Sticking to plans: capacity limitation or decision-making bias?
Meij, G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
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 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 
depended on 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 refraining 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.
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 ($t = 1.21, df = 26, \text{n.s}$).

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.