demonstrate the phenomenon of cognitive lockup in a supervisory control task that included all three task characteristics: an environmental change, multiple tasks and time costs. We operationalized cognitive lockup in this experiment as completing the first fire before detection of the second fire (and consequently réfraining from reassessing priorities).

The critical condition is therefore the situation in which two fires are presented sequentially. In this condition, a change in the environment takes place (a second fire occurs at the moment participants are diagnosing the first fire), (part of) the attentional resources are utilized by the first fire, and time investments are made at the moment the second fire starts. This condition is compared with a simultaneous condition in which a plan can be made for both fires at the beginning of the scenario. For this planning process there is therefore no reaction needed to an environmental change, all attentional resources are available and no investments have been made.

We predicted that cognitive lockup is stronger for the sequential condition than for the simultaneous condition.

Method

Participants

Twenty seven participants voluntarily participated in the experiment. They were all first year students at the University of Utrecht. The experiment lasted about two hours and they were paid Dfl. 70.

Experimental task

The experimental task was a simulation of a ship on which fires could occur. Participants had to monitor the ship on fires that they had to fight. For monitoring purposes, there was a two-dimensional representation of the ship that consisted of four layers. Participants had to monitor this part of the system for the detection of fires. When a fire occurred, a small red triangle popped out somewhere in the representation of the ship, accompanied by a high beep. Over time, a fire expanded which was shown by small red triangles fanning out from the fire symbol.

For the purpose of fire fighting, there were windows available in the subpart of the screen. In this part of the system the diagnosis process of a fire took place. For each fire detected a list of seven questions and seven possible treatments were provided. Answers could be requested by clicking the button that represented a particular question. The first three questions were for the assessment of priorities. The last four questions were to determine which treatment to select. In order to realise costs for requesting information, we built in time delays. Answers to questions for priority assessment were provided with a one-second delay and answers to questions for treatment selection were provided with a four seconds delay. During this period the system was blocked and participants couldn't perform any action in the system. Participants could chose one out of seven possible treatments, also by mouseclicks.

Not every fire needed all four questions to be answered. To simulate uncertainty with regard to the amount of time needed to solve disturbances, participants didn't know in advance how many questions had to be answered in order to determine the appropriate treatment. For example, when 'removing smoke' was the correct treatment, participants needed to click all questions, but when 'sending a large casualty team' was the appropriate treatment, only the answers to the fourth and fifth question had to be asked (see Appendix A).

Selecting the appropriate treatment immediately extinguished the fire. The red triangles disappeared from the screen and the participants heard three short beeps. Selecting an incorrect treatment shut down the system for 7 seconds.

During the experiment, participants were confronted with scenarios in which either one or two fires occurred. For this latter category, fires could differ in priority. The fire with a higher priority had a higher expansion rate. A relative high priority fire expanded at a faster rate than a low priority fire, and would sooner end in a burn down.

When two fires had the same priority, they should be solved within 50 seconds after onset. When the priority of two fires differed, participants had 35 seconds to solve the fire with the highest priority and 50 seconds to solve the fire with the lowest priority. When there was only one fire on the ship, this fire always had to be solved within 35 seconds after its onset.

Procedure

Before the actual experiment, participants were trained in the assessment of priorities and the selection of the correct treatment. For the assessment of priorities, participants were presented with three questions that could be answered by clicking on them. Under this list of questions were four buttons that represented the four different states of priority. Participants were handed out a tree-structure that aided them in asking the appropriate questions and choosing the correct priority (see Appendix B for this tree-structure). They were instructed to determine the priority of 20 fires. When incorrectly prioritised, a fire was presented again. When the priority of all 20 fires were

correctly assessed the training was finished.
The training for selecting the appropriate treatment consisted of a training in walking through a tree of questions. Again, they were handed out a tree-structure that could help them asking the relevant questions and determine the correct treatment (see Appendix A). There were four questions and seven possible treatments that could be selected. For 28 fires the correct treatments had to be provided. Fires for which an incorrect treatment was selected were presented again later.

After the training-session, participants were given the instructions for the experimental task. After 16 practice scenarios, the actual experiment began.

Participants were seated in front of a screen that showed the representation of the ship in the upper part of the screen and the fire control task in the lower part of the screen. Figure 2.1 shows the two parts of the system after two fires have been detected.

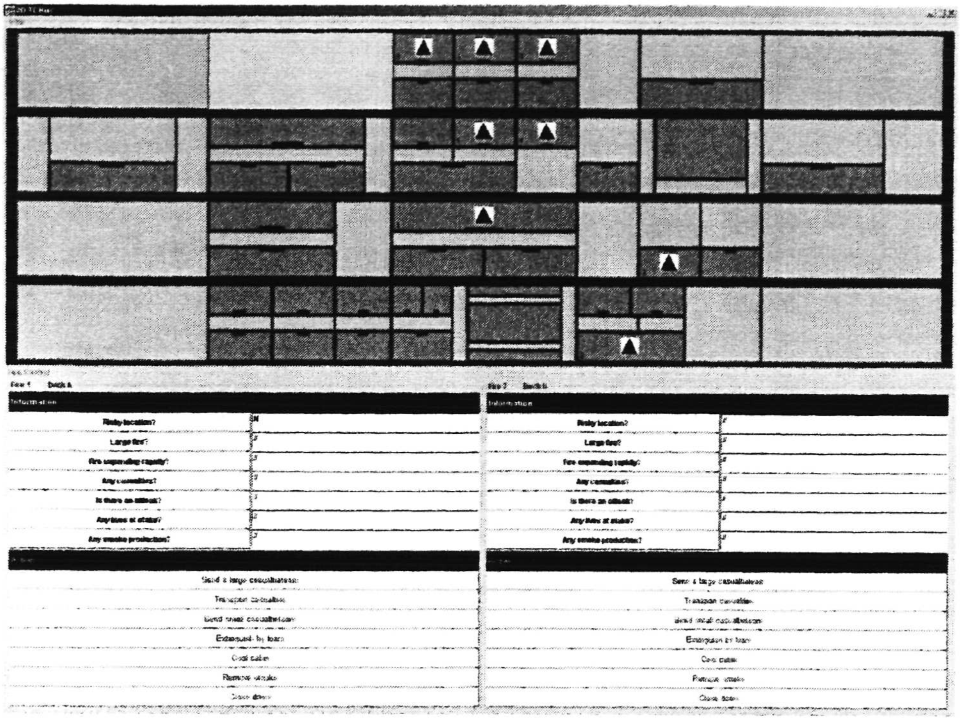


Figure. 2.1: *An overview of the system at the moment two fires have been detected.*

The main goal for the participants was to detect fires and to select the appropriate treatment as soon as possible. Participants had to detect a fire by clicking with the mouse on the red triangular icon in the representation of the screen. After detection, a window appeared in the bottom part of the screen in which the questions and possible treatments were presented. A number that showed up with the icon indicated to which fire the window referred. In case participants didn't solve a fire in time, the ship burned down which was represented by a blank screen.

To assess a fire's priority, participants needed the answers of only two of the three priority questions. Depending on the answer to the first question, the subsequent question to be asked was either the second or the third one (see appendix B). A question was answered by yes or no. Answers to the first three questions were generated by the system after a one-second delay. A fire could have a priority that ranged from 1 (the highest priority) to 4 (the lowest priority).

Design

The manipulation in this experiment was the presentation mode of two fires. Fires could start at the same time - forming the simultaneous condition - or the second fire started after one had clicked on a question of the first fire - forming the sequential condition.

Each condition contained 54 scenarios. For each condition there was an equal number of scenarios where the first fire had priority, the second fire had priority and the first and second fire had equal priority. (Note that in the simultaneous condition there was no 'first' or 'second' presented fire, since both fires started at the same time.) The conditions were also balanced for the number of questions that needed to be answered in order to determine the correct treatment of a fire.

To the 108 'two fires' scenarios, 18 scenarios were added in which only one fire occurred.

Dependent variables

- Performance. Overall task performance is defined by *the number of burn downs*. A burn down implied that the participant did not solve the fire(s) in time.
- Strategy. Two variables indicated participants' strategic behavior. One variable, *switch moment*, measured whether participants made the switch to the second fire before or after completion of the first fire. The second variable measured whether *priority information was requested* for the first fire detected.

Results

The description of the results is divided into two sections. First, performance data are presented, indicating to what extent participants were able to solve the fires in each experimental condition. The second section describes the strategy participants chose to deal with the fires. When did they switch to the second fire and how often did they ask priority information?

In the simultaneous condition the two fires were presented at the same time, whereas in the sequential condition the second fire started after the first question had been answered. This implies that participants had more time to solve the second fire in the sequential condition as the onset was later in time than in the simultaneous condition. In order to make the sequential and simultaneous condition comparable we subtracted this extra time from the time that was left after the selection of a correct treatment for the sequential scenarios where participants solved the second fire after the first fire. If the subtraction resulted in a negative value it was counted as a burn down.

Performance

Figure 2.2 presents the mean percentage of scenarios that ended in a burn down for the sequential and simultaneous condition.

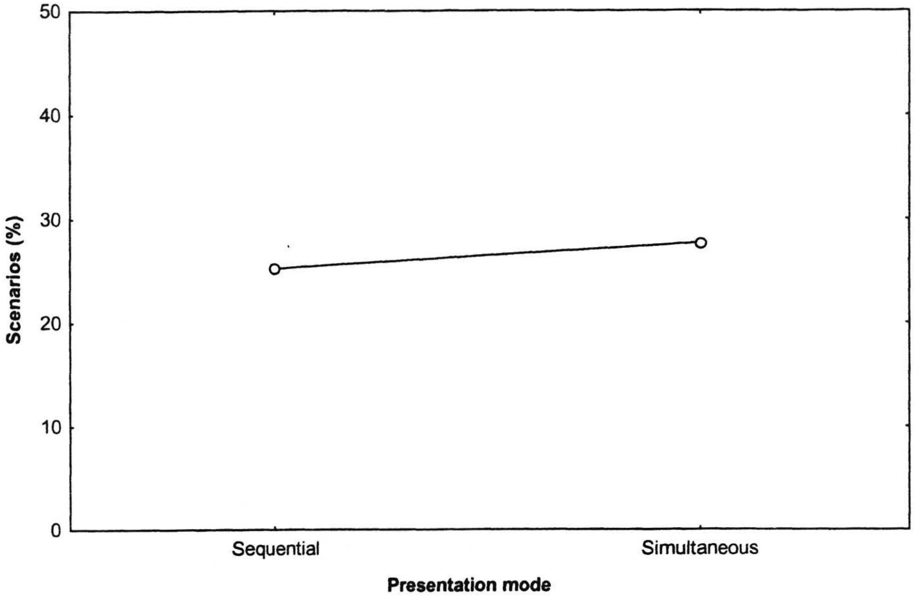


Figure 2.2: *The mean percentage of scenarios that ended in a burn down as a function of presentation mode.*

A dependent t-test showed that there was no difference in the percentage of burn downs between the sequential and simultaneous condition ($t = 1.21, df = 26$, n.s).

Strategy

Figure 2.3 depicts the mean percentage of scenarios in which the second fire was detected after the first fire was solved.

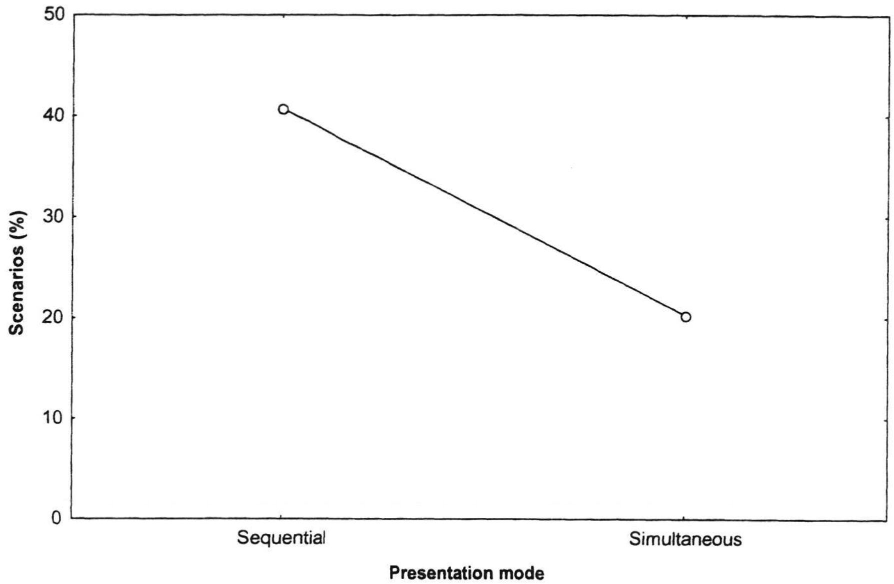


Figure 2.3: *The mean percentage of scenarios in which the second fire was detected after the first fire was completed as a function of presentation mode.*

Figure 2.3 shows that in the sequential condition the detection of the second fire after completion of the first fire was more frequent than in the simultaneous condition. This effect is significant ($t= 5.47, df= 26, p< 0.01$) and implies that cognitive lockup is stronger for the condition where the second fire was presented after one had started executing the first fire.

Figure 2.4 presents the mean percentage of scenarios in which priority information was requested for the first fire for both the sequential and the simultaneous condition.

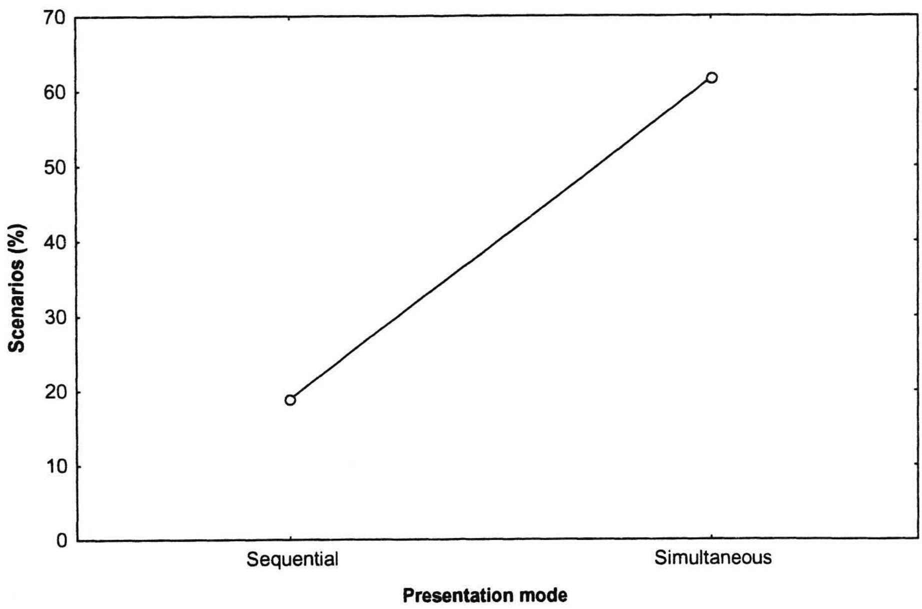


Figure 2.4: *The mean percentage of scenarios in which priority information was requested for the first fire as a function of presentation mode.*

The figure shows that there was an effect of presentation mode for the number of times priority information was requested for the first fire. In the simultaneous condition significantly more priority information was requested than in the sequential condition ($t= -5.48$, $df= 26$, $p< 0.01$). The effect supports the notion that in the sequential condition the tendency to refrain from reassessing task priorities is much stronger than in the simultaneous condition.

Discussion

The main goal of the pilot experiment was to demonstrate the phenomenon of cognitive lockup in a supervisory control task which included the three task characteristic which we identified earlier (namely an environmental change, multiple tasks and costs in terms of time). In the experiment we compared two modes of presentation – the sequential presentation and the simultaneous presentation – and we expected cognitive lockup to be stronger for the sequential condition, the condition in which participants needed to reassess the situation.

The experiment successfully demonstrated the effect of cognitive lockup for the sequential condition. The tendency to detect the second fire only after completion of the first fire is much stronger in the sequential than in the simultaneous condition. In the same line, the tendency to (re)assess priority information is far more prominent in the simultaneous condition than in the sequential condition. To put it in other words, participants were less inclined to switch to the level of assessing priority when they were already involved in solving the first fire.

It was somewhat surprising that a difference in strategy between the sequential and simultaneous condition was not reflected in the performance data. The number of burn downs was almost identical for the sequential and the simultaneous condition. A possible reason may be that the determination of priority required more time than we had expected. The extra information that was obtained by executing the task of asking priority questions did perhaps not always outweigh the additional costs in time. Asking priority information may have required so much time that even if one started solving the most urgent fire, the remaining time was often too short to complete both fires in time.

In all, the first experiment clearly showed people's tendency for cognitive lockup. The experimental paradigm of the fire control task can therefore be adequately used to study psychological explanations for cognitive lockup.

Back to the explanations

The main goal of this thesis is to find the most plausible explanation for cognitive lockup. In the introduction we summed up a number of possible explanations that we inferred from the three adjacent paradigms. In this section we will return to these paradigms and we will indicate for each paradigm whether and, if so, how the accompanying hypotheses will be investigated in the present work.

Planning

The process control task that is used in the present thesis is too restrictive to investigate the aforementioned hypotheses of planning. In our fire control task, participants start off with an initial plan (solve the first fire) and when a change in the environment takes place (the occurrence of a second fire), the question is whether they set out a new plan (first solve the fire with a higher priority) or not. Since the opportunity for planning is so limited we consider the present task unsuitable to test the specific planning hypotheses.

Nevertheless, the planning literature has pointed out the importance of an adequate reaction to an environmental change. We therefore included in every experiment of this thesis a condition where such an environmental change occurred and we recorded whether participants adapted their initial plan. The experiment described above contained so-called 'sequential' scenarios which included such an environmental change. This in contrast with 'simultaneous' scenarios which did not include such a change. A comparison of these two kinds of scenarios demonstrated people's inability to provide an adequate reaction to an environmental change. We tried to replicate this finding in the second and third experiment of this thesis which also included both sequential and simultaneous scenarios.

Task switching

Task switching research identified two explanations for cognitive lockup: (1) a lack of resources and (2) high costs of switching. According to the first explanation cognitive lockup is due to limitations in our information processing capacities. Operators can not deal with the second fire because not enough resources are left for making a switch to the second fire. The workload of the first fire is so high that all resources are needed.

In order to test this workload hypothesis we manipulated the degree of complexity for the first fire. If this hypothesis holds, being locked up in the first fire should be affected by complexity. Fires requiring a more complex diagnosis process are expected to demand more resources. And since more resources are allocated to the first fire, there will be less resources left for the second fire, resulting in a greater degree of cognitive lockup. We tested the workload hypothesis in the second experiment of this thesis.

The second explanation from task switching literature is that the costs of switching are perceived as too high. All task-switching studies demonstrated that switching between tasks inevitably results in switch costs. In the supervisory control task of solving fires, there are, besides the expected switch costs of interference, more explicit switch costs in the form of reassessing task priorities. The start of a second fire requested such a reassessment of priorities.

People may trade-off the costs and benefits of a reassessment and because they perceive the costs as high they may refrain from making an reassessment. This explanation is investigated in the third experiment of this thesis. In this experiment we varied the costs of reassessing priorities. According to the switch cost hypothesis, cognitive lock up would decrease as the costs of switching would be lower.

Decision-making

With respect to the decision-making paradigm, we concentrated on the sunk-cost and task completion explanation, not on the explanation of loss aversion. This because the prospect theory and loss aversion are more concerned with discrete choices and sunk-cost and task completion explanation take into account the dynamics of the situation. As a consequence these explanations seem to be more suitable to investigate cognitive lockup in supervisory control than an explanation in terms of loss aversion.

So, first, there is the explanation in terms of sunk costs. Following this explanation, it is predicted that the more participants have progressed on the first fault, the less they will be inclined to abandon the first fault. Our experimental task setting provides the opportunity to manipulate the degree of sunk costs by varying the number of questions that are asked at the time a second fire begins. If the sunk-cost hypothesis holds it is expected that cognitive lockup increases as more questions have been asked at the moment the second fire occurs.

The second explanation emanating from decision making research is task completion. Our fire control task offers the opportunity to vary the degree of completion, independent of the degree of investment. As just noted, the degree of investment can be manipulated by varying the number of questions asked at the moment the second fire started. The degree of completion can be manipulated by the number of questions that still needs to be answered at the moment a second fire starts.

According to this hypothesis, as the number of questions to be answered decreases the tendency for being locked up in the first fire will be stronger. The hypothesis of sunk costs and the hypothesis of task completion will be tested in experiment four and five of this thesis.

3. Cognitive lockup: Capacity limitation or decision-making bias?

Previous research has indicated that operators have a tendency to continue with an ongoing task and ignore additional tasks (cognitive lockup). In this chapter we report two experiments in which we examined whether this phenomenon is caused by capacity limitations or by a decision making bias. Participants are required to monitor a simulated fire control task and to deal with fires that occurred either sequentially or simultaneously. Results show that the tendency to continue with an ongoing task is not affected by the workload of the diagnosis process but rather by the perception of the costs making of a reassessment. We conclude that cognitive lockup is due to a deliberate decision rather than capacity limitations. We provide some suggestions for future system support.

Introduction

Technological developments have shifted the role of humans from direct control to supervision of complex automated systems. Rather than single persons who are responsible for several subparts of the system, one or two operators supervise the entire system. The operator monitors (a large part of) the system and intervenes when a non-normal situation occurs.

Automation undeniably has a number of advantages. Computers can, for example, assist the human operator in highly complex tasks, such as mathematical operations. However despite these opportunities, automation has also resulted in major problems. One of the main problems of automation mentioned with regard to supervisory control is that operators experience a greater distance to the system (Wiener, 1985). Because of automation, the operator has less overview of what exactly is going on at a detailed level. In case an intervention is required it may be difficult for the operator to build up an accurate model of the situation (Sheridan, 1988).

Intervention is mostly required when a disturbance occurs. A disturbance is a non-routine situation and because subsystems are often related, a disturbance is likely to propagate through the system. As a consequence, the operator faces a complex situation, where there are multiple disturbances to be dealt with. The system deteriorates fast and since the consequences are often very negative (think of a power station) the operator experiences time pressure.

In complex situations where operators have to deal with more than one disturbance at the same time the dynamics of the system are highly important. Due to the system dynamics the system can change in such a way during fault handling that in other parts of the system disturbances occur that are more urgent. In that case the operator should interrupt his current activities to switch to a disturbance with a higher priority. However, operators are often reluctant to stop ongoing activities and to focus on another task.

Moray and Rotenberg (1989) used the term 'cognitive lockup' for this phenomenon, which they referred to as the tendency to focus on only a limited part of the system while ignoring the rest of it. Operators tend to zoom in on a single fault and do not provide enough attention to other parts of the system. However, when during the handling of this fault other faults emerge that appear to be more urgent they remain unattended.

There are many examples, especially in shipping and aviation, where, in hindsight, accidents could be ascribed to cognitive lockup. One such example is the air crash of flight 173 of United Airlines in December 1978. The captain of the flight focused entirely on a landing gear malfunction and ignored monitoring the aircraft's fuel state. As a consequence, the captain did not react to a decreasing fuel supply. The plane ran out of fuel and crashed (National Transportation Safety Board, 1978).

In addition, cognitive lockup has been found in experimental studies where the experimental task consisted of a simulation of a supervisory control system. Moray and Rotenberg's data (1989) showed that participants were reluctant to attend to additional faults. This finding was replicated by Kerstholt, Passenier, Houttuin and Schuffel (1996) who found that the first fault received attention much sooner than subsequent faults.

Even though cognitive lockup has been successfully demonstrated in experimental studies, no founded explanations for this phenomenon have been provided to date. A possible explanation that has a strong intuitive appeal is that of human capacity limitations. People are simply unable to deal with more than one task at the same time and consequently will attend to the environment only after the task has been ended.

This explanation has particularly been investigated in the PRP (Psychological Refractory Period) paradigm where two simple S-R tasks overlap: the second stimulus is presented after the first stimulus but before a response is given. The typical finding in this paradigm is that the second task overlapping the first task takes more time to finish as compared to the situation where the second task is executed in isolation.

Within the PRP research, researchers have distinguished two different notions for this so called PRP effect. The 'classical' notion assumes that processing the first task fully "captures" human processing capacity so that processing the second task can only begin after the first one is terminated. The claim on the information processing system made by the first task prevents the second task from being attended, at least till the moment the first task is completed and the information processing system is no longer occupied. As a result, the reaction time of the second task increases.

A more recent notion assumes that there is also a strategic component involved in dual task performance. That is, people seem to have flexible control over the order in which tasks are executed. De Jong (1995), for instance, showed that under some conditions participants were able to process the second stimulus presented first. Other studies, also conducted within the same PRP paradigm, provided additional evidence for flexible control over the moment a second task is processed (Meyer and Kieras, 1997; Meyer, Kieras, Lauber, Schumacher, Glass, Zurbriggen, Gmeindl and Apfelblat, 1995 and Schumacher, Lauber, Glass, Zurbriggen, Gmeindl, Kieras and Meyer, 1999).

These two notions on people's capability to handle tasks simultaneously have been investigated with simple reaction-time tasks. For more complex tasks such as supervisory control this issue has never been addressed. The purpose of the present research is therefore to investigate which of the two notions can explain the findings of cognitive lockup in supervisory control tasks. Are operators unable to deal with more than one fault simultaneously due to capacity limitations, or do they make a deliberate choice not to deal with a second fault before ending the first fault?

We addressed the first explanation - cognitive lockup is due to capacity limitations - in the second experiment of this thesis. In this experiment we manipulated the mental workload of the first fault by increasing diagnosis complexity. A limited capacity explanation implies that cognitive lockup occurs when the processing system is saturated. In this line, it is expected that as the complexity of the first fault increases, the system will be closer to saturation and less 'residual' resources will be available for dealing with subsequent faults. As a consequence, the increase of the first fault's complexity results in a stronger tendency to deal with faults in a strictly sequential manner.

The second explanation - a controller that deliberately allocates resources to a specific task - is investigated in the third experiment. According to this notion, the controller makes a trade-off between the anticipated costs and benefits of reassessing the situation the moment a second disturbance occurs. The controller decides on the basis of the outcome of this trade-off whether to reallocate attention or not. So, operators refrain from a reassessment not because they are not able to, but rather because they think the costs of a reassessment do not outweigh the benefits. To test this hypothesis we decreased the costs of reassessing the situation in order to make the outcome of the trade-off clearly in favour of reassessment. If operators indeed make a trade-off between the anticipated costs and benefits, the tendency for cognitive lockup will decrease as the costs of a reassessment are less than the benefits.

Summarized, in the present chapter we will investigate two different explanations for cognitive lockup: limitations in human information-processing capacity and a controller that refrains from allocating resources to the reassessment of priorities because he considers the costs of reassessment too high. In order to test these explanations we used the simulation of the shipping control task.

Experiment 2

In this experiment we investigated whether cognitive lockup can be explained in terms of limited information-processing capacity by manipulating the complexity of the diagnosis process. Increasing the complexity of diagnosing the cause of a fire places a higher claim on the human information processing system. As a consequence, we expected that there would be less residual capacity to deal with additional fires, increasing cognitive lockup.

The most interesting scenarios in this respect are those where a fire starts at the point the operator is already involved in a diagnosis process. In these cases the human information processing system is occupied the moment a second fire starts. If people's tendency to deal with faults sequentially can indeed be ascribed to a limited information processing human capacity, we would expect cognitive lockup to increase when the diagnosis process becomes more complex.

This in contrast with scenarios where two fires start simultaneously. In these cases the operator can first assess priorities and set out a strategy before entering the diagnosis process. In these cases, we would not expect any effect of complexity of the diagnosis process on cognitive lockup.

Method

Participants

Twenty-seven participants voluntarily participated in the experiment. They were all first year students at the University of Utrecht. The experiment lasted about two hours and they were paid Dfl. 70 (approximately € 32).

Experimental task

The same experimental task was used as in the first experiment

Procedure

The procedure was the same as in the first experiment.

Design

A 2 * 3 (Presentation mode, Complexity) factorial was used. Both factors were manipulated within subjects. The experiment was made up of three blocks, each block representing a different level of complexity. The order of blocks was counterbalanced across participants. Participants were given 5 minutes rest between blocks. One manipulation in this experiment was the presentation mode of two fires: a sequential presentation and a simultaneous presentation. The second manipulation was the complexity of the diagnosis process, which comprised three conditions. In the least complex condition, one tree of questions had to be walked through to determine the appropriate treatment. In the middle condition, there were two different trees: one for fires starting at the two upper decks and one for fires starting at the two lower decks. In the most complex condition there were four different trees, one for each separate deck.

The resulting design consisted of 6 cells each containing 18 scenarios. In each cell there was an equal number of scenarios where the first fire had priority, the second fire had priority and the first and second fire had equal priority. (Note that in the simultaneous condition there was no 'first' or 'second' presented fire, since both fires started at the same time.) The cells were also balanced for the number of questions that needed to be answered in order to determine the correct treatment for a fire.

108 scenarios were constructed in this way. In order to prevent participants from anticipating a second fire, 54 scenarios were added in which only one fire occurred.

Dependent variables

As in the previous experiment the following variables were measured:

- Performance: *number of burn downs;*
- Strategy: *moment of switch* and *request of priority information*

Results

Performance

Figure 3.1 presents the mean percentage of scenarios that ended in a burn down for the sequential and simultaneous scenarios as a function of task complexity.

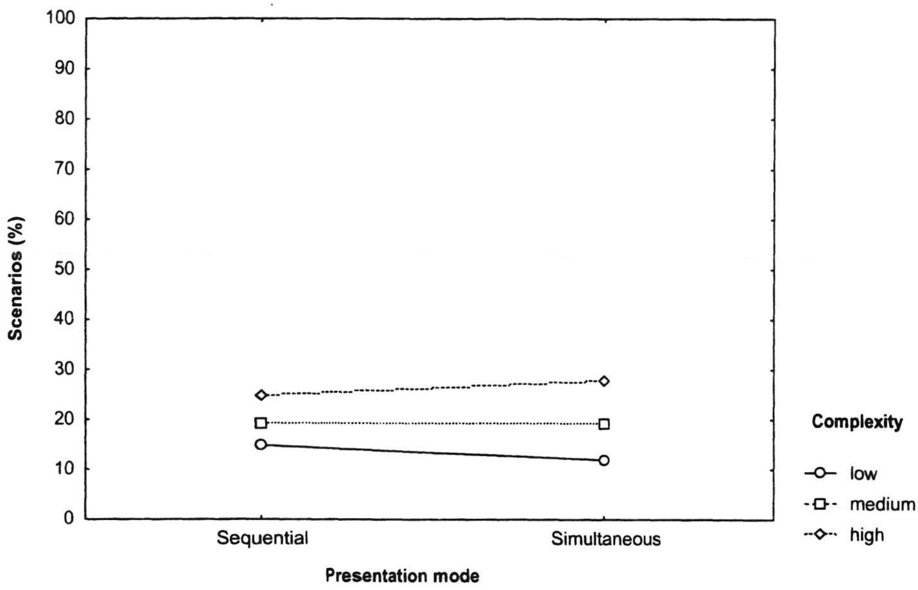


Figure 3.1: *Mean percentage of scenarios that ended in a burn down as a function of presentation mode and task complexity.*

There was no effect of presentation mode ($F(1,26) < 1$). There was an effect of complexity: as the complexity of the diagnosis process increased, more scenarios ended in a burn down ($F(2,52) = 8.54$, $p < 0.01$).

There was no interaction between Presentation Mode and Complexity ($F(2,52) < 1$), implying that for all levels of task complexity the number of burn downs was equal for the sequential and simultaneous condition. An increasing level of task complexity did not have a stronger effect on performance in sequential scenarios.

Strategy

Figure 3.2 shows the percentage of scenarios in which the second fire was detected after the correct treatment of the first fire had been selected.

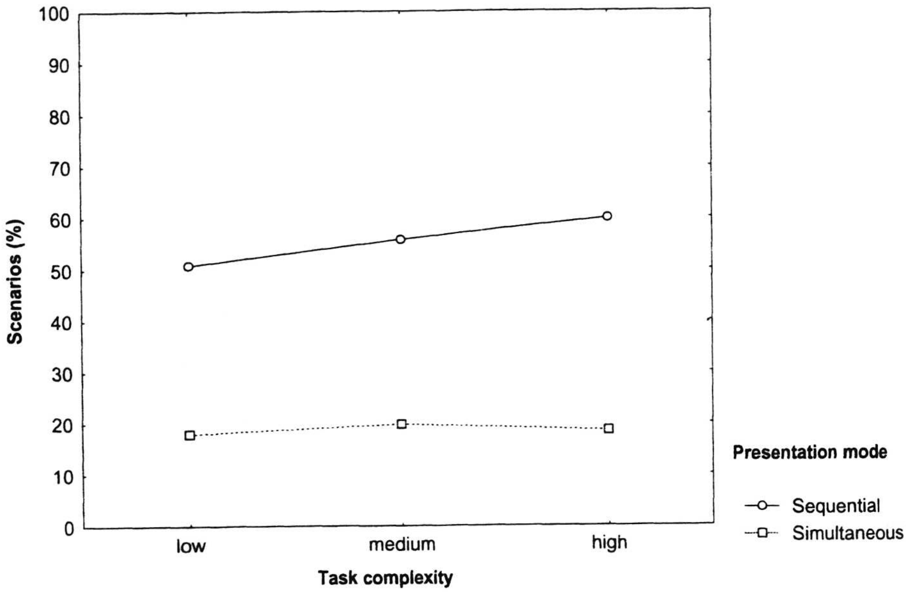


Figure 3.2: *Mean percentage of scenarios in which the second fire was detected after completion of the first fire as a function of task complexity and presentation mode.*

The figure shows a difference in strategy for the sequential and simultaneous condition ($F(1,26) = 40.76$), $p < 0.01$). For the sequential scenarios the second fire was detected in most cases after the first one was completed. For the sequential scenarios the second fire was detected *before* the first fire in half of the cases and *after* the first fire in the other half of the cases. In other words, the tendency to handle faults one after another is much stronger in the sequential condition.

There was no effect of Complexity ($F(2,52) < 1$) nor an interaction effect between Complexity and Presentation mode ($F(2,52) < 1$). This implies that an increasing degree of complexity did not have any effect on participants' strategy.

Figure 3.3 presents for each level of complexity the percentage of scenarios priority information was requested for the first fire.

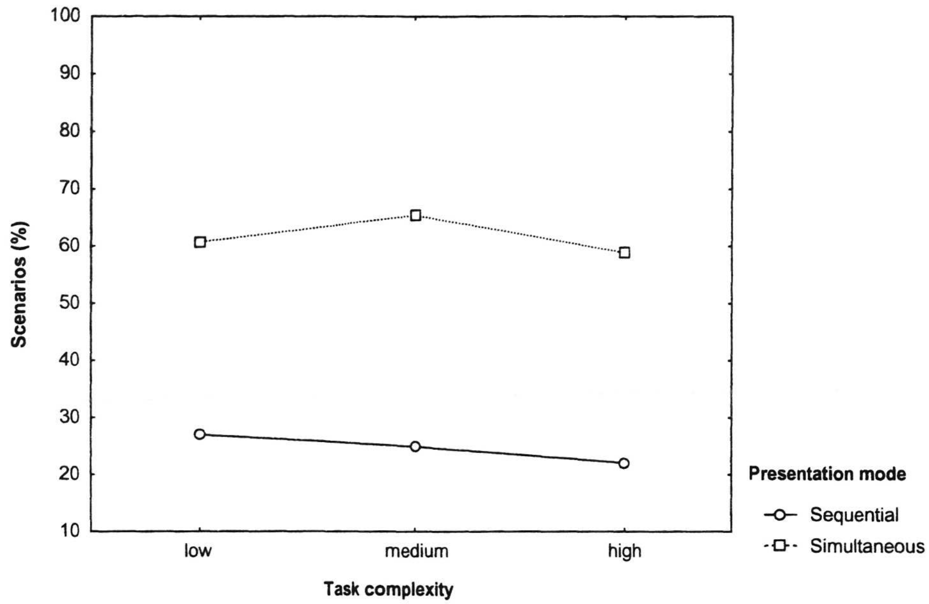


Figure 3.3: *Mean percentage of scenarios in which priority was assessed for the first fire detected as a function of task complexity and presentation mode.*

In the sequential scenarios participants requested significantly less priority information than in the simultaneous scenarios ($F(1,26) = 25.74$, $p < 0.01$).

There was no effect of complexity ($F(2,52) < 1$) nor an interaction between complexity and presentation mode ($F(2,52) < 1$). For each level of complexity much more priority information was asked in the simultaneous condition than in the sequential condition.

Discussion

The purpose of this experiment was to investigate whether cognitive lockup could be explained in terms of limited information-processing capacity. The workload of dealing with the first disturbance is assumed to be so high that no resources are left for reassessing the situation. In this experiment we varied the claim on the operators' information-processing capacity when dealing with the first fault. We did so by manipulating the complexity of the diagnosis process. We reasoned that if people's tendency for cognitive lockup could be ascribed to limited information-processing capacity, higher levels of complexity would result in more instances of cognitive lockup.

The data showed however that an increasing level of complexity had no effect on the moment the second fire was detected nor on the frequency of priority assessment in either mode of presentation. Since different levels of complexity in diagnosing the first fire didn't affect participants' tendency for cognitive lockup, we can not attribute the lockup phenomenon to limits in the human information processing system.

As in the first experiment we found that a difference in strategy between the sequential and simultaneous condition was not reflected in the performance data. We ascribed this effect to the fact that the assessment of priorities required more time than we had expected so that the assessment of priorities did not always outweigh the information that was obtained. Task complexity did have an effect on the performance data. This finding can be considered as a manipulation check. Since more complex tasks resulted in more burn downs, we may assume that more complex tasks did indeed demand more resources.

In all, the experiment replicated people's tendency for cognitive lockup that was found in the previous experiment. However this tendency can not be explained in terms of limitations in human information-processing capacity. Therefore, we turned to another explanation for cognitive lockup in the next experiment.

Experiment 3

As stated earlier, recent research in the realm of the PRP-paradigm, demonstrated that people are able to process a second task presented first, at least in certain conditions. This finding suggests that people have strategic control over the order of processing. The third experiment was conducted to investigate whether this notion was also valid for more complex tasks such as process control. This would mean that in case of cognitive lockup, operators make a deliberate choice to start with the second fault after having completed the first fault.

In the third experiment of this thesis we investigated the explanation that participants make such a deliberate choice on the basis of a trade-off between costs and benefits. The reluctance to reassess the situation could result from the anticipated costs for reassessing the situation. In our fire control simulation, a reassessment is required the moment a second fire occurred. Participants then are assumed to make a trade-off between the anticipated costs and benefits of priority assessment. If cognitive lockup occurs because the benefits of assessing priority do not outweigh the costs, we expected that as the anticipated costs of assessing priority would decrease, participants would become more inclined to interrupt the ongoing task and enter the task of priority assessment.

In order to test this explanation, we designed an experiment in which we manipulated the costs of priority assessment. We reasoned that if operators indeed make a trade-off between costs and benefits of priority assessment, the tendency to reassess the situation increases when the costs are lowered while the benefits remain the same.

For that reason we replicated the first experiment of this thesis. This condition formed the baseline condition. We added two conditions where the costs of assessing priority were substantially lower. In the second condition the costs were only minimal. Participants needed to click only one single button to assess a fire's priority. In the third condition participants didn't even need to take action themselves. The system reassessed the situation, which implied that there were no costs for the participants. Compared to the baseline condition, we expected a decline of cognitive lockup for the second condition and a further decline for the third condition.

Method

Participants

Thirty participants voluntary participated in the experiment. They were all first year students of the University of Utrecht. The experiment lasted about two hours and participants were paid Dfl. 70 (approximately € 32).

Experimental task

The same experimental task was used as in the previous experiments.

Procedure

As in the previous experiments there were two training sessions: a training session for the assessment of priority and training session for the selection of the correct treatment. Since participants in the Button and Window condition did not need to ask questions to assess priorities, this training session was left out for participants in these conditions.

Design

A 2 * 3 (Presentation, Priority assessment) factorial was used. The first factor was manipulated within subjects and the second factor between subjects. The presentation of fires could either be simultaneous or sequential. There were three between-subjects conditions that each comprised of ten participants: the Questions condition, the Button condition and the Window condition.

The second factor that was manipulated in this experiment was the way priority could be assessed. There were three conditions of Priority Assessment: the Questions condition, the Button condition and the Window condition.

1. Questions: as in the first experiment answering a fire's priority could be assessed by answering questions. A tree-structure determined which questions had to be asked and, at the end of the tree, which priority a fire had. Answers were provided with a one-second delay.
2. Button: a fire's priority was generated by clicking one single button that provided immediate priority information.
3. Window: priority information was presented on a window This window popped up out on the screen the moment two separate fires raged the ship. In this condition priority information was therefore provided when additional fires occurred. By clicking on it the window closed, enabling participants to proceed solving the fires.

Each condition of priority assessment consisted of 72 scenarios: 27 sequential scenarios, 27 simultaneous scenarios and 18 single fire scenarios. The sequential and simultaneous conditions contained an equal number of scenarios in which the first fire had priority, the second fire had priority and the two fires had equal priority. The conditions were also balanced for the number of questions that were required in order to diagnose a fire.

Dependent variables

As in the previous experiments we measured the *number of burn downs, moment of switch* and *request of priority information.*

Results

Again, we subdivide the results in two sections: in the first section performance data are provided and in the second section strategy data.

For each condition of priority assessment, participants had 35 seconds for the high priority fire, 50 seconds for the low priority fire and 50 seconds for both fires when priorities were equal. The assessment of priorities required more time in the Questions condition than in the Button and Window condition. Because the assessment of priorities takes relatively little time in the Button and Window condition, more time is left for the diagnosing process and, as a consequence, more scenarios will be completed successfully. In other words, the burn downs in the Questions condition are disproportional high. In order to make the three conditions comparable for the number of burn downs, we computed the overall time on priority assessment in the Questions conditions and added this amount to the total time participants spent on each trial in the Button and Window condition. If for a scenario the total time exceeded the time limit, it was counted as a burn down.

Performance

Figure 3.4 presents the percentage of burn downs for the three different conditions of priority assessment for both the sequential and simultaneous scenarios.

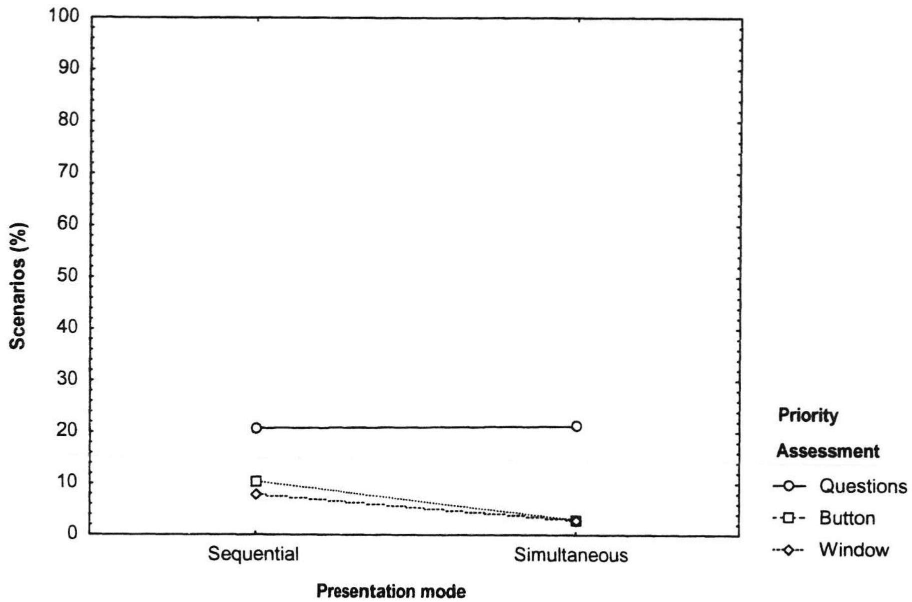


Figure 3.4: *Mean percentage of scenarios that ended in a burn down as a function of presentation mode and priority assessment.*

An ANOVA showed that there was a main effect of Presentation Mode ($F(1,27) = 6.27$, $p < 0.05$). Overall, in the sequential condition more scenarios ended in a burn down than in the simultaneous condition.

There was also a main effect of Priority Assessment ($F(2,27) = 14.33$, $p < 0.01$). A post hoc analysis revealed that in the Questions condition significantly more scenarios ended in a burn down than in the Button and in the Window condition ($p < 0.01$). A decrease in costs for the assessment of priorities resulted in better performance.

Though the figure shows a different pattern for the Questions condition in comparison with the Button and Window condition, there was no interaction between Priority Assessment and Presentation Mode ($F(2,27) =$

2.10, p>.1). A reduction in costs did not have a differential effect in the two conditions of presentation.

Strategy

It was recorded whether the second fire was detected before or after completion of the first fire and how much priority information was requested.

Figure 3.5 presents the percentage of scenarios in which the second fire was detected after the first fire was completed. This was done for the sequential and simultaneous scenarios, and for the different conditions of priority assessment.

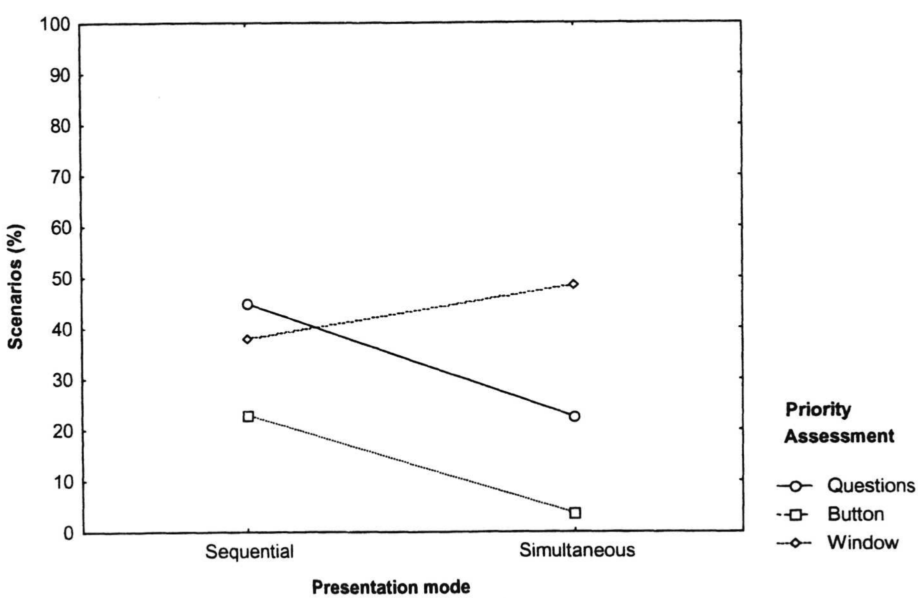


Figure 3.5: *Mean percentage of scenarios in which the second fire was detected after completion of the first fire as a function of presentation mode and priority assessment.*

First, there was a main effect of Presentation mode ($F(1,27) = 5.43$, $p < 0.05$). In the sequential condition there were more scenarios in which the second fire was detected after completion of the first fire than in the simultaneous condition. So, overall, there was more cognitive lockup in the sequential presentation mode than in the simultaneous presentation mode.

Second, there was no significant effect of Priority Assessment ($F(2,27) = 2.30$, $p > 0.1$). However, there was a significant interaction between the different ways in which priority could be assessed and the presentation mode ($F(1,36) = 3.82$, $p < 0.02$). The effect that in sequential scenarios participants more often detected the second fire after completion of the first fire than in simultaneous scenarios was present for the Questions condition ($F(1,9) = 5.26$, $p < 0.05$) and the Button condition $F(1,9) = 11.08$, $p < 0.01$) but not for the Windows condition ($F(1,9) = 2.47$, $p > 0.1$). There is, in other words, a tendency for cognitive lockup in the Question and Button condition, but this tendency is absent in the Window condition.

The second variable indicating participants' strategy was the percentage of scenarios for which priority information was requested for the first fire detected. Figure 3.6 presents the percentages for two different conditions of priority assessment. Since participants in the Window condition did not need to request priority information actively, this variable could only be registered for the Questions and Button Condition.

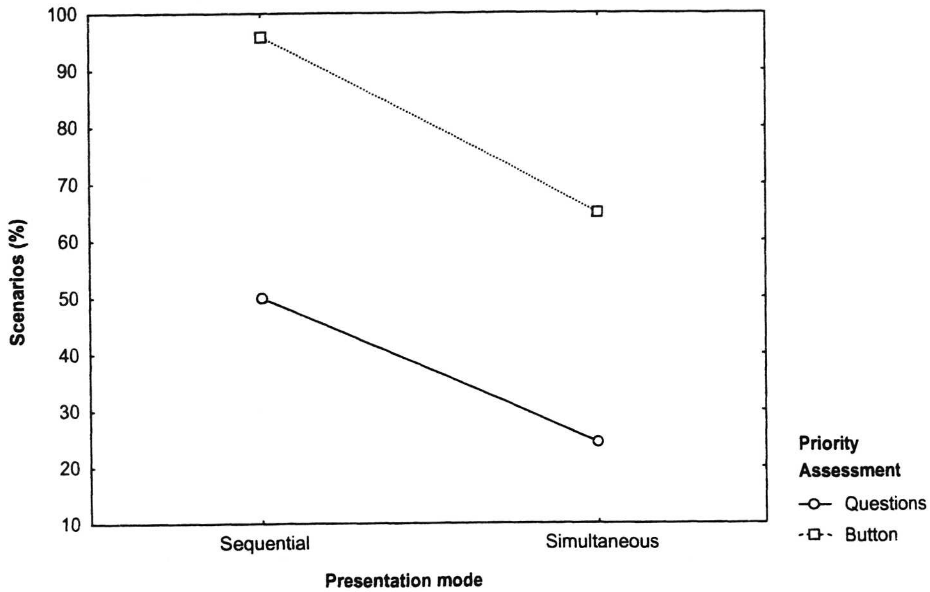


Figure 3.6: *Mean percentage of scenarios in which priority was assessed for the first fire detected as a function of presentation mode and priority assessment.*

Again, there was a main effect for Presentation Mode ($F(1,18) = 15{,}50$, $p < 0.01$). In the simultaneous condition more priority information was requested than in the sequential condition. There was also a main effect of Priority Assessment ($F(1,18) = 15.25$, $p < 0.01$) implying that in the Button condition significantly more priority information was requested than in the Question condition.

There was no interaction effect between Priority Assessment and Presentation Mode ($F(1,18) < 1$). The reduction of costs from asking multiple questions to clicking a single button had the same effect on either mode of presentation: a decreasing tendency to process fires in a serial way. This implies, as can also be gathered from figure 3.6, that for the Button-condition participants' tendency to process fires sequentially and to refrain from assessing priorities is still more prominent in the sequential condition.

Discussion

The purpose of the third experiment was to investigate whether cognitive lockup could be explained by a controlling function that allocates attention on the basis of a trade-off between costs and benefits of reassessing the situation. Following this line of reasoning, cognitive lockup is due to a perception of high costs. To test whether participants refrain from reassessing the situation because they consider the costs too high, we manipulated the costs that accompany the reassessment. We hypothesized that if participants indeed make a trade-off between costs and benefits of a reassessment, a decline in costs would result in a weaker tendency for cognitive lockup.

The data of the third experiment provided support for the notion that the costs of reassessing the situation affect operators' switching behavior. When the costs were reduced to clicking a single button, cognitive lockup decreased as well. In case there were no costs and the system reassessed the situation for them, cognitive lockup appeared to be absent. So it seems that operators indeed make a trade-off between the costs and benefits of reassessing the situation. A decrease in costs results in an outcome of this trade-off that is more in favor of reassessing the situation.

A closer look at the separate conditions of priority assessment provides some additional information on cognitive lockup. First, the effect of cognitive lockup in the first and second experiment was replicated in the present experiment for the condition where priority information could be assessed by asking multiple questions. Again, in the sequential scenarios, operators showed a stronger tendency to deal with faults sequentially than in simultaneous scenarios. Cognitive lockup in other words, is a robust finding.

In case the costs of assessing priorities merely consisted of clicking a single button, we found a decline in the number of scenarios in which operators dealt with fires in a sequential order (even though we still found some degree of cognitive lockup in the sequential scenarios). Apparently, operators decide to reassess the situation only when the benefits clearly outweigh the costs. Evidently, there is an asymmetry between the way the costs and benefits are evaluated.

The finding that participants sometimes incorrectly estimate certain costs is in line with earlier findings reported by Kerstholt (1994). In dynamic environments such as process control, people generally use a strategy to request information rather than apply actions, also in conditions where an action-oriented strategy is optimal. In another study, Kerstholt (1995) found that people can use a action-oriented strategy but only when the costs of information clearly outweighed the costs of applying actions.

It is an interesting question why operators still do not reassess the situation when the costs for a reassessment are lower than the benefits. There are a number of plausible reasons. One possible reason is people's difficulty in estimating time durations. The costs and benefits in the present task are expressed in terms of time. Reassessment of the situation costs valuable time, but there are also benefits in terms of time. A reassessment of the situation provides information about a possible change in priorities. Participants can apply this information to effectively rearrange their available time. However, since people find it difficult to estimate time durations, they can not adequately make a trade-off between costs and benefits.

Another possible reason is that participants may be aware that, with the present design, the odds are only one against three that the second fire has higher priority. Participants may be less prepared to spend effort into reassessing the environment when the probability that it pays off is only

0.33. In general, people choose a strategy so that a sufficient level of accuracy is reached for the lowest possible level of mental effort (Payne, Bettman and Johnson, 1988). In other words, because there is a high chance that one's effort will be in vain, one is less prepared to invest energy in reassessing the situation.

For the other condition of priority assessment, there are practically no costs to the assessment of priorities. In this condition, in which participants are forced to interrupt their ongoing activities in order to take notice of priority information, cognitive lockup is absent. The fact that participants in this condition do not return to the diagnosis process of the first fire when this fire has low priority suggests that they do not strictly hold on to the first fire. The moment they are detached from the diagnosis activities of the first fire, they are able to make a new decision that takes into account the change in the environment.

General Discussion

In supervisory control tasks, human decision making can become very complex when disturbances occur. Since disturbances have to be dealt with within a certain time limit and the consequences of exceeding this limit are often dramatic, there is a very high level of time pressure. Moreover, because disturbances often propagate through the system, operators have to deal with multiple tasks at the same time. Accident analyses have reported that in these critical situations operators have the tendency to focus on a single disturbance and ignore the rest of the system. A consequence of cognitive lockup is that other, more urgent disturbances remain unattended, resulting in a break down of the system.

The rationale of the research in this chapter was to investigate the reasons why operators are locked up in a subpart of the system. We suggested two possible explanations. One explanation was that operators lack sufficient information-processing capacity to deal with subsequent disturbances. The other explanation was that cognitive lockup is the outcome of a trade-off between anticipated costs and benefits of reassessing the situation. The former explanation was investigated in the second experiment and the latter explanation in the third experiment.

The data provided evidence for only the latter explanation. The second experiment showed that a higher workload, realized by an increase of complexity, did not result in a stronger tendency for cognitive lockup. In the third experiment cognitive lockup appeared to be affected by the anticipated costs of a reassessment. When we lowered these costs, cognitive lockup decreased substantially. This implies that - when confronted with a second fault – operators make a trade-off between the costs and benefits of a reassessment.

Cognitive lockup occurs because operators perceive the costs of a reassessment too high relative to the benefits. Operators seem to be able to break through the tendency for cognitive lockup when the benefits of a reassessment are very high relative to the costs. Nevertheless, in the sequential scenarios cognitive lockup is still present whereas in the simultaneous scenarios participants practically always decided to reassess the situation, even when the priorities are absolutely clear.

Why do operators in sequential scenarios still decide not to reassess the situation when the benefits are high and the costs low? In other words, what drives operators to continue with the first disturbance in spite of the fact that they know it is better to reassess the situation? The answer may be found in the realm of human decision making.

In decision making literature there is a similar class of phenomena, all reflecting people's tendency to stick to their initial plan, even if the outcome of that plan is clearly negative (e.g. the sunk cost bias and escalation of commitment). A number of explanations have been suggested to explain these phenomena (e.g. Arkes and Blumer, 1985; Brockner, 1992; Kahneman, Knetsch and Thaler, 1990). However, these explanations have been investigated with static, highly hypothetical scenarios. In the next chapter we will investigate to what extent these explanations can account for cognitive lockup as observed in a dynamic supervisory control task.

The results of the experiments in this chapter have important implications for the design of decision support systems. To date, most support systems are constructed to relieve operators from a high workload. The overall idea is that operators make mistakes because of limitations in their cognitive capacity. The data of the experiments make clear that sub-optimal performance due to cognitive lockup can not be overcome by reducing workload. Operators do not refrain from reassessing priorities because they have reached their limits of information-processing capacity.

A decision support system that seems more productive is one that makes the costs and benefits of switching more explicit. At the moment a second fault turns up, operators make a trade-off between the costs and benefits of making a reassessment of the situation. These costs and benefits are not apparent in most supervisory control tasks. Operators often lack information concerning the costs and benefits of a reassessment, which makes it difficult for them to make a trade-off. Therefore operators should be assisted assessing the costs and benefits of a reassessment in order to make an accurate trade-off.

To conclude, in order to prevent operators from becoming locked up in a subpart of an automated system, we that recommend designers of decision support tools to change their focus from relieving cognitive load to providing operators with means to facilitate a reassessment of the situation.

4. Behavioral Entrapment in a Dynamic Environment: Sunk Costs or Task Completion?

In decision making literature behavioral entrapment has primary been explained in terms of sunk cost but recent studies have shown that task completion can provide an alternative explanation. Both explanations were examined in the context of the real time simulation of a fire control task. Participants were required to handle multiple fires that occurred sequentially. Results of the fourth experiment showed a reversed sunk cost effect that we ascribed to high subjective time pressure. In a fifth experiment we added a static condition in order to identify the attribution of a real time component. Behavioral entrapment appeared to be stronger in static scenarios, probably because participants lacked the opportunity to adjust their strategy. In the static condition behavioral entrapment could be explained by task completion. In the dynamic condition there was a reversed sunk cost effect but only when the task was not near completion.

Introduction

People have a tendency to stick to their initial plan even if a change in the environment would require a revision. Several phenomena have been identified that reflect this behavioral entrapment, such as escalation of commitment (Staw and Ross, 1989), the sunk cost effect (Arkes and Blumer, 1985) and task completion (Garland and Conlon, 1998).

Hitherto, people's reluctance to switch to an alternative course of action has generally been explained in terms of amount of investment. The sunk cost effect, for example, indicates that there is a greater tendency to pursue a course of action when investments are made, such as time, money or effort, even when these costs are irrelevant to the current decision (Arkes and Blumer, 1985). Thaler (1980) illustrates the sunk cost effect as follows: a family pays $40 for tickets to a basketball game to be played 60 miles from their home. On the day of the game there is a snowstorm. They decide to go anyway, but note in passing that had the tickets been given to them, they would have stayed home. Even though the correct trade-off should just involve the costs of defying the snowstorm versus the pleasure of the game, people do take prior investments – the costs of the tickets – into account.

Several psychological mechanisms have been suggested for the sunk cost effect, either motivational – e.g. a desire not to appear wasteful (Arkes and Blumer, 1985) or cognitive – e.g. risk seeking behavior in the domain of losses (Whyte, 1986). Almost all mechanisms for behavioral entrapment focus on sunk costs. Recent findings, however, point to explanations in terms of termination of a course of action rather than prior investments.

Boehne and Paese (2000) pointed that out that the degree of investments in a course of action is confounded with the degree of completion. Putting money or effort into a project not only implies that investments are made, but also that the project comes closer to completion. About the effects observed in studies on prior investments, the authors state, "were due to project completion rather than sunk costs and any attempt to explain these results in sunk-cost terms is therefore moot"(p.179).

Conlon and Garland (1993), Garland and Conlon (1998) and Boehne and Paese (2000) therefore conducted experiments in which they disentangled investment and completion. They crossed small versus large sunk costs with low versus high project completion and only found evidence for the completion factor, not for sunk costs. In addition, Moon (2001) found evidence for both task completion and sunk costs the latter being present only when the task was nearly completed. In all, it can be concluded that there is no agreement concerning the psychological mechanism underlying the tendency to stick to a course of action: sunk costs or task completion.

All studies on behavioral entrapment used scenarios that were highly hypothetical. As also noted by Boehne and Paese (2000), this seems to be a serious limitation. As they argue "real-world investment situations are likely to be more involving"(p. 192) than decision scenarios where participants have to imagine a previous decision (for example, having invested 10 million dollars into a research project for constructing a plane) and to reconsider the decision after new information has come up (another firm can build a better plane)'. On the one hand, this involvement concerns the personal value of the decision, a motivational limitation inherent to laboratory studies. On the other hand, persons are also less involved in the decision process as they have to imagine the environmental change, rather than experiencing it.

The general procedure in experiments on behavioral entrapment is that participants are given a description of a project in which they have invested time, money or effort. At a certain point in time the situation changes. Based on new information participants have to make a decision whether to continue investing in the project or not. Although the scenario includes history information, the dynamics of the task are not taken into account explicitly. Participants have to provide a reaction to an environmental change, but the dynamic development of the situation has to be imagined, rather than experienced.

Research with dynamic tasks has also demonstrated people's tendency to stick to a course of action. An example of a dynamic environment is supervisory control. In supervisory control tasks an operator has to monitor a system and has to intervene whenever disturbances occur. The main conclusion from studies that have examined behavioral entrapment was that disturbances were handled in a strict sequential order (Kerstholt, Passenier, Houttuin and Schuffel, 1996; Kerstholt and Passenier, 2000 and Meij and Kerstholt, submitted). Only after a disturbance had been dealt with, participants assessed the situation again. As a consequence the overall state of the system was not noticed during fault handling which could result in negative consequences. Most importantly, these studies show that participants reacted inadequately to environmental changes.

Whereas behavioral entrapment in decision making research has generally been explained in terms of sunk costs or, recently, task completion, in dynamic environments like process control, this tendency has mostly been explained by limited attention (Moray and Rotenberg, 1989). All resources are needed for handling the first fault. In a recent study, however, Mey and Kerstholt (submitted) showed that the tendency to deal with faults sequentially was not affected by workload, but rather by the anticipated cost of reassessing task priorities. This would suggest that operators make a deliberate trade-off between the costs and benefits of a reassessment. However, it still has to be examined which factors affect this trade-off. Perhaps in these supervisory control tasks factors like sunk costs or task completion do play a role as well. In the present chapter we will investigate behavioral entrapment in the dynamic environment of the fire control task.

The purpose of the fourth experiment was to investigate whether behavioral entrapment in a dynamic task environment can be explained by prior investments or by expectations about the future. The purpose of the fifth experiment was to investigate to what extent the inclusion of a real time component in a task attributes to behavioral entrapment. Is behavioral entrapment in a dynamic task environment stronger than in a static environment? And, can behavioral entrapment in a dynamic condition be ascribed to the same psychological mechanism as in a static environment?

Experiment 4

The aim of this experiment was to examine the influence of sunk costs and task completion on behavioral entrapment in a dynamic environment. We manipulated the degree of investment and task completion independently.

Method

Participants

Twenty-four participants voluntarily took part in the experiment. They were all first year students at the University of Utrecht. The experiment lasted about one hour and a half and participants were paid Dfl. 70 (approximately € 32).

Experimental task

The experimental task was identical to the fire control task used in the previous experiments with a few exceptions.

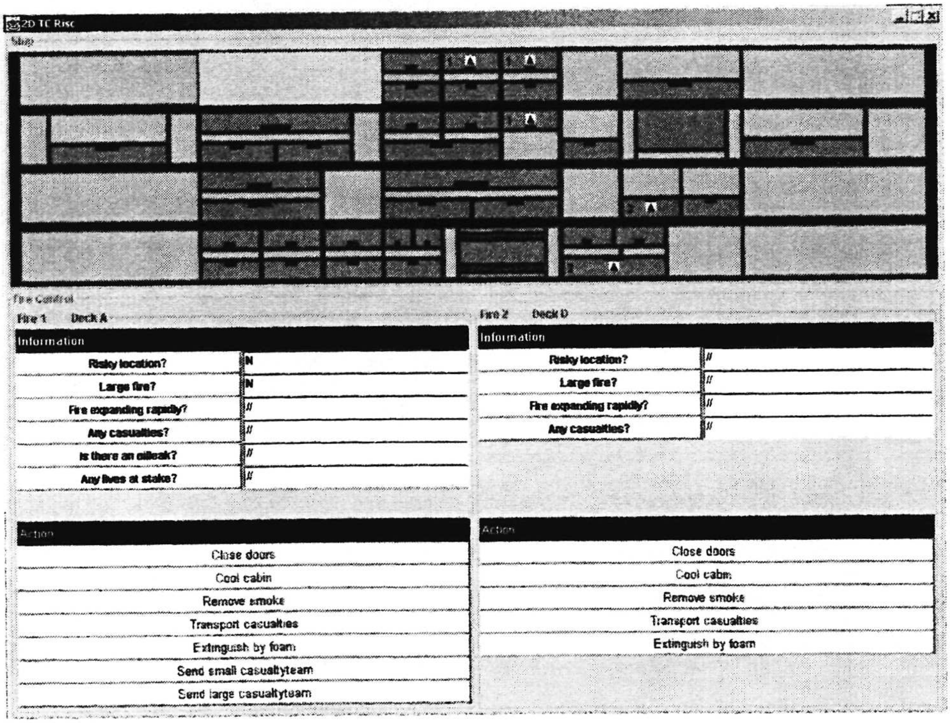


***Figure 4.1**: An overview of the system at the moment two fires have been detected.*

As in the previous experiments, there were windows available in the subpart of the screen for the purpose of fire fighting (see figure 4.1). These windows showed questions that could be asked in order to select the correct action. The number of questions that was presented in the lower part of the screen ranged from four to six. This number constituted the maximum number of questions. Participants had to work down the row of questions until a conclusive answer was provided. Participants didn't know in advance how many questions had to be answered.

Maximum problem solving time was dependent on the number of questions that had to be requested. In case the maximum number was four, participants had 30 seconds to solve the fire, in case of five questions they had 35 seconds and in case of six questions they had 40 seconds. (Appendix C presents the tree structure to determine the appropriate action in case the maximum number of questions to be asked is four.) Answering a question closed down the system for 4 seconds.

At several points during fire handling a second fire symbol could pop out somewhere on the ship. This fire symbol could either be a rapid evolving fire (to be solved within 15 seconds) or it could be a false alarm. In case of a real fire, this fire always had priority meaning that this fire had to be dealt with first. To find out whether the symbol was a fast spreading fire or a false alarm, participants could click the symbol. If it represented a false alarm, the symbol simply disappeared from the screen. If it indeed represented a fire, a list of four questions was presented on the right part of the lower part of the screen. The structure of the question-and-answer tree was identical to the trees for the solving of the first fire in case of four questions, with the exception that the time delay was one instead of four seconds.

Procedure

Before the actual experiment, participants were trained in the selection of the correct treatment. Participants were handed out three trees of questions that could help them to ask the relevant questions and determine the correct treatment. There was one tree in case of a maximum of four questions, a tree for a maximum of five questions and a tree for a maximum of six questions. Participants were allowed to use these trees throughout the experiment.

After the training-session, participants were given the instructions for the experimental task. Before the experiment started, participants received nine scenarios for practice purposes.

Design

Behavioral entrapment was operationalized as detecting the second fire after completion of the first fire. So, for each trial we recorded whether participants detected the second fire before or after completion of the first fire.

A 2 * 2 factorial was used. Both factors - investment and completion – were manipulated within subjects. The level of investment was equal to the number of questions that had been answered at the moment a second fire started. A second fire started after either one or two questions had been answered. The level of completion was equal to the maximum number of questions that still needed to be answered in order to select the correct action. Either three or four questions still needed to be answered when a second fire occurred. To accomplish sufficient uncertainty, we added some additional scenarios that did not fall in any condition of the experimental design. First, we added a number of scenarios without additional fires. Second, we included a number of scenarios in which the second fire occurred at different points in time.

In total, there were sixteen scenarios with two fires that we expanded with twenty-one filler scenarios. For the analysis we only used the sixteen scenarios that were part of the experimental design.

Results

Figure 4.2 shows the mean of scenarios in which percentage the second fire was detected after the first fire was completed.

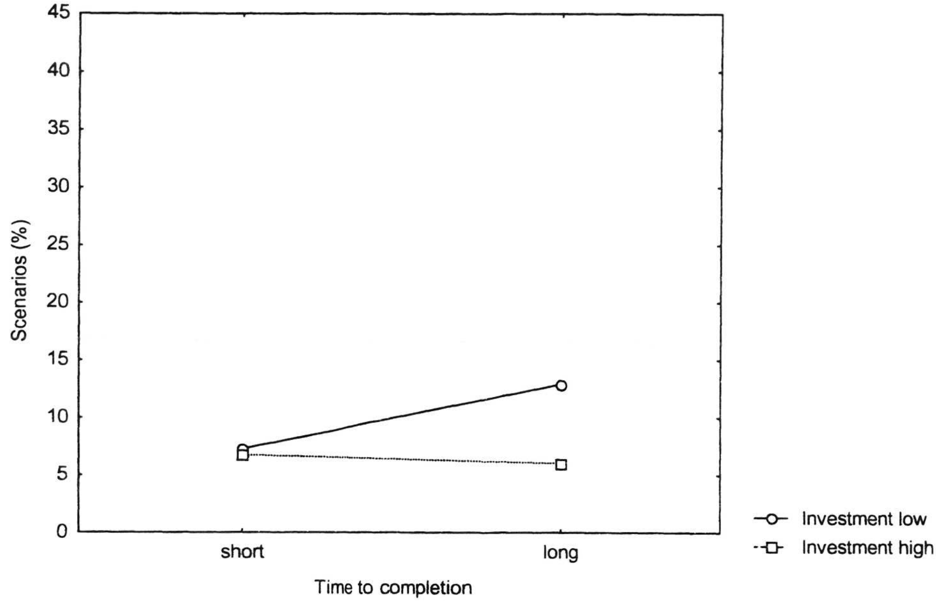


Figure 4.2: *Mean percentage of scenarios in which participants detected the second fire after completion of the first fire as a function of investment and completion*

There was a main effect of investment, which was in the opposite direction as we expected ($F(1,23) = 9.68$, $p < 0.01$). This implies that the inclination to detect the second fire only after completion of the first fire, became less strong when more prior investment were done.

There was no effect of task completion ($F(1,23) = 1.65$, $p > 0.1$) nor an interaction between investment and task completion ($F(1,23) = 1.79$, $p > 0.1$).

Discussion
Results showed that - in a dynamic environment - participants do not seem to have a strong inclination for behavioral entrapment. In most cases they detected the second fire before determination of the first fire. Moreover, prior investments nor the degree of task completion can account for behavioral entrapment. For prior investments we even found an effect in the opposite direction, implying that people are more inclined to abandon an ongoing task when more investments have been made.

This reversed sunk cost effect may be explained by the perception of time pressure. It is a general characteristic of dynamic tasks that time is often limited and as the task develops there is less time available and time pressure increases. Although the available time to complete fires is equal for both conditions, participants in the high investment condition may perceive time pressure to be higher than in the low investment condition. This because the available time in relation to invested time is lower in the high investment condition than in the low investment condition. As a consequence, participants in the high investment condition may be more inclined to switch.

The present findings suggest that behavioral entrapment is less prominent in scenarios that include a real time component. However, a closer look at the data revealed that in 25,6% of the scenarios where there actually was a second fire, participants chose to solve the first fire first and then to switch to the second fire. In other words, in 74,4% of the cases participants made a switch to the fire symbol to check whether the symbol represented a fire or a false alarm. And, when there actually appeared to be a second fire, they did not always continue with that fire. It seems that at the moment they detected the fire symbol and knew the symbol indeed represented a fire, they made a new decision whether to continue with the first fire or not. In making this decision they tended to finish the first fire, in spite of the fact that the second fire always had a higher priority.

Since problem solving was not independent of detection in this experiment, it was not possible to analyse the data from a problem solving perspective only. We therefore decided to conduct a fifth experiment in which participants had to decide whether to solve the first or the second fire. Detection of the second fire was no longer required.

Experiment 5

The purpose of this experiment was identical to the previous experiment, namely to examine which psychological mechanism behavioral entrapment could be ascribed: prior investments or task completion. The experiment was a replication of the fourth experiment but with a dependent variable that is more related to problem solving. Instead of recording whether participants detected the second fire before or after the first fire, we recorded whether participants solved the second fire before or after the first fire.

A second question of this experiment was to what extent the inclusion of a real time component in a task attributes to behavioral entrapment. Is behavioral entrapment in a dynamic task environment stronger than in a static environment? In order to investigate this question, we compared two conditions of the shipping control task: a dynamic condition and a static condition. In the dynamic condition participants had to solve the fires in real time whereas in the static condition they had to make a decision on the basis of snapshots of the dynamic developments.

Method

Participants

Sixty-nine first-year students from the University of Utrecht and the University of Amsterdam participated in the experiment. They were randomly assigned to an experimental condition and paid for participation.

Experimental task

We used the same experimental task as in the previous experiment. In addition, we introduced for each fire a time indication. From the onset of the fire(s), the available time to select the correct treatment for that fire was shown. For instance, if the maximum number of questions is equal to four, it is indicated that participants have 30 seconds of their disposal to solve that particular fire. Furthermore, the system is continuously updating for each fire how much time has elapsed so that participants always know how much time is left for the diagnosis process.

Procedure

The procedure was identical to the procedure of the previous experiment, with the exception that a second fire symbol always represented a high priority fire and never a false alarm. In addition, participants no longer needed to detect a second fire: at the moment a second fire started, the system automatically presented a list of four questions and five possible actions.

Participants in the dynamic condition were instructed in the same way as in the first experiment. Participants in the static condition where presented with snapshots of scenarios the moment the second fire started. For each snapshot they had to indicate which fire they would solve first: the first fire or the second one. Participants received the same practice session as the participants in the dynamic condition.

Design

The design was identical to the previous experiment with the exception that the dependent variable was operationalized as solving the second fire after completion of the first fire. So, for each scenario we recorded whether participants solved the second fire before or after completion of the first fire.

Results

Figure 4.3 shows the number of times participants solved the second fire only after the first fire was completed for the dynamic condition (left panel) and for the static condition (right panel).

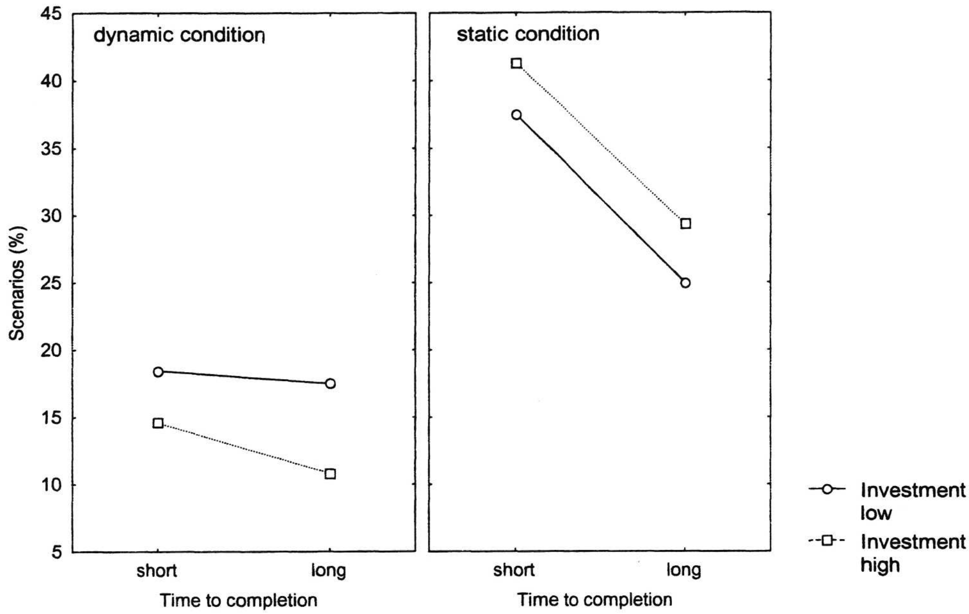


Figure 4.3: *Mean percentage of scenarios in which participants solved the second fire after completion of the first fire as a function of investment and completion. This was done for the dynamic condition (left panel) and the static condition (right panel).*

The tendency to complete the first fire first is stronger in the static than in the dynamic condition ($F(1,68) = 11.25$, $p < 0.01$). For each condition we conducted a separate analysis of variance.

In the dynamic condition there was a main effect of completion ($F(1,23) = 4.29$, $p = 0.05$), a main effect of investment ($F(1,23) = 4.29$, $p = 0.05$) and an interaction between completion and investment ($F(1,23) = 4.53$, $p = 0.05$). The interaction implied that there was a reversed sunk cost effect only for the condition in which there was a relatively long time before the task would be completed ($F(1,23) = 4.78$, $p < 0.05$). For the condition in which the task was near completion there was no effect of sunk costs ($F(1,23) = 1.93$, $p > 0.1$). So, participants tend to abandon an ongoing task after high investments have been made but only when this task was not near completion.

For the static condition there was an effect of task completion ($F(1,45) = 7.89$, $p < 0.01$). The tendency to continue with the first fire became stronger when less questions still needed to be answered. There was no effect of investment ($F(1,45) = 1.05$, $p > 0.1$) nor an interaction between completion and investment ($F(1,45) < 1$).

Discussion

The main purpose of the present study was to examine behavioral entrapment in a dynamic task environment, that is, when individuals have to make decisions in a continuously changing environment.

The present data showed that even though behavioral entrapment was present for both the static and the dynamic condition, it was stronger for the static one. A plausible explanation for this result is that participants in the dynamic condition are in an interactive mode with the system and, in contrast to a static condition, received feedback concerning the consequences of their actions. When participants decided to stick to the first fire it almost always resulted in a shutdown of the system. As a consequence, participants had the opportunity to adjust their strategy in order to prevent shutdowns for subsequent scenarios. In the static condition, on the other hand, participants did not learn the consequences of their choices and did probably not adjust their strategy accordingly. Over all trials, this resulted in more instances of behavioral entrapment in the static condition. Still, learning is not the sole explanation as there was still a bias in the dynamic condition as well.

In the dynamic condition we found a reversed sunk cost effect, but only for scenarios in which the first fire was not near completion. In the present task environment two different mechanisms seem to interact. On the one hand there are prior investments. As we argued earlier, prior investments may have increased subjective time pressure that induced participants to withdraw from the current fire. On the other hand there is an effect of task completion: participants continued with the task because it was near completion. The data suggest that participants tended to withdraw as subjective time pressure increased but only when the task was not near completion. The tendency to withdraw was reduced when the task came closer to completion.

The results of the static condition provided clear evidence for an explanation in terms of task completion and not for an explanation in terms of sunk costs. It seems that when sunk costs and task completion are disentangled more evidence is found for task completion as an explanation for behavioral entrapment than sunk costs (Boehne and Paese, 2000; Conlon and Garland, 1993; Garland and Conlon, 1998). So, it is plausible that, because prior investments and task completion have been confounded, previous findings that were ascribed to sunk costs were actually due to task completion.

In the first experiment behavioral entrapment was much less prominent than in the second experiment. A main difference between the experiments is the dependent variable that was used. In experiment four we focused on detection of the second fire and in second five on solving the second fire. The fact that participants' inclination to abandon the first fire in experiment four is much stronger implies that behavioral entrapment is not due to not detecting a subsequent fault but rather to the decision not to invest time and effort on it yet.

The difference in dependent variable can also explain the fact that no effect of task completion was found in the fourth experiment. Taking task completion into account is only relevant when there is an intention to indeed invest time to complete the task at hand. Detection of the second fire, however, does not necessarily imply that one is going to invest in it, which may explain that task completion is not taken into account in the fourth experiment. So, task completion is only relevant to problem solving and not to detection.

In all, we can conclude that behavioral entrapment is a problem that is mainly present during the phase of solving disturbances rather than during the phase of detecting them. Furthermore, for a static task environment behavioral entrapment can be entirely explained in terms of task completion. For a dynamic environment the explanation is slightly more complex. In a dynamic environment there is time pressure that accumulates over time. After each investment (subjective) time pressure increases. On a problem solving level, the effect of time pressure seems to be contingent on the degree in which an ongoing task is completed. In high-pressure situations where considerable investments are still required people are more likely to abandon the ongoing task than in situations where this task is nearly completed.

5. Conclusions

In the fifth chapter we discuss the explanations of the three research paradigm – planning, task-switching and decision making – in the light of the results of the present thesis. Next, we return to the example of flight 401 and try to apply the findings of this thesis to this example. Finally, we give some recommendations for future system support.

In the introduction we discussed three paradigms that described phenomena that are similar to cognitive lockup: planning, task switching and decision making. Each paradigm provided us with a number of possible explanations for cognitive lockup. Table 5.1 presents an overview of these explanations.

Table 5.1	An overview of possible explanations for cognitive lockup.
Planning	
1.	People commit themselves too early to a detailed plan
2.	People refrain from monitoring the environment
3.	People generate future scenarios that are too optimistic
Task-switching	
4.	People lack sufficient resources to switch to a second disturbance
5.	The costs of switching are perceived as too high
Decision-making	
6.	Sunk costs: prior investments are taken into account
7	Task completion: people have a desire to fulfil a task

In the following section we will discuss each explanation in the light of the findings of the present thesis.

1. People commit themselves too early to a detailed plan

The experimental task that was used throughout the thesis was not designed with the intention to examine planning activities. In the present task, there was only one plan available that in some scenarios had to be substituted by another plan. Nevertheless, we did manipulate commitment in the present thesis. All experiments in the present thesis contained conditions in which participants were committed to a fire when a second fire occurred, the so called the sequential condition. In the first three experiments this condition was compared to the simultaneous condition in which fires started at the same time and participants were not committed to a fire. Overall, participants in the sequential condition tended to hold on to their initial plan and as a consequence fires where not always dealt with in the correct order. In the simultaneous condition participants did not need to revise plans and fires were generally handled in the correct order. So, it seems that once people have committed themselves to a course of action, they are inclined to continue with it. We therefore found the results in favor of the explanation that a commitment to a plan prevents participants from adopting an alternative plan.

2. People refrain from monitoring the environment

According to the monitoring explanation for cognitive lockup would occur because people do not actively scan environment on changes. In our fire control task this would imply that during the process of diagnosis, operators neglect their monitoring function. The fourth experiment of this thesis provided data that do not comply with this explanation. Overall, during the diagnosis of the first fire, participants detected a second fire *before* the first fire was dealt with but often solved it *after* the first fire was dealt with. So, operators did scan the environment and notice any changes, but nevertheless decided to continue with the first problem. In other words, cognitive lockup in our experiment could not be explained by a neglect of the environment.

3. People generate future scenarios that are too optimistic

Another explanation for cognitive lockup that originates from planning research is that people generate scenarios that are too optimistic. As a consequence, they overestimate the available time necessary to complete a task (Hayes-Roth, 1981; Hayes-Roth and Hayes-Roth, 1979). So a possible explanation for cognitive lockup in our fire control task was that participants overestimated the remaining time to deal with two fires. And, because operators overestimate the available time, they could have assumed they had sufficient time to take care of the first fire and then the second fire.

This explanation may hold for the first four experiments of the present thesis but not for the fifth experiment. In this experiment we provided participants with a time index, which indicated how much time was left before the fire caused a total burn down. Making this time explicit would rule out the possibility that operators make incorrect estimations of the remaining time. The fact that participants still continued with the first fire, makes it implausible that cognitive lockup can be ascribed to a too optimistic view of the time that is left. We therefore did not find support for the explanation that cognitive lockup is due to people generating future scenarios that are too optimistic.

4. People lack sufficient resources to switch to a second disturbance

This explanation was examined in the second experiment of this thesis. In this experiment we varied the complexity of the diagnosis process of the first fire. We reasoned that if cognitive lockup is due to a lack of resources, there would be more cognitive lockup as the diagnosis process of the first fire was more complex and consequently required more resources. However, manipulations of complexity did not affect the degree of cognitive lockup.

So, no support was provided for the notion that cognitive lockup is due to limited information processing capacity.

5. The costs of switching are (perceived as) too high

This explanation was examined in the third experiment of the thesis. In this experiment we manipulated the costs of reassessment when a second fire was introduced. We found that cognitive lockup was reduced when it was obvious that the benefits of a reassessment were higher than the costs. On the basis of these results we concluded that individuals decide whether to reassess the situation by trading off the costs and benefits of such a reassessment and that the benefits clearly have to outweigh the costs before participants decide to switch. The fact that the costs of making a reassessment have to be disproportionally low before participants decide to abandon the current task suggests that they are biased in their decision-making process.

6. Sunk costs: prior investments are taken into account
7. Task completion: people have a desire to fulfil a task

Since the sunk-cost bias and task completion were examined in the same experiments we discuss these explanations together.

The sunk-cost effect is the overall finding in decision-making studies that people are inclined to continue with an ongoing task once they have invested time, money or effort in it. However, the fourth experiment of this thesis showed a reversed sunk-cost effect: participants were less inclined to continue with the first fire when more investments were done. We ascribed this effect to a higher perception of time pressure in scenarios with high investments. Although the actual time pressure is equal for both conditions of investment, the experience of time pressure may have been stronger in scenarios were more investments were done. The results further indicated that participants generally detected a second fire before but solved it after the first fire was dealt with. As we could not distinguish a problem-solving phase in this experiment we conducted a fifth experiment in which there was no need to detect a fire explicitly and participants were only required to solve the second fire. We also included a static condition that comprised snapshots of the dynamic condition. To examine the attribution of a real time component we compared this static condition to the dynamic condition.

In the static condition we found an effect of task completion: participants decided to continue with an ongoing task when it was closer to completion. Furthermore, we found more cognitive lockup in the static condition than in the dynamic condition, probably because the absence of feedback in the static condition made it impossible to adjust an initial strategy. As in the previous experiment there was a reversed sunk cost effect but only for the case the task was not near completion. In that case the assumed high perception of time pressure in the high investment scenarios urged participants to withdraw. We reasoned that in case the task was near completion, the subjective time pressure had an opposite effect: participants tend to complete the ongoing task.

To conclude, sunk costs and task completion do explain cognitive lockup but not entirely in the way as expected.

Conclusion

The main purpose of the present thesis was to find a plausible explanation for cognitive lockup. In all, a main conclusion from the present work is that cognitive lockup is a matter of commitment. The findings support the planning explanation that once people are committed to a fault they are inclined to stick to it. No support could be found for the other planning explanations, that is, that people neglect to monitor the environment or that people generate too optimistic scenarios of the future.

Neither could cognitive lockup be explained by limited information processing capacity. An explanation that does seem to hold is that operators make a trade-off between the costs and benefits of making a reassessment of the situation when a second fault starts. This trade-off seems to be biased because the costs or a reassessment have to be considerably lower relative to the benefits in order to reduce cognitive lockup. And even then cognitive lockup is still present. Only in the condition in which participants were detached from the diagnosis process they were always capable of revising their initial plan and cognitive lockup was no longer present.

Nevertheless, as long as operators are working on a fault they are biased in their decision to continue with this fault. Sunk costs and task completion affect this decision to continue. In our real time fire control task we found that as more prior investments were done, participants were less inclined to continue. We also found this effect when we looked at problem solving rather than detection, but this effect was mediated by the degree of task completion. There was a reversed sunk cost effect but only for the case the ongoing task was not near completion. In case the task was near completion participants tended to complete the task.

Flight 401 of Eastern Air Lines

We started the present thesis with the example of flight 401 of Eastern Air Lines, where cognitive lockup was the cause of a dramatic plane crash. We asked ourselves the question which psychological mechanism was responsible for the pilot being locked up in the problem of the landing gear. Now, at the end of the thesis, we return to this example and - although we have to be cautious in extrapolating the present findings - apply the knowledge we acquired to explain the pilot's behavior.

An important conclusion from the present work is that cognitive lockup cannot be exclusively ascribed to a bottleneck in human information-processing capacities. For the case of flight 401 this would imply that the preoccupation with the landing gear problem couldn't be entirely explained by a shortage of resources.

The pilot's decision to continue with the landing gear problem and to refrain from dealing with the problem of the descending altitude could be caused by a combination of high time pressure and the expectation that the problem was nearly solved. Because the pilot experienced a high level of time pressure and because he considered the problem of the landing gear to be nearly solved, he decided to continue with that problem.

Which decision support tools might have prevented the pilot from holding on the malfunctioning landing gear? To date, most support systems are designed with the purpose to relieve operators from a high workload. Since the findings from the present thesis suggest that cognitive lockup is mainly due to a combination of subjective time pressure and task completion, we recommend a change of focus.

First of all, the findings of the present thesis imply that support in the form of an auditory alarm is not likely to prevent cognitive lockup. Our data suggest that operators are aware of the occurrence of additional faults (and are able to detect them) but decide not to deal with them until after the ongoing task is dealt with. The report on the accident showed that the low-altitude alarm did not urge the pilot to abandon the current task and since the present findings demonstrated that participants do seem to notice subsequent faults, we do not expect that a more salient alarm will result in less cognitive lockup.

Our findings showed that cognitive lockup can be reduced by lowering the costs of making a reassessment. To stimulate that operators reassess the situation we therefore recommend that system information is presented in such a way that it is not difficult to determine task priorities. Information about task priorities should be easily accessible for the operator.

However, although our results showed that the tendency for cognitive lockup was considerably reduced in case the costs of a reassessment were low, the tendency was still present. The question then is "how can we entirely break through this tendency?"

The only condition in the present series of experiments where cognitive lockup was no longer present, was in the condition in which priority information was presented on a separate window and participants needed to interrupt their ongoing activities to close that window. Since we found that operators are biased in their decision whether to continue with an ongoing task we consider that detaching operators from this task is an efficient way to debias them. Apparently, irrelevant factors like sunk costs or task completion are eliminated when operators are detached from their ongoing task and are placed in the position to make a new choice between continuing on the ongoing task and start working on an additional task. We therefore suggest that future system support should focus on how operators can be detached from their ongoing activities so that they can make an unbiased decision which fault to deal with first.

Summary

The purpose of the present thesis was to find an explanation for cognitive lockup: the tendency to focus on a subpart of a system and ignore the rest of it. The first chapter of the thesis started with the example of flight 401 of Eastern Air Lines. In this example the pilot continued with a problem on the landing gear while there was a more urgent problem of the descending altitude. The dramatic result was a plane crash which, in hindsight, could be ascribed to cognitive lockup. The crash could have been prevented if the pilot had reassessed the situation and had dealt with the problem of the descending altitude first. Two specific experimental reports on cognitive lockup were discussed (Moray and Rotenberg, 1989 and Kerstholt, Passenier, Houttuin and Schuffel, 1996). However, neither of these studies provided a theoretical explanation for the lockup phenomenon and for that reason the present thesis aimed at finding a plausible explanation for cognitive lockup.

In order to identify possible explanations for cognitive lockup we discussed research of three paradigms that examined phenomena similar to cognitive lockup: planning, task-switching and decision making. Recent definitions of **planning** have incorporated the notion that efficient planning requires a revision of plans when the environment has changed. From a planning point of view there were three different explanations for cognitive lockup: (1) people commit themselves too early to a detailed plan, (2) people refrain from monitoring the environment, and (3) people generate future scenarios that are too optimistic. **Task-switching** studies could be subdivided into two categories: successive task-switching, where the second task is presented in close succession to the first task, and concurrent task-switching, where the second task is presented before the first one is completed. In the last category we also incorporated interruption studies. In these studies, the second task also starts before the first task is completed, but the tasks are

more complex than in the simple reaction times experiments in the traditional task-switching paradigm. Overall, the task switching literature provided two explanations for cognitive lockup: (1) limited information-processing capacity and (2) perception of high switching costs.

Decision making literature contains a number of phenomena that also reflect people's tendency to continue with an ongoing task. Three possible explanations were provided: (1) the sunk cost effect: people are inclined to continue with a course of action because investments such as time, money or effort are made; (2) task completion: people continue with a course of action because they want to complete the task, and (3) loss aversion: the tendency to weigh potential losses larger than potential gains.

In the second chapter we related the explanations provided in the first chapter to the main characteristic(s) of the tasks that are used in each paradigm. As we were mainly interested in supervisory control tasks, such as flying an airplane, we also identified its relevance to supervisory control. For the planning paradigm we identified the reaction to an environmental change as a main characteristic. A main characteristic of the task-switching paradigm is that participants have to deal with multiple problems at the same time. Decision-making studies indicated the importance of prior and future investments in the first task the moment a second task is introduced. All these factors are of importance in supervisory control tasks.

All task characteristics that we identified as relevant to supervisory control were implemented in the experimental task, which we used in five experiments. This task was a simulation of a shipping control task. Globally, there were two modes of control: monitoring the system and fault diagnosis. The system was in a steady state until a fire breaks out. At that moment, participants had to detect the fire and start diagnosing the cause of the fire in order to select the appropriate treatment. When there were two fires at the same time, the situation had to be (re)assessed in order to find out which fire was the most urgent and had to be dealt with first.

In the first experiment we compared a sequential presentation of fires (a second fire starts while the participant is working on the first fire) with a simultaneous presentation (both fires start at the same time). In the sequential scenarios there was an environmental change of the situation during fault handling which was absent in the simultaneous condition. Results of this study showed that the performance level was equal for both conditions, but that cognitive lockup (operationalized as completing the first fire before detecting the second fire) was stronger in sequential scenarios. So, participants were less inclined to assess priorities when they were already involved in fire fighting.

In the third chapter we examined the explanations from the task-switching paradigm. Is cognitive lockup due to limited information-processing capacity or to a deliberate decision? We investigated the explanation of limited human information-processing capacity in the second experiment of this thesis. By varying the complexity of the diagnosis process we tried to manipulate the claim on the human information-processing system. It was assumed that a more complex diagnosis process would claim the system to a higher degree. As a consequence, in case of additional fires, participants would be less inclined to reassess the situation in scenarios in which the diagnosis process of the first fire was more complex. Results showed, however, that task complexity did not have an effect on participants' tendency to continue with the first fire. So, participants' tendency to solve the first fire first and to refrain from reassessing the situation is independent of the task load of the first fire.

In the third experiment we tested the notion that cognitive lockup is a deliberate decision, resulting from an explicit trade-off between costs and benefits of making a reassessment. Participants may decide to continue with the first fire because they anticipate that the costs of making a reassessment are too high relative to the benefits. We therefore varied the costs of making a reassessment, that is, assessing the priorities of both fires. There were three different conditions of priority assessment, in order of declining costs: (1) asking questions; (2) clicking a button, and (3) reading priority information from a separate window. Results showed that cognitive lockup decreased when the costs of making a reassessment were lower. In all, cognitive lockup seems to be due to a trade-off between costs and benefits of making a reassessment rather than limitations in human information processing. However, when the costs are evidently lower than the benefits, participants still decided to continue with the ongoing task. Apparently, participants are biased in their decision to continue.

In the decision making paradigm people's tendency to continue an ongoing task is referred to as behavioral entrapment. In the fourth chapter we examined two explanations from this paradigm that may account for cognitive lockup in supervisory control: the sunk cost bias and task completion. Apart them the purpose of finding a plausible explanation for cognitive lockup, the present task environment provided the opportunity to test the validity of sunk costs and task completion in a dynamic task setting, instead of a static setting as used in previous studies. Hitherto, sunk cost and task completion effects have always been found in a static task setting in which an environmental change had to be imagined. In the present task setting of a fire control task, the environmental change is actually experienced.

In both experiments, the level of sunk costs and task completion was manipulated by respectively the number of questions that had already been answered and the number of questions that still had to be answered, at the moment a second fire was introduced. Results showed that there was a sunk cost effect in a dynamic task setting as well, but opposite to what was found in static tasks when investments in the first fire increased participants detected the second fire more rather than less often. This effect could be explained by the presence of time pressure in a dynamic task. We reasoned that participants might experience more time pressure when relative more by investments were done.

A closer look at the data indicated that in many cases participants did detect the second fire before they had completed the first fire, but still continued with solving the first fire. For that reason, an additional experiment was conducted in which it was recorded whether the second fire was solved before or after completion of the first fire. Furthermore, to identify the effect of a real time component we added a static condition in which participants were presented with snapshots of the dynamic condition. The results of this experiment showed that behavioral entrapment was stronger in the static environment. This effect may be explained by the presence of feedback concerning the consequences of the decision, which was only present in the dynamic task condition. This feedback enabled participants in the dynamic scenarios to adjust their strategy. As in the previous experiment there was a reversed sunk cost effect, but this effect was only present when the task was not near completion. When the task was near completion, this effect was not present.

In the fifth and final chapter of this thesis, we returned to the explanations that we identified in the second chapter and we discussed them in the light of the present findings. The present task environment was not designed with the purpose to examine planning explanations, but mainly to examine individuals' reaction to an environmental change. The first planning explanation was an overall indication for cognitive lockup. The first experiment of this thesis clearly demonstrated that when participants were engaged in fire fighting and when an additional fire was introduced, they often did not deal with fires in the correct order. This in contrast with scenarios in which participants had to deal with two fires from the start. In that case participants nearly always dealt with fires in the correct order. This effect was replicated in the second and third experiment of the thesis.

A second explanation from the planning literature was that people neglect the monitoring task. The fourth experiment of this thesis demonstrated that participants were able to interrupt the diagnosis process of the first fire to detect the second fire, but nevertheless chose to solve the first fire first. This finding does therefore not support the explanation that cognitive lockup is due to a neglect of monitoring the environment. Another explanation - people overestimate the available time - was refuted in the fifth experiment. In that experiment a time index exactly indicated how much time was still available. The finding that participants still continued with the first fire makes it implausible that they overestimated the time they had at their disposal.

For the task-switching paradigm, no support was found for the explanation of a bottleneck in the human information-processing system. We did find support for the second 'task-switching' explanation, namely that people consider the costs of making a switch - in this case a reassessing the situation – as too high. Even in cases where the costs of a reassessment were clearly lower than the benefits, participants decided to continue with the first fire. With respect to the explanations derived from decision making literature we found a reversed sunk-cost effect in a dynamic environment: participants were less inclined to continue with the ongoing task when more investments had been made. This effect was found on both the level of detection and on the level of problem solving. However, on the level of problem solving, the reversed sunk-cost effect was mediated by the degree of task completion. There was a reversed sunk-cost effect but only in case the ongoing task was not near completion.

Compared to the static environment we found that in the dynamic environment people's tendency to continue with the ongoing task was less strong. This effect may be explained by the fact that the dynamic task provided participants with feedback on their decision that enabled them to adjust their overall strategy. This feedback was absent in the static environment.

We ended this chapter with the example of flight 401. We tried to apply the findings of this thesis to this example and provided some suggestions for future design of system support tools.

Samenvatting

Het doel van dit proefschrift was een verklaring te vinden voor *cognitive lockup*: de neiging van mensen om zich te concentreren op een onderdeel van een systeem en de rest van het systeem te veronachtzamen. Het eerste hoofdstuk van het proefschrift begon met het voorbeeld van vlucht 401 van Eastern Air Lines. De piloot van deze vlucht ging verder met het oplossen van een probleem met het landingsgestel terwijl er zich een urgenter probleem voordeed (een dalende hoogte). De dramatische afloop was een vliegtuigongeluk dat, achteraf gezien, toegeschreven kan worden aan cognitive lockup. Het ongeluk was te voorkomen geweest als de piloot een nieuwe inschatting van de situatie had gemaakt en eerst het probleem met de dalende hoogte had opgelost. Twee specifieke experimentele studies werden besproken waar cognitive lockup zich voordeed (Moray & Rotenberg, 1989 en Kerstholt, Houttuin & Schuffel, 1996). Geen van deze studies kwam echter met een theoretische verklaring voor het verschijnsel. Het doel van het huidige proefschrift was derhalve om een plausibele verklaring te vinden voor cognitive lockup.

Om mogelijke verklaringen voor cognitive lockup te identificeren beschreven we onderzoek uit drie paradigma's waar soortgelijke verschijnselen als cognitive lockup zijn waargenomen: planning, taakwisselingen en besliskunde. Volgens recente definities van **planning** is planning efficiënt wanneer een plan wordt herzien in het geval de omgeving verandert. Vanuit de planning literatuur zijn er drie verschillende verklaringen voor cognitive lockup gegeven: (1) mensen committeren zich te snel en in te veel detail aan een plan, (2) mensen zien af van het bewaken van de omgeving, en (3) mensen genereren toekomst scenario's die te optimistisch zijn. Studies met betrekking tot **taakwisselingen** kunnen worden onderverdeeld in twee categorieën: opeenvolgende taakwisselingen, waarbij een tweede taak wordt aangeboden kort nadat de eerste taak is afgerond en gelijktijdige taakwisselingen, waarbij de tweede taak wordt aangeboden *voordat* de eerste taak is afgerond. Tot deze laatste categorie rekenden we ook studies over interrupties. In deze studies begint de tweede taak ook voordat de eerste is afgerond, maar de taken zijn complexer van aard dan de taken welke gebruikt worden in de simpele reactietijd experimenten uit het traditionele paradigma van taakwisselingen. Vanuit de literatuur over taakwisselingen onderscheidden we twee verklaringen: (1) men beschikt over te weinig informatieverwerkingscapaciteit om naar een tweede taak te switchen, en (2) de kosten van een taakwisseling worden als te hoog waargenomen ten opzichte van de opbrengsten.

Ook literatuur op het gebied van **besliskunde** bevat een aantal verschijnselen dat de neiging van mensen laat zien om door te gaan met een taak waaraan men reeds is begonnen. Er zijn drie mogelijke verklaringen genoemd: (1) de sunk-cost effect: mensen zijn geneigd door te gaan met een taak omdat zij al in de taak hebben geïnvesteerd; (2) task completion: mensen zijn geneigd door te gaan met een taak omdat zij de taak willen afronden, en (3) loss aversion: de neiging om mogelijke verliezen zwaarder te wegen dan mogelijke winsten.

In het tweede hoofdstuk legden we een verband tussen de verklaringen uit het eerste hoofdstuk en de belangrijkste taakkenmerken van ieder paradigma. Aangezien we vooral geïnteresseerd waren in supervisietaken, zoals het besturen van een vliegtuig, keken we daarbij naar de relevantie van de taakkarakteristieken voor dergelijke supervisietaken. Het planning paradigma wordt gekenmerkt door een reactie op een verandering in de omgeving. Een hoofdkenmerk van het paradigma met betrekking tot taakwisselingen is dat proefpersonen met meerdere problemen tegelijk worden geconfronteerd. Studies met betrekking tot besluitvorming toonden het belang aan van eerdere en toekomstige investeringen in de eerste taak op het moment dat de tweede taak van start gaat. Al deze factoren zijn van belang voor supervisie taken.

Alle taakkenmerken die we van belang achtten voor supervisietaken (namelijk een verandering in de omgeving, meerdere taken en tijdskosten) werden geïmplementeerd in de experimentele taak, die we gebruikten in de vijf experimenten van het proefschrift. De taak was een simulatie van een brandweertaak op een schip dat ruwweg bestond uit twee doelen: het bewaken van het schip en het diagnosticeren van branden die zich voordeden. Het systeem verkeerde in een stabiele toestand tot het moment waarop er een brand uitbrak. Op dat moment moesten proefpersonen de brand detecteren en de oorzaak van de brand diagnosticeren om zo de juiste "behandeling" te kunnen selecteren. Op het moment dat er twee branden tegelijkertijd optraden moesten proefpersonen de situatie (opnieuw) inschatten om uit te vinden welke brand het meest urgent was en als eerste moest worden opgelost.

In het eerste experiment vergeleken we een sequentiële presentatie van branden (een tweede brand begon terwijl men bezig was met de eerste brand) met een simultane presentatie (beide branden begonnen gelijktijdig). In sequentiële scenario's was er dus sprake van een verandering in de omgeving terwijl men bezig was een storing op te lossen. Deze verandering in de omgeving was niet aanwezig in de simultane presentatie van branden. Resultaten van dit experiment lieten zien dat de prestatie in beide condities gelijk was, maar dat de neiging tot *cognitive lockup* (geoperationaliseerd als het detecteren van de tweede brand pas nadat de eerste brand is opgelost) sterker was in de sequentiële scenario's. Klaarblijkelijk waren proefpersonen minder geneigd om prioriteiten vast te stellen wanneer ze bezig waren met het oplossen van brand.

In het derde hoofdstuk onderzochten we de verklaringen afkomstig uit het paradigma met betrekking tot taakwisselingen. De vraag die in dit hoofdstuk centraal stond was of *cognitive lockup* toe te schrijven is aan een beperking in de informatieverwerkingscapaciteit of aan een weloverwogen beslissing. We onderzochten de verklaring van een capaciteitsbeperking in het tweede experiment van het proefschrift. We manipuleerden de belasting op het menselijk informatieverwerkingssysteem door de complexiteit van het diagnoseproces te variëren. Daarbij werd aangenomen dat hoe complexer het diagnose proces was hoe groter de belasting van het informatieverwerkingssysteem. Dit zou betekenen dat naarmate het diagnoseproces complexer werd, proefpersonen minder geneigd zouden zijn de situatie opnieuw in te schatten als zich een tweede brand voordeed. De resultaten lieten echter zien dat de complexiteit van het diagnoseproces geen effect had op de neiging van proefpersonen om door te gaan met de eerste brand. Met andere woorden, de neiging van proefpersonen om de eerste brand eerst op te lossen en niet opnieuw de situatie in te schatten was onafhankelijk van de werkbelasting van de eerste brand.

In het derde experiment onderzochten we of cognitive lockup het gevolg is van een bewuste beslissing. We veronderstelden dat deze beslissing wordt genomen door een expliciete afweging te maken tussen de kosten en baten van het opnieuw inschatten van de situatie. Proefpersonen zouden beslissen om door te gaan met de eerste brand omdat ze van tevoren veronderstellen dat de kosten voor het maken van een nieuwe inschatting te hoog zijn in verhouding tot de opbrengsten. Om die reden varieerden we de kosten voor het maken van een nieuwe inschatting van de situatie, in dit geval door het vaststellen van prioriteiten. Er waren drie verschillende manieren om prioriteit vast te stellen, in volgorde van afnemende kosten (1) vragen stellen; (2) een knop aanklikken, en (3) het aflezen van prioriteiten informatie van een apart venster. Resultaten lieten zien dat *cognitive lockup* minder werd naarmate de kosten voor het maken van een nieuwe inschatting afnamen. *Cognitive lockup* lijkt dus te worden veroorzaakt door een afweging tussen kosten en opbrengsten van het maken van een nieuwe inschatting van de situatie. Echter, wanneer de kosten duidelijk lager waren dan de opbrengsten, bestond bij proefpersonen nog steeds de neiging om door te gaan met de taak. Klaarblijkelijk werden proefpersonen bij het nemen van hun beslissing beïnvloed door factoren die niet relevant waren.

In het vierde hoofdstuk onderzochten we verklaringen vanuit de besliskunde voor cognitive lockup: de sunk-cost bias en task completion. We hebben in dit hoofdstuk tevens onderzocht of de mate waarin proefpersonen bij een taak blijven hangen afhangt van de dynamiek van de taak. Onze brandweertaak is dynamisch in de zin dat de situatie continu verslechtert. In voorgaand onderzoek zijn echter alleen statische omgevingen gebruikt: proefpersonen moesten zich weliswaar voorstellen dat de situatie was veranderd, maar het continue verloop in de tijd ontbrak.

In twee experimenten werd het niveau van sunk-costs (de hoeveelheid investeringen die reeds gedaan zijn) en task completion (de investeringen die nog gedaan moeten worden) gemanipuleerd. De sunk-costs bestonden uit het aantal vragen dat reeds beantwoord was en task completion bestond uit het aantal vragen dat nog beantwoord moest worden op het moment dat de tweede brand van start ging. Resultaten van het eerste beslisexperiment lieten zien dat er een effect van sunk-costs was, maar in de tegengestelde richting dan verwacht: Proefpersonen waren eerder geneigd om de tweede vraag te detecteren naarmate er meer vragen van de eerste brand waren beantwoord. Dit effect kan wellicht verklaard worden door het feit dat proefpersonen meer tijdsdruk voelden naarmate er meer investeringen waren gedaan.

Een nadere beschouwing van de data liet zien dat in veel gevallen proefpersonen de tweede brand weliswaar detecteerden voordat de eerste brand was afgerond maar deze pas oplosten nadat de eerste brand was afgerond. Daarom voerden we nog een experiment uit waarin we registreerden of de tweede brand werd opgelost voor of na oplossing van de eerste brand. En verder, om het effect van een real time component te kunnen vaststellen, voegden we een statische conditie toe. Deze conditie bestond uit snapshots van de dynamische conditie. De resultaten van dit experiment lieten zien dat, in tegenstelling tot onze verwachting, cognitive lockup sterker was in de statische conditie. Dit kan verklaard worden door de aanwezigheid van feedback over de consequenties van acties in dynamische conditie. Deze feedback gaf proefpersonen in de dynamische conditie de gelegenheid hun strategie aan te passen. Net als in het vorige experiment was er een tegengesteld sunk-cost effect, maar dit effect werd beïnvloed door de mate van *task completion*. Het omgekeerde sunk-cost effect was alleen aanwezig wanneer het nog lang zou duren voordat de taak voltooid was en niet wanneer de taak wel bijna voltooid was.

In het vijfde en laatste hoofdstuk van het proefschrift keerden we terug naar de verklaringen uit het tweede hoofdstuk. De eerste planning verklaring was een meer algemene indicatie voor *cognitive lockup*. Het eerste experiment van het proefschrift liet duidelijk zien dat als mensen eenmaal bezig zijn met het oplossen van een brand en zich een tweede brand voordoet, zij deze branden vaak niet in de correcte volgorde afhandelen dat wil zeggen dat zij een brand met een lagere prioriteit eerst oplosten. Dit in tegenstelling tot scenario's waarin proefpersonen vanaf het begin met twee branden geconfronteerd werden. In dat geval handelden proefpersonen de branden bijna altijd in de juiste volgorde af. Dit effect werd gerepliceerd in het tweede en derde experiment van het proefschrift.

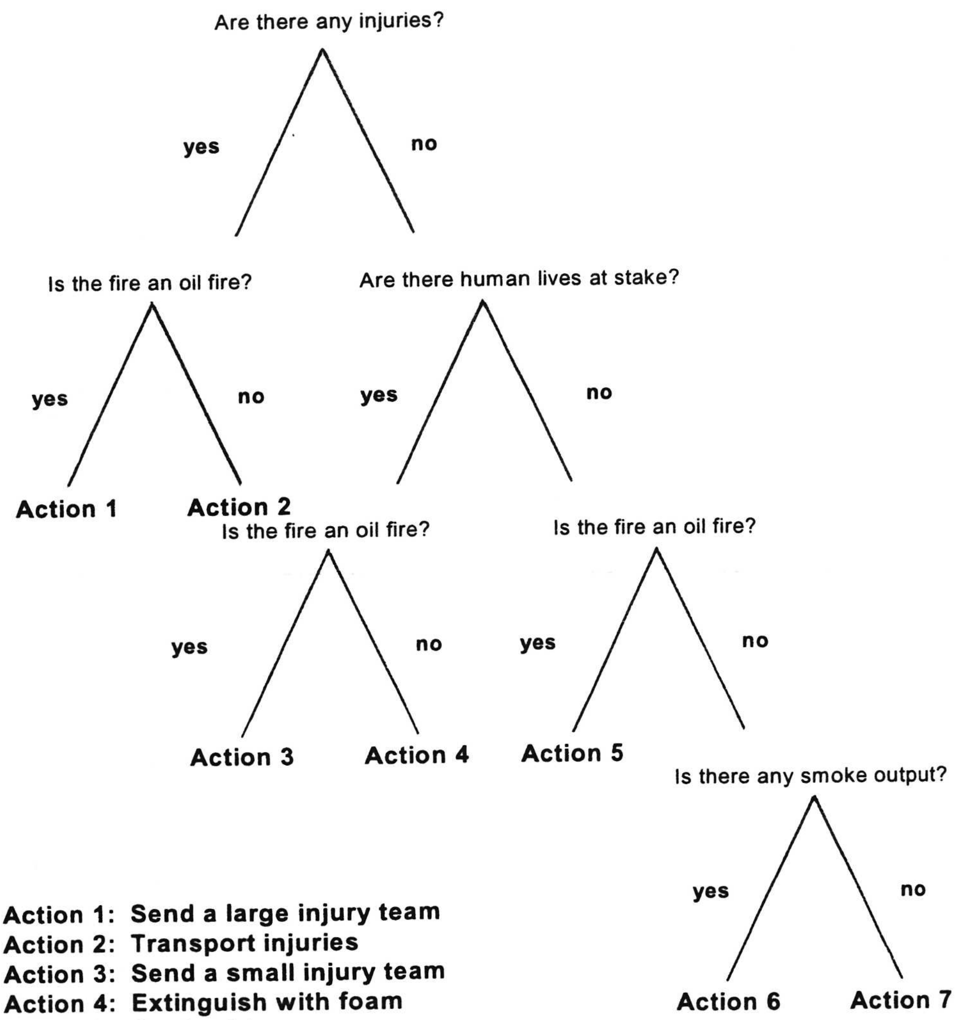
Een tweede planning verklaring was dat mensen nalaten de omgeving te bewaken. Het vierde experiment liet zien dat proefpersonen in staat waren het diagnose proces van de eerste brand te onderbreken om de tweede de brand te detecteren. Echter, tevens bleek dat men wel besloot om vervolgens brand één eerst op te lossen. Dit resultaat strookte dus niet met de verklaring dat mensen, tijdens het oplossen van een probleem, de rest van het systeem nalaten te bewaken. Een andere verklaring – mensen overschatten de beschikbare tijd – werd weerlegd in het vijfde experiment. In dat experiment liep een tijd mee die exact aangaf hoeveel tijd er nog beschikbaar was. Het gegeven dat proefpersonen in deze conditie nog steeds door gaan met het oplossen van de eerste brand maakte de verklaring voor *cognitive lockup* dat mensen hun beschikbare tijd overschatten niet plausibel.

Wat betreft de verklaringen uit het taakwisselingen paradigma, kon geen steun gevonden worden voor de verklaring dat *cognitive lockup* het gevolg is van een beperking in het menselijk informatieverwerkingssysteem. We vonden wel steun voor de verklaring dat mensen de kosten voor het maken van een wisseling – in dit geval het maken van een nieuwe *assessment* van de situatie – als te hoog inschatten. Naarmate de kosten voor het maken van een *(re)assessment* lager waren, waren proefpersonen eerder bereid te switchen. Echter, zelfs wanneer de kosten van een *reassessement* duidelijk lager waren dan de baten besloten proefpersonen nog om door te gaan met de eerste brand. Wat betreft de verklaringen die voortkwamen uit de besliskunde vonden we een tegengesteld *sunk-cost* effect: proefpersonen in een dynamische omgeving waren minder geneigd om door te gaan met de taak waaraan ze werkten in het geval er meer in de taak geïnvesteerd was. Dit effect werd gevonden op zowel het niveau van detectie van branden als op het niveau van oplossen van branden. Echter op dit laatste niveau was er een interactie met *task completion*. Er was een tegengesteld *sunk-cost* effect maar alleen in geval het relatief lang duurde voordat de taak opgelost zou zijn.

Vergeleken met een statische omgeving vonden we dat in een dynamische omgeving de neiging om door te gaan met de taak minder sterk was. Dit effect kan verklaard worden door het feit dat proefpersonen in de dynamische omgeving feedback kregen over hun beslissing en hun strategie gedurende het experiment konden aanpassen. Deze feedback was niet aanwezig in de statische omgeving.

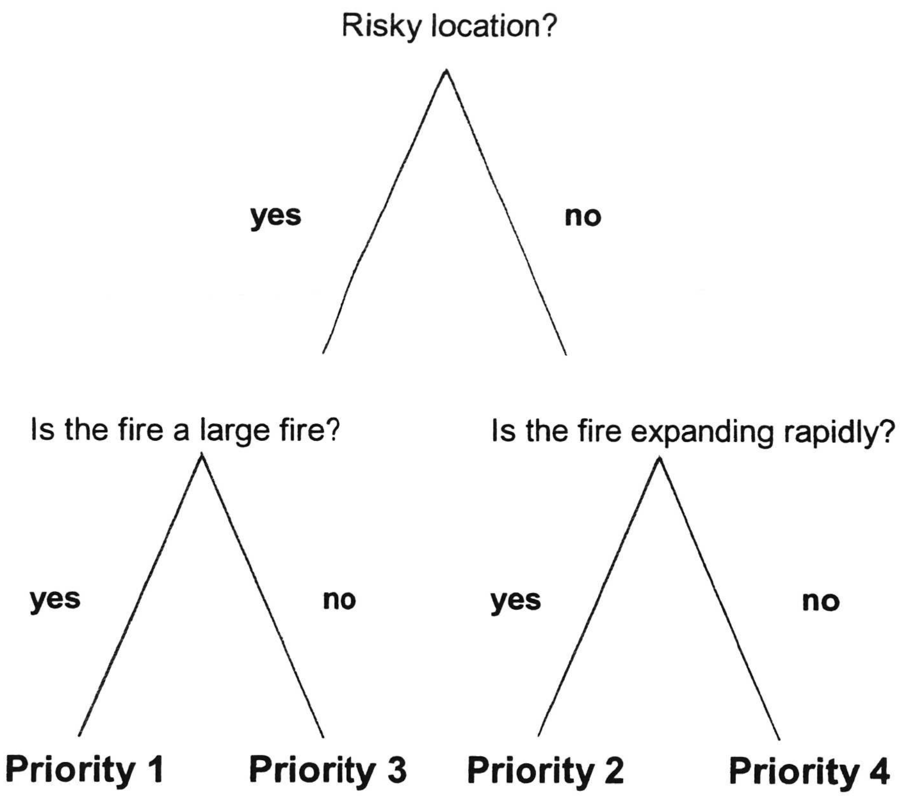
We sloten het hoofdstuk af met het voorbeeld van vlucht 401. We probeerden de resultaten van het proefschrift toe te passen op dit voorbeeld en we gaven enkele suggesties voor toekomstig ontwerp van systeem ondersteuning.

APPENDIX A

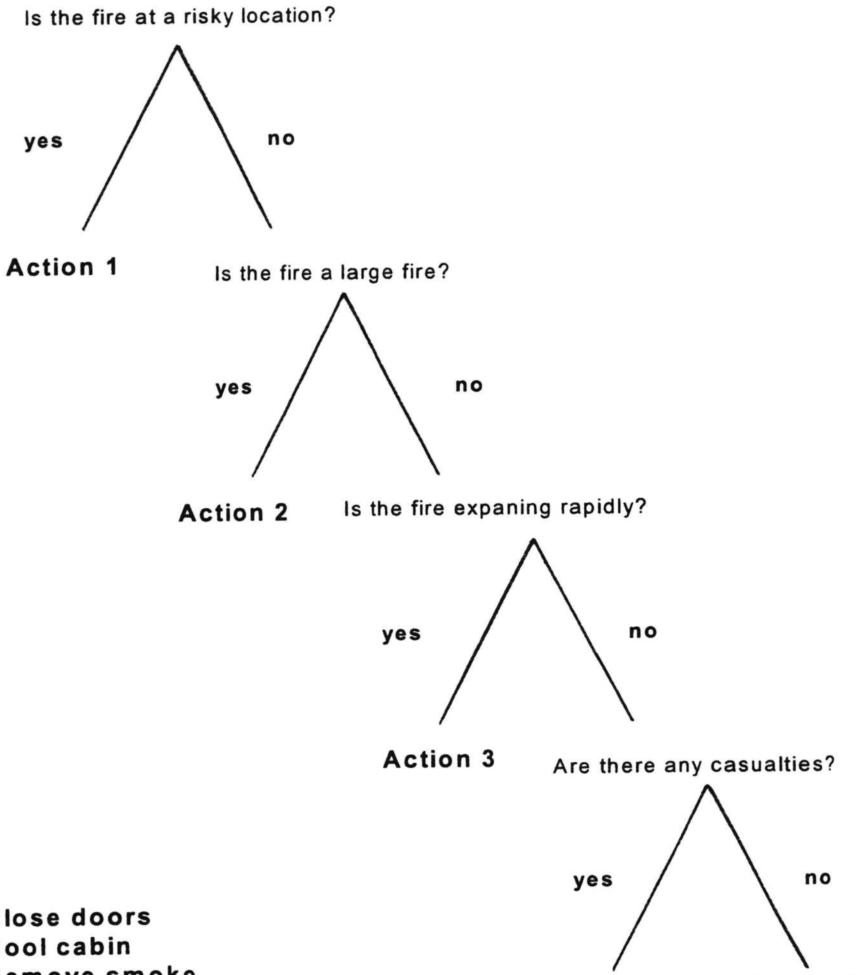


Action 1: Send a large injury team
Action 2: Transport injuries
Action 3: Send a small injury team
Action 4: Extinguish with foam
Action 5: Cool cabin
Action 6: Remove smoke
Action 7: Close doors

APPENDIX B


Risky location?
yes
no
Is the fire a large fire?
Is the fire expanding rapidly?
yes
no
yes
no
Priority 1
Priority 3
Priority 2
Priority 4

APPENDIX C



Action 1: Close doors
Action 2: Cool cabin
Action 3: Remove smoke
Action 4: Transport casualties
Action 5: Extinguish by foam

References

Allport, D.A., Styles, E.A. and Hsieh, S. (1994). Shifting intentional set: exploring the dynamic control of tasks. In C. Umiltá and M. Moscovitch (Eds.), *Attention and Performance XV*, Cambridge, MA: MIT Press.

Arkes, H. R. and Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, **35**, 124-140.

Bainbridge, L. (1997). The change in concepts needed to account for human behavior in complex dynamic systems. *IEEE Transactions on Systems, Man, and Cybernetics*, **27**, 351-359.

Berg, C. R., Strough, J., Calderone, K S., Sansone, C. and Weir, C. (1998). The role of problem definitions in understanding age and context effects on strategies for solving everyday problems. *Psychology and Aging*, **13**, 29-44.

Berger, C. R., Karol, S. H. and Jordan. J. M. (1989). When a lot of knowledge is a dangerous thing: The debilitating effect of plan complexity on verbal fluency. *Human Communication Research*, **16**, 91-119.

Boehne, D. M. and Paese, P. W. (2000). Deciding whether to complete or terminate an unfinished project: A strong test of the project completion hypothesis. *Organizational Behavior and Human Decision Processes*, **2**, 178-194.

Brichin, M. and Rachadzo, R. (1995). Intentional vs. "artificial" planning of actions. *Studia Psychologica*, **37**, 57-61.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review*, **17**, 39-61.

Brockner, J., Rubin, J. and Lang, E. (1981) Face-saving and entrapment. *Journal of Experimental Social Psychology*, **15**, 492-503.

Brown, J. (1948). Gradients of approach and avoidance responses and their relation to motivation. *Journal of Comparative and Physiological Psychology*, **41**, 450-465.

Buehler, R, Griffin, D. and Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, **67**, 366-381.

Byram, S. J. (1997). Cognitive and motivational factors influencing time prediction. *Journal of Experimental Psychology: Applied*, **3**, 216-239.

Conlon, D. E. and Garland, H. (1993). The role of project completion information in resource allocation decisions. *Academy of Management Journal*, **36**, 402-413.

De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology*, **48**A, 2-25.

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal*, **32**, 543-576.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (p. 97-101). Santa Monica, CA: Human Factors and Ergonomics Society.

Ensdley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, **37**, 32-64.

Garland, H. (1990). Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Social Psychology*, **75**, 721-727.

Garland, H. and Conlon, D. E. (1998). Too close to quit: The role of project completion in maintaining commitment. *Journal of Applied Social Psychology*, **28**, 2025-2048.

Gillie, T. and Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity and complexity. *Psychological Research*, **50**, 243-250.

Goldin, S. E. and Hayes-Roth, B. (1980). *Individual differences in planning*. San Diego, CA: Rand.

Hayes-Roth, B. (1981). *A cognitive science approach to improving planning*. Proceedings of the Third Annual Conference of the Cognitive Science Society. Berkeley, CA: Cognitive Science Society.

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

Hine, D. W. and Grifford, R. (1997). What harvesters really think about in common dilemma situations: A grounded theory analysis. *Canadian Journal of Behavior Science*, **29**, 180-194.

Janis, I. L. and Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York: The Free Press.

Jersild, A. T. (1927) Mental set and shift. *Archives of Psychology*, Whole No. 89.

Kahneman, D., Knetsch, J.L. and Thaler, R.H. (1990). Experimental tests of the endowment effect and the Coase- theorem. *Journal of Political Economy*, **98**, 1325-1348.

Kahneman, D., and Tversky, A. (1979). Prospect theory. An analysis of decision under risk. *Econometrica*, **47**, 263-291.

Kahneman, D. and Tversky, A. (1982). The psychology of preferences. *Scientific American*, **246**, 160-173.

Kerstholt, J. H. (1994). The effect of time pressure on decision making behavior in a dynamic task environment. *Acta Psychologica*, 86, 89-104.

Kerstholt, J. H. (1995). Decision making in dynamic situations: The effect of false alarms and time pressure. *Journal of Behavioral Decision Making*, 8, 181-200.

Kerstholt, J. H. and Passenier, P. O. (2000). Fault management in supervisory control: the effect false alarms and support. *Ergonomics*, **43**, 1371-2389.

Kerstholt, J. H., Passenier, P. O., Houttuin, K. and Schuffel, H. (1996). The effect of a priori probability and complexity on decision making in ship operation, *Human Factors*, **38**, 65-78.

Kleinmutz, D. N. and Thomas, J. B. (1987). The value of action and inference in dynamic decision making. *Organizational Behavior and Human Decision Processes*, **39**, 241-364.

Krech, D. Crutchfield, R. and Livson, N. (1967). Elements of psychology (2nd ed.) New York: Alfred A. Knopf.

McDermott, D. (1978). Planning and acting. *Cognitive Science*, **2**, 71-109.
McLenan, J. and Omodei, M M. (1996). The role of prepriming in recognition-primed decision-making. *Perceptual and motor skills*, **82**, 1059-1069.

Meyer, D. E. and Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic Mechanisms. *Psychological Review,* **104**, 3-65.

Meyer, D. E. and Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of Psychological Refractory Period Phenomena. *Psychological Review,* **104**, 749-791.

Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E. Glas, J., Zurbriggen, E. Gmeindl, L. and Apfel, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica,* **90**, 163-190.

Monsell, S. (1995). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial Essays in Cognition*. Erlbaum (UK): Taylor & Francis.

Moon, H. (2001). Looking forward and looking back: Integrating completion and sunk-cost effects within an escalation-of-commitment progress decision. *Journal of Applied Psychology,* **1**, 104-113.

Moray, N. and Rotenberg, I. (1989). Fault management in process control, *Ergonomics,* **32**, 1319-1342.

Mumford, M. D., Schulz, R. A., and Van Doorn, J. R. (2001). Performance in Planning: Processes, Requirements, and Errors. *Review of General Psychology,* **3**, 213-240.

Newby-Clark, I., R. Ross, M., Buehler, R. Koehler, D. J. and Griffin, D. (2000). People focus on optimistic and disregard pessimistic scenarios while predicting task completion times. *Journal of Experimental Psychology: Applied,* **3**, 171-182.

National Transportation Safety Board. (1978). United Airlines, Inc., McDonnel Douglas DC-8-61, N8082U, Portland, Oregon.

Osviankina, M. (1928). Die Wiederaufname unterbrocherner Handlungen. *Psychologische Forschung,* **11**, 302-379.

Oswald, S. C., Mossholder, K. W. and Harris, S. G. (1997). Relations between strategic involvement and managers' perception of environment and competitive strengths: The effect of vision salience. *Group and Organization Management,* **22**, 343-365.

Payne, J. W., Bettman, J. R. and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **14**, 534-552.

Patalano, A, L. and Seifert, C. M. (1997). Opportunistic planning: Being reminded of pending goals. *Cognitive Psychology*, **34**, 1-36.

Read, S.J. (1987). Constructing causal scenarios: A knowledge structure approach to causal resoning. *Journal of Personality and Social Psychology*, **52**, 288-302.

Ritov. I. and Baron, J. (1992). Status-quo and omission biases. *Journal of Risk and Uncertainty*, **5**, 49-61.

Rogers, R. and Monsell, S. (1995). The costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**, 207-231.

Rubinstein, J., Meyer, D. E. and Evans, J. E. (2001). *Journal of Experimental Psychology: Human Perception and Performance*, **27**(4), 763-797.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, **1**, 7-59.

Schiffman, N. and Greist-Bousquet, S. (1992). The effect of task interruption and closure on perceived duration. *Bulletin of the Psychonomic Society*, **30**, 9-11.

Schumacher, E.H., Lauber, E. Glass, J. Zurbriggen, E., Gmeindl, L, Kieras, D. E. and Meyer, D. E. 1999 Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology*, **25**, 791-814.

Schweizer, M. (1994). Disentangling status-quo bias and omission effect: an experimental analysis. *Organizational Behavior and Human Decision Processes*, **58**, 457-476.

Sheridan, T. B. (1988). Task allocation and supervisory control, In: M. Helander (Ed.), *Handbook of human-computer interaction*. Amsterdam: Elsevier.

Simon, D. J. and Galotti, K. M. (1992). Everyday planning: An analysis of daily time management. *Bulletin of the Psychonomic Society*, **30**, 61-64.

Spector, A. and Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology*, **89**, 669-679.

Staw, B. M. (1981). The escalation of commitment to a course of action. *Academy of Management Review*, **6**, 577-587.

Staw, B. M. and Ross, J. (1987). Behavior in escalation situations: antecedents, prototypes, and solutions. In B. M. Staw (Ed.), *Research in organizational behavior* (Vol.9, pp. 39-78). Greenwich, CT: JAI Press.

Staw, B. M. and Ross, J. (1989). Understanding behavior in escalation situations.

Teger, A. I. (1980). *Too much invested to quit: The psychology of the escalation of conflict.* Elmsford, NY: Pergamon.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, **1**, 39-60.

Thomas, J. B. and McDaniel, R. R. (1990). Interpreting strategic issues: Effects of strategy and the information processing structure of top management teams. *Academy of Management Journal*, **33**, 286-306.

Wiener, E. L. (1985). Beyond the sterile cockpit. *Human Factors*, **27**, 75-90.

Whyte, G. (1986). Escalating commitment to a course of action: A reinterpretation. *Academy of Management Review*, 11, 311-321.

Xiao, Y., Milgram, P. and Doyle, D. J. (1997). Planning behavior and its functional role in interactions with complex systems. *IEEE Transactions on Systems, Man and Cybernetics*, **27**, 313-325.

Zeignarik , B. (1927). Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung*, **9**, 1-85.

Zijlstra, F. R. H., Roe, R.A., Leonora, A.B. and Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, **72**, 163-185.

Dankwoord

Allereerst wil ik mijn copromotor, José Kerstholt, bedanken. Zij is gedurende het project een voortdurende bron van inspiratie voor mij geweest. Zonder haar enthousiaste en deskundige begeleiding zou het proefschrift nooit zo zijn geworden zoals het nu is.

Natuurlijk gaat ook mijn dank uit naar mijn promotor, Jeroen Raaijmakers, die het project op afstand volgde maar er altijd was wanneer dat nodig was.

Verder wil ik mijn ex-collega's van Technische Menskunde bedanken voor de prettige werksfeer waarin ik mijn onderzoek heb mogen uitvoeren. Mijn dank gaat in het bijzonder uit naar Maarten Abbring en Bert Bierman voor het vele programmeerwerk.

Ook gaat mijn dank uit naar twee mensen die mij in de eindfase van het project hebben geholpen: Jacqueline van Beek voor de lay-out van het proefschrift en Camiel Claus voor het ontwerp van de omslag.

Last maar zeker not least wil ik mijn ouders bedanken voor hun steun en voor het vertrouwen dat het uiteindelijk allemaal wel goed zou komen.

Gerard Meij
Monnickendam
november 2003

epos

EXPERIMENTEEL-
PSYCHOLOGISCHE
ONDERZOEKSCHOOL