The maritime pilot at work

The maritime pilot at work

evaluation and use of a time-to-boundary model of mental workload in human-machine systems

PROEFSCHRIFT

ter verkrijging van de graad doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op 11 mei 1999 te 16.00 uur
door

Fulko C. van Westrenen

doctorandus in de psychologie

Dit proefschrift is goedgekeurd door de promotor:
prof. dr. Andrew R. Hale

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
prof. dr. A.R. Hale, Technische Universiteit Delft, promotor
prof. dr. ir. R.E.C.M. van der Heijden, Technische Universiteit Delft
prof. dr. ir. J.A. Mulder, Technische Universiteit Delft
prof. dr. ir. H.G. Stassen, Technische Universiteit Delft
prof. dr. P.T.W. Hudson, Rijksuniversiteit Maastricht
prof. dr. T.F. Meijman, Rijksuniversiteit Groningen
prof. dr. N.P. Moray, University of Surrey (U.K.)

Ir. T. Heijer heeft als begeleider in belangrijke mate aan het totstandkomen van het proefschrift bijgedragen.

TRAIL Thesis Series nr. T99/2, The Netherlands TRAIL Research School

ISBN 90 5166 703 6

Photo cover: Niels van der Hoeven, ©Het Vrije Volk
(Collectie Maritiem Museum Rotterdam)

Acknowledgements

This thesis would not have been possible without the help of many people. First of all I would like to thank the Rotterdam Pilots, in particular the four pilots who participated in the experiments. I would also like to thank all the captains who welcomed me on board and allowed me to videotape all the activity on the bridge.

While preparing for the experiment I was introduced into the maritime world by spending some time at sea on the Royal Clipper, Dutch Mate, Engelina Broere, and the Norsun. I would like to thank the captains and the shipping companies for their hospitality and all the information they gave me.

The process simulator experiment was conducted at the Cognitive Engineering Laboratory of the University of Toronto. They were very helpful in introducing me to the field of ecological displays, and helping me with the equipment, subjects and analysis.

I am grateful to The Organisational and Work Psychology Group at the TU Delft for lending me recording equipment and helping me analyse the numerous videotapes. I must also thank many staff members of various research groups for helping me with various specialized topics in this research.

Writing a thesis in a language other than your native language is difficult. This thesis was rewritten and debugged by several people in order to turn it into correct English, after which I took the liberty to reintroduce some of my typical errors.

The Safety Science Group gave me the opportunity to do this research, for which I am very grateful. Finally, I would like to thank my family and friends. Without their continuous support this would not have been possible.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

People have proven to be flexible and reliable in many control tasks, such as car driving and ship navigation. Much effort has been invested into automating these tasks but the benefits have so far been limited and the problems enormous. Other tasks, such as plant control, where complicated systems are tightly coupled to obtain large volumes of high quality products with very strict production demands show a much higher level of automation. This automation makes control of these complicated systems possible, relieving the controller of many tasks, improving the production quality, and reducing the operators workload. However even with very sophisticated control systems human operators are needed to offer the flexibility needed to deal with unforeseen events. In the first group of tasks control is so complex that we depend completely on human control. In the second group (plant control) human control is not required continuously but its presence is needed in order to deal with a large group of events. In both cases, in order to fit the work to the operator and to ensure that his capacity is used to the maximum, that system safety is optimal, and that working conditions meet human long term needs, extensive knowledge of operator abilities and limitations is required. Because these tasks often have major safety consequences if they go wrong there are constant attempts to improve the task design of these tasks, to decide on shifts of the boundary between human and automated conduct of the tasks, and to support the tasks with use of the developing technologies, particularly telematica. Evolution of such proposed changes poses heavy demands on modeling how the human operator actually functions in practice.

All control tasks have four major components: observing the state of the system, comparing this state to the desired state, deciding on necessary changes, and implementing these changes. The first two steps of the control process are the subject of this thesis: observing the state of the system and comparing this to the desired state. This part of the task is better known as monitoring. Much has been written about monitoring. Monitoring is the activity of examining the system's displayed status information to decide whether or not the system is in a normal state and under control (Moray, 1986, p. 40.2). Most of the studies in the literature on monitoring focus on relatively simple tasks that can be carefully studied in laboratory settings, using covered instruments or eye-movement recordings. For the development of detailed theory and models of the human operator this is fine. However, in order to make the step from these laboratory settings to practical applications in real situations, such as in safety science, it is also necessary to study operator behaviour in these real settings. In this study monitoring is studied in its full context of a control process. The operator not only had to monitor but also decide on and implement the corrective actions as well. This was done in two situations: a relatively complex process simulator in which the operator had to perform a realistic control task, and in the real situation of maritime pilots on board sea ships. The operator's monitoring activity is measured using mental workload measures, in particular heart-rate variability.

There are many types of control tasks in which people are involved. The type of task that is the centre of focus in this thesis is a continuous task that is done by a single controller. The field experiment described in this thesis involved a maritime pilot on the bridge. This choice was governed by the demand of a study into the feasibility of transferring parts of the task of a harbour pilot from ship to shore. When navigating a ship very complex information about the situation must be retrieved from the environment and based on this information decisions are made on heading, speed, and manoeuvring. This is a very complex task that is not well understood. The reason for choosing a maritime pilot is personal interest in shipping, more important the fact that it was a study of a real task, not a simulated, and above all it was made possible because I was given the opportunity for this study as a follow-up of a study on shore based radar support. In order to verify the theory a second study was done using a rather complex process simulator. This simulator made it possible to experiment in a very controlled environment, and to develop and test a model which proved useful for understanding the pilot's behaviour in the real task.

The pilot's task is a complicated real world task with all the elements that come with such tasks. They perform it in a highly variable environment. Wind, currents, waves, and other traffic make each voyage unique. This requires a high degree of adaptation by the pilot in order to cope with all possible situations. In addition all ships are unique in their manoeuvring characteristics. Ships react very slowly to control changes and these therefore require careful planning by the pilot. And because of their large mass and often dangerous cargo the consequences of an accident are far reaching. This all makes the work of the pilot very interesting.

The simulated process on the other hand is a well controlled situation that is much more adequate for experiments. In contrast to the real world situation of ships it is fully under the experimenter's control. The process is much simpler, all process variables are known and the simulation can be started at any situation wanted. In addition the disturbances in a laboratory situation are very limited compared to the bridge of a ship. It is not claimed that the simulator situation used is an ideal experimental model of the navigator task. The opportunity to study it was far more a question of chance. However, the simulator shares enough characteristics with the pilot task to provide a good base for developing and testing an appropriate model.

The approach taken in this study is to concentrate on monitoring. During process control the operator must observe the process variables to obtain the information needed to come to control decisions; changing the set points of the controls in order to keep the variables within specifications. How frequently this process must be executed depends on the characteristics of the variables and the specifications: highly changeable variables need more attention then very slow ones, and variables with very strict specifications need more attention than those with loose ones.

Many studies have been done on monitoring. Most of these studies focussed on rather simple laboratory experiments with participants who were trained for a short time to perform that specific task. There was not the time, the will, or the money to train a group of participants to the level of skill that is considered normal with real control tasks, in order to do a useless task just for a scientific experiment. However it is known that skilled operators behave differently from novice. The skilled-operator's knowledge of the system makes his sampling behaviour different. For practical applications it is knowledge of the behaviour of skilled operators that is most needed. This is a very important reason that, in this study, well trained participants were used in the laboratory experiment and experienced pilots were used in the field experiment.

Different theories on monitoring (or sampling behaviour) have been developed depending on the system being studied, the precise research questions and the various theoretical backgrounds. Most techniques for measuring monitoring behaviour are based on covered instruments, so when the operator needs to know a value he needs to uncover the instrument, which in turn is recorded as 'taking a sample'. Another group of experiments is based on eye-view monitoring, with which the direction of gaze can be followed from moment to moment. This technique has even successfully been used on bridge simulators (Itoh e.a., 1998). This can be done in many situations but is not a preferred method when studying the pilot's work on, for example, a supertanker in real life on a crowded waterway.

If we restrict ourselves to visual sampling it was shown by Senders that there is a strong relationship between the visual sampling behaviour of a driver and the curvature of the road (Senders, 1983, appendix B) with more curvature tending to lead to more sampling. A firm conclusion could not be made by Senders because the width of the road and how the driver 'negotiated' the curves had a large influence on the outcome of the sampling characteristics. A similar experiment by Godthelp confirmed these findings (Godthelp et al., 1984). I observed a similar effect during the experiments with maritime pilots: the pilots became very alert when the ship was moving towards a physical boundary. This seemed to be irrespective of the fact that some pilots sailed rather close to objects and other pilots preferred to stay away as far as possible from objects, and was also irrespective of the different speed regimes. This suggested the idea that not distance and speed were dominant in their visual sampling behaviour, but time (i.e. time to boundary). This finding was later confirmed in the simulator experiment, where between subjects differences in strategy were very clear ('walking the boundaries' versus 'keeping the middle of the road'). This notion forms the basis of the model developed in the next chapter.

In many situations studying actual visual sampling directly is extremely hard. This is certainly the case in the maritime environment. Therefore another approach is taken. In this study mental workload is used as a measure for monitoring and control. It is indisputable that monitoring puts a mental workload on the operator. This thesis aims to develop a simple, lawful relationship between the state of the system under control (the operator's performance), monitoring behaviour, and the mental workload of the operator. The result is of importance for determining the possible risks and opportunities when changing existing or designing new control situations. To be able to choose the optimum type of control it is necessary to have detailed knowledge of the relationship between process characteristics, the control task, performance, and operator behaviour. Without this knowledge it may be that the system design or changes to it are based on the wrong criteria.

Although human operators are successful in performing many control tasks very reliably, they are fallible and do make mistakes. The consequences of such events can be devastating, and due to the scale of modern production processes disruption of this process has become increasingly unacceptable. The study of human error has therefore become a natural element in the design of large and critical systems. One possibility of studying these "human errors" is to focus, usually retrospectively, on the events that lead to the error and try to generalize from these events, each of which in itself is in many ways unique. This knowledge can then be used to change the working situation is such a way that these human errors are unlikely to occur. However, the scale and costs of even minor accidents of shipping, and the public reaction to actual disasters makes such a 'trial and error' approach inefficient and unacceptable. Another, and according to Rasmussen (1997) a more productive, possibility is to study control behaviour in normal situations and try to understand human control behaviour there. This knowledge can then be used to improve the working situation and reduce the probability of making errors. This study of sampling behaviour and mental workload fits into this last approach.

In brief, chapter 2 will describe the theory on monitoring by operators.

These models form the basis for the relationship between the performance of the operator and his information collection. In that chapter the model which will be the subject of the two studies is derived and explained. The theory which is developed and tested is that the operator's monitoring behaviour (and with that his mental workload) depends largely on the period of time before a change in set points is needed, also known as time-to-contact or time-to-boundary. This means that the process goals and the tactics applied by the operator/navigator determine his workload. In chapter 3 the background of mental workload is discussed, together with techniques to measure mental workload. Chapter 4 discusses the results of the simulator experiment. In chapter 5 mental models are discussed, as they are an integral part of decision making and therefore of information collection, particularly in the real world situation of pilots. In chapter 6 the real life study and the task of the navigator is introduced, and chapter 7 describes the results of the maritime-pilot experiment to evaluate the predictions of the model. Chapter 8 gives the conclusions and discussion of the outcomes.

The results of the experiments will show that in both situations (process simulator and maritime navigators) there is a strong relationship between the system state and the operator's workload. The workload was high on very specific locations and in very specific situations which were predictable using the time–to–boundary model.

CHAPTER 2

Monitoring behaviour

When controlling a system the operator has to monitor the system. This is done against a background of the operator's knowledge of the system, how it works and what goals he and it have. However, I will postpone discussions of a number of these aspects until later in this and subsequent chapters and focus first on the narrower monitoring part of the task. Monitoring is observing the system's displayed status information to decide if the system is in a normal state and under control (Moray, 1986). Particularly in a complex task or system, an operator cannot observe all displays simultaneously. There are simply too many. Therefore he must continuously decide which variable he must observe or monitor. Monitoring consists of several elements. The display, which can be a view of the real world or an instrument panel or other interface, consists of a number of sources of information, each with their characteristics and level of importance. The operator must schedule how to sample each variable. These samples are then used to retrieve information on the current state. The current state is not only the actual state at this moment in time but also includes information about changes, the rate of change, and possibly the acceleration of the change. This information then has to be compared to the desired state to decide on whether the operator has to intervene or not.

In the past several authors have developed models on monitoring behaviour, on how operators sample variables and how they distribute their attention over the different variables. Each of these models focuses on different elements of monitoring behaviour. One group of models is based on the characteristics of the variable. A second group is based on the operator, on what information he retrieves from the samples and how this is related to the need for intervention.

In this chapter a number of these monitoring models are discussed. The overview presented in this chapter begins with models that are source-oriented and moves to those that are observer-oriented. It will be argued that monitoring cannot be separated from the operator and his control tasks when the operator is human. The reason for this is that the human operator attempts to make optimal use of his capabilities and limitations and changes strategy in different situations. At the end of this discussion a model for this research is developed, based on a combination of elements out of the models discussed. This model in itself is very simple but can potentially account for a number of elements of monitoring behaviour in a complex environment, such as navigation. The central element of the model is that the operator focuses on the variable that requires intervention first.

2.1 Monitoring and control

2.1.1 Source oriented models

The first group of models do not take operator characteristics into consideration but model only characteristics of the information source. An important representative of this group is Senders (Senders, 1964, 1983). Senders studied sampling behaviour of flight instruments by pilots during an instrument landing approach. Senders postulated that an operator is a sampled-data, single-channel processor, and that the human processor attempts to reconstruct the associated time-signal from the sampled data. In terms of attention allocation, the allocation variable f_i is given by

$$f_i = 2c_1\omega_i log_2(\frac{\sigma_i}{E_i}) + 2\omega_i c_2 \tag{2.1}$$

where ω_i is the bandwidth of a signal y_i, σ_i is the standard deviation of y_i, and E_i is the permissible error of y_i. If the operator allocates his attention according to this model he can fully reconstruct all the sampled signals. Characteristic for this model however is that the operator's behaviour is not included in the model. The signal characteristics determine his sampling behaviour fully. Operator characteristics such as memory limitation or the actual value of the variable play no role in this model. The strong points of this model are that it is simple and accurate in uncomplicated situations with uncorrelated signals. The model explains why and how operators sample system variables and for this reason it is included here. However, it is unable to predict monitoring behaviour in more complex situations adequately (Senders et al., 1969). An attempt to cope with this shortcoming is the development of 'conditional sampling', discussed briefly later in this chapter. Another shortcoming is that in this model the operator is only expected to reconstruct the signal and make predictions of the variables based on this reconstruction. In control situations predicting is only one step of control. Other steps are the selection and implementation of control actions. These process variables are not random but they are the result of the process characteristics. The operator changes the parameters of the process, thereby changing the process characteristics and making the sampled data of limited value. For these reasons it is decided that source based models on their own are not sufficient for this study.

2.1.2 Queuing based models

A group of models that is based on a performance metric (and is therefore more operator oriented) are queuing models. Carbonell proposed a visual sampling model based on queuing theory (Carbonell, 1966). The basic assumption of this model is that the operator can only attend to one signal at a time. Several signals compete for the attention of the operator, and the model describes the dynamic assignment to the priorities of the different signals. This priority is based on the probability that a signal will exceed a threshold and the costs of exceeding such a threshold. The probabilities and cost are related by:

$$C(t) = \Sigma_{i=1}^{M} \frac{C_i P_i(t)}{1 - P_i(t)} \tag{2.2}$$

in which $C(t)$ is the total cost of not scanning or not looking at any instrument, M denotes the number of instruments, t the observation time, C_i the cost associated with exceeding a threshold of instrument i, and P_i the probability that instrument i exceeds threshold L_i at instant t.

This model is much more flexible for incorporating operator behaviour. It considers relative importance of signals, the risk of not paying attention to a signal, time related properties of the signals, and the time needed to process a signal. These elements will be incorporated and more closely linked to each other in the model proposed later in this chapter. However, this model has some limitations too. Correlations between signals are not considered, analytic predictions of the future state of the signals by the operator are not part of the model, and control is not incorporated into the model. Since each of these three play an important part in the control of complex systems a more advanced model is required. A step forward in this respect is the expected value model by Sheridan.

2.1.3 Expected value model

Smallwood (1967) recognized the need to incorporate an internal model of the operator's environment when modelling human instrument monitoring behaviour. The operator forms this mental model on the basis of past observations of the process, and uses this model to optimize his decisions. This model makes it possible for the operator not only to know the past state but also to predict the future state of the process, and to use this knowledge to sample more adequately and come to better control decisions. Based on the idea of a mental model Sheridan (1970) considered the question of how often the operator should sample a dynamic process. He introduced the concept of 'information value', the difference between the added information of taking a sample and the costs involved. An operator always has

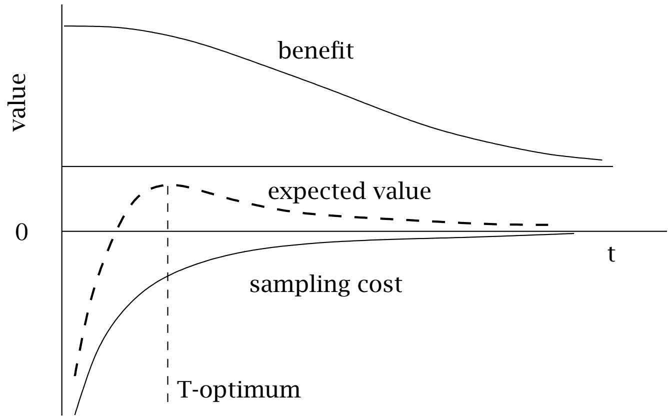

Figure 2.1: *The relationship between the information obtained by choosing a sampling interval t, and the cost associated with that interval. (reproduced from Sheridan, 1970)*

knowledge of the system variables based on past samples. Without taking any new samples the accuracy of this knowledge deteriorates. According to Sheridan the operator balances the benefit of taking a new sample to update his knowledge and the cost involved with this. A major difference with the former two models is that Sheridan gives the operator control over the process. The operator is not only monitoring the process, but also interfering in the process. He modelled this behaviour for a deterministic one-dimensional control process. The optimization of information and costs is explained with figure 2.1. In this figure the costs of sampling combined with the benefit of sampling to the operator provides the expected value, represented with a dotted line. The sampling interval is optimal when this expected value reaches its maximum.

Bainbridge (1978) discussed various aspects of the relationship between sampling and and control behaviour. According to Bainbridge sampling behaviour may not so much be related to predetermined intervals but is undertaken as the next activity after a sequence of other activities, when sampling becomes the behaviour with the highest expected value.

This expected value model approach assumes that the operator has an internal representation of the process, which is a major difference with the source-based and queuing-based models. Using the last observation available and his knowledge of the process the operator estimates the present situation. This estimation is not perfect because of an imperfect mental representation, errors in the samples, and disturbances working on the process.

Experiments by Sharidon & Rouse (1971) showed that operators do indeed apply this strategy to some extent. Because the operators could not predict the state of the system over longer periods of time, they chose shorter periods “over which they felt their skills more closely resembled optimal". However, the operators also showed risk aversion in avoiding large sampling times which might lead to large errors but which would keep the sampling cost low, and so give high profit. Another limitation was that the theory requires the operator to estimate the mean of the stochastic process signals, something the operators had great difficulty with.

A limitation of this model is that it is based on theoretical assumptions of operator behaviour, taking only the cost criterion into consideration. When controlling a system the operator applies various skills to optimise system performance and the various skill may be based on different criteria. In this perspective it is important to study actual operator behaviour.

2.2 Difference theory - laboratory

The theoretical models presented so far are mathematical models based on strict mathematical criteria. Because mathematical models are formal systems, all elements must be defined strictly. They are effective only in situations where all their assumptions about human operators are met. In most real situations these assumptions are not met. Operators ignore instruments for long periods of time. Relative importance, variable characteristics, and covariation of signals ensure such behaviour usually has no adverse effects. Skilled operators manage with very poor instruments although they tend to depend on "informal displays" in such situations. A proportion of data displayed to the operator is in many situations superfluous, but the operator has no problem ignoring many instrument readings during his work. These are indications that the models presented have a limited validity in real life situations.

The assumptions made in the mathematical models are generally time-invariance, perfect memory, no co-variation of the signals, gaussian sampling-errors, no a-priori knowledge, optimal decision-making, and goal independent sampling. These are not entirely valid for human operators who, of course, have both limitations and opportunities.

If the assumptions and constraints of the models are lifted somewhat, experiments show results that cannot be explained by the mathematical models presented. Crossman et al. (1974) found some remarkable differences between the theoretical sampling strategies and studies on operators in laboratory situations. Studying visual sampling behaviour using an eye-view monitor showed that the operator's sampling behaviour was consistent with the uncertainty analysis, but the variables scanned were of an "amplitude great enough to cause excursions exceeding the allowed tolerances". When the variables were within the specified range but the system was not correctly adjusted, the sampling rate was "controlled by the rate of drift, and rises when the variable is near either of the limits of its tolerance range." When a variable is outside its tolerance band and the operator makes a large stepwise control change, "a sample is taken after each control change at a time when response is expected to have reached some 80% of its final value.” This suggests a major difference between human operators and control theoretical models. The control theoretic models are built assuming only one objective. The experimental results presented by Crossman et al. suggest that the objective of the operator changes over time, and so does the sampling strategy. So, since the objective of the operator changes over time with the situation at hand, the actual sampling tactic cannot be predicted and decided upon completely beforehand. In other words, the operator’s behaviour is adaptive to the system state. If not all model assumptions are met by the user interface, the operator has several simple strategies available to handle the situation. These can be modelled using somewhat rougher models.

2.3 Monitoring and time-to-boundary

If we bring this knowledge to a real situation of maritime pilotage the problem becomes even larger. Calculating the monitoring parameters with the mathematical models at hand is almost impossible in such a situation. The task is too complex and, moreover, the maritime pilot scans the entire environment of the ship through the window. Which elements he chooses to observe and which characteristics of these elements is very hard to retrieve because the possibilities are almost endless in the different situations of his work. It is for this reason that we try to come to a slightly different monitoring model that can be applied in complex situations such as shipping.

With a closer look the results of the experiment by Carbonell are very logical if we project them onto reality. Imagine that we are sailing in the middle of the ocean with no other ships in sight. Then there is no need for frequent observations of the ocean around us. If another ship comes in sight the monitoring strategy depends very much on the possibility of a conflict. The closer the ship comes, the more frequent the observations which must be made. But this depends also very much on other factors determining the possibility of a conflict. If the other ship is faster and sailing ahead, monitoring needs to be less frequent than in the situation that our ship is faster and is overtaking the other. When the courses are crossing the monitoring becomes more and more frequent when the ships come closer together and a collision becomes feasible. In short, the monitoring strategy depends very much on the time that is left for intervention.

This is what is known as “conditional sampling”: the sampling frequency depends on the situation at hand. Sanders (1983) tried to model this, but excluded control behaviour, although his observations on driver behaviour do include this element. An experiment by Godthelp et al. (1984) showed that the time a car driver would voluntary close his eyes was proportional to the “time to line-crossing” (TLC). Carbonell (1966) showed that the sampling frequency depended largely on the expectations about the result of an in-

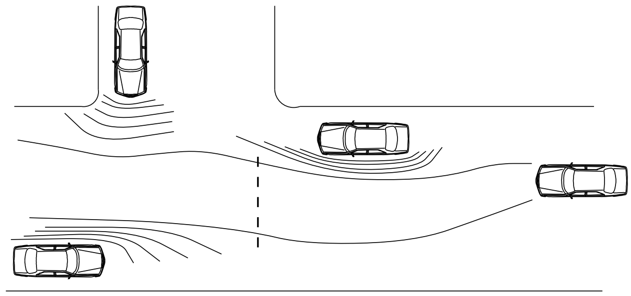

Figure 2.2: *The field of safe driving and the minimum stopping zone of the driver of the rightmost car. Adapted from Gibson (1938).*

tervention and that the sampling frequency rose significantly in situations where the operator felt that something unusual might be happening. Leermakers (1995) confirmed the idea of conditional sampling in an experiment on a four-instrument sampling task with Gaussian signals using the observing response method[1]. His conclusions were: " Monitoring behaviour is strongly dependent on the actually observed signal state. Especially near the event region, monitoring behaviour is more governed by the actually observed state than by global characteristics of these variables." (p. 119). All these studies lead to the time to boundary model developed in the next section.

Another line of study which leads in the same direction is that of visual pattern recognition. Humans are very good in visual pattern recognition. This visual information in turn is used to direct behaviour, but how this is done precisely is largely unknown. An answer may be found in an old article by Gibson & Crooks where they discuss a theoretical concept known as "fields of safe driving" (Gibson & Crooks, 1938). Each car driven on the road generates a field of minimum stopping zones. This is represented by graphs around each car of 'isostoptimes' (see figure 2.2). If this is done for each car on the road, one can find a trajectory through the traffic with a given minimum stopping zone level, which can be used to direct car-driving behaviour. This formulation generates boundaries within which the driver must manoeuvre. This rather simple idea has far-reaching consequences. Using very sophisticated pattern recognition techniques the driver can rely on rather simple tactics, which do not require the the sophisticated mathematical processing incorporated into the models discussed earlier.

Gibson's fields of safe driving are a function of time. Time is also a major variable of monitoring behaviour. When the minimum stopping-zone times decrease, the same number of observations and decisions needed to avoid collision must be made in this shorter time, resulting in a higher workload.

[1] Whenever the operator wants to sample a variable he must select that variable by pushing an associated button. The response to this request is considered as one sample.

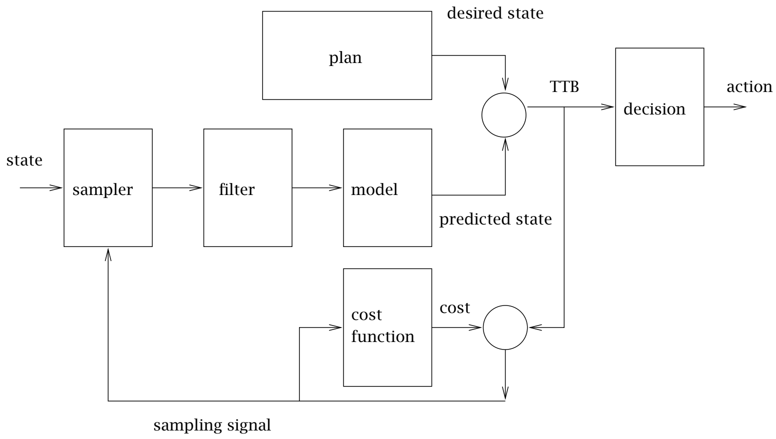

Figure 2.3: *Structure of the alarm-clock model.*

Although the examples given by Gibson always involve objects that travel at rather high speeds, his is not essential for the theoretical concept. The fields are extrapolations of the current situation and this extrapolation is something that is learned through experience. Therefore there is no reason to reject this theory in the situation of maritime pilots. The reason for introducing this old idea of "fields of safe driving" is that this includes the notion of changing barriers or boundaries. In navigation pilots have to deal with other ships that may have an effect on what is a safe trajectory.

These fields of save driving are almost identical to the idea of time–to–boundary (TTB). The major difference is that TTB is a single variable, whereas these fields are a two-dimensional representation of the environment of the car or ship. This representation shows the relationship between a location and the time–to–contact. These ideas will be combined in the next section into the model used for the rest of the study.

2.4 Alarm-clock model

The monitoring model proposed here is based on the control of a continuous process. It draws upon the single-server queuing-model by Carbonell and the expected value process by Sheridan.

Many variables need to be monitored. This process requires selection of the variables needed to control the process. This selection, or sampler, is the first stage in the model shown in figure 2.3. This sampler, or input selector, takes a sample from one of the state variables that is requested by a reference model. This request is in turn based on the expected value, derived from the cost assessment block.

These samples are fed into a filter needed to reconstruct the sampled signal before it is fed into a reference model. This reference model provides the operator's view of the actual state of the system based on these reconstructed signals. The actual state is compared to the desired state to come to control decisions. The desired state is retrieved from a plan. The origin of this plan will be discussed in later chapters. The result of the comparison of the actual state and the desired state is a time–to–boundary (TTB). It is this TTB that is used to make decisions about intervention. A second outcome of the reference model is an expected value, a indicator for the amount of information contained in the TTB-prediction. This signal is used to modulate the sampling signal by combining it with the cost associated with the state variable. The reference model is a mental model applied by the operator to make a prediction of each of the system variables. It is this part that allows the operator to turn away (temporarily) from one variable and pay attention to other variables in order to keep all system variables under control.

All variables in the system (e.g. for the pilot position, speed, rudder angle, rotation) have different expected-value functions and sampling costs associated with them. For each of the variables there is an optimal time for resampling which determines which variable it is that receives the momentary attention of the operator. Resampling will not only trigger the sampling process but at the same time the operator will compare the predicted state to the desired state. The decision on resampling depends on two variables: the expected value, obtained from the reference model, and the cost of sampling. When a maximum cost-benefit is reached a new sample is taken.

This process in itself is very simple when only one variable has to be monitored and controlled. This becomes different when studying a complex process such as navigation. In such situations many variables must be monitored and controlled. Each of these variables has to take its turn based on the expected-value theory. As a result a queue of variables "to be sampled" will build up under certain conditions. Not all sampling requests can be handled simultaneously. The result will be a degradation in performance due to poor predictions and maybe even the loss of an activated mental-model associated with any variable which is not sampled for a long time.

The outcome of the reference model provides the predicted actual state which is compared to the desired state for this variable in this situation. When the estimated time–to–boundary drops below the desired TTB intervention is required. This intervention involves the selection of a tactic. This tactic is retrieved from the plan which is part of the knowledge base of the operator. This tactic is then implemented by the operator. The sampling frequency depends in this way on the time left to the operator for intervention. When several variables come to a minimum the variable associated which the shortest time or the highest expected value will receive attention. If the required sampling frequency is higher than the maximum sampling frequency the sampling request will be dropped, resulting in a lowered performance. In this way the system state, workload, and performance are predicted to be linked in a relatively simple way. If the sampling frequency does not require the regular attention of the operator the model will not engage the operator, leaving him free to undertake other (mental) tasks, and hence his mental activity level may bear no relationship to what this limited model predicts.

The name alarm-clock model is chosen because the monitoring is driven by the model generated sampling signal, an alarm telling the operator that it is time to retrieve a new sample. This process is activated for all variables monitored, resulting in series of alarms.

2.5 Monitoring and workload

If this model is to be verified in a maritime situation the possibilities for measuring and experiments are very limited. The pilot's task must not be disturbed, so only 'passive' measurements can be taken to monitor his work. For this reason the use of a mental workload measure is chosen in preference to eye movement recordings or direct intervention in the task. Since the pilot's task is critical to the study its restrictions determine the choice of the measurement technique. The initial confirmatory experiment, described in chapter 4, must use the same technique for reasons of comparability.

I define workload as the number of observations (or samples) per time unit required to obtain the dominant parameters needed to control the system. Since task execution relies on skills and skills are automated single-step processes, the assumption that workload is a function of the sampling frequency is not unreasonable. This relates directly to skill based models of workload (see the next chapter).

The relationship between task and workload is the centre of much debate. In this debate two major groups can be recognized: those who attribute mental workload to limited capacity and those who relate mental workload to limited skills. Both these groups have provided conceptual models, helping to understand phenomena associated with mental workload. However, they are not performance models.

The capacity-based models include resource and effort based models and are all derived from the single-channel model by Broadbent (1958). This model is based on a single processing unit that behaves as a bottle neck for the processing capacity, leading to reduction in performance if the load demanded exceeds this maximum. This model of Broadbent was succeeded by a single resource model by Kahneman (1973) and multiple resource models by Kantowitz & Knight (1976) and Wickens (1980). Other successors are based on knowledge of the cerebrum (Polson & Friedman, 1988 and Kinsbourne & Hicks, 1978), or are based on arousal, effort, and activation such as Sanders (1979) and Navon & Gopher (1979). All these models focus on the source of information and predict an increase in workload with increased information. It was discussed by Duncan (1979) that the interpretation by the operator of that information can lead to completely different outcomes than those predicted by the capacity based models. It is precisely this element that is crucial in the alarm-clock model and these models are therefore not appropriate for explaining the relationship between the task and mental-workload per se. Techniques for workload measurement based on such models are therefore, a priori, less appropriate for measuring workload in this study, though they will be reviewed briefly in the next section for the sake of completeness.

The alternative group of models are the skill based models. Time and time again it has been shown that human capacity can be stretched further by training. To explain this phenomenon a theory was developed based on the fundamentals of skills. The basic assumption for this model is that the human brain does not have a limited capacity but that all limitations stem from incompatible applications of a shared schema or skill (Neisser, 1976). Skills and schemata are treated as one and the same. First, a skill is an acquired control structure that is stored in long-term memory. Second, skills are nested hierarchically (Schmidt, 1975). These skills are structured to reach an action goal, and these action goals are nested in a similar way. This would mean that if we practice enough our capacity should be unlimited. This is not the case and a theory for these limitations was proposed by Allport (1980). His model ascribes attention problems to interference. This interference may take place on the information side due to cross-talk and on the action side due to activation competition. Logan (1988) describes mental processing in terms of automaticity. Many memory phenomena can be explained by theories that depend on direct-access, single-step retrieval from memory. Novice performance is not limited by the lack of resources but by lack of knowledge (thus of skills).

If we accept the fact of input selection (we cannot look at two things at once), motor limitations, and attention, the skill-based workload model fits very well in the proposed monitoring model.

The hypothesis is that workload is a linear function of the number of alarm-clock alarms. This assumption is only valid if the system's information is available continuously, i.e. the information is not the result of a sampling process itself. This model accounts for workload variations due to variations in predictability, variations due to the number of variables, and the effect of overload. The next chapter will discuss how this mental workload can be measured.

CHAPTER 3

Measuring mental activity

In the discussion of monitoring behaviour it was argued that monitoring depends very much on the task at hand and how operators deal with this task. Because it is impossible to measure directly how operators use the information we need a way around in order to be able to measure sampling behaviour. In the last chapter it was argued that monitoring must be related to mental workload. In this chapter is discussed how this mental workload can be measured.

This is then used to choose a measuring technique that is able to record the elements of mental workload that are considered important. A very important criterion in this selection is the applicability of the technique in a real-life maritime environment. Heart-rate variability will be chosen as a best compromise.

3.1 Mental workload

Monitoring and control activities result in workload, or better still: mental workload. The state of the process must be recognized, the desired control activities decided upon, and implemented. All this requires the application of various mental resources and skills. The application of these is what is called workload. This may look simple, but it is deceptive. Much has been written on workload over the decades.

To facilitate a discussion on workload we need some definitions. Curry at al. (1990) provide some definitions. The first definition is on performance: *system performance is an objective measure of the system state variables relative to the desired characteristics of these variables* The measure of performance is derived from the goal or objective which the system is set to achieve. To accomplish this goal the operator will show activity. *The operator activity is the behaviour of the operator in executing the required task.* In addition to these two, some other definitions can be formulated. Operator performance is similar to system performance, but focused only on the operator in the system: *operator performance is an objective measure of the operator behaviour relative to the desired characteristics of his behaviour. The task demand load* or required activity *is the sum of activities required of the operator by the system to meet the performance requirements.* These activities include monitoring behaviour, decisions on control interventions, and the implementation of these interventions. This task demand load will result in workload. *The workload is the task demand load relative to the operator's abilities and the available time.* Since I will only consider the mental aspects of a task in this thesis, wherever I mention workload this should be read as mental workload. How much mental workload the operator invests in the tasks is determined by his effort. *Effort is conscious attention focusing, leading to higher sampling frequencies, and higher workload.* Greater effort results in higher workload which in turn results in a change in behaviour. In summary, *workload is a psychological construct which expresses the operator activity imposed on him by his tasks, resulting in system performance.* (see also Curry et al., 1979; Senders, 1983)

The hypothesis in this study is that the mental workload is the result of sampling and decision making and is proportional to the frequency of this cycle (sampling and decision making). How often this sequence is executed depends on how far in time the operator can predict the future with adequate accuracy and how much time is left for intervention. When the environment is highly variable the predictions will hold for only a short time and when the mental model is not very refined the uncertainty will grow rapidly in time.

3.2 Workload and performance strategy

In this study, the use of highly trained operators is an important factor. If skills are not yet well developed two processes will take place simultaneously: controlling the process, and learning how to control the system. Because untrained operators do not have adequate models to predict system behaviour and do not have appropriate tactics and control skills the relationships between performance, monitoring, and workload will not necessarily follow the alarm-clock model in predictable ways. It is generally accepted that the result of training is the development of various skills. Rasmussen defines skills as automated routines, based on subconscious time-space manipulations of objects or symbols in a familiar scenery (Rasmussen, 1986, p. 113). Once skills have been formed, the higher cognitive levels of control are not needed anymore, and the rules and knowledge on which they are based may deteriorate. As a result of the acquisition of skills the workload will reduce with the same, or even better, performance. The relation between performance and workload is then largely dependent on the available skills of the operator. It is important to notice that all cognitive levels of control are present all the time, but it is the particular strategy chosen and the skills available that will determine the relation between workload and performance.

Typical for skilled behaviour is that it is not open to introspection, unlike goals and strategies which can be obtained using various verbal techniques. One might even say that skilled behaviour is that part of behaviour that is shown to the observer but never verbalized. Thus in situations where highly trained operators are used, verbal techniques will only reveal a small part of the entire process. It is this skilled part of behaviour that is of most interest in this research; it governs the vast majority of the task or work of navigators – and of the system operators in this experiment.

It is generally accepted that any operator will try to achieve his goals with minimal effort by applying skills which enable him to do complicated tasks with high accuracy for longer periods of time. This is achieved through anticipation and the minimization of the number of control actions. The operator will try to stabilize the process and reduce the error slowly, avoiding an unstable process which in turn leads to a high workload. The operator will not necessarily keep all process variables on the setpoints but use the margins allowed for the setpoints in such a way that workload is minimized. This is a direct result of the time–to–boundary theory. If operators were not to strive for minimal effort, a relationship between performance and mental workload could not be established due to the unknown changes in effort.

3.2.1 Deterioration of performance

It has long been noted that task demand load and performance are related. When task demand is low performance is high, deteriorating with increasing levels of task demand load. Meister (1976) describes a hypothetical relationship between task demand load and performance. He distinguishes three regions of task demand load. In the first region the performance is invariant under low levels of demand. In the second region the performance declines with increased demand. And in the last region the performance remains at a constant low level under high levels of demand load. Later it was recognized that the performance may decrease also at very low levels of workload. This later relationship is commonly known as the inverted U-shape.

The relationship described by Meister stems from a long history of experiments going back to the beginning of this century (Yerkes & Dodson, 1908). Although often presented as the relationship between workload and performance they originally showed the relationship between arousal and performance in a learning experiment. When later used to describe the relationship between arousal and performance, the variables were ill-defined and the units missing. In this way the curves depicted have not only become uncalibrated, as noted by Rasmussen[1], but describe no experimentally val-

[1] Personal communication

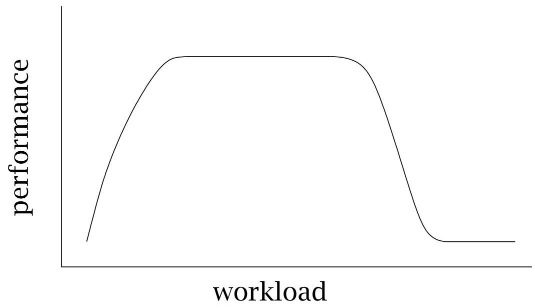

Figure 3.1: *The hypothetical relationship between workload and performance.*

idated relationship at all. The only reason why this 'law' has been around for so long is that it seems so logical. The existence of some deterioration of performance under high levels of task demand has been demonstrated, but attempts to show the left hand side of the inverted U-shape, where performance deteriorates under low levels of task demand have failed (Riley et al., 1984). Attempts to show a relationship between performance and workload have been numerous, usually with the result that a clear relationship could not be established (Yeh & Wickens, 1983). High levels of workload carry risks, but how operators adapt to such situations is largely unknown. From this we can conclude that workload and performance are not directly related and that either of them cannot be used to predict or explain the other directly.

3.3 Measuring workload

As was discussed in chapter 2 we are going to study monitoring behaviour using mental workload measures. To measure workload various techniques have been developed, which can be classified into three groups of techniques: secondary task performance, self-rating, and physiological measures. Within each group there are various techniques based on different theoretical models.

The many possibilities and limitations encountered when using these techniques are widely reported and will be discussed in this section. All techniques that have been developed over the years show a high degree of validity and reliability but their use is not without limitations. The technique that will be used in the navigator and process control studies must meet the demands for that particular type of research. This choice is discussed in the sections to come.

3.3.1 Secondary technique

A resource-oriented model for workload suggests the use of a secondary task to measure the workload of a primary task. The assumption is that the operator has a limited capacity and that his spare capacity can be used to perform a secondary task. The performance on the secondary task then gives an indication of the workload induced by the primary task. This approach also assumes that there exists a mathematical relationship between errors and the effort invested in the secondary task. Under a number of assumptions it is possible to draw a performance operating characteristic or POC, which plots the performance on one task as a function of the performance on a concurrent second task (Navon & Gopher, 1979).

In a simple task a secondary task can be a valuable tool, but with a complex one this secondary-task technique is much more difficult. In order to be an effective measuring tool the spare capacity for each of the elements of interest must be used by a secondary technique. Damos recognizes six task characteristics that must be decided on (Damos, 1991): the input separation, the input modality, the correlation between the stimuli, central processing, the output separation, and the response modality. Each of these factors will have its own influence on the POC and therefore on the specific implementation of the secondary task.

Designing a secondary task for navigation would be very difficult given the complexity of the navigation task combined with the secondary task demands discussed above. However, there is a more fundamental, theoretical reason for rejecting this approach. In the previous chapter it was discussed that resource-based models were not suitable for the tasks being studied. This type of model was rejected in favour of skill-based models. An equally important reason for excluding a secondary task technique is the fact that it has to be applied with maritime pilots during their normal work. Loading these pilots to their limits with an additional task was considered unacceptable from the standpoint of safety.

3.3.2 Self-rating technique

The assumption in self-rating techniques is that the operator is the best evaluator of the mental effort he invests in the performance of the required task. This technique is per definition operator-centred and subjective. From this point of view various techniques have been developed to enable collection of reliable data. This can be done using a linguistic approach, a numerical approach or a mixture of the two to come to a reliable and sensitive instrument of workload. In the design of the instrument the objective has been to develop “a rating scale that provides a sensitive summary of workload variations within and between tasks that is diagnostic with respect to the sources of workload and relatively insensitive to the individual differences among subjects” (Hart & Staveland, 1988). In the NASA-TLX this is achieved by combining six rating-scales using a weighted average based on ranking of these six scales. The SWAT-scale obtains a similar outcome by ranking of subtasks (Reid et al., 1981), and the RSME scale (Zijlstra, 1993) asks subjects to rate the workload on a single numerical scale.

A strong point of the subjective scales is their high face value. Also, they have been thoroughly tested and validated. A very logical argument for subjective rating is provided by Johansson: "If the person feels loaded and effortful, he is loaded and effortful. Whatever the behavioural end performance measures may show." (Johansson et al., 1979, p105).

However, changes in subjective workload ratings can have several causes and do not necessarily reflect changes in the task-demand load. Erikson & Simon (1980) report that subjective ratings are accurate in reporting the processing of information that is done consciously, but subconscious processing is not available for report. This would make subjective rating effective when the task is at the rule or knowledge based level and largely involves conscious processing of data, but ineffective with unconscious processing of highly practised skill based tasks.

A single task scale is most appropriate in situations where the goals are very focussed towards only one objective. In more complicated situations multi-dimensional techniques are more appropriate. SWAT and the TLX are the best known representatives of this group. For SWAT the work must consist of a set of well distinguished tasks that are separated in time. In navigation most tasks exist in parallel. Therefore TLX would be preferred, though even here the time period of the task must be divided into discrete blocks for judgements to be made.

Taking all of these points together, the use of subjective ratings in the case of navigators does not seem promising. The task is continuous, and most of the processing is at an unconscious level. A subjective rating is not sensitive for such a task. However, in order to learn more about the task of the navigator and because of its high face validity the TLX was still used, but not as a primary source of information.

3.3.3 Physiological techniques

This leaves the third type of measure: physiological. For some time now, reliable techniques have been available to calculate physical workload in terms of energy consumption. Oxygen consumption, heart-rate, and blood pressure can be used to estimate the subject's energy expenditure in performing the task. Blood pressure and heart-rate are known to be influenced by the effort required to perform a mental task. Other cardio-vascular parameters are more hidden. As early as 1876 Mayer described variations in the heart-rate that were slower than the respiratory cycle (Peňáz, 1978), and as early as 1963, Kalsbeek and Ettema found decreased heart-rate variability with increased task complexity (Kalsbeek & Ettema, 1963). The key element in the influence of mental activity on physiological variables is arousal, a state of preparedness of the body due to an increased activity in the nervous system. The cardiovascular effects of mental tasks are also described to be similar to a defence reaction (Mulder, G, 1980).

The mechanism that regulates arousal can be found in the part of the brain that controls homeostasis, the regulation of bodily functions under changing internal and external conditions. This mechanism can be found in the autonomic nervous system, a part of the peripheral nervous system. This part of the brain controls the heart, secreting glands, and involuntary muscles. The other part of interest is the central nervous system, the nervous system within the skull and spinal column, that includes the brain, the brain stem and the spinal cord. Activity in these structures can be measured using various techniques and can provide highly valuable information about the mental activity involved in executing a task. The variables that can be obtained include brain activity, muscle tension, muscle tremor, pupil diameter, and eye blink rate. Unfortunately, these techniques have no applications outside a laboratory due to the high sensitivity for environmental noise or the need for huge measuring equipment.

The functional systems of interest for this project are therefore only those influenced by the autonomic nervous system which can potentially be measured with sufficient accuracy in the working environment of the pilot, i.e. on the bridge of a ship, where the physical movements are limited, almost all information is presented visually, speech is infrequent and happens in short bursts, voyages take one to three hours, and the time constants of the system are in the order of minutes. The next sections review these body systems in turn: cardiovascular, respiratory, and hormonal.

The cardiovascular system: heart-rate, heart-rate variability, blood pressure

Several variables of the cardiovascular system are associated with workload, fatigue, arousal and stress, to which these variables all react differently. Workload, fatigue, arousal and stress are all closely related to the concept of mental workload. Fatigue is the result of a prolonged period of high levels of workload without periods of recuperation. Arousal is physiological readiness and stress is a physiological reaction to high levels of arousal for a prolonged period of time without recuperation. Blood pressure has been shown to be an indicator for arousal, stress and fatigue in aviation settings (Balke, 1966; Blix, 1974) but is not a good indicator for mental workload (Wierwille, 1979). Heart rate is affected by physical workload, the readiness for bodily movements (Lysaght et al., 1989) and general arousal. For the subject of mental workload heart-rate variability is of more interest. This variable has been recognized for some time now, and has on numerous occasions been shown to correlate well with mental workload (Aasman et al., 1987; Hanson et al., 1993; Hartman & McKenzie, 1979; Jorna, 1992; Middleton et al., 1994; Mulder et al., 1992; Mulder, 1988 & 1992, Orlebeke et al., 1985; Vicente et al., 1987; Wiethoff, 1997). Heart-rate variability (HRV) decreases when mental load increases, in particular in the region of 0.1 Hz. The causal relation between HRV and workload is largely unknown, but heart-rate variability is known to be part of the blood pressure regulating mechanism, and affected by heart-rate, respiration, skin temperature and age.

Respiratory system: respiratory rate, respiratory rate variability and ventilation

Respiration is an automatic process, regulated by a respiratory centre in the hind-brain. Connections with the cerebral cortex make limited voluntary control possible. Respiration rate and ventilation have been found to be affected by emotional states, stress and arousal (Thackray, 1969; Grossman, 1983). Due to connections between the respiratory and the cardiovascular system the HRV shows a component around 0.40 Hz, the phenomenon known as respiratory sinus arrhythmia. The correlation between respiratory rate variability and mental load has been suggested, but extensive research has not been conducted. Respiratory measures are severely modulated by speech. This makes it inappropriate for use in a working environment where speech is essential.

Hormone secretion: catecholamines and cortisol

Hormone levels are directly associated with stress, and have received extensive attention in the field of stress research (e.g. Marchbanks, 1958; Krahenbuhl et al., 1977; Frankenhaeuser, 1986). The central medulla secretes a group of hormones known as catecholamines, which are related to short-term stress. The adrenal cortex secrets a group of hormones known as corticosteroids, which are related to long-term stress. The level of catecholamines in the blood is of most interest when researching workload. The hormones may be collected from blood, urine, and saliva from the parotid glands. The secretion of hormones varies greatly with the circadian rhythm and food intake. The increase of hormone levels can be fast but the excretion of hormones from the blood is slow. The collection of samples has,however, severe practical limitations on the bridge of a ship with a crew present.

Each of the hormone secretions described is governed by mechanisms of the nervous system. The functional relationship between any of these secretions and mental activity is largely unknown, arousal excepted. Some secretions have shown high correlations with task-demand load parameters and subjective rating, but this correlation is largely explained by task characteristics such as modality, dynamic parameters and elements such as stress (e.g. Frankenhaeuser, 1986; Meijman, 1993; Rose & Fogg, 1993; Wiethoff, 1997). Taken together these drawbacks rule hormonal measures out for this study.

Drawbacks and advantages of physiological techniques

What criteria are of importance when selecting a workload-measuring technique? The selection lists for workload measurement techniques are many, and all techniques have been evaluated on these criteria. For instance, O'Donnel & Eggemeier (1986) evaluate on sensitivity, diagnostics, intrusiveness, implementation requirements and operator acceptance, whilst Curry et al. (1979) use sensitivity to primary variable, minimal interference with quantity being measured, insensitivity to other variables, known (and suitable) bandwidth and dynamics of transducer system, repeatability and reliability. Much attention is given to the first items mentioned in these list of criteria, but from a viewpoint of systems engineering very little attention has been given to bandwidth and dynamics of instruments. If the selected instrument is to be useful it must have a bandwidth to match the variable dynamics of the interval under consideration. This bandwidth in turn is derived from the systems dynamics and the theoretical model underlying the experiment. Rouse et al. (1993) modelled the dynamic relationship between subjective workload and performance. From earlier experiments Rouse reports a severe time-lagged relationship in moment-to-moment variations between workload, behaviour, and performance. Also, operators tend to persevere in their successive subjective workload ratings. The phenomenon of perseverance has received little attention in experimental research, but might have a large influence on human decision making.

The relationship between mental activity and physiological measures and the relationship between physical activity and physiological measures make these measures very useful for acquiring a deeper understanding of the task of the navigator or operator. However, by and large physiological techniques do not distinguish very well between physical effort and mental effort. Therefore the application requires a technique for filtering the effects caused by physical effort. This is most often achieved by minimizing physical effort or keeping it at a constant low level. A second and very serious drawback is the large amount of data which becomes available with physiological techniques. Data reduction can be bothersome if simple stochastic relationships between performance and operator behaviour or between operator performance and the recorded physiological variable do not exist. Due to large variances of the signals resulting from physiological processes, task-induced variance and noise, statistical significance requires large numbers of data. The scaling problems reported (e.g. Kantowitz & Weldon, 1985) are not different from secondary task and self-rating technique problems, the system dynamics being equal.

Ease of recording a number of variables is an additional practical advantage to the theoretical advantages mentioned earlier. It enables selective recordings of workload elements. Some of these recording techniques are almost unobtrusive and make continuous recording during normal task execution possible. This is particularly true for the cardiovascular methods. The dynamic response of the different techniques varies widely. Techniques such as heart-rate respond in the order of seconds, heart-rate variability in tens of seconds, to several minutes for certain hormone secretions. The correct choice of tools depends largely on the situation at hand. Each single technique is only valid within its own particular set of constraints. However where, as with the navigator we are dealing with a task which is highly skilled and the task structure is unclear, a physiological technique seems the only viable choice.

When applying this to the navigation task the bandwidth of the instrument must fit the dynamics of the ship. The time constants of a ship are in the order of tens of seconds to several minutes. This in the same order of magnitude as heart-rate variability. Considering this and all this other issues of sensitivity, unobtrusiveness of the measurement and the relative insensitivity to the incidental characteristics of the pilot task heart-rate variability seems an appropriate choice for this study.

3.4 Heart-rate variability

The heart-rate shows fluctuations of about 10% of the mean heart-rate. This is especially true under conditions of complete rest; with increased heart-rates this fluctuation generally decreases. These fluctuations are known as heart-rate variability or HRV. The magnitude of the HRV depends on the physical load, the mental load, and some unknown factors (e.g. Mulder, 1988, Aasman et al., 1987, Orlebeke et al., 1985, Grossman, 1983).

When analysing the HRV, three components can be distinguished: one with a period of about 50s, one of about 10s,and one of about 3s. Since all components have a periodic characteristic, analysis is often done in the frequency domain where each of the components are clearly observable. The lowest frequency component is concentrated around 0.02Hz and is related to slow regulatory processes such as thermo-regulatory activity (Kitney et.al., 1976). The second component, which is related to the blood pressure regulating mechanism, focuses around 0.1Hz (0.07–0.14Hz is a practical spectrum, see Mulder 1988). And the third component, which results from the respiration mechanism, focuses around 0.3Hz (0.20–0.40Hz is often used in research, Mulder 1988). There is an increase of heart-rate during inspiration and a decrease during expiration. The HRV has shown sensitivity to mental load. Mental load decreases the HRV in both the middle and high frequency band but particularly the area around 0.1Hz (e.g. Mulder, 1988). Hence subsequent analysis in this thesis uses this band.

Transforming the heart-rate from the time domain into the frequency domain is not trivial. Since the heartbeats are not a continuous signal, but are made up of a series of events in time, direct transformation is impossi-

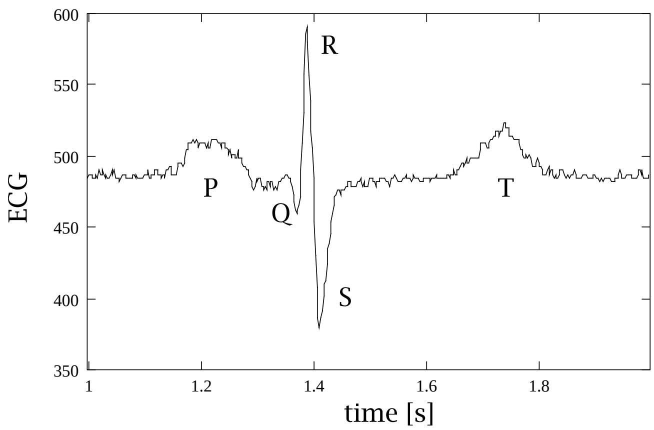

Figure 3.2: *The ECG of one contraction.*

ble. The event of interest is the moment of occurrence of the QRS-complex in the ECG. The contraction is initiated with the P-wave, and the contraction results in the QRS-complex (see figure 3.2). Ideally the technique would use the P-wave since the P-wave initiates the heart contraction, but the P-wave is much harder to detect than the QRS-complex due to the low signal-to-noise ratio. Since the time difference between the P-wave and the QRS-complex only depends on the conductivity of a node of nerve cells on the heart (the AV-node) the fluctuation in time for the time-difference is regarded to be small enough to ignore (< 5ms). In this research the R-peak of the QRS-complex will be used to record heartbeat (Rompelman, 1987).

3.4.1 The Fourier transformation

The time series of the heartbeat is the basis of the HRV-technique. An ECG provides a continuous signal from the electrical activity of the heart. An example may be seen in figure 3.3.

From this signal the occurrence of the R-peak is extracted and these peaks are used to calculate the frequency spectrum. This transformation is done using the IPFM-model, described extensively by Rompelman (1982, 1986, 1987). This model allows us to transform the heart-rate time series directly into a Fourier series based on a model of the heart using a Discrete Fourier Transform.

When the time series is transformed into the frequency domain it is generally represented as a power-density function. This power-density function shows all the frequency components present in the time-signal, with their respective energy. An example is shown in figure 3.4. The three characteristic elements can easily be recognized: the peak at very low frequencies ($f < 0.05$Hz), a distinctive peak around $f = 0.10$Hz, and a small peak around $f = 0.3$Hz. Because the Fourier transform is done using a short

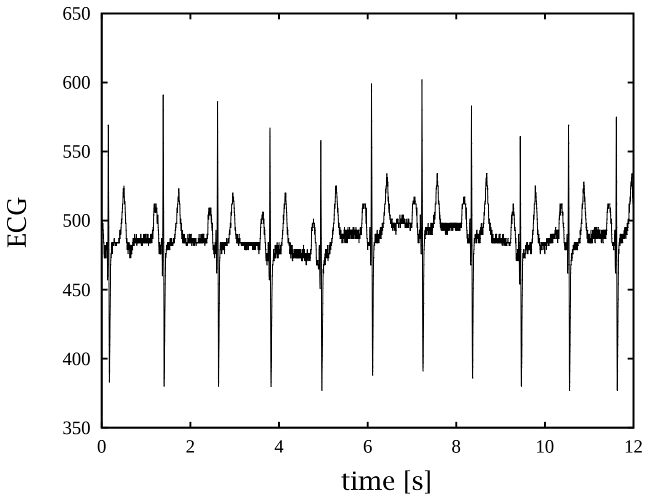

Figure 3.3: *Continuous ECG.*

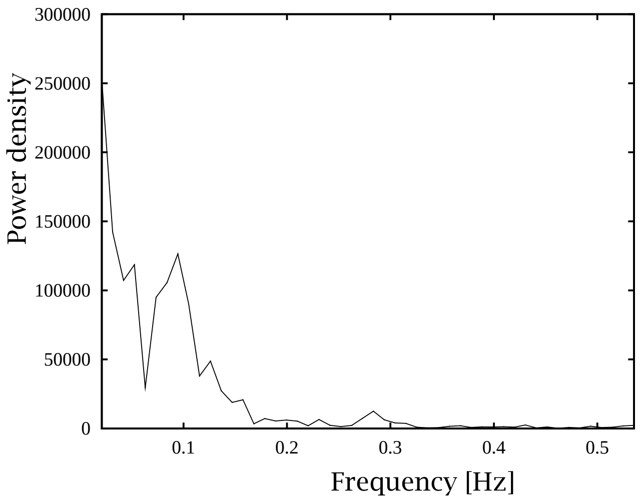

Figure 3.4: *Power density function.*

time period there is a cut-off near $f = 0$Hz.

Figure 3.4 shows all information of the time series in the frequency domain. Unfortunately this is not helpful when studying changes in the heart-rate variability. For that we need a different presentation.

There is an inevitable conflict when studying spectrum changes over time. The change-over time is included in the transformed data: no information is lost. One possibility is to make time-windows, and transform the data within that window into the frequency domain. Unfortunately, frequency resolution is the inverse of the time-window length: the shorter the period of time analysed, the less information in the frequency domain

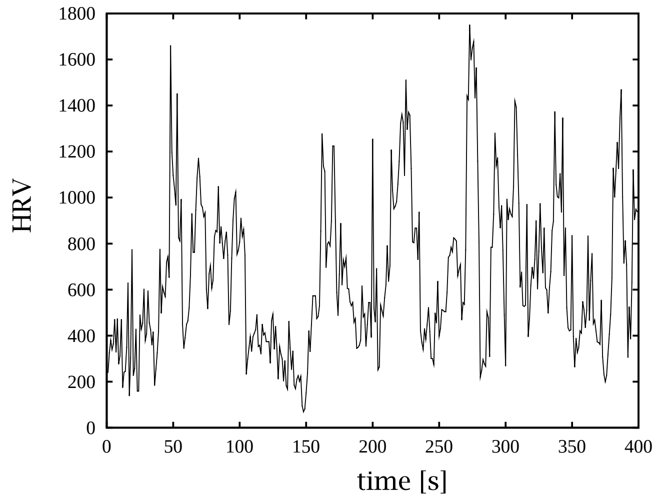

Figure 3.5: *Power as a function of time within the* 0.1 *Hz band.*

becomes available. An additional problem is that with a Discrete Fourier Transform, the variance of a value for one frequency component equals that value (which is the mean). As a result, when a better reliability is required it is necessary to average over a number of frequency components. For instance, if the time window is taken as $t = 100$s, then $\Delta f = 1/100$Hz, and when we are interested in the frequency range $f = (0.07 - 0.14)$Hz, the variance of the mean is $\sigma_{\overline{x}} = \overline{x}/49$, approximately 2%. In this study windows of $t = 50$s will be applied, giving a $\sigma_{\overline{x}} = 10\%$. This time is short enough to show time variations and long enough for acceptable variance. Shifting the window with $\Delta t = 5$s provides a smooth graph.

If we apply this technique we can generate a spectrogram where time, frequency, and power are presented in one graph. These graphs turn out to be very hard to read. It is better to select the frequency range of interest, and present a graph that shows the power in that frequency range as a function of time. An example of this is shown in figure 3.5. In this figure the power in a frequency area is calculated, in this instance $f = (0.07 - 0.14)$Hz, and presented as a function of time. Periods of low power, or low HRV, represent periods of concentration or high workload. These are the parts of the HRV data which will be related in the experiments to the periods of high workload according to the alarm-clock model.

3.4.2 Artefacts

The heart is a physiological system that does not always behave as predictably as one would like. A well known problem in heart-rate variability analysis are 'random' contractions, so called extra-systoles (see figure 3.6). Following the IPFM-model these extra-systoles must originate from large

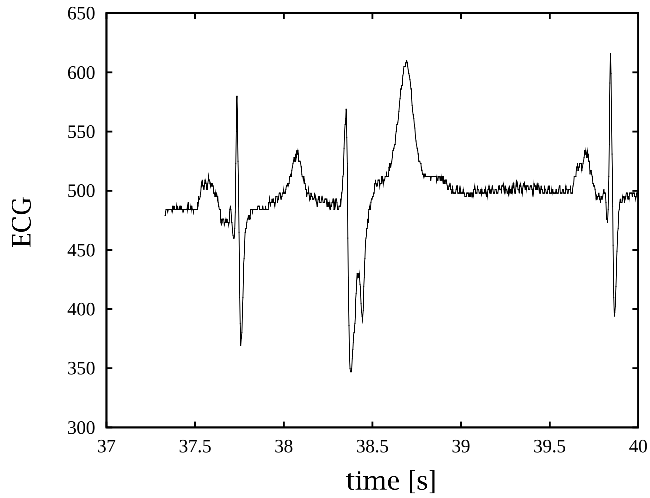

Figure 3.6: *An example of an extra systole (t = 38.5s).*

control signals. And this is exactly what will appear as a result of the Fourier transformation: large (extraneous) peaks that are the result of these extra contractions. The reasons for these extra-systoles are purely physiological, and they have no meaning from our theoretical point of view. So they are treated as artefacts, which means that they must be removed from the data or ignored.

An artefact that has not been removed will seriously distort the outcome of the timeseries analysed. For instance, removing one heartbeat from a timeseries is equivalent to the effect of a Dirac-function, making the outcome useless due to a very low signal-to-noise ratio that results from this. Reducing the window size does limit the effect of such artefacts. However, to deal with this artefact problem more effectively, in most analysis packages the heartbeat data is preprocessed to minimize the number of artefacts. The package used in this research is Carspan 1.99, described by Mulder (Mulder, 1988). In this preprocessing a moving window is applied, and within this time window the IBI's (inter-beat intervals) are calculated. From these IBI's, based on statistical parameters, outliers are identified. Next, attempts are made to remove or shift the outliers or insert an extra heartbeat without losing the overall time relationship. This may be successful in a number of cases, but not when a heartbeat is slightly delayed. The reason for this is that the correction is based on a different model than the IPFM-model. Corrections should be made in accordance with the IPFM-model. To realize this one would have to rewrite the complete analysis package. However, it proved impossible to write a separate correction program without losing the time relationship and plans to write a new analysis package were therefore stopped. In this research the original Carspan package was used and the remaining artefacts were ignored during the interpretation of the data.

3.5 Discussion

The use of mental workload can have many objectives. The use in this research is to provide a measure that can be compared to monitoring behaviour. There are many theories about the sources of workload, depending on the processing models used for cognitive decision making. The approach chosen here is that of unlimited resources in a skill based model. The source of workload depends on the number of control decisions, which depends on the number of (sampling) observations, which in turn depends on the quality of the prediction of the system state over time. How this works in a simple laboratory task is discussed in the next chapter.

How one measures workload largely depends on the actual situation at hand. In this research I have chosen physiological measures because they comply best with the assumption of a skill based model, which fits the situation of a skilled operator best. Of the many physiological techniques available the heart-rate variability appears to be the most promising to apply in the situation of a maritime navigator. Heart-rate variability is chosen because of its ease of recording, sensitivity, and dynamic characteristics. The remaining drawbacks are the absence of a fully researched and firm theoretical base, the high levels of noise in the data, and the interference from physical activity. It was felt that the last two could be overcome by careful recording and data processing. The first drawback cannot be dealt with here. Further theoretical work is needed. However, the empirical link between HRV and workload was felt to be sufficient to justify its use, especially when weighted against the practical drawbacks of other more firmly based measuring techniques.

CHAPTER 4

A process control experiment

The theory on monitoring behaviour is given an initial test in an experiment described in this chapter. The objective of the experiment is to confirm the theory contained in the alarm-clock model on the relationship between performance and mental workload in a man-machine environment during manual process control. In the previous chapters it has been argued that the workload is directly related to the time-pressure laid on the operator, the time the operator has until a decision must be implemented. This hypothesis is tested in this chapter using a simulated process (DURESS): the production of hot water with a boiler system.

The reason for choosing this simulator was threefold. The first reason is that DURESS is a rather complex simulator in which the operator is really in control. Second, it is a well tested simulator with highly skilled operators available. Third, the characteristics that were found to be important from a theoretical point of view in relation to workload were available in this simulator.

The operator performance will be expressed in different variables of the simulator status. The mental workload will be derived from the heart rate variability. It will be shown that heart-rate variability is a reliable technique to measure monitoring behaviour and that the alarm-clock model works well in this relatively controlled environment. In addition, several elements of control behaviour in relation to mental workload are discussed.

4.1 DURESS

DURESS is an acronym for DUal REservoir System Simulator, the process simulator used for this experiment. It is an experimental implementation of the Ecological Interface Design (EID), developed by the Cognitive Engineering Laboratory of the University of Toronto. The EID-concept was proposed by Vicente and Rasmussen in an attempt to design man-machine interfaces

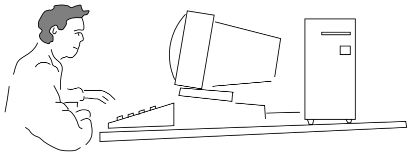

Figure 4.1: *The experimental layout.*

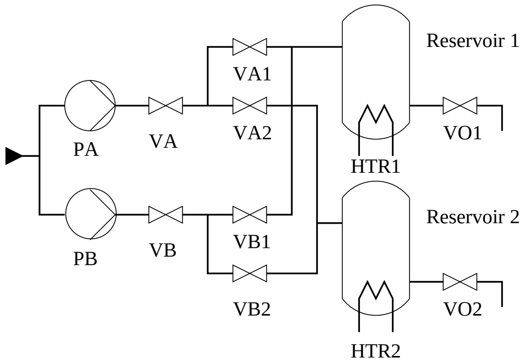

Figure 4.2: *The physical structure of DURESS.*

for complex work environments (Vicente & Rasmussen, 1992). It is based on the Rasmussen skill, rules, and knowledge taxonomy of levels of cognition (Rasmussen, 1986). The DURESS-system used in this experiment was actually DURESS II, a successor of the DURESS used for earlier research (Bisantz & Vicente, 1994). A longitudinal experimental study with the EID interface of DURESS was described by Christoffersen et al. (1996).

DURESS is a thermal-hydraulic process control simulation. It represents a number of elements that can be found in complex man-machine systems: it is a real-time, dynamic system with time lags in the controls. The dominant element is an integrating system with long time constants making it suitable for this particular experiment since ship navigation (the main focus in this thesis) has similar time constants for the controls (the ship is much slower with respect to inertia). The simulator consists of several interacting subsystems. It has multiple goals that can interact in a conflicting manner, and dysfunctional operator behaviour will lead to equipment failure.

The system, see figure 4.2, consists of two reservoirs ($1.5m^3$) that are fed by two feedwater streams. These streams can be configured to supply water to the two reservoirs in different ways. The general goal is to produce two streams of water, each of an externally determined quantity and temperature. This demand can change over time. To meet the demands the operator must maintain a water quantity in each boiler, he must satisfy the water output flow demand, and he must keep the reservoirs at a

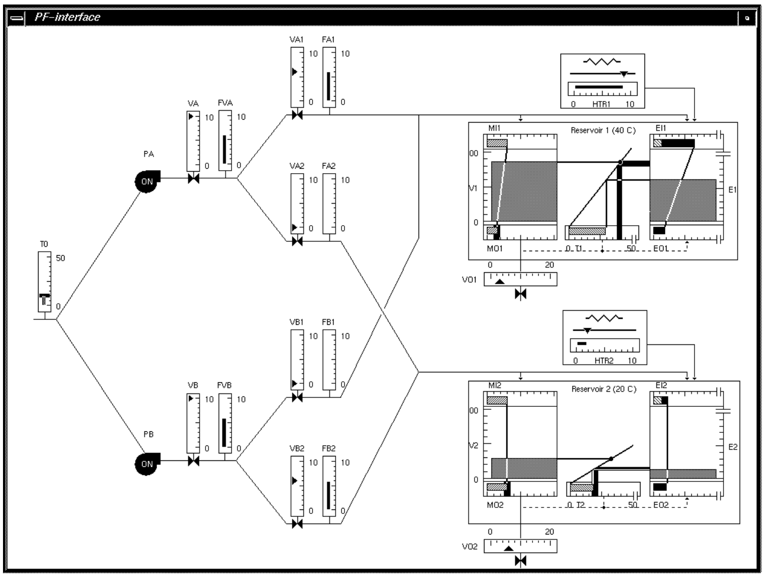

Figure 4.3: *The EID interface of DURESS (the graphical interface is in color, but has been converted to black-and-white for ease of reproduction). For a description see the text.*

prescribed temperature (20°C and 40°C). The input flow is controlled by the pumps PA and PB (20kg/s max.), and the valves VA, VB, VA1, VA2, VB1, and VB2, the output flow is controlled by the valves VO1 and VO2, and the heat transfer is controlled by the heaters HTR1 and HTR2 (2MW max.). All these components are subjected to first order lag dynamics, 15s or 30s for the heaters and 5s or 10s for all other components. The feedwater has a constant temperature of 10°C. Instruments are accurate and show no delay.

Equipment failures will occur when the pumps are switched on but the valves are closed, when the reservoir is filled beyond maximum level, and when the reservoir runs empty with the heaters on.

To control this system, DURESS has a computer interface based on ecological principles. The EID interface, represented in figure 4.3, shows a physical representation of the system with all control settings, mass and energy flows, temperatures and mass-energy balances. Together with the control display a clock is visible. DURESS is implemented on a Silicon Graphics Indigo workstation.

Apart from a set of conventional instruments to measure flows and temperatures, the EID interface shows higher order functional information. This information has a direct relationship with the process goals that the operator has to achieve. The box in the upper right corner is marked “Reservoir 1”. It has a graphical representation of the mass and energy balance. The mass balance is shown on the left, where the shaded area represents the water level in the boiler, the bar on the top the water inflow, and the bar on the bottom the water outflow. The inflow and outflow bar are connected with a line, representing the first derivative of the water-level. The graph on the right hand side of the box represents the energy balance. The energy stored in the boiler is represented as a shaded area, and the energy inflow and outflow are represented by a bar on the top and bottom of the graph, just as with the mass balance. The energy in the boiler is the mathematical product of mass and temperature. This multiplication is represented by the lines in between the two balances. The one line is the line with a slope. The slope of the line is determined by the mass present in the boiler. The second line is the vertical line leaving from the temperature representation. The intersection of these two is a value of the energy.

The actual outputs are the three horizontal bars near the bottom of the reservoir box. The outputs demanded are the vertical elements in two of the three instruments: the water output flow and the temperature. The required temperature is also projected on the energy graph, and when the actual temperature corresponds with the requested temperature the energy level is within this temperature projection, as can be seen in the box of reservoir 2 in figure 4.3.

The control settings are operated with a mouse pointer, and can be changed by either clicking at the requested value, or by dragging the marker to the requested value. The settings are represented with triangles, and the actual value is represented by the thermometer scales next to the settings.

4.2 The experiment

The experiments consisted of six trials. The first two were training trials, and the last four were experimental trials. The first two trials (1 & 2) consisted of one part, where the operator had to start the simulation and maintain one set of output demands for five minutes. The five minutes period started as soon as the output demands were met. The trials 3 to 6 each consisted of two parts. In the first part the operator first had to meet one set of output demands for five minutes after which the setpoints changed and he had to meet a new set of output demands for another five minutes. In trial 3 and 4 the simulator characteristics were the same as in trial 1 and 2. In trial 5 and 6 all time constants of the controls were doubled (the pumps, valves, and heaters), making the controls feel more ‘sluggish’. The last trial (trial 6) was special. In the first five trials the demands could be met with normal control settings, but the last trial the second set of demands requested an output flow higher than the maximum input flow. To meet this demand the operator had to use about 2/3 of the maximum capacity of the boiler in addition to the maximum input flow. This was something the operators had never done before and required the subjects to go beyond a skilled level of control.

Trial 3 was initially an experimental trial. However, the analysis showed that at the start of the third trial the heart-rate was sometimes not yet returned from an initial high level at the start of the experiment to its normal level. It was therefore decided to treat trial 3 as a training trial and use only trial 4 to 6 as experimental trials. The whole experiment took about 1.5 hours.

During the experiment all control activities were recorded by the simulator, together with all system variables. This could later be used to produce data files with various new variables, and the recorded data could be used to play back each trial.

During the experiment the heart-beat (R-peak) of the operators was recorded with a separate recorder. The recording was done with a portable, micro-processor controlled recorder with a 1ms accuracy. AgCl–electrodes were used to obtain the heart-signal, that was fed to an R–wave trigger to obtain the heart-beat signal. Markers were added to the recording to indicate different stages in the experiment. These data were later used to calculate HRV values as a measure for mental workload.

4.2.1 Operators

The operators were seven graduate students and one staff member of the engineering department, all male, in the age range of 23-31. Half of them had participated in previous experiments and they had used the system one hour every day for several months in an earlier experiment (Christoffersen, 1996). The other half had been involved in the development of that experiment. All were therefore skilled operators. They were paid 20 $Can for participation in this experiment. All subjects were non-smokers and in good physical condition.

4.2.2 Procedure

Since all operators were well-trained there was no need to instruct them about the DURESS-simulator. They were given written instructions about the experiment. The instructions explained that they had to do six trials, first two with one setpoint, and then another four with two setpoints, and they had to maintain those setpoints for four or five minutes. They were also told that no random component failures would occur, but all parameters of the system could change in between the trials. The whole experiment would take about 1.5 hours. After these written instructions the subjects were informed by the experimenter about the heart-beat recorder and were hooked up to the system. They had been asked not to take any coffee or Coke for one hour before the experiment.

All six trials were recorded so they could be used for later analysis. After the six trials the subjects were asked to remain seated and relax for another 10 minutes to obtain a baseline for their heart-beat. During the different trials the operators were asked not to speak because of the possible interference with the 0.1 Hz signal.

4.2.3 Hypothesis

The alarm-clock model predicts for this type of process and simulator that:

- When the system is under control the mental workload is a linear function of the time–to–boundary.
- When the time–to–boundary becomes large the mental workload becomes independent of the TTB.
- Control is largely based on skilled control sequences, especially during familiar situations.
- The physical activity for these control sequences has no effect on the mental workload.

The first hypothesis is the direct result of the theory on monitoring. To prove this relationship the correlation function must show a clear relationship between the two variables HRV and TTB. The correlation function is given by

$$r_{xy}(\tau) = \frac{R_{xy}(\tau)}{\sqrt{R_{xx}(0)R_{yy}(0)}}$$

The function $R_{xy}(\tau)$ denotes the cross-covariance function with "$x(t)$ leading $y(t)$", and $R_{xx}(0)$ and $R_{yy}(0)$ denote the auto-covariances.

The second hypothesis is a direct result from the first one. When TTB becomes very large no further monitoring is required and the subject is free to do what he wishes. The third hypothesis is based on the assumption that the operators manipulate the controls entirely on a skilled level.

4.3 Data and analysis

4.3.1 Variables

From this experiment three types of data were collected. First, the DURESS system automatically logged all subject control actions with their respective times. Second, it logged all variables also shown by the EID interface, but stored them in a numerical format instead of the graphical representation used in the EID, with one second intervals. The third type of data consisted of heart-beat data. Markers were placed in the data files to synchronize the data later. From the data logged by the DURESS system the performance is calculated.

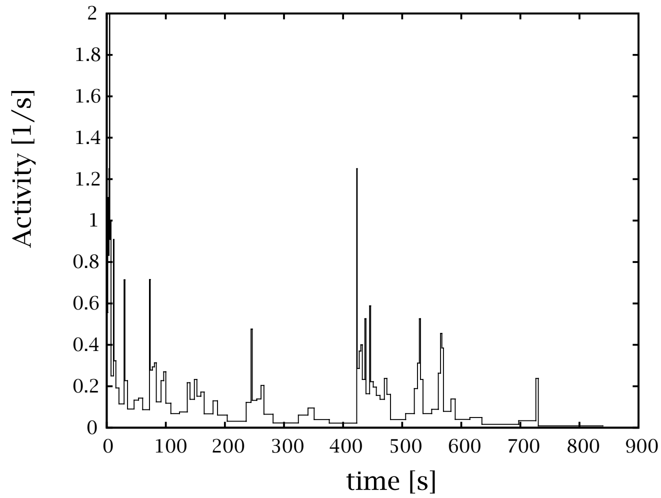

Figure 4.4: *Activity rate of one trial, activity [1/s] against time [s].*

The data were analysed in different ways by extracting various types of performance from the DURESS log-files. First the different types of control actions and their respective times were extracted. From this the distribution of control actions could be calculated. Second the different output variables were extracted of the production of the boilers, the temperature of the two reservoirs and the output flow, together with the demands. The third type of information proved its use during analysis: the water level in the reservoirs. A water level error was the only type of failure that affected operator behaviour during the trials. Other types of failure had no influence on the operator behaviour.

If the distribution of the control actions is plotted as a function of time it appears as in figure 4.4. The activity rate is calculated as

$$a(t) = \frac{1}{T_{act}(i+1) - T_{act}(i)}, T_{act}(i+1) \leq t < T_{act}(i)$$

where T_{act} is the moment a setpoint is changed using the controls. In figure 4.4 at $t = 410$s the setpoints are changed, and the trial finishes at $t = 870$s. It shows clearly a very high rate of control actions at the start (up to $t \simeq 25$) of the experiment up to 2 actions/s that reduces very rapidly to stabilize at a low level of activity. When a change of setpoint is given after the first goal is met ($t \simeq 400$), a second peak occurs, not as high as the first one (figure 4.4).

The second type of performance data is the error from the output demand, i.e. the difference between the actual output and the demanded optimal output, added across all four demands (two output flows and two temperatures). Since two types of demands exist the data have to be normalized. This is done by summation of all output errors normalized to the same (arbitrary) level. The result of a trial is shown in figure 4.5. The first

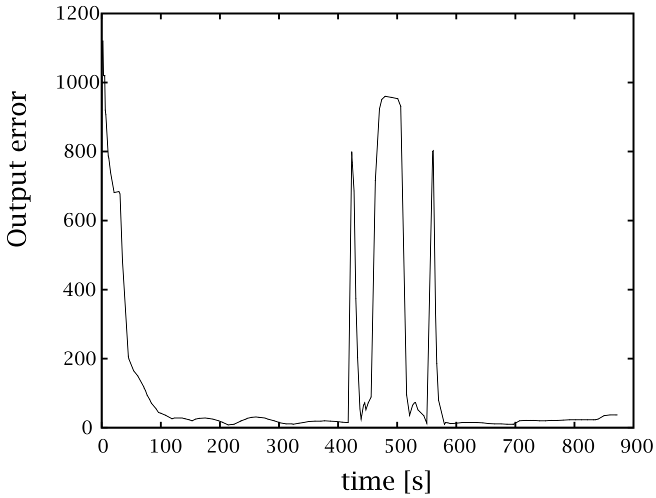

Figure 4.5: *An example of the output error [1] against time [s]. The same trial is shown as in 4.4.*

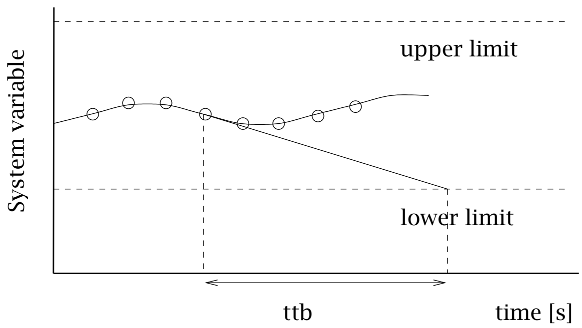

Figure 4.6: *Calculating the boundary time.*

100s shows the capture phase and achievement of the required output. At $t \simeq 400$s the setpoint is changed, followed by a new capture phase between $t = 420$s and $t = 575$s, after which the demands are again met.

Visual inspection of the performance data suggested that operators were not so much optimizing the output error as making sure that the output stayed within the specifications. Some very experienced operators even preferred to stay near the limits. This was later confirmed by interviewing some of the subjects. The maximum error for the temperature was 2°C, the maximum error for the output flow was 1kg/s. From the actual output the first derivative from the sampled data was calculated and used as an estimate of the time it would take to violate the specified limits for each of the four output demands (figure 4.6).

Additionally the operator had to take care that the water level in the two

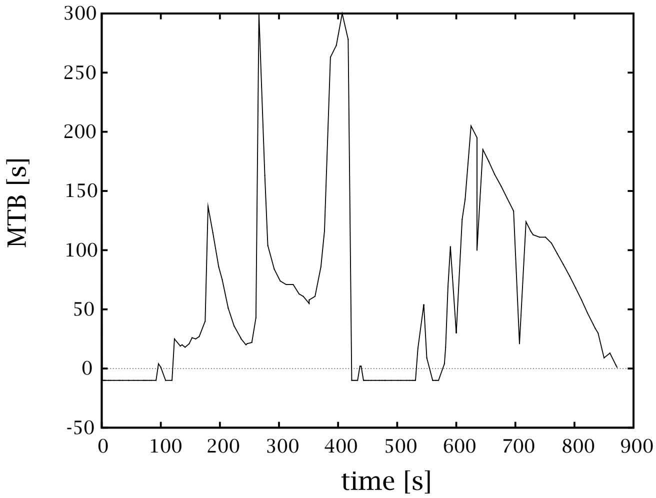

Figure 4.7: *An example of the minimum time to boundary against time, using the same trial as in figure 4.4.*

reservoirs remained within the limits. The time to reach the two reservoir water level limits was calculated in the same way. From these six values the minimum time was taken as a variable of the process stability, the ‘minimum time to boundary’ or MTB. Calculations using an average time to boundary[1] and a mean time to boundary[2] were also analysed. Because the displays have limited resolution it was estimated that the maximum ’time-horizon’ was 300s. When the demands are not met this is marked as –10s in figure 4.7. The downhill slope near the end in figure 4.7 is the result of the steadily reducing water level in the reservoirs.

The heart-rate recordings were analysed using the Carspan analysis software package (version 1.96). This package has many features to analyse the heart-data, including heart-rate and frequency analysis (Mulder, 1988). Outliers were removed based on a time window of 50s and a maximum deviation of the RR-interval of 4.0 standard deviations (default). If an outlier is detected the program tries to remove it by inserting or removing a heartbeat or by moving a heartbeat. These corrected data was then transformed into HRV signals using the IPFM model. The frequency bands were 0.02–0.05Hz (low), 0.07–0.14Hz (middle), 0.15–0.50Hz (high) (all default). For the anal-

[1] average time to boundary is defined as:

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} T_i$$

[2] mean time to boundary is defined as (harmonic mean):

$$\bar{t} = \frac{n}{\sum_{i=1}^{n}\frac{1}{T_i}}$$

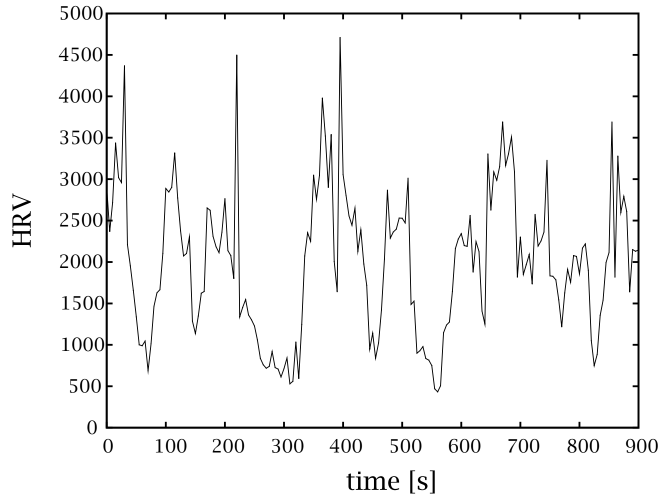

Figure 4.8: *The 0.1Hz power (smi), using the same trial as in figure 4.4.*

ysis only heart-rate and the middle-frequency band ($f = (0.07 - 0.15)$ Hz) were used. The power in the 0.1Hz band was used as a measure for the mental workload of the operator[3]. Harcor diagnosed about 0.37% of the heartbeats as artefacts and about 75% of these artefacts was corrected by Harcor. The DURESS-HRV data required no additional corrections.

In the Duress experiment the MTB was available for periods of 300s (this was the time the operators had to keep the system stable). Therefore correlation functions could be calculated with 300s signals only. There were 54 of such 300s trials available. These 300s were copied from the complete HRV signals. If the subjects managed to obtain a $MTB > 200$s the trial was used up to that point and the signals were padded with the border values. The reason for this procedure is that if the system became completely stable the relationship between process variables and HRV was lost.

In order to reduce the noise level the HRV signals were filtered using a low-pass filter with $F_c = 0.03$Hz, 6dB/oct, using FFT, so that the spectrum of the HRV and the MTB had comparable characteristics. An 'ideal' filter was rejected for its poor impulse response. These filtered HRV signals and the calculated MTB were re-sampled on 1024 points using linear interpolation. The correlation functions were calculated using these HRV and MTB signals. For more information on this see section 7.4.1. Figure 4.8 shows an example of the HRV signal.

The information that proved to be the most interesting for understanding operator behaviour in this experiment are the control settings, the minimum time to boundary and the HRV (0.1 Hz). These data can be combined in one graph to achieve better understanding of the relation between them

[3]Please note that workload and HRV are reversely related: The power of the HRV decreases when the mental workload increases.

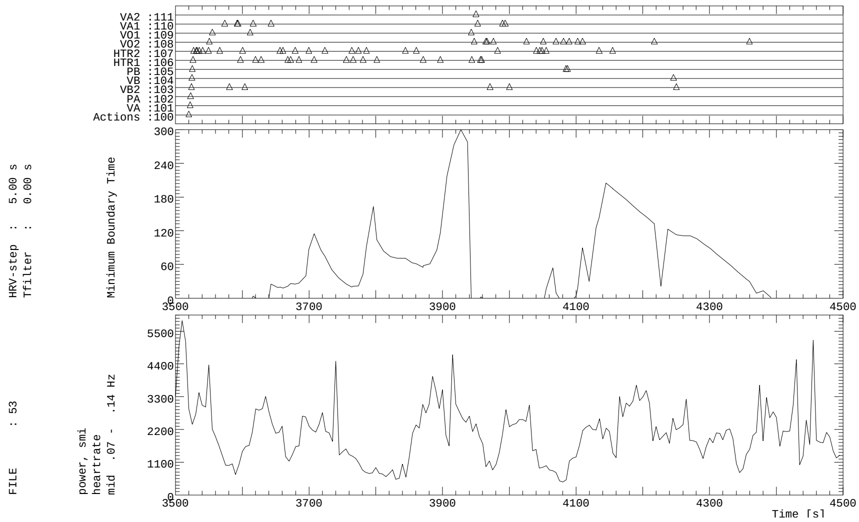

Figure 4.9: *Combined graph, using the same trial as in 4.4: control actions, HRV, and process stability. See page 36 for coding of controls.*

(figure 4.9). The top graph shows all control actions, marked by their control name. It is easy to see that the heater settings are used most (HTR1 and HTR2), together with one output valve in the second part of the trial (VO2). The middle graph shows the HRV during the trial, and the bottom graph shows the minimum time to boundary (process stability). The whole trial took about 15 minutes, and started at $t = 3520$s. The subject first stabilized the simulator at $t = 3640$s for the first setpoint, and again at $t = 4100$s for the second setpoint after losing control for 200 s. All displayed data are unfiltered.

4.4 Results

4.4.1 Heart-rate

All subjects showed an increased heart-rate at the start of the experiment, which lasted for about 15 minutes (one trial). There are two possible causes for this. One is the experiment with the simulator itself, the other one is the fact that the subjects are connected for the first time to a heart-beat recorder. These data were therefore not used for the workload analysis.

All subjects were asked to relax and do nothing for ten minutes after the final trial. Some subject were better in doing this than others, but they all provided a stable high HRV which served as a baseline and that was significantly higher than the HRV during the experiment (i.e. showed less workload).

4.4.2 Correlations

Various correlations were calculated from the data. The most important correlation is that between HRV and MTB, that is discussed in the next section. Other correlations are between output error and activity rate, workload and activity rate, and workload and output error.

Correlation between HRV and Minimum Time to Boundary

The monitoring theory predicts a causal relationship between the minimum time–to–boundary and the heart-rate variability. For this the correlation function between the MTB and the HRV was calculated. The MTB exists only when the system variables are within the target zone: a small area around the demanded output. As a result of the experimental setup each trial provides two series of measurement: one for each setpoint. This results is 54 correlation functions (9 subjects, 3 experimental trials, 2 setpoints per trial). The relationships outside these two setpoints are discussed later.

It was expected that the relationship between the mental workload and the minimum time–to–boundary was lost when the MTB became large. In the experimental results this effect was found at a MTB of about t= 180s. From that point the measurements of HRV indicated that the workload increased. The debriefing confirmed that in these rare situations the operators started to focus on other aspects of the experiment as soon as they had stabilized the process (e.g. quality of the graphics, events in past trials). In the further analysis the calculation were therefore stopped when the process was stable (MTB= 200s).

Five correlation functions with $r_{xy}^{max} < 0.10$ were excluded from this calculation because the HRV was severely affected by events outside the experimental setting (i.e. telephone ringing). The results after filtering and using only data that had a MTB < 200s are shown in figure 4.10. This figure shows that the estimated mean correlation-function between HRV and MTB has a maximum value of $r_{xy} = 0.72$ and that this maximum is reached at $\tau = 0$. In addition a 95% confidence interval for the mean was calculated in order to check the significance of this mean correlation function. The shape of the correlation function and size of the correlation were the indicative factors for the significance of the correlation function. The short time period of the analysis (300s) compared to the first crossing of $r_{xy} = 0$ at $\tau = 80$ was not considered a problem.

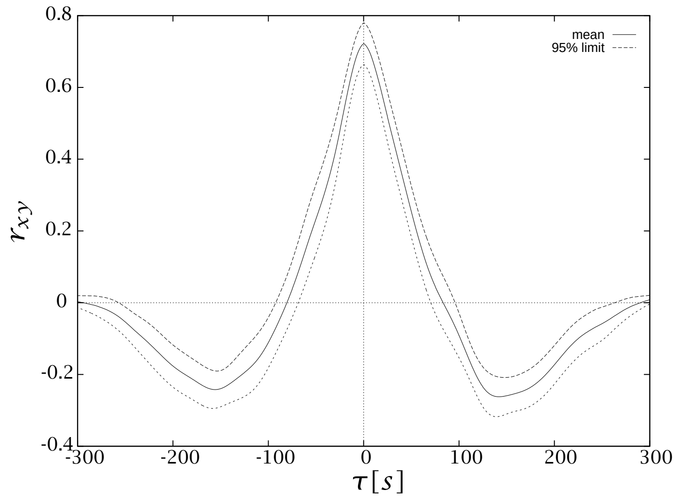

Figure 4.10: *The correlation function between HRV and MTB. From all correlation functions the mean and a 95% percent confidence interval was calculated.*

To see if indeed the minimum time and not the mean time was the best predictor for the mental workload the correlation functions between the the HRV and the harmonic mean of the time to boundary was calculated. The overall correlation between the harmonic mean of the time to boundary and HRV was
$r_{xy}(\tau = 0) = 0.58[0.49, 0.71]_{p=0.95}$
slightly worse than with the MTB, which was
$r_{xy}(\tau = 0) = 0.72[0.66, 0.78]_{p=0.95}$ (figure 4.10)
The correlation between the average time and HRV was far worse
($r_{xy}^{max} < 0.40$).

The introduction of the unfamiliar situation in the second setpoint of trial 6 required special attention during the analysis. The subjects started the second half of trial 6 similar to all other trials. However, as the the system variables came near the setpoints it became clear to the subjects that the normal strategy would not work. At this point a new strategy had to be developed. As soon as this new strategy was found the subjects started the second phase of trial 6 and were able to complete the task. The analysis of this phase could be done as in all other trials. The time the subject needed to find this strategy was excluded from correlation functions.

With respect to the timeshift of the correlation functions the hypothesis was that a maximum correlation would be found for $\tau = 0$. The maxima of the correlation functions were determined if such a maximum was within a maximum time shift of $\Delta t = \pm 60$s and if this was found to be acceptable on visual inspection of the MTB and HRV[4]. This was not immediately confirmed.

[4]A change in lead to lag or vice versa during a trial seriously decreased the correlation.

Although the mean correlation function did show one clear maximum at $\tau = 0$, the maxima of the correlation functions varied considerably between trials. With some trials the HRV was ahead of the MTB, and with some this was reverse.

It was calculated that over all trials there was a minor significant difference between subjects over the time-shifts ($F_{8,1} = 1,81, p < 0.10$). All subjects closely involved in the development of DURESS had a positive average lead-time, and all other subjects a negative one (the same result was obtained when the median instead of average was taken), suggesting that the developers had another quality of mental model and therefore showed another type of control than those subjects who were involved in the original experimental trials.

Visual analysis and interviews revealed a second difference between these two groups. When confronted with the unfamiliar situation in trial 6, the developers were puzzled by the situation and tried to discover what was happening which took time and effort, whereas the operators trained normally simply tried to achieve their goal.

Other correlations

Various other correlations were obtained from the data in order to test the hypotheses. If there is a causal relationship between MTB and HRV there cannot be a relationship between the output error (deviation from the desired optimal setpoint) and the HRV. To check this the correlation functions were calculated between output error and HRV during the control phase. The maximum value of the correlation functions was:
$r_{xy}(\tau = 0) = 0.07 \quad [-0.08, +0.19]_{p=0.95}, n = 27$
This result shows that there is indeed no relationship found.

A second derived hypothesis is that because of the skilled behaviour of the operator and because there is no additional monitoring needed for executing his skills, there is no relationship between the activity rate and the HRV. For this the correlation functions between activity rate and HRV in the control phase were calculated. The maximum value of the correlation functions was:
$r_{xy}(\tau = 0) = -0.05 \quad [-0.18, +0.13]_{p=0.95}, n = 27$
This result shows that there is indeed no relationship found.

Based on the idea of pre-programmed sequences of actions to deal with common situations the hypothesis is that operators show high activity rates when the process is in a normal out–of–bounds situation, such as during startup or when changing setpoints. From this it is hypothesized that the output error correlates with the activity rate during the trials. In this way the operator keeps the out of range time as short as possible, and attempts

No attempts have been made to correct for this. The visual inspection was done to see if the maximum correlation calculated and a lead or lag time estimate were approximately the same. In three cases the difference was so large that an estimate could not be established.

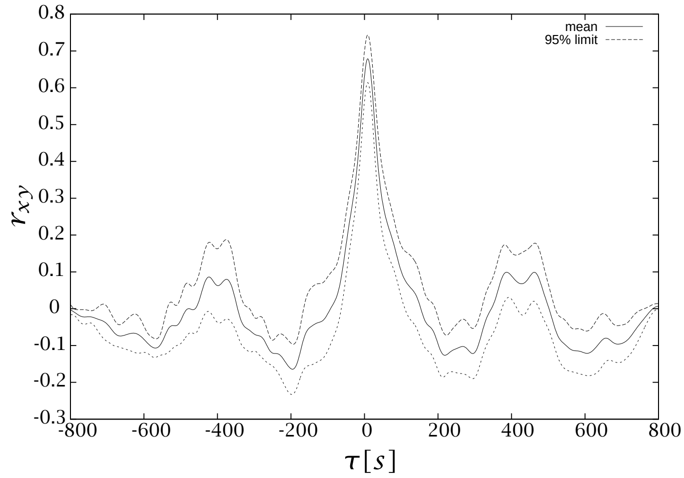

Figure 4.11: *The correlation function between activity and output error. From all correlation functions the mean and a 95% percent confidence interval was calculated.*

to stabilize the system during normal operation. As can be seen in figure 4.11, there is a correlation between activity rate and output error:
$r_{xy}(\tau = 0) = 0.68 \quad [0.61, 0.74]_{p=0.95}, n = 27.$
This result shows that there is indeed a relationship between the output error and the activity rate. The hypothesis of preprogrammed action sequences cannot be rejected.

4.4.3 HRV pattern

The theory on the relationship between MTB and HRV is only valid when the system is under control. When the system is 'out of boundaries' the MTB has no meaning, and the result of the relationship between output error and activity rate show that in these situations control behaviour seems to depend largely on preprogrammed action sequences. The hypothesis is that during these action sequences the mental workload is low (i.e. the HRV is high), and that the mental workload increases rapidly the moment the system is brought within control ranges. In order to throw some light on this the pattern of the HRV through the whole of the trial can be examined. What we are looking for are the coarser changes in workload (HRV) within and between the phases, rather than the detailed relationship between HRV and MTB within these phases when the process is under control.

This analysis was done using trial 4 because this was the trial that was 'standard'. In trial 5 the time-constants were doubled and in trial 6 a new problem was presented. For the start-up phase and change in setpoint the average HRV values were calculated to obtain HRV patterns. Figure 4.12

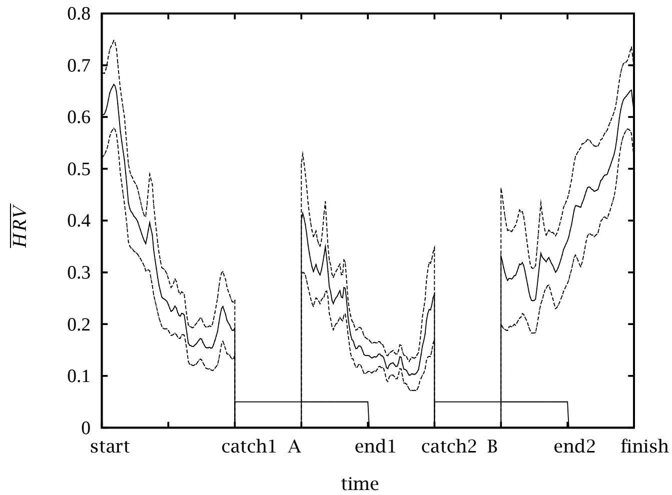

Figure 4.12: *Average HRV for all subjects during trial 4.*

shows the pattern of trial 4, averaged over all subjects during three phases: the start-up, the change in setpoint, and the end of the trial[5]. The controlled part of the trial, when the process falls within its boundaries, is deliberately left blank to focus attention on the workload in the other phases. Figure 4.12 is a composition of the characteristic phases in the process: start-up, change, and finish. From *start* to *catch1* is the start-up phase, where the process is under control at *catch1*. From *end1* to *catch2* is the change in setpoint, where a new setpoint is given at *end1* and the process is again under control at *catch2*. The time before from point A on the graph to *end1* is a preparation period. The time-period is chosen to be as long as the time-period between *end1* and *catch2*. The same period of time is chosen for *end2* to *finish*, the end of the trial, and and for the preparation of this phase (point B to *end2*. The actual times depended very much on how the trial was executed.

Around the mean HRV a confidence interval ($p = 0.95$) for the mean is drawn. The horizontal bars near the axes are the control phases. Both time and HRV values were normalized for each subject. What can be seen is that the workload is low at the start, and reaches a maximum just before the system is within the specified range. Near the completion of the first setpoints the workload increases again, to reach a second maximum when the process is brought within the specified range for the second setpoints. Near the finish of the trial the workload decreases, and shortly after the trial the workload returns to a minimum level as at the start of the trial. The workload during the two control phases was explained in the previous sections.

[5]Please note that the HRV is inversely related to the mental workload.

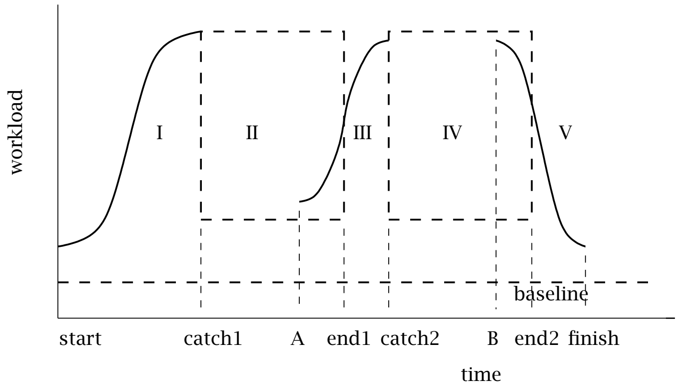

Figure 4.13: *Basic workload pattern.*

4.4.4 Basic pattern

If we combine all these characteristics of the HRV data it shows a basic pattern across subjects which is presented in a stylized way in figure 4.13. At the start of a trial the operator has a standard series of actions to start the process and shows a high level of activity; however, the workload is low, but climbs to a maximum when the output levels approach the demand levels, the 'catching'-phase (phase I). It is likely that the alarm–clock model describes behaviour during this stage but the short duration (10–20s) does not allow any analytical analysis. From that point there is a relation between the process state (MTB) and the workload (phase II). In this phase the workload cannot be predicted beforehand but can be predicted from the MTB. The dotted lines represent the area in which the workload can than be found. When the setpoint changes, the workload climbs to a peak much more rapidly than at the start, and after that shows a similar pattern as the first setpoint in the trial (phase III & IV). Points A and in this graph correspond to those on figure 4.12, indicating that, on average, the curve follows the solid line within the dotted area, according to the results of this experiment. The end of a setpoint is special. When the operators 'feels' that he will meet the output demand for the requested time, the workload decreases sharply (phase V). The workload during rest is represented by the baseline. Compare this also to figure 4.12. The major difference is that figure 4.12 shows HRV, whereas figure 4.13 shows mental workload, which is the inverse of the other.

4.5 Discussion

The data from this experiment show that the time-to-boundary (TTB) approach is successful in explaining the results of this DURESS experiment.

The TTB measure correlated well with the HRV during the control phase, which confirms the theory on monitoring behaviour. This means that there is a direct relationship between the time left for the operator to intervene and his sampling frequency. This relationship was lost when all TTB's were above a threshold. In this case no relationship between system variables and mental workload could be found. The threshold for this experiment was about 3 minutes. When the MTB was above 3 minutes the operators directed their attention to other elements then operating the simulator. This also confirms our monitoring behaviour theories. Measures such as deviation from an optimal setpoint showed no relationship to workload. In the unfamiliar situation of the second half of trial 6 the workload remained high until a strategy was developed to deal with the problem at hand. In this situation the workload remained high, together with the activity rate, as long as no sufficient strategy was found.

The activity rate showed a strong relationship to the error from the optimal setpoints. This supports the idea that there is a relationship between error and task-demand load. However, this result together with the findings in the previous paragraph means that a relationship between task-demand-load and mental workload cannot be established.

The experiment showed that HRV provided valuable information on cognitive activities during process control. The control of the process in this experiment can be divided into four phases:

- The start-up phase: A well trained activity pattern is executed to start the simulation. Due to the high training level, workload is low but activity is high. The theory on monitoring does not apply.
- The catching phase: All simulator variables must be brought within the specification using various techniques. The activity rate is rather low but the workload is high. The theory on monitoring may apply, but the collected data cannot test whether this is so.
- The controlling phase: When all variables are within the specifications within which they must be controlled. In this phase the activity rate is lowest and the workload may be accurately predicted from the minimal time-to-boundary, a stability measure, as predicted by the alarm-clock model.
- The end-phase: When the operator knows he will end the trial successfully with the variables within their boundaries, the workload falls rapidly towards the start-up minimum. The theory on monitoring does not apply.

The workload is lowest during the start-up phase and highest during the catch phase. Workload during the control phase always moves between these extremes. Only when the process stability exceeds a certain threshold of the MTB (in this experiment approx. 180s) can the workload not be predicted by process variables any more. Other activities than those imposed by the simulator then determine the workload. These findings confirm largely the findings of Crossman, as discussed in chapter 2 (Crossman et al., 1974, pp.44-45).

Regardless of operator strategies, techniques, and skills, workload in general could be predicted by the process variables. The theory would predict that this only applies when the operator is well trained. He must have the required strategies, techniques, and skills.

When the operator is skilled there is also a clear relationship between output error and activity rate: activity rate minimizes with the output error. The output error correlated well in this experiment with activity rate for all subjects. All subjects were well trained, except for the change in setpoint during trial 5 and 6, when the system was brought into a state unfamiliar to them, and a new strategy had to be developed. The analysis showed no relationship between activity rate and mental workload for these skilled operators, as the theory predicted.

In the experiment the operators were forced to show all types of control. The experiment began with well known sequences taken from previous training sessions, which could largely be done using their skills. It progressed in a number of trials to situations the operators had never experienced before and did not know how to handle in advance and which needed to be dealt with on a knowledge-based level of control. In these last situations the hypothesized skilled relation between error, activity rate and workload was lost. Because of the complicated nature of this relation no attempt was made to model this process. However, during the experiment it seemed that the operators who were involved in the design of the simulator had more problems meeting the task requirement when they were confronted with the new setpoint in trial 6. It was a situation made possible for this experiment only and it took these 'experts' a lot of effort to discover what was going on. They were puzzled by the situation because of their deeper understanding of the system and hence could not believe it could act outside their experience, whereas the normally trained operators 'simply did their job'.

The experimental results with respect to the lead-lag times between HRV and MTB are an indication that two strategies might exist: 'proactive' vs. 'reactive' control. For proactive control the operator needs a model that enables him to make a good prediction of the future. Those most familiar with DURESS showed this characteristic. When the system is in a situation where the predictive model is not valid, or when the rate of change is so high that the predictive model cannot be used any more, the operator must switch to another type of control. His actions cannot be proactive with reference to the system but must follow the events and keep the system within a specified range by applying more simple control rules. This type of control will be referred to as reactive control. This pattern was shown more often by those operators trained for the DURESS task, but not having a profound mental model of how the set-up was designed.

Distinguishing different types of control has implications for workload. In proactive control the mental activity associated with the task execution will take place before the task-demand load, system performance, or time-pressure; workload precedes performance. Reactive control shows the opposite: workload lags behind system performance. These two types of control can result from different mental models. The results can be interpreted as in indication that the subjects closely involved in the development of DURESS had another mental model of the simulator which enabled them to apply different control strategies. Interpretation of the control actions is outside the scope of this thesis but there may be more information in those data. The HRV data gave no indication that the group with positive lead times had a lower mental load, just that they applied a different strategy with some obvious advantages. Note that deeper understanding in this instance has positive connotations, whereas above (page 53) it led to negative results when they had to deal with “impossible system behaviour” according to their mental model.

The threshold effect in process stability is a clear indication for the distinction necessary between workload and taskload. Up to this threshold the mental load is very much determined by the task imposed. From that point on, mental load is determined by other elements in the working environment or maybe even in the life of the subject. Distractions of various nature play a role in this process. The alarm-clock model can only be used to predict mental workload when the operator is sufficiently engaged in the task to absorb the major part of his attention.

This experiment showed that measuring mental workload can provide information about the way operators react to systems under their control. The proposed alarm-clock model proved to be useful in understanding operator monitoring behaviour. In the next chapters a similar approach is taken to study navigator behaviour but first the subject of mental models, introduced in this discussion section, is given some further attention to link it better to the alarm-clock model and the subsequent task to be studied, that of the maritime pilots.

CHAPTER 5

Knowledge representations

When controlling a system knowledge on that system is essential. It is this knowledge that gives the system variables a meaning and relates the system variables to each other. In this chapter, in which this knowledge is discussed, it will usually referred to as a mental model. Mental models are knowledge structures that enable us to interpret the state of the system and decide on what to do in order to reach the goals. Monitoring by the operator depends on this knowledge in order to select and scan meaningful information that is needed at that moment in time. From this it will be clear that mental models depend very much on the task at hand. We can turn this around by inferring something about the mental model and task/goal from knowledge on monitoring behaviour. Both assumptions will be used in this thesis. Some elements of a navigator's mental model are discussed in the second half of this chapter. The notion of mental models will be used extensively in the interpretation of the maritime pilot experiment.

The DURESS experiment showed that some more discussion of mental models is needed. The DURESS operators relied on their mental models of the process, built up in their training, to know how to use the controls to bring the system within the set demands and keep it there. They also used them to predict the lags between actions and responses and to interpret system behaviour (including unexpected behaviour). In DURESS the mental models were relatively simple and determined by the experimental set up and the instructions. When we turn to navigators the models are likely to be much more complex and we need to take a more sophisticated view of them. Hence this chapter goes more deeply into the notion of mental models and applies that to the navigator task. In chapter 6 the task of the pilot will be discussed in detail so that the components of the mental model become clear.

5.1 Mental models

When a car driver turns his steering wheel he knows that his front wheels turn, although he cannot see this. This knowledge is essential for successful parking. The operator's knowledge of the system is what we call a mental model. What the driver expects is also important for his monitoring: when turning he will direct his gaze differently from driving straight. This simple example shows that some knowledge about mental models is needed when studying monitoring behaviour.

The use of mental models has been shown successfully in many experiments and we found it necessary to bring them into the discussion of the DURESS experiment in order to understand the results. From this evidence, we may safely assume that operators have explicit mental models, such as the mental model allowing him to reason about a system and the workings of it, plus an autonomous model that directs his behaviour. It has been shown by Conant & Ashby (1970) that efficient manual control of a dynamic system requires a model of the system's dynamics. Such a model may be applied as a reference model, required to perform feed-forward control behaviour, as may be seen in slow-response systems. Feed-forward control can be seen in such relatively simple systems as reaching for something with one of the limbs and in very complicated systems such as that of steering super-tankers at sea. What is not clear at present is what mental models look like and exactly how they are applied.

The use of mental models is threefold: to describe, explain and predict. Which of these three properties is used depends very much on the purpose and application of the model and the theoretical background of the observer. In this study three groups of mental models are used. These groups were very briefly introduced with the description of the alarm-clock model. The first group of models are those needed to predict the future state of the process. From this prediction the time to boundary is calculated and the moment of resampling is determined. In the situation of navigation this is a model that describes the movements of the ship. The second group is a representation of the desired state of the process, the plan that has to be followed in order to arrive at the goals. The navigator needs some kind of mental map of the sailing environment with navigable areas, boundaries, objects, headings, course lines, turning points, etc., to which he can compare the state of the ship. The third group of models are tactics that can be applied to reduce the difference between the actual state and the desired state. The navigator has knowledge on the effect of rudder changes and speed orders. He applies this knowledge to come to orders that can be implemented by the crew.

Veldhuyzen & Stassen (1977) concluded in their article on controlling large ships that a mental model is the basis for state estimation, developing and adopting control strategies, selecting proper control actions, determining whether actions lead to desired results, and understanding unexpected phenomena that occur as the task progresses.

Rasmussen (1983) limits the use of mental models to the prediction of future events, the process of explaining system behaviour in unknown situations and the evaluation of actions in new situations. Most of this predicting, explaining, and evaluating is done using the model in 'thought' experiments, conscious and subconscious. Rasmussen recognises the existence of different types of mental models for different types of problems, and orders them in an hierarchy of increasing abstraction, going from 'physical form' to 'functional purpose' in five steps. Rouse & Morris define mental models as: "... the mechanisms whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states." (Rouse & Morris, 1986, p. 351).

There is an important difference between Rasmussen and Rouse & Morris in their approach that results from two distinct concepts of mental models. The first sees them as the operator's knowledge of the process used to predict the outcome of his actions. The second considers them to be the mechanisms that generate behaviour. In the case of mental models as a knowledge structure, these models are structures of task-related information that enable the operator to discover what causes events, or these models are mental simulations that allow the operator to select an appropriate action. In the case of mental models as mechanisms the models referred to are embedded in processes which result in behaviour.

The Rasmussen model is a knowledge structure that may be consciously manipulated by the operator to evaluate the effect of his control actions. Rouse & Morris see mental models as an integrated mechanism, containing both knowledge and its manipulation, acting autonomously. One reason for this difference can be found in the origin of these mental models. The Rasmussen mental model is part of his skill, rule, and knowledge (abbreviated to SRK) classification of operator behaviour, and was intended to provide designers of complex systems with a reference when designing operator support. The SRK-classification is a normative element in the design of displays. This classification is therefore a normative description, it is not a model of actual operator behaviour (see Sanders, 1988). Rouse & Morris use a mental model to describe operator behaviour during actual process control. The latter use parallels the context of this research and hence I will use the term 'mental model' in a way analogous to Rouse & Morris.

Most mental models presented in literature begin with the assumption that an operator makes rational decisions and attempts to strive for optimal performance. He does this by evaluating predictions of the system state and selecting an optimal strategy. However, this sequence does not necessarily underlie what surfaces when we observe effective control in practice. Effectively catching a ball coming towards you with a curved motion may very well be the result of the correct application of a sophisticated mental model, but McBeath et al. (1995) have shown that clever strategies avoid the need for complex mental models in certain situations. In this case the question was how an outfielder baseball player would catch a ball. The research showed that the movement of the player was such that the player maintained a linear optical trajectory relative to the base-plate and the background. This does not require a very sophisticated model but a simple and effective linear model derived from practice.

Such observations can also be done in knowledge dominated fields. Extensive research has shown that most people, even physics students have a very rudimentary, pre-Newtonian mechanical model about the world. Cooke & Breedin (1988) concluded on the base of experiments that their answers were affected by display features of the problem and that the subjects used very little theory in solving the problem, but generated explanations 'on the fly'.

The prediction model block, which is part of the alarm-clock model, used in this thesis is such a mental model and contains, itself, sub models. It includes knowledge about the systems's dynamics, goals, and disturbances. The hypothesis tested in the previous, and in the next chapters is that the skilled operator has such a good model of the system under control that operating the control leads to virtually no mental workload, and that all workload is the result of the decision cycle to decide on when intervention is required. This means that there is a strong relationship predicted between the the process that is controlled, the goals set for the operator, his control actions and his monitoring behaviour. In the next section some literature on the expert knowledge of navigators is discussed. It will be shown that besides their skill to navigate ships they also have explicit mental models available. These models have qualitative and quantitative elements. It is unknown if these models are generated 'on the fly' from implicit mental models, but their presence is clearly demonstrated.

5.2 Mental model used by pilots

The navigator has a set of mental models that enables him to do his task. These include the plan, a representation of the situation that includes a number of groups of model: geography, weather and currents, traffic and ship movement. The first group is the geography: a set of maps of the harbour. The navigator must decide what route to take to the destination. The model must make it possible for him to compare the outside view to these maps, so he can see where he is. The maps contain special recognition points such as buoys and characteristic objects, headings for important stretches, names of quays and companies, and the dimensions of the fairway, width and depth. The second group are those model elements which describe the wind and the currents in the area that result from tide and rivers. Wind plays an important role in the manoeuvring of the ship. Currents have an instantaneous effect on the ship's movements and when disregarded will strand the ship. A third group of model elements required is a representation of other ships in the harbour: sailing, anchored, or doing maintenance work on the harbour. The navigator updates these models by reading the day-to-day notices and by listening to the VHF-radio to hear other ships' reports on where they are. The fourth group of model elements help the navigator to predict the movements of a ship. He needs this to decide when to initiate a turn, what rudder angle to apply, and how to use the engine. For this it is also necessary that the pilot has a representation of how the ship interacts with other ships and nearby constructions; the dynamic relation between these. All these models exist in parallel and information goes to and from each to enable control and tactical and strategical decision-making. This is used to come to a plan, a set states that make up the desired state during the entire voyage.

The first three groups of model elements are based on map like representations of the environment. In fact, different types of maps available to the navigator do provide this information. Maritime pilots are provided with navigation maps, fairway profile maps based on actual depths, current maps, maps with all harbours, quays, and companies, publications of advisories for pilots, weather charts and weather predictions, short term tidal predictions, lists of expected traffic conditions, maps with radar stations, and synthesised radar images of the entire navigation area, and sometimes the radar on board. These are all combined to build a representation of the navigation area.

The last group of model elements (the ship movements) can for a small part be found in books but they are largely formed during work. It is this group that is described in the next section.

5.2.1 Form of a navigator model

Within this broad framework, based on what is logically needed to do the task, the actual mental models which pilots have are very hard to discover. Different models may exist for different circumstances, and the models may be different for each and every navigator, based on training, past experience, mental capacities, and personal preference. My observation showed that individual pilots did apply different tactics to reach the same goal.

Despite the differences between pilots there must be important similarities because the system is largely governed by physical laws and it has some inherent limitations in the control possibilities. This means that the models must be based on the physical entities. Controlling a ship is largely based on visual information derived from looking outside to the environment of the ship. From this information it must be decided whether intervention with the ongoing process is needed or not. For this the pilot needs to have a model of the manoeuvrability of the ship. The manoeuvrability describes which area of course-lines can be covered by the ship. The navigable area must be covered by this manoeuvrability. Therefore the mental model must contain two elements: one is knowledge about the manoeuvrability of the ship, and the other is a representation of the navigable area that can be compared with the manoeuvrability.

5.2.2 Paper and pencil

As has been argued earlier, a vast discrepancy may exist between what people actually do and what they tell someone else that they do. Pfeiffer (1987) conducted several experiments in an attempt to elicit from navigator's knowledge on how to manoeuvre a ship. These data were later used by Papenhuijzen (1994). Two groups participated in the paper–and–pencil experiment: pilots, and experienced mates[1]. Pilots have extensive experience in manoeuvring a large variety of ships in confined waterways. The experienced mates have extensive experience with manoeuvring at sea with a limited number of ships, but they have little experience manoeuvring in confined waterways; in confined waterways the task is taken over by the captain and/or a pilot. The questions asked were all at a Newtonian level of ship manoeuvring characteristics; the effects of rudder and engine settings on speed and course were to be answered through drawing tasks and a questionnaire.

The first task was to draw the path of a ship sailing at constant speed in a straight line, which then had a rudder angle set at 15° or 35° at a given point. The subjects could select a ship type they were familiar with. They had to draw the resulting turning-circle. The drawings were judged on qualitative and quantitative aspects. Among the quantitative aspects were the initial shift of the centre of gravity to the port side when starboard rudder is given, due to the rudder force ('it goes to the wrong side first'), and the ship ending inside the spiral after one turn. Qualitative aspects were, for instance, the advance, the distance the ship progresses before a 90° turn is made, and the tactical diameter, the distance between the initial course-line and the course-line after a 180° turn. (see figure 5.1). The two spirals for 15° and 35° do not cross.

The results for this experiment are presented in two groups. In table 5.1 are the percentages of correct results on the qualitative questions. In table 5.2 are the results of the quantitative results of these drawing tasks, normalised for the real values for the type of ship chosen by the participants. For each of the three measures the estimates are normalised with the actual measures, and with these a mean and variance (between brackets) for both groups are estimated.

The results are surprising at a few points. Observations on board tell us that pilots know very well that the stern of a ship goes to port when ordering starboard rudder. But the drawing task yielded a mere 56 percent

[1] The study actually included a third group, junior mates, but this group is not interesting for this thesis

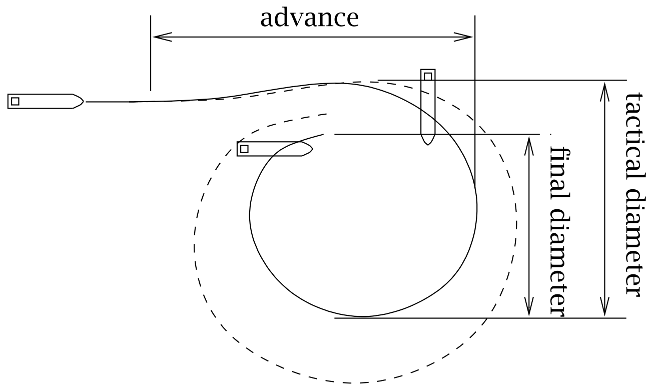

Figure 5.1: *The spiral test. This shows the track of ship when the rudder makes a step from zero to a selected angle (generally 15°, dotted line, and 35°, solid line).*

Table 5.1: *Percentage correct answers on qualitative aspects of the turning-circle by pilots and mates. (Data taken from Pfeiffer, 1987, p. 19-20)*

	pilots	experienced mates
correct shape	100	31
'goes to wrong side first'	56	23
15 and 35 degree spirals do not cross	56	27
bigger ships react slower	89	100

Table 5.2: *Normalised data on the quantitative aspect of the turning- circle by pilots and mates. (Data taken from Pfeiffer, 1987, p. 21)*

	advance	tactical diameter	final diameter
pilots	1.0 (0.1)	0.7 (0.1)	0.6 (0.0)
experienced mates	1.3 (0.5)	1.2 (0.6)	1.2 (0.7)

correct answer. When it comes to qualitative data, the pilots show a remarkable coherence: the deviations are very small. The advance was perfect, a measure required to enter any harbour. Both diameters are too small. This may be because pilots never turn without increasing thrust. When this is considered the values are about right. The mates do not perform very well on the qualitative aspects, except for the recognition of the slow reaction of a big ship. They show acceptable values on average on the quantitative aspects but there are large variations between mates.

In the results of the questionnaire about 2/3 of the pilots gave correct answers on the quantitative relationship between speed and rudder-angle, speed and advance, natural deceleration, forced deceleration, relationship between rudder-angle and rate-of-turn, etc. One series of questions was on the deceleration: making a natural stop. There is a quadratic relationship between resistance on the hull and the ship's velocity. As a result the deceleration decreases with time. Because of this, the stopping distance of a ship that has twice the speed is much less than twice the distance. With a forced stop due to reverse turning engines the deceleration is almost constant. Pilots failed to recognise this. They thought, by and large, that deceleration increased with time. Pfeiffer explained this as follows. When applying a forced stop the pilot allows for the time it takes to start the engine and to build up power. This was not included in the question. Questions based on measures such as the degrees per minute for the rate of turn were poorly answered by the pilots. This is not surprising since these measures have hardly any value on board. On the qualitative relationships on manoeuvring the pilots did much better.

In his discussion Papenhuijzen (1994) concludes that pilots probably "do possess accurate quantitative knowledge of the ships' dynamics after all, only for them a questionnaire might not be the most adequate tool to trace this knowledge.' (p. 23). Pilots are not interested in the exact numerical measures of various parameters, but whether the ship is having the right rate-of-turn in order to enter a harbour or the adequate speed in a given situation. This experiment illustrates the distinction made earlier between effective knowledge (mechanisms) and knowledge available for interrogation. The earlier discussion related this to the measurement of workload. The conclusion drawn there was that the elements of workload were not likely to be open to interrogation since they operate at a skill level. The conclusion here is similar. The mental model on which control is based is also not directly available for interrogation.

The models studied here are models needed to make up a plan. For making a plan they are sufficient. For actual sailing and navigation additional models are required: a prediction of the near future based on the current state and a model to implement required changes.

5.3 Discussion

In order to be able to control a system some sort of mental model of the system is needed. Depending on the type of control, different types of mental models are needed. Mental models are mechanisms that generate behaviour in order to control the system. Mental models may also have characteristics of knowledge structures but these are not the most important for system control. For each of the types of control separate models will exist. Following the alarm-clock model three types of models are required: plan, prediction, and control. It is likely that within each of these three types separate mental models exist for different states of the system. Depending on the research objective different characteristics of the operator mental-models may be retrieved.

The definition of mental models by Rouse & Morris requires that mental models are retrieved by studying behaviour, not interviewing. Interviewing might provide information about some of the elements of the mental models, but they cannot be the sole information base. Theoretical considerations and observations must provide the main basis for an operator mental-model. The mental models referred to here are models that are internalised at the time-space manipulation level, a level generally referred to as skill.

Experiments by Pfeiffer showed that pilots and navigators have explicit knowledge about the manoeuvring characteristics of the ship but not all of their models are subject to direct interrogation. This is one part of the planning process. This plan is essential for effective control in general and for effective navigation in particular. Only by conducting experiments of pilot behaviour will is be possible to discover which elements of a model are essential for control, or rather which parameters of the task and system appear to be linked more closely to pilot workload. In the next chapter the working environment of the maritime pilot is described, together with some other elements of the navigator's mental models. That chapter will lay the foundations for the experiment in chapter 7 in which the theories on monitoring and mental models are combined and tested.

CHAPTER 6

Maritime pilots

This chapter describes the work done to arrive at a sufficient understanding of the maritime pilot's work to form the basis for interpreting the experimental results to be discussed in chapter 7. For this three elements are discussed: the work of the maritime pilot, a model of the pilot, and the relationship between monitoring and the pilot's work. This chapter is meant as an introduction to the next chapter where maritime pilots participated in an experiment. Most of this information is based on informal interviews during five days of introductory voyages with the Rotterdam pilots, four weeks of sailing with a number of ships on the North Sea, and from data collected during the recorded voyages which are described in a later chapter.

In this chapter the very complex task of a maritime pilot is modelled in such a way that the heart-rate variability can be used as a measure to predict and explain monitoring behaviour and control. It will be argued that the pilot's workload is dominated by prediction. It will also be argued that prediction cannot exist without planning and control. It is the plan which facilitates the interpretation of the situation, and makes the construction of a prediction. The control level allows the pilot to implement the plan without requiring effort to do so. These three types of mental models are part of the monitoring theory discussed in chapter 2 and discussed in more detail in chapter 5.

In order to come to the elements required for a pilot model, an introduction of the daily work of the Rotterdam pilots is laid out in the next sections.

6.1 The world of the navigator

Life on board a ship is different from that ashore. The aspect that seems to dominate all others is the isolation: a ship is an isolated autonomous system. The crew has to solve all problems themselves and can hardly rely on any help from ashore. Navigation, maintenance, repairs, accidents, all problems must be solved by the crew. Only near harbours is help available, but ships are far away from any harbour most of the time. The only regular help the ship receives is the pilot who comes on board near every harbour. It is the pilot who assists the crew with the navigation when accurate navigation is crucial.

A pilot is an experienced officer who is trained in navigation in the area for which he is certified. So he is a seaman with extensive navigational knowledge of the sailing area. Within the pilot organization there are specializations such as deep draught pilots specialized to navigate very large ships through the canal in the approach of a harbour, and docking pilots specialized in slow speed manoeuvres with tugs.

To learn more about the work of pilots I approached the Rotterdam Pilots, and subsequently I made several voyages during five days (or better: five shifts) with different pilots. Ships differ largely. Some are very old, some brand new. Crews come from all over the world and some of them speak English perfectly, others hardly. Some bridges still have almost the same instrumentation as when they were built 25 years ago, other ships have state-of-the-art instrumentation. Some of the ship's exteriors and interiors make you very suspicious about the state of the engine room, let alone the galley. But there is always a friendly welcome by the captain. It is in this varying environment that the pilot has to do his work. The following is a general description of this work. It is intended as a reference to understand the models involved in maritime pilotage for those who have no knowledge of this world.

A ship's navigation officer usually works on a single ship for a period of one to nine months, after which he gets some time off. During this working period he sails the ship to various harbours. When the ship offers a liner trade the number of different harbours is small, but with tramping any harbour in the world can be the point of destination. A pilot on the other hand works in a restricted area, the area he is licensed for. All ships in that area that want or require pilot service will receive a pilot on board and this pilot officially advises the captain on matters of navigation. The captain is according to law and regulations fully responsible for navigating the ship and the pilot is only responsible for the quality of his advice, but a captain is permitted to allow the pilot to be the navigator, which makes the pilot liable for his navigational decisions. In practice the captain almost always hands over control to the pilot, leaving the entire navigation task in the pilot's hands and lets him be the navigator. The reason for this is mainly that the navigation task is an integrated task of planning and control, based on knowledge of the environment and manoeuvring characteristics of the ship. Since the pilot knows far more about the environment than the crew, the pilot can do this integrated task more easily. The captain will make sure that the pilot does indeed do the task at the level the captain considers appropriate for that situation. The crew is likely to know far more about the manoeuvring characteristics of the ship, but the pilot has more experience in manoeuvring, especially when tugs are involved. It may be important to differentiate between big harbours such a Rotterdam, with 35.000 ships per year that require a pilot, and a small harbour with 200 ships per year. In the big harbours the pilots are much more experienced. The captain will also make sure that the bridge team is complete and functioning, that the ship is ready to manoeuvre in confined waterways, ready to anchor, and ready to moor. The bridge team generally consists of a helmsman (a sailor) who controls the helm, the chief mate who gives the orders to all crew on board, orders the engine thrust, logs all events in a logbook and on the map, and controls the ongoing process, and the captain who acts as supervisor. Due to the continuous character of the work on board, the size of the crew, and the complexity of the ship it may happen that there are minor differences in the bridge team. If needed, there is a lookout on the bow of the ship who reports all traffic and navigational marks to the bridge, e.g. when visibility is poor. In the engine room the chief engineer and his crew are responsible for the correct functioning of all mechanical systems on board, in particular the main engine and all auxiliary systems such as the generators and pumps, and for delivering the required thrust.

6.2 The Rotterdam pilots

Pilots work around the clock. On average just over one hundred ships enter (and leave) the harbour of Rotterdam (figure 6.1) every day, more or less evenly spread in time, requiring pilot assistance both inward and outward bound. At the time of this study the pilots worked five days in a row and then would have four days off. During their five working days the pilots work for eight to twelve hours and take ten hours of rest. This means that they do not have a normal day but on average make working cycles of about 20 hours.

There were approximately 310 pilots working in the Rotterdam region. The pilots are specialized in region, draught and ship-length. Historically the pilots were specialized for sea, river or dock, requiring a regular change of pilot for the ships. Changes in shipping and modern management led to an integral service for the ships: one pilot from sea to quay. The advantages are that fewer pilots are required on station, the pilots are more flexible, they have more time to learn the ship's characteristics, and the difficult pilot changes are not needed anymore. A disadvantage may be that the pilots become less specialized. The normal procedure is that as pilots become more experienced they receive additional training to be certified for bigger ships, additional areas, or other types of specialization. As a result, the biggest ships are navigated by the oldest pilots. The certification is organized in such a way that a pilot can maintain the required level of skill for his specialization: the number of ships within a category and the number

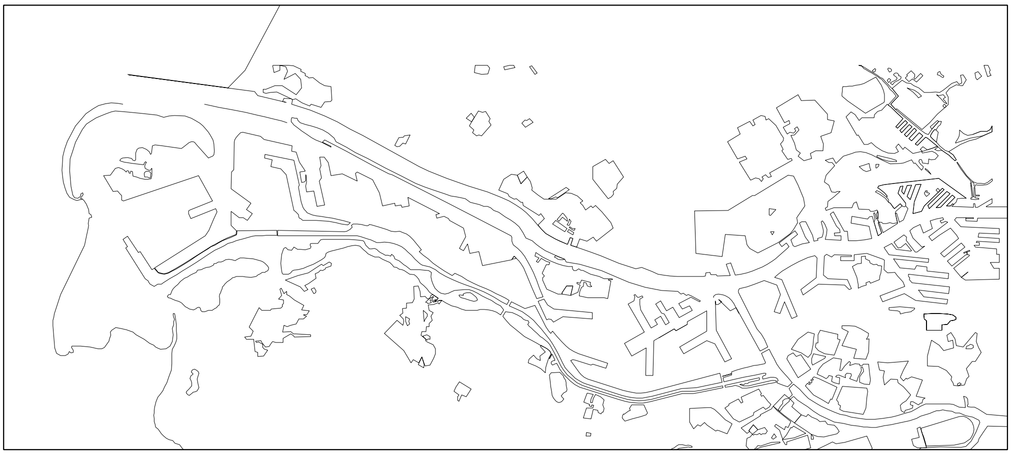

Figure 6.1: *A map of the harbour of Rotterdam.*

of specialized pilots must be in balance. It is his experience skills that are the most important features of the pilot.

In Rotterdam the pilot station is some 6 n mile off the coast, 1 n mile south-west of a point known as Maas Center. The pilot station (figure 6.2) is a ship stationed there to deliver pilots to the ships heading for the harbour of Rotterdam. The Maas-Center buoy marks the approach of the fairway into the river Nieuwe Waterweg. It is also a bend in the dredged channel Eurogeul that has its entrance 34 n mile west of Maas Center and has to be used by deep draught vessels. Ships who need to use that channel will always receive a pilot further out so sea, landed by helicopter. Pilots can also be delivered by a tender, a large fast boat that delivers pilots from a station ashore so pilots do not have to wait at sea. Recently there is a tendency to deliver more pilots by tender and helicopter. There are also North Sea Pilots, assisting the ships on the entire North Sea, with the exception of most harbours. The place where the pilot boards depends partly on the local regulations and partly on the captain who is responsible.

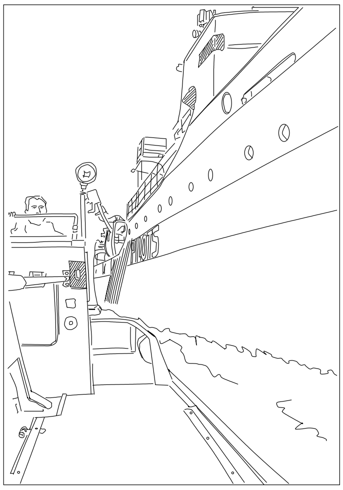

Figure 6.2: *Pilot station.*

At the time of this study (1994) most of the pilots were delivered to the ship from the pilot station. Pilots have a rating for ship size and sailing area, and depending on the demand the pilots are assigned to the ship on a first in - first out basis for the pilots and a first come - first served for the ships.

Ships that require pilot assistance give 24, 6, 2, and 1 hour notices[1]. This enables the pilot house to have a pilot available at the time of arrival. The notices are also needed to plan the ship with respect to the tide, spread the arrival and departure of ships, to check if there is place at the quay, order tugs and linesmen, give notice to the harbour authorities, etc. So when the ship arrives at the pilot station all requirements have been fulfilled. The pilot is brought on board with a fast speed boat that is lowered from the pilot station. From this speedboat the pilot will board the ship with a pilot ladder (figure 6.3). For this it is generally needed that the ship reduces speed and 'makes lee', turning the ship in such a way that the ladder is at the lee side of the ship. Swell and waves require some acrobatics from the pilot to step onto the pilot ladder. The pilot service is suspended when it becomes too dangerous to step on/from board or when the speedboat cannot be taken on board the pilot cutter anymore. Smaller ships can make less lee therefore pilot service is suspended for smaller ships sooner. When the pilot service is suspended ships have to wait at sea, although Rotterdam Pilots provide radar based pilotage under certain strict conditions.

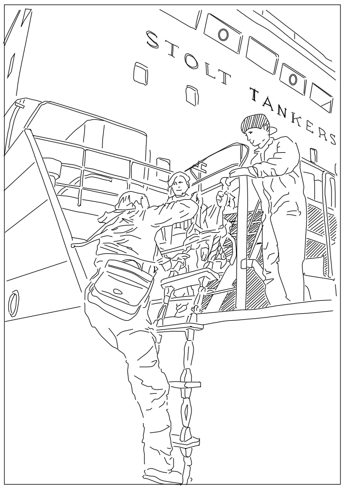

Figure 6.3: *On the pilot ladder.*

Depending on the size of the ship and destination a pilots will be selected. Up to a certain size ships can enter around the clock. When they are over this size the time of entrance must be calculated carefully for optimal currents and tide. The same goes for non-standard ships that require tug assistance for entering. When no such deviations from standard procedures are expected the pilot is provided with information about the overall dimensions of the ship and its destination about one hour in advance. With each voyage the pilot is given a slip with the name of the ship, type of ship, cargo, tonnage, length, beam, draught, bow/stern thruster, call sign, destination, and nationality. In addition it contains the time of boarding, last/next port of call, quay, cargo company, mooring side, agent, confirmations for tugs and linesmen, special information, and time-ports when needed. This provides the pilot with the essential information to prepare the voyage. Many ships which regularly visit Rotterdam are known by the pilots for their handling characteristics, quality of the captain and crew, the state of the ship, and the drinks and food on board. If the ship is unknown then he must prepare with the slip alone, and the information he can collect when he goes to the ship and comes on board. At the moment the pilot leaves the pilot station he is informed about his current position, the traffic situation, expected

[1]The 2-hour notice is known as "birth notification", the moment all actions are initiated

changes in the weather, and special situations. A radar on board the pilot station provides the pilot a reference to start the navigation task on board the new ship.

When brought to the ship with the speedboat the pilot will often verify the draught of the ship. This is for two reasons. The draught is one of the most important measures for navigation, and the pilot fee depends on the draught. After that the pilot will climb on board and will be welcomed by a mate. The mate will bring the pilot to the bridge where the captain awaits his appearance. The pilot is welcomed on board by the captain, and after verifying the destination the pilot will verify his position, give his first orders for course and speed, and he will report to the Vessel Traffic Service his ship's name, position and destination and will give them an expected time of arrival. This last information is needed by the tugs and boatmen for their planning. He will then be informed by the captain about the type and characteristics of the propulsion, rudder, manoeuvring characteristics, and the number of required tugs will be verified. The pilot will inspect the compass, rudder indicator, the steering machine when no helmsman is present, the radar when available and needed, and other equipment available. A very important part of this exchange of information is to get a 'feel' for the ship and her crew: how predictable is the ship, how professional is the crew? The more professional the crew is, the more the pilot can concentrate on navigation alone. If not, the pilot extends his duties to ensure safe sailing. The captain does the same for the pilot: if he is found trustworthy, which is generally always the case, he will become an "active traffic participant" and receive full control, otherwise he is restricted to giving advice only.

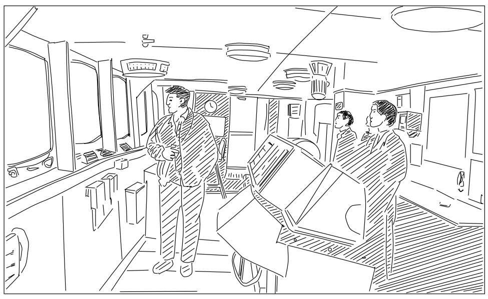

Figure 6.4: *Pilot working on the bridge.*

This initial phase is very important (figure 6.4). The pilot must decide if he can berth the ship without complications. Ships vary greatly in type, size, age, state of maintenance, crew, etc. For instance, the labelling of instruments can be in a completely unfamiliar language, some crews have an extremely poor understanding of English, the level of training and experience of the crew varies largely, some new ships have completely computerized bridges requiring specialized knowledge to operate them, instrument readings can show large deviations from the true values or do not function at all, equipment can fail during the voyage, and the state of the engine room can only be guessed by the pilot. The captain has to tell the pilot the condition of the ship, but many captains are reserved about the true condition of their ship. So pilots generally depend very little on the bridge instruments, which does not mean that they do not use them. They are used for information and verification, but pilots select a sailing strategy appropriate for the estimated condition of the ship.

From sea the first objective is to bring the ship between the piers that mark the entrance of the river (figure 6.5). The tidal streams are the primary disturbance, the wind a secondary. Because of tidal streams headings of $-10°$ to $+20°$ are required with respect to the course. This is also known as crabbing. With drift due to wind this angle can increase to well over $-15°$ to $+30°$. The tidal current will vary greatly near the coast. Constantly moving sand dunes on the sea floor will have their effect on the tidal currents. Especially near a river there will be a clear distinction between the tidal current at sea and the current from the river. These two cross currents are divided by a sharp line extending into the sea or into the river entrance, depending on the tide. The stream at sea will usually run parallel to the coastline, and the stream of the river will be in the direction of the river, i.e. these two are generally perpendicular to each other. This means that a major correction can be required when crossing this line. For a good verification of the correct heading an alignment of transits is needed. This can be done by using the piers, a series of buoys, or preferably, a light-line consisting of 2 lights some distance apart[2], with a different height, that appear on top of each other when on the correct positional line. Swell and waves can seriously hinder steering the ship. The waves will rise with the sea-bed, putting large forces on the stern of the ship.

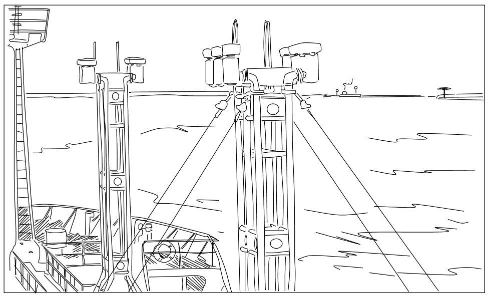

Figure 6.5: *Approaching the estuary.*

Entering the river will change the environment greatly. From open sea, where the navigable area is very large, via a channel into the sea, the ship enters a waterway that has very clear visual boundaries. Where at sea it was possible to turn the ship around in a large circle to avoid collision, the manoeuvring space in the river is very limited. Most sea going vessels will not be able to turn. The minimal required space for a ship to turn is about four ship-lengths, where most of the river is no more then 300-400 m wide (navigable). The under-keel clearance will also reduce during the voyage up the river because the river becomes shallower when moving upstream. This will make the ship less manoeuvrable, also know as the shallow-water effect, and it can even cause the ship to become unstable, thus requiring

[2]In Rotterdam the lights of the major light-line (112°) are 0.6 n mile apart and the lights of the minor light-line (106°) are 450m apart, visible from respectively 21 n mile and 18 n mile (nominal values).

good helmsmanship to keep her under control. A second characteristic is that a river is curved. In effect the profile of the current in the river is not evenly spread, but is at an almost standstill near the banks and has a maximum in the middle. In curvatures of the river the current will set to the outside, taking the ship with it to the outside too.

On the river there will be other traffic: sea ships leaving for sea, barges, dredgers, and tugs in all sizes and shapes. At sea the ships keep a large distance to other ships. A ship at 1 n mile is at sea very nearby, and most captains impose a standing order to start avoidance manoeuvres long before other traffic can come closer than 1 n mile in order to avoid the possibility of collisions. In the river this tactic is impossible (figure 6.6). Ships sail behind each other at 3 ship-lengths distance, and they overtake each other with only 2 ship-breadths separation. When a ship anchors at sea they have a large empty area around the ship. When the ship has to anchor in the river the ship cannot swing around its anchor without hitting the river bank or obstacles. The scale is completely different, the problems are of another dimension than those at sea.

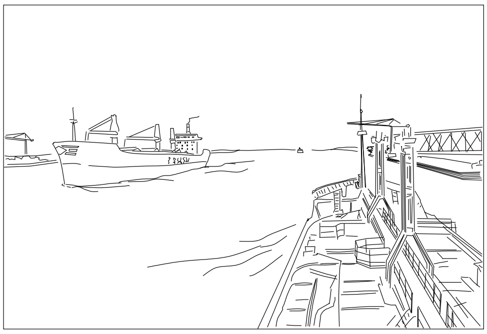

Figure 6.6: *At the high-water barrier.*

In the river there are many features that show the shape of the river. The banks, groynes marked with lights, jetties, light-lines, trees alongside the river, buildings, chimneys, and other fixed objects that are easy to recognize help to fix position. The official navigational marks provide an adequate set of elements to navigate, but the additional unofficial marks can largely enhance the quality of navigation. At night each light has its local character to enhance recognition, but recognizing a light requires careful observation for a number of seconds (light periods are 5–15 seconds) and provides only a local confirmation because only a limited number of light characters is available. Finally, where at sea a course can generally be maintained for many miles, in the river an almost continuous alteration of the course is required.

Most ships have manoeuvring characteristics that are good at sea but are insufficient in harbour situations. They require tug assistance: from one to five tugs are used around the ship depending on ship and situation. Manoeuvring with this requires a different approach from normal manoeuvring. The own ship is connected with the tug by a rope, allowing it only to pull. Pushing is possible, but only very limited. The tug is restricted in its pulling by the connection point on the ship and the speed of the ship. The connection point must be chosen well in advance with respect to the expected manoeuvres. Reconnecting a tug takes considerable time. The pulling direction is limited by the speed of the ship: when the speed is very low almost all directions are possible, but when the speed increases a sideways pull becomes very limited. In contrast to the ship, the pilot cannot control the tug directly. The tug is controlled by its own captain. The pilot will not tell the tug captain how to manoeuvre it, but will tell the tug what has to be done to the ship, and the tug's captain will interpret this to manoeuvre the tug. The pilot and the tug must work together to manoeuvre the ship, the pilot being the 'conductor' of this 'performance'. This requires a different type of control than the normal orders given by the pilot. The captain of the tug must understand what the goal of the operation is in order to do his work right. Because the tug is relatively slow the captain of the tug must anticipate and be ahead of the situation to make precise control possible.

On ships that have manoeuvring characteristics adequate for entering and docking without assistance the pilots are faced with another type of problem. By now they are familiar with the control of the ship at normal speed with small rudder angle, but they have no experience at low speeds with large rudder angles yet. There can be a large difference between these two, causing unexpected surprises for the pilot.

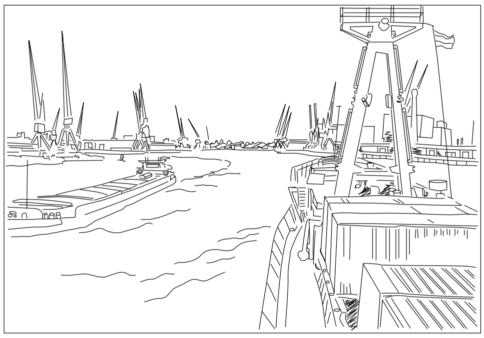

Figure 6.7: *The harbour.*

In the harbour area (figure 6.7) the navigation is quite chaotic, especially compared to the situation at sea. The layout of the dock is obscured by ships, barges, and other floating objects. The quay is used for storage, making its appearance change rapidly with cargo of various types. At night this can become even more prominent. The official marks are obscured by a large number of objects. During the night the lights from factories and buildings make it hard to find the navigation lights. Other sailing ships are hidden by the many lights and objects around the harbours, requiring an extensive knowledge about the position of the navigational marks and the movements of other ships and boats. The manoeuvring space in the docks is very limited, and often the dock can not be overseen by the navigator until the ship is almost inside it.

Mooring is the final part of the voyage. The ship must contact the quay, and it is at this stage that the ship can easily be damaged. Due to the enormous mass of the ship even low speeds can result in considerable damage. Wind is a factor that must be taken into account very seriously. Often tugs are needed to make safe mooring possible and as explained earlier the use of tugs is complex due to the many variables and slow proceedings. Linesmen, tugs, the crew using the winches, and the engines on board all work together to position the ship at the quay to make loading and unloading as easy as possible.

6.2.1 Some experiences of pilots

A pilot once told me: “Nothing really surprises me anymore”, but something unexpected happens every day”. I have only experienced the pilot’s work for a number of days, but even in this short time I have seen strange things. Ships of all ages and types come from all over the world. Often they behave unexpectedly, and crews vary largely in education, experience, and culture. Language can be a serious barrier. Pilots rely on language. All pilots speak English, and crews are expected to speak a limited amount of English. Additionally most pilots will speak some German and French, sometimes Spanish. Although it is expected from crews that they speak a little English, this is not always the case. To overcome this problem partly the I.M.O. (International Maritime Organisation) have adopted Seaspeak, essential English for international maritime use. It was developed primary for VHF-use, but it is enough for the pilot to do his work. Unfortunately not all crews master English even to this level and they sometimes have great difficulty understanding anything outside very straightforward sailing orders. So it can happen that the helmsman has been given a course, the helmsman repeats the course and then the chief mate has to translate the course before the helmsman can implement it. This is not as dangerous as it may seem. At least one officer who speaks sufficient English is present on the bridge at all times, and the pilot verifies correct implementation. The rudder is the most important control, and the pilot always checks the rudder indicator after giving an order. This prevents a lot of problems. Then again, pilots are not infallible: on a Russian ship outward bound one of the pilots ordered a wrong course. The helmsman repeated the course and started to implement it. The pilot saw that the movement was not according to his expectations and rectified the course. The system of feedback, check and double check, prevented progression of the error.

On a ship there is a potential conflict between the bridge and the engine room. The engineers have to keep the machines running, around the clock, for months. For this they do continuous maintenance which means that the engines cannot always perform up to factory specifications. When the situation affects readiness or sailing performance the engineers will tell the captain and the bridge, but they often keep their problems to themselves and so as not to ‘bother the bridge’. This is all according to a ‘need to know’-principle. The pilot gets to hear even less. Serious problems are reported as minor engine problems and only when essential subsystems, such as bow-thrusters, are not functional is the pilot told. Someone ashore will be told almost nothing since he cannot do anything anyway. On a ship I visited a high-pressure oil conduit broke loose when we left the harbour.

This is a very serious condition because there is severe threat of fire, one of the worst events imaginable. So the engineer stopped the engine almost immediately after informing the bridge, with the ship still in the channel. The radar station was told that we would anchor just outside the channel because of a 'minor engine problem'. This 'need to know' principle can have more serious effects. A ship left for sea and the pilot was told that everything was in good working order. What he did not know was that he did not have full power and that changing ahead-astern was very slow. The ship had to turn inside the harbour. In order to do so he sailed astern in a curve, and then set the engine ahead to complete the turn. But because it took such a long time and little power was available the change took much longer than the pilot had expected. Only by using the assisting tugboat to the maximum could it be prevented that the ship sailed into a quay. Had the pilot been told, he would have switched earlier, or maybe he would have used a bigger or second tugboat. The pilot is often not informed early enough about such problems because of economic reasons. The interesting thing was that during this entire event everyone remained very calm and quiet and continued doing their job as usual.

6.3 Work analysis

The impressionistic description in the section 6.1 will now be presented more systematically in the form of a work analysis. For this project it was not necessary to do a full task analysis. It was sufficient to gain a clear idea of the types of tasks to be performed and the factors relevant to them, in particular monitoring and workload.

6.3.1 Entering the harbour

When coming from sea the harbour is somewhat like a funnel (or fyke). At sea the sailing area is very wide, but this rapidly decreases to a rather narrow fairway because of the channel present. This fairway will narrow slowly to the river, and it will end in a dock, where the ship comes to berth. In addition the river becomes more shallow, the current becomes stronger, and the traffic density increases. This is not only specific to Rotterdam; most harbours that are part of a river show the same picture (figure 6.8). In this harbour model three phases can be distinguished: sea, river, and harbour. At sea the manoeuvrable area rapidly decreases, but this cannot be observed directly. This decrease is only known from maps and experience. In the river the ship must carefully follow the centreline of the river, while taking care of other traffic; other seaships also follow the centreline, whereas barges follow a route near the river bank. In the harbour the ship must carefully manoeuvre in between ships and come to a stop at the quay. The effect of such a voyage on workload is discussed in the next section.

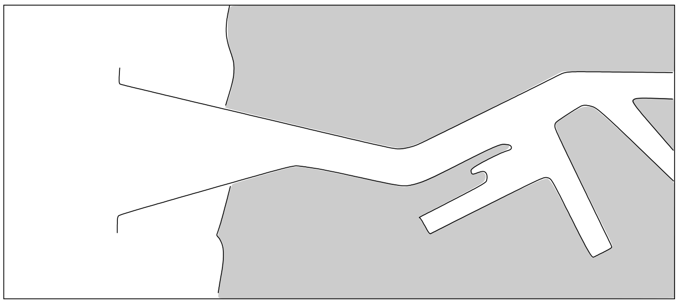

Figure 6.8: *A schematic representation of a harbour alongside a river.*

6.3.2 Tasks on the bridge

From the previous description of the pilot's activities one can derive the tasks which the pilot performs during his work. Various tasks have to be done. The maritime pilot does not control the ship directly. He is part of a team of sailors responsible for effective ship operation. It is the captain's work to make sure all tasks are allocated, all team members are fit for their task, and the ship is safe at all times. During the voyage the tasks change. Some tasks disappear, other tasks emerge, necessitating constant reallocation of tasks. The pilot monitors the state of the ship and the environment, compares this to a plan, and makes decisions to change setpoints based on this. The control of the various subsystems is left to special operators or automated control systems: for instance the helmsman or an auto-pilot. The helmsman will implement the orders and manipulate the helm accordingly, or when the helmsman is replaced by an auto-pilot, the desired heading is maintained by a machine. Also, requested changes in engine thrust are not set directly but are transmitted to the engineers, who will change the engine settings accordingly. To make a decision the navigator must know the 'state of the ship'. This is a general term which includes the position, heading, speed, yaw and roll motion, engine thrust, fuel economy, and rudder movements, relative to waves, wind, and current, the sea bed, sailing plan, destination, and time schedule. Some of these variables need frequent monitoring, others only need to be updated once in a while. For all these variables there are physical, imposed, or agreed limits. Navigation, or pilotage, is arriving at compromises to keep all these variables within acceptable ranges coupled to the overall objectives. It is these limits and variables which are potential candidates for inclusion in the alarm-clock model as boundaries in relation to which the pilot exercises control. The overall objectives are to arrive, to avoid loss of life, ship and cargo, to make life on board comfortable, and to operate the ship at minimal cost, which generally means sailing on schedule thereby balancing energy cost and time.

When analysing the work on the bridge three components may be observed: collecting of information, setting the controls, and allocating tasks. The result for a small size ship may be seen in appendix C. This task analysis is based on the observations on four ships done over a period of over four weeks sailing.

6.3.3 Navigator control

In this section three levels of navigation will be recognized: planning, prediction, and manual control. Planning is about the series of way-points that must be passed to reach the destination, prediction is about the near future state of the ship, and manual control is about the selection and implementation of decisions made on heading and speed. These three levels are distinguished on the basis of the theoretical considerations on monitoring and mental models described in earlier chapters.

Planning

Planning is about selecting the goals for navigation. It may have no direct relationship with monitoring but it is essential for the prediction of what will happen and what control action should be initiated. The planning information available to the navigator is largely contained in maps and tables. The map gives geographically variable information and the tables give time variable information.

Interviews showed that all pilots have detailed geographical knowledge of the sailing area with fairway layout, locations of geographical markers, headings, and turning points (way-points). An example of such a standard plan is shown in figure 6.9. This is a standard plan of one pilot for sailing from sea to the Eemshaven that he drew at home on a map.

The information on which the plan is based is regularly updated. The plan is based on the latest information available on weather, tide, and 'bottlenecks' because of work or traffic. This information is retrieved from pilot tables and news bulletins. Together with the ship's characteristics it yields the constraints within which the ship must move. The most dominant constraints are the fairway, under-keel clearance, and the current. The fairway must be unobstructed, the under-keel clearance must be large enough for manoeuvring, and the current small enough with respect to manoeuvring tolerances. That the ship can be sailed to its destination is taken for granted here. The constraints are functions of time and place and can often be visualized as such, yielding manoeuvring areas and time windows. If the manoeuvring characteristics of the ship do not comply with the constraints of the voyage planned, additional measures must be taken, such as the use of tugs. Also, support for mooring and unmooring must be arranged.

An important part of the plan is the selection of available navigational marks with respect to the expected conditions. In shipping, one must be ready to handle situations where 'things go wrong'. Marks may not be avail-

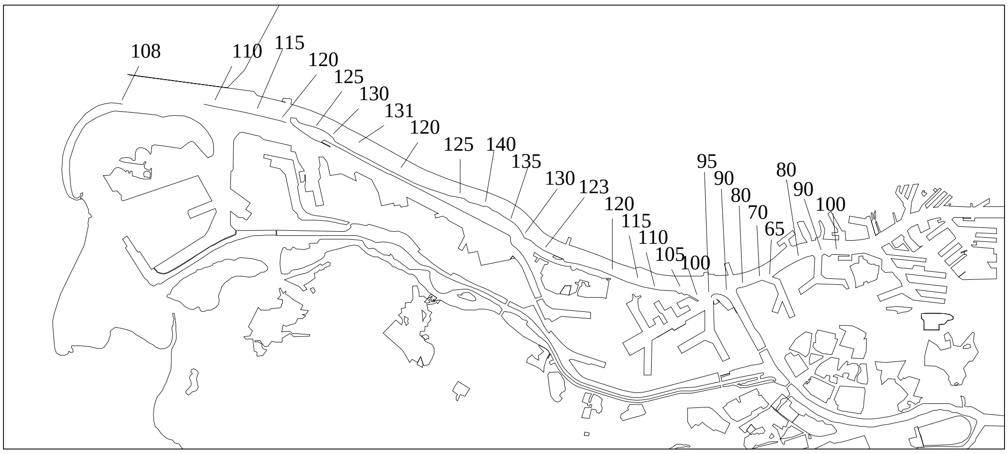

Figure 6.9: *A map of a long term plan as drawn by a pilot for sailing from sea to the Eemshaven. The numbers indicate the headings ordered at different locations.*

able, instruments may fail, systems may break down, and weather conditions may change dramatically. This means that, in the plan, the navigator has different sets of navigational marks, can fix position in a number of different ways using different displays and instruments and plans for engine failure, rudder-system failure, and a ship 'black out', all based on his primary goal: to bring the ship to its destination undamaged.

During sailing this plan becomes more refined when up–to–date information is integrated into the standard plan. To illustrate this process the harbour of Rotterdam is chosen. When approaching Rotterdam the pilot boards some 6 nautical miles off the coast. At the pilot station the pilot checks the ship's position and the position of other ships in the area and their destinations. From this he calculates the first course the ship has to steer, correcting for wind and current. The initial corrections are a rule of thumb, $+20°$ for rising tide and $-10°$ for falling tide, with additional corrections needed for wind. The resulting course will be the pilot's first heading order when arriving on the bridge. With this heading the pilot will check his drift angle, making sure the drift angle is small and well within his fairway area and according to his expectations. The optimal course line with respect to the fairway and other traffic is just south of the locally known red light line: a navigational aid. Coming close to the entrance of the river Maas, the pilot will start steering by a series of navigational marks with known headings. Interviews and video-recordings showed that each pilot has a standard series of headings that are adapted to fit the situation. Pilots seem to prefer round numbers as is shown in figure 6.10. Courses that end on '0' or '5' are preferred over all other values, rather independent of the fairway. This presents no problem because the fairway is always slightly curved. He will also make sure that his transverse distance from a number of locations on

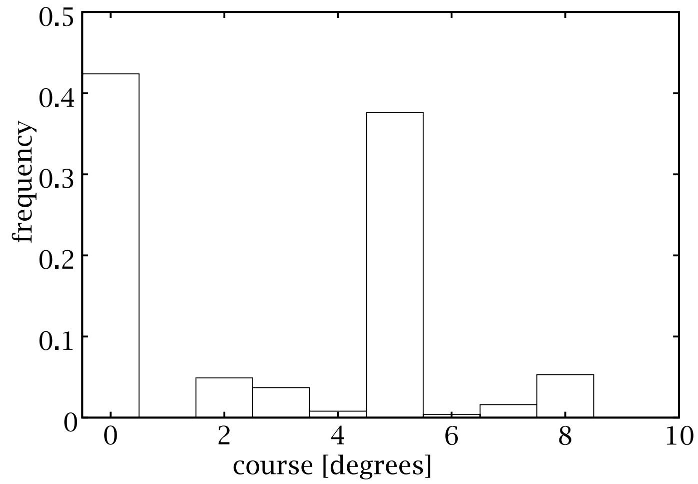

Figure 6.10: *The ordered headings (modulus 10) when sailing on the Nieuwe Waterweg (Rotterdam).*

his way is appropriate; the pilot is trained to check and cross-check. The number of headings used by the pilot generally depends on the size of the ship. With small ships bends in the river require fewer heading orders than with large ships, the latter changing with steps of $2° - 3°$. (see also figure 6.13)

The pilot has a number of checkpoints distributed over the river separated by about 1 n mile, where he will check for his heading and the transverse position of the ship in the river. For each of the turns the pilot has a series of marks from which he can choose. Different sets exist for high tide and low tide, depending on the current in the area, and the size and type of ship which he is sailing. This will be illustrated in the section on information collection (page 88).

In planning the general objective is to sail from way-point to way-point, where the way-points result from a plan made previously. The way-points are chosen such that they mark significant changes in course or they are at a comfortable sailing distance away from each other. The actual course planning between these sailing points is done during sailing. This is the planning of the actual track for a short period of time, some 10 minutes, also known as the 'horizon'. This estimate is based on the answer given by all pilots, and includes a series of about three way-points. It is this element of planning that will become the center of focus in the experiments that are described later.

As an example of the on board planning process a map of an area in the river is shown in figure 6.11. In this map the current position is drawn, together with the plan. The dotted line is the planned course line to the ship to the right of the figure and the bars on this course-line represent the

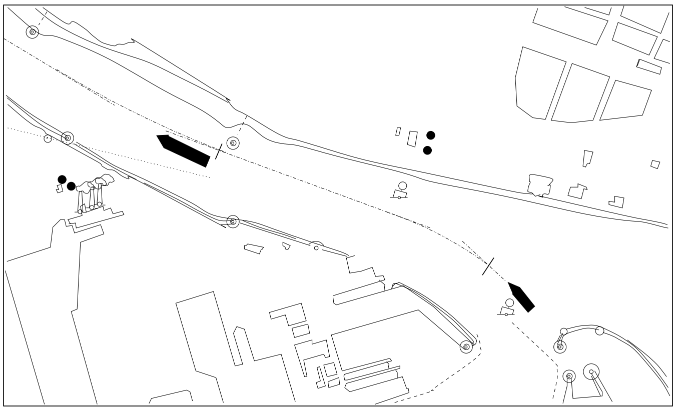

Figure 6.11: *A map of a small section of the river. The positions of the ship and that of another ship are drawn on the map, together with a short term plan. (See also figure 6.15)*

way-point. The bigger ship on the left will remain ahead. This map can be compared to figure 6.15 in the section on information, later this chapter.

Entering harbours is a difficult element in short-term planning. Harbours are discontinuities in the track, contrary to normal bends and curves.

6.3.4 Manual control

The pilot depends on his control skills when the difference between the actual state and the desired state becomes too large. By applying these skills the discrepancy between the plan and the prediction can be resolved. For this resolution the pilot can choose between various strategies using the two controls at his disposal: the rudder and the engine. How these controls can be used is the topic of this section. Two types of control are distinguished: heading control and rudder control. With heading control the pilot orders the helmsman to steer a heading and leaves the control of the rudder to the helmsman. With rudder control the pilot orders the helmsman the rudder setting, the actual angle of the rudder. Rudder control is generally combined with changes in engine thrust to optimize manoeuvrability of the ship and is done when a high rate of turn of the ship is required.

In heading control the control task is generally recognized as a pursuit tracking task. The helmsman has to steer the set heading $\phi_d(t)$ (demanded heading) as closely as possible. The present heading $\phi(t)$ is displayed as a compass heading.

The difference between demanded heading and current heading $\phi_e(t)$ must be kept to a minimum. For this the helmsman has the use of a rudder. By setting a rudder angle δ he can make the ship rotate. A major problem is that a ship is a slow non-linear system, difficult to control. Veldhuyzen &

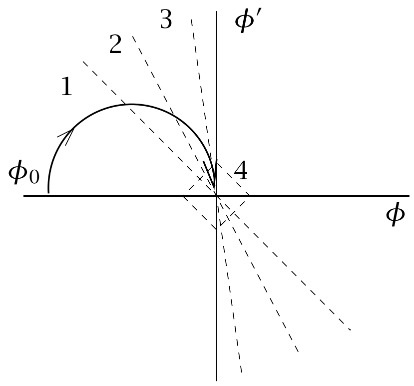

Figure 6.12: *Phase plane description of course correction. Adapted from Veldhuyzen, 1976.*

Stassen (1977) have modelled the helmsman. Central to their model is the use of a mental model derived from a helmsman. The following description covers how a course change is effected. This is a four-step manoeuvre. In step one the helmsman initiates ship rotation by applying a rudder angle. In the second phase the rudder angle is reduced and the ship is kept rotating with a constant rate-of-turn. In the third phase the helmsman stops the rotation of the ship by applying a reverse rudder angle and in the fourth phase, when the desired course is achieved, the rudder-angle is reduced to zero. This process is repeated constantly, experienced helmsmen being more successful than novices, who are generally not able to make a course change in one cycle and have to keep correcting. This manoeuvre can be visualized in a phase-plane, where the rate of turn ϕ is plotted as a function of the heading error ϕ_e. A plot is shown in figure 6.12. The four stages recognized by Veldhuyzen & Stassen are marked numbers $1-4$ and separated by dotted lines. In the phase plane a single course correction has been drawn. It starts with a course error ϕ_0 and no rate-of-turn (ϕ'). Then the rate-of turn increases, and the error is reduced. After reaching a maximum the rate-of-turn is reduced until the $|\phi_e| \leq \epsilon$, marked as phase 4, the area where the values are below a threshold. When the course state, represented by the line, goes from one area to another the rudder setting is changed accordingly. This manoeuvre is a tradeoff between time–accuracy and energy costs.

Sutton & Dowill proposed a model on course keeping (Sutton & Dowill, 1987) and on course changing (Sutton & Towill, 1988) based on fuzzy control. The results were very similar to Veldhuyzen.

When the pilot changes from heading control to rudder control a similar pattern can be seen. First the ship is made to rotate by ordering a rather large rudder angle (phase 1), often helped by ordering increased thrust. Then this rotating motion is maintained (2), and at the end the rotation

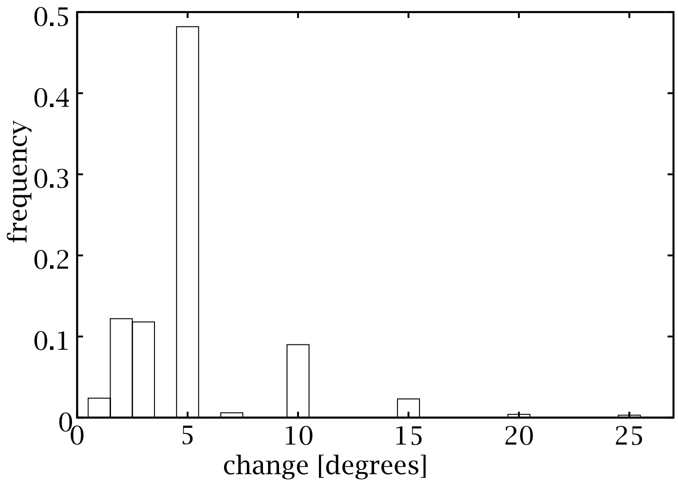

Figure 6.13: *The changes ordered in headings when sailing on the Nieuwe Waterweg (Rotterdam).*

is stopped by ordering a reversed rudder angle (3). The final corrections are made by the helmsman (4). Direct comparison with the model above is difficult because these manoeuvres are almost always part of a complex manoeuvre.

The two types of control can easily be recognized, but when do pilots change from one type to the other? My observation showed that the maximum rate at which headings are called is about one ship's length progress, and headings are ordered with 2°, 3°, and 5° updates, depending on ship size, fairway layout, and steering skill of the helmsman. The distribution of the updates recorded in the voyages I documented on the Nieuwe Waterweg (Rotterdam) can be seen in figure 6.13. If the required accuracy is such that a new heading is ordered every 5° during turning and the heading is updated at a maximum rate of one ship's length progress, then the ship must make a radius of 12 ship's length. When a rudder-setting of 10° is ordered, the resulting radius will be around four ship's length with adequate under-keel clearance. Rudder orders for angles smaller than 10° were not observed, so this may be another indication for the transition from heading control to rudder control by the pilot. From these observations it can be concluded that the transition from heading to rudder control occurs near radii of ten times the ship's length. The actual measure may vary due to the required accuracy, 'ground effect', in which the ship seems to 'stick' to the river bed thereby reducing the manoeuvrability, and 'squat', an increase of the draught of the ship due to the speed.

Although the rudder and engine offer an almost infinite number of possible settings, the number of settings actually ordered is very limited. Seven rudder settings and nine engine settings are all the pilot uses in my obser-

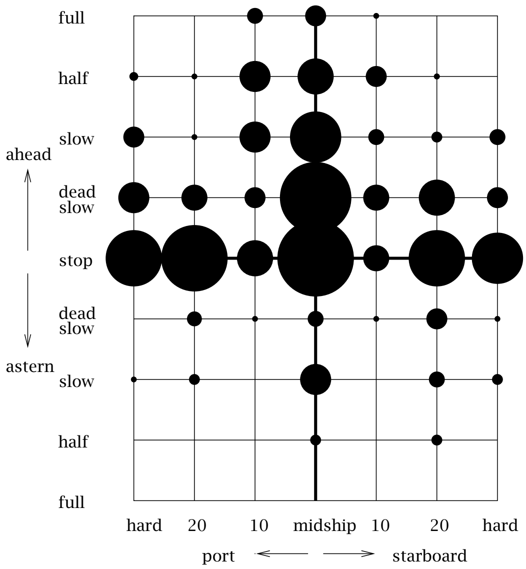

Figure 6.14: *The recorded rudder-thrust combinations.*

vations. The rudder settings are midships, 10°, 20°, and "hard to port" or starboard, where 'hard' is between 30° and 45°. As for the engine, the settings are 'stop engine', and then 'dead slow-', 'slow-', 'half-', and 'full-ahead' or '-astern'. It is the timing of the settings that allows the pilot very fine control. Of the 63 combinations about 1/2 is used, as can be seen in figure 6.14. This figure is an inventory of all rudder-thrust combinations used during the experiment described in the next chapter. The diameters of the dots is proportional to the number of that particular combination ordered. The combinations can be divided into four groups: turning left and right at very low speeds to bring the ship to the quay, bringing the ship to a full stop, and turning the ship at normal sailing speeds.

In the timing of the orders several delays must be taken into consideration. The engine is never directly under the pilot's control. The chief officer will either set the thrust directly on the engine controls or else the order is transmitted to the engine room where it is implemented. Depending on the type of engine it takes time to build up thrust. A special action is to reverse the engine: the engine is stopped, started up again in the opposite direction, and then the engine must build up thrust again. Smaller engines are usually much faster than large ones at this. Reversing an engine from slow ahead to slow astern will take from 20 seconds to about 2 minutes.

The rudder is not set with a proportional control, but by using a so-called 'tiller', a switch that activates hydraulic pumps that in turn rotate the rudder. This system means that the rudder has a constant angular velocity, in the order of 4°/s, so from hard-to-port to hard-to-starboard requires about 20s, varying greatly between ships. The increase of angular velocity of the ship depends on the size of the rudder and its maximum angular velocity. With a large and fast rudder the ship can pick up angular speed very readily, thereby reducing the need for additional engine thrust.

When the pilot arrives on board he must categorize the control characteristics of the ship. By giving a few orders the pilot verifies the ship's response: is it normal, slow or reactive? His conclusion depends largely on the time delay between a rudder order and a significant rate-of-turn.

6.4 Navigator mental workload

Pilots are highly experienced process controllers with highly developed models of the system they are controlling and the areas in which they work. The most important contribution to their workload is therefore expected to come from monitoring/prediction. The prediction for this study is that the dominant criterion for this workload is the time–to–boundary. When the time–to–boundary is above a threshold value no planning is required, the shorter the time is below that threshold, the more pressure is put on the navigator to change the state of the ship. Two groups of elements directly affect workload. One is the shape of the environment relative to the dynamics of the ship. The narrower and the more curved the environment to manoeuvre, the higher the workload is predicted to be. The bigger the ship is, the narrower the tolerances within the fairway, which in effect is the same as a narrow fairway with a high curvature. From the funnel model of the piloting process (see section 6.3.1), a low workload is expected at sea and a high workload in the dock. An important extra element in this are disturbances working on the ship: current and wind. The result of these is that the fairway is effectively narrower. A second group is traffic. Traffic is not only a varying change of the navigable fairway but there are also strong hydrodynamic interactions between two ships. Both groups of factors in workload are discussed in this section, together with some remarks on control in order to clarify the predictions which can be made about workload variation.

6.4.1 Fairway

If we use the funnel as harbour representation in our pilot model and study its effect on monitoring (and workload), a number of predictions may be made. The plan covers the major characteristics of the harbour. It contains information and decisions on where the funnel narrows at sea, the width of the fairway at sea, the ground-course that it takes on all stretches, the positions of the bends in the river, the position of the entrance of the dock, and the quay of destination. All this information is used to come to such a detailed plan. As mentioned previously, the horizon of detailed plan is not unlike an inchworm: it goes as far as the upcoming manoeuvre. As long as the course is within these set boundaries and the time–to–boundary of all variables is above a given threshold, the prediction is that the ongoing process is left uninterrupted. When the time–to–boundary drops below a threshold, actions are initiated to interfere in the process. At sea, where there are no direct physical boundaries, the time–to–boundary criterion will be largely based on headings. On the river and in the channel the criterion will be largely based on the physical boundaries of the fairway. It is expected as a first approximation that this boundary can be considered to be pilot-independent. Because monitoring behaviour is a result of the actual state of the ship, the workload associated with the manoeuvrable area will be called the 'fairway-related workload'. Some general remarks on the workload can be made.

At sea the manoeuvrable area is large, resulting in a low workload. Approaching the channel, the manoeuvrable area decreases, resulting in increased workload. In the river the fairway will gradually decrease in width, resulting in a slow increase in workload. When the fairway bends, the time to boundary reduces as a result of this. This means that the workload increases because of bends in the fairway. When there are uncertainties regarding the position of the ship or the fairway, this has the same effect as a reduced manoeuvring space: it results in increased workload. Uncertainties regarding the position of the ship are always relative to the layout of the fairway, making transverse uncertainties generally dominant over longitudinal uncertainties.

6.4.2 Traffic

In the previous section only the effect of the fairway on monitoring and workload were discussed. Another factor contributing to the need for decision making is the presence of other traffic or ships. Often ships are modelled as floating islands. It is then assumed that these objects travel in a straight line at a constant speed. This is a very simple model that can be used to describe the effect of other traffic on the plan. If traffic is modelled as a floating island it can be treated as a part of a time-varying fairway. Unfortunately modelling traffic as floating islands fails to take some fundamental features of that traffic into consideration. The other traffic is in fact a ship similar to the one under control. The crew on its bridge plans, decides and controls the ship, so it does not travel in a straight line, but reacts to disturbances and other traffic. What manoeuvres are made depends largely on the manoeuvrability (the size) of the ship. Smaller ships have a larger area in which they can manoeuvre and react faster to the helm and the engine and therefore often have to give way to larger ships. Also, there is a hydrodynamic interaction between ships that makes them behave differently from floating islands.

In this study three categories of encounters are recognized. These are ships in a crossing situation, a head-on situation, and an overtaking situation. When ships are crossing, the problem is the decision if there is the possibility of a conflict and how to avoid the conflict with minimal manoeuvring. The formal rules are very clear, but deciding if a potential conflict exists is often difficult in a harbour where ships manoeuvre, slack down, and accelerate. Head-on encounters are complicated because the time to contact, and the workload, are very sensitive to small variations in heading. When the courses of both ships are varying this increases the uncertainty. When the ship is overtaken at a small lateral distance there is a hydrodynamic interaction between the two ships: i.e. there is a force towards each other. At the control level this must be compensated for, or there will be a serious risk that the ships will collide board on board, or that the overtaking ship crosses the other ship just after overtaking. The smaller ship is the ship that is most likely to change course because of these forces on it.

6.4.3 Control

The pilot communicates with the crew, with other ships, and with the radar operators. In general this communication will have little effect on workload if it passes quietly within the plans made. However, if there are communication problems, or when unexpected events happen, communication will have a significant effect on the workload. The same happens when equipment fails, control actions do not result in the effect that was expected, and with many other 'minor' problems during the execution of the task.

All of the situations described in this section indicate when workload can be expected to increase, following the time–to–boundary criterion. In the experimental chapter these predictions will be compared to the data collected for the various voyages with pilots.

6.5 Information collection

To allow it to be controlled the state of the ship must be known. This is done through assessing the state variables by regularly observing or monitoring the formal and informal displays. A display in this context is an abstract description of an information source that can be sensed by a human, with no distinction between any of the sense organs. So the total display is the assembly of instruments and other information sources, such as the window, for the navigator. More difficult are the notions of 'formal' and 'informal'. Formal display elements are displays formally designed for the job, e.g. a compass, a VHF-radio, a temperature meter, and the view from the window. Informal display elements were not intended by the designer for the use of the operator, but are used nonetheless, e.g. the vibration of the engine as an indicator for thrust, the clicking of the compass repeater used as a rate-of-turn indicator, using the bridge symmetry as a centreline indicator etc. Based on all this information the navigator has to decide whether or not to intervene in the ongoing process. If intervention is required he will decide to change the setpoints, the tactics, the strategy, or even the plan. Changing the plan involves changing the goal, e.g. deciding to direct the ship to another harbour because of the weather conditions.

We have recognized three mental models for monitoring and control. These three models are not completely independent models, but are strongly interconnected and dependent on each others' outcome. From the navigator's point of view the number of control inputs on a ship is very limited. He has only a rudder and propeller at his disposal. The highly variable environment causes the large variety in decisions made by any one navigator sailing the same environment: no two passes are executed the same way. Most of the actual control tasks with respect to the rudder and propeller are left to special operators or to automatic control, leaving the navigator to set heading and speed. The task of the navigator is really supervisory control. He has to schedule the sampling of the various variables. On the basis of these samples and the information derived from previous samples he estimates the true state of the system. He then has to decide if the system is in a non-optimal state or requires intervention. When it is in an abnormal state he must decide whether to intervene in the process and if so, how this intervention must take place, or whether to change the plan and start deciding from this new plan. Based on this decision he will choose an appropriate level of control.

6.5.1 Planning information

The information needed for updating the plan is largely derived from the actual layout and the current situation of the fairway over the next way-point. A way-point is a point at which a new course is set. Under normal conditions this planning is largely based on the visual scenery. Due to the constantly changing perspective and the deviations due to disturbances such as wind and current the plan is regularly updated.

From the visual scenery the navigator extracts features such as width, length, and curvature that are compared to the ship's size, course, and heading. An example of the navigator's view is presented in figure 6.15 (corresponding to the plan in figure 6.11). In this figure the 'own' ship is following the ship further down the river (and will keep on doing this to sea). The own ship has to go to port in order to follow the river (and the other ship ahead). The 'other' ship is expected to change her heading to the starboard side in order to follow the river. This view is compared to the ship's expected state in planning. The navigator has limited references to obtain the current state of the ship and so he is also limited in the accuracy with which he may determine this state.

The navigation plan is largely based on a series of straight lines that follow the centreline of the river closely. As was argued earlier pilots very

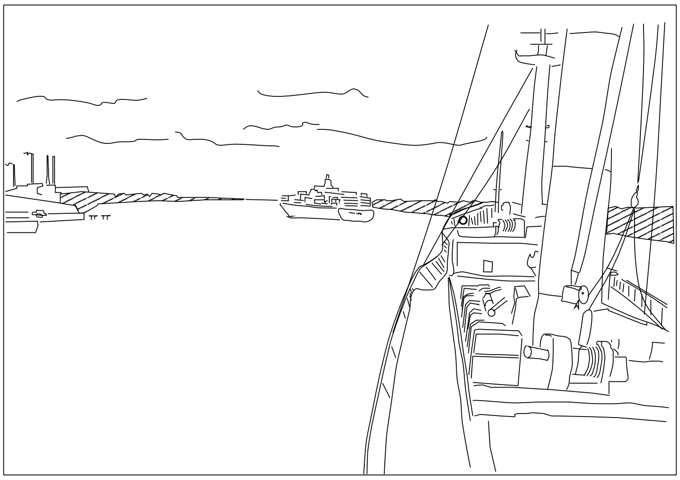

Figure 6.15: *The visual information for planning on the river (the situation can be seen in plan in figure 6.11).*

much depend on recognition points for their navigation. Any object that has a fixed position and can easily be recognized can be used for this purpose. Commonly used are the numbered moorings along the fairway, buoys, harbour entries, buildings and chimneys, landings, etc. Objects for navigation must be available during day and all night because shipping is done around the clock. For navigation using radar, objects must be available that can easily be recognized on radar and are distant enough not to be obscured by clutter and do not suffer from reflections. The recognition points are spread over the river in such a way that they can be used as markings to initiate new headings and as transits to verify heading and position. Schraagen (1993) analysed the use of these objects by different pilots using a 'think aloud' technique on board. What can be seen from the frequency tables obtained by him is that the frequency of use very closely resembles the availability of these objects as can be seen in table 6.1: $\chi^2_{(4)} = 0.185, p < 0.005$. The availability was derived from counting these objects from a navigational map.
Considering that these objects are spread over the river with a 5 minute interval and the planning horizon is about 10 minutes it can be concluded that all pilots more or less use the same way-points. If we include the information on how frequently pilots update the requested headings and with shape of the river the assumption can be made that the course-lines chosen by the pilots are essentially the same.

The long-term plan provides information on which navigational marks are available and the ship's required position in relation to them. So, the navigator monitors the scenery for marks and decides on how to manoeuvre to the next way-point position to meet his long term plan.

Table 6.1: *The use of object by pilots, as studied by Schraagen, compared to the availability of these objects.*

object[a]	pilots	available
moorings	54%	53%
buoys	16%	21%
harbours	16%	17%
landings	4%	7%
other	5%	1%

[a]the original number of categories of 9 was reduced to 5 because sailing in the docks was excluded in this analysis

6.5.2 Control information

At the control level all information sources are monitored that are needed to implement the short term plan. The dominant element is the transverse position and the heading of the ship. Together with the speed these three determine the time-to-boundary, i.e. the time left before grounding. In a variable situation such as navigation in a river the TTB is changing continuously. All information used by the pilot can be seen in figure 6.16. This is a simple control model of the ship in the fairway, showing the relationship between the different control variables. This figure shows four feed-back loops: order (verbally), rudder angle, ship's heading, and ships position. The inner loop is the fastest, the outer loop the slowest.

The navigator requires three groups of information. The first is the position in the fairway relative to the fairway (and objects in that fairway). For this the navigator can use the visual reference frame provided by the ship's hull, superstructure, and windows. The strict symmetry of the ship is a reference for the estimation of the relative position of the ship from his position on the middle of the bridge. From the bridge the hull and bow also provide a very good reference for the relative heading of the ship. The compass is used in open water (sea), and as an instrument for instructing the helmsman where to go. For a course a series of positions is normally needed, though two fixed objects in the horizontal plane provide a very accurate course, which is preferred in most confined fairways. Rotation around the vertical axis or yaw is by observing the changes in heading, sometimes supported by a rate-of-turn indicator. The rate-of-turn depends on the rudder angle. In confined waterways rotations are very important, and therefore the rudder indicator has a dominant place on the bridge, providing the rudder's angle. The information on compass course and rudder angle can be read from instruments and are called by the helmsman the moment the setting ordered by the navigator has been reached.

When manoeuvring becomes too complex using heading orders only, the tactic for controlling the ship is changed. The navigator will not use heading

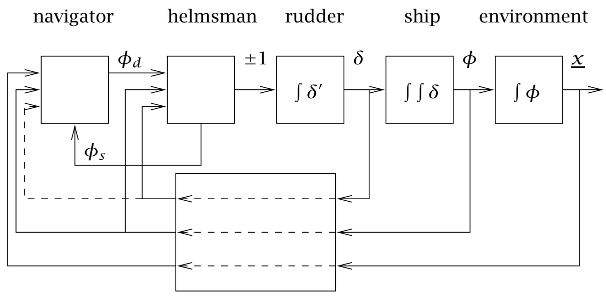

Figure 6.16: *A model of heading control of a ship. The lower block represents the display of the instruments and the view of the environment from which different variables are extracted.*

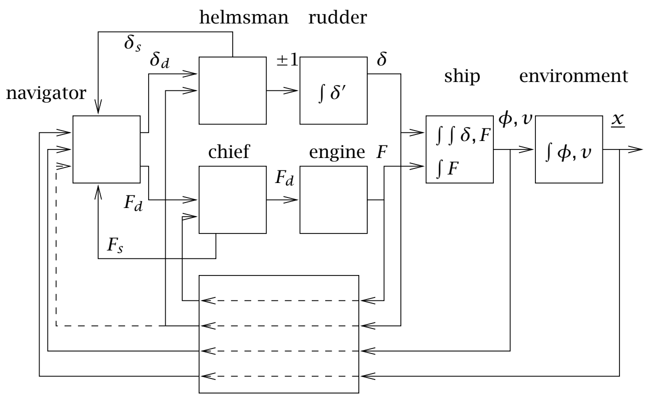

Figure 6.17: *A model of rudder control of a ship.*

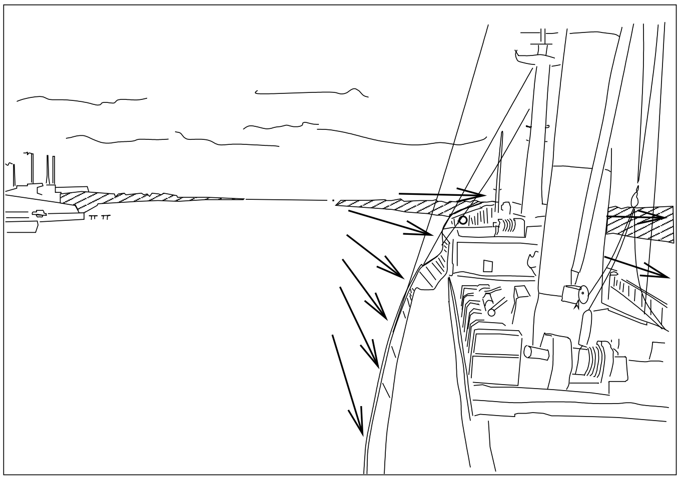

Figure 6.18: *The 'relative shift' at the river as perceived by the navigator.*

orders to manoeuvre the ship, but will change to rudder order to obtain more direct control (see figure 6.17). In this situation an officer takes control over the engine setting, and the helmsman is restricted to implementing rudder-angle orders. The task of the helmsman is then taken over by the navigator to a large extent. Together with the direct control of the rudder the navigator will also start using the engine for extra manoeuvrability. Besides the extra control information on rate-of-turn, the speed of the ship must be kept under control, making the task more complicated.

Because of the large response times of a ship, accurate prediction is needed on the transverse position. For this the heading, yaw, rudder angle, and the change in rudder angle can be used. The order of these information sources represent the order of rate of change, from slow to fast (compare this to figure 6.17). This is inversely proportional to the scanning frequency needed to monitor them, but also to their information content as a measure of the variable which the navigator really wants to control: the transverse position. The longitudinal position is slightly different because it is judged based on movements in the direction of the fairway which is considered to be infinitely long at this stage and therefore generally not of importance.
Some variables used by the pilot are retrieved from the changes relative to the own ship, and shifts relative to various elements in the environments. The ones considered most important are rotation and sideway movements. These are directly available to the pilot in the form of relative shift (figure 6.18).

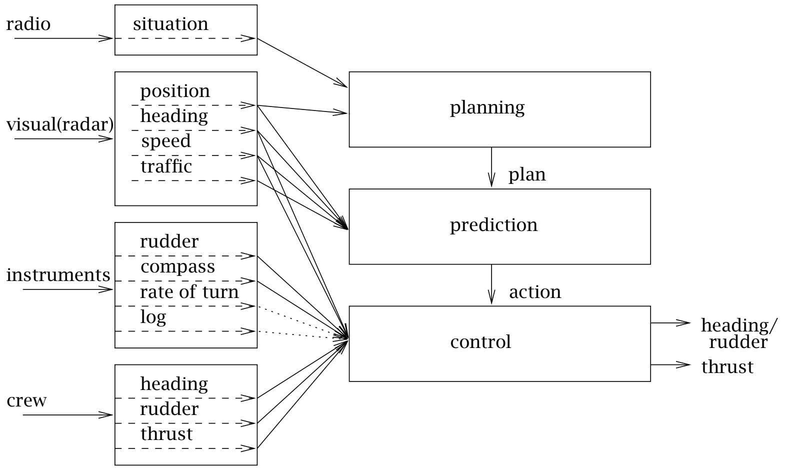

Figure 6.19: *Detailed navigator model. The dominant loop is the prediction providing the TTB-criterion. The dotted lines represent information used very rarely.*

6.6 A navigator model

A more detailed model of the navigator as was presented in figure 6.16 can be made using the information of the last two section to come to a more detailed model, presented in figure 6.19.

The model shows the three mental models with their information sources and how they might be related to each other. The lowest level of decision making is the control level. This level is dedicated to fast changes in the order of seconds. The heading and speed decided on at higher levels of decision making are monitored and adjusted. The most important type of information is the rate-of-turn of the ship. This is generally derived from the outside view. The relative shift of the bow to the environment is an accurate and reliable rate-of-turn indicator. The rate–of–turn instrument is hardly ever used (as the dotted line in the figure indicates). Corrections are made using rudder control or by providing the helmsman new headings.

The second level is the prediction. Prediction concentrates on implementing the general headings provided by planning using the TTB as a criterion for intervention. It does this by keeping the time–to–boundary acceptably long. The prediction is dedicated to processes with times of a few minutes. The idea is that the pilot does not interfere in the ongoing process if the TTB is above a threshold, and decides to a change in heading (and/or speed) when the TTB becomes too low.

The third and highest level of control is planning. Planning is dedicated to processes in the order of tens of minutes. This plan is the set of way-points and the headings that connect the way-points. The standard plan is largely completed before the voyage starts and this plan is continuously updated with information collected by the navigator. So the initial plan provides general headings for the planning process and the planning process in turn provides changes in heading or rate-of-turn for control.

It is the prediction phase with the TTB that is the centre of focus in this study. It is at this level that the actual on-route control decisions are made.

6.7 Discussion

In this chapter various elements of the navigator's task have been discussed. In order to navigate a ship information is needed. For different types of decisions, or different levels of control, different types of information are needed. The pilot is in a special position because his knowledge about the ship he is navigating is limited and his knowledge about the environment is very rich. This is different from a normal navigation situation where the navigator is on a familiar ship in unfamiliar waters.

In decision making different types of control can be recognized. The primary level is manual control. The pilot does this only to a limited extent. Two types of manual control can be recognized: heading control and rudder control. Which one of these two is used by the pilot depends on the environment in relation to the dimensions and manoeuvrability of the ship. This level of control has been often researched and described in literature. Considering the time constants involved, it will largely depend in skilled behaviour.

The next level of control is prediction. At this level of decision making the use of controls is not of importance but the state of the ship in relation to the environment is. It is at this level that the pilot decides how to change the setpoints for the state of the ship. This level will be the focus of the next chapter.

The highest level of control is planning. At this level an initial plan is made up once, before the voyage starts. The most important outcomes are the constraints in which the ship must stay and the way-points that mark the changes in straight stretches and curves in the fairway. During the voyage this is continuously updated with new information.

Navigation is dominated by the environment of the ship in relation to the manoeuvrability of the ship. The environment is therefore very important when studying the navigator during his work. It is also expected that prediction dominates the navigator's monitoring behaviour. It is at this level that control decision must be made continuously. Because of the time-to-boundary hypothesis and the importance of prediction the navigators mental workload is expected to be a function of the time-to-boundary. In the next chapter this hypothesis is tested in an experiment with pilots doing their normal work in and around the harbour.

CHAPTER 7

The Rotterdam pilots at work

In this chapter the theory described in the previous chapters is put to the test in a navigation situation. Four Rotterdam pilots participated in an experiment in which they were recorded on video, their heartbeat was recorded, and their voyage was logged on maps, all during their normal work. In this chapter it will be shown that prediction, as has been defined and explained in the previous chapter, dominates mental workload. From this it will be concluded that monitoring and control behaviour in the complex task of maritime pilotage is dominated by the prediction of what has to happen in the near future and that the theory on monitoring can adequately account for the vast majority of workload. This is shown based on the reactions to the fairway and traffic from other ships.

7.1 Experiment

The main hypothesis of the alarm-clock model to be tested in this experiment are the following:

- The workload (HRV) will covary with the time–to–boundary (TTB) of the ship in the fairway.
- This covariation of the HRV and the TTB is independent of the individual pilots.
- Workload related to shipping traffic, in particular meeting, overtaking, and crossing, can also be predicted using the TTB criterion.
- The HRV pattern will lead the TTB pattern by a significant distance in order to compensate for the large inertia of ships.
- The workload during phases of rudder control will be significantly higher than during phases of heading control.

- The visual sampling frequency depends on the fastest control on the ship (i.e. the rudder).
- The TLX provides an average subjective workload with information about the attitude of the pilots towards the task, whereas the HRV provides a moment to moment measure of their mental workload.

7.2 Data collection

The primary measure chosen for mental workload for this study is heart-rate variability. Its most important feature is the bandwidth of this measuring technique. Another important element in this choice was the ease with which data could be collected and acceptance by the participants or subjects. The subjects are pilots doing their normal work. If the recording technique is inconvenient they start feeling like guinea pigs. On ships one has limited possibilities for recording. First of all the equipment must be able to be worn conveniently because the pilot is brought on board by a speedboat, and from there he has to climb up a nine meter high ladder. The HRV-recorder we used had the size of a hand-held tape-recorder, weighed 0.5 kg and was carried in a belt. Similar considerations were made for the other recordings. Complicated and extensive recordings that may be done easily in a laboratory are very difficult to set up in a few minutes on the bridge of an unknown ship after having to obtain permission from the captain. Receiving permission from the captain proved to be no problem at all. As long as my presence would not give rise to any conflicts with ‘the office’ they were very willing to help me with everything. Work at sea can be very boring, so every event that gives something to talk about during dinner seemed to be welcome.

All activity on the bridge was recorded using a small video camera (“handicam”) with the signal lights covered. The voyage was plotted on a set of maps. A notebook computer was used to collect the data from the heartbeat recorder after the voyage. All fitted easily in a small backpack. A stopwatch was used for time-synchronization.

Four experienced Rotterdam pilots participated in a series of recordings. These pilots were selected from a group of eight pilots that assisted during the introduction and the initial task analysis for this research. Each pilot was asked to do at least two inward-bound and two outward-bound voyages; most did more. This resulted in a series of 25 voyages, with a large range of types and sizes of ships and destinations. The ships were selected for maximum size from the ships available at the start of the shift of the pilots participating. The reason for choosing the large ships was that the smaller ones are usually “feeders”, ships that tramp between the various ports so that the large ships can restrict themselves to a limited numbers of ports. As a result the crews of these feeders are very familiar with the harbour and the task of the pilot on such ships tends to be more that of an adviser and supervisor than that of a navigator. The most important criterion was that the voyage would be from sea to quay, or vice versa. Sometimes this was impossible due to the pilot demand at that moment; the pilots participating were doing their regular work during regular shifts. Ships for the river (leading to the city of Rotterdam) were preferred to ships for Europort because the voyage then included the river. All pilots were videotaped with a camera placed on the window of the bridge, facing into the bridge. In addition the voyage was plotted on a map by hand, marking the route, ships passing, and special events.

Normally there were two voyages during one shift, with one outgoing and one incoming ship if we were lucky. From all ships the date, time, pilot, ship name, type of ship, length, size, draught, wind, propeller, rudders, and additional thrusters were noted . The voyages ranged from 1.5 hour to just over 3 hours. Ship sizes ranged from 59 m to 294 m, with draughts from 4.7 m to 12.1 m. There were 13 ships bound for sea, 10 ship coming from sea, and two ships from quay to quay. (see appendix A for details)

The electrodes for the HRV recording were attached to the chest of the pilot at the pilot station when we were at sea, or on the ship before departure from the quay. Careful cleaning and placing of the three pregelled disposable long-term ECG-electrodes of the silver/silver-chloride type provided a reliable signal over the entire shift. During one voyage the heart-beat monitor failed and all data were lost. The ECG was recorded with the electrodes on position V6 (according to Wilson; see Keidel, 1973) and at the sternum, in a straight line across the subjects heart. From this ECG the heart-beat was extracted using a R-wave trigger (accuracy < 3ms), and this signal was fed into the recorder. The recorder could store 28520 heart-beats with 1 ms accuracy, making is possible to record for 5 hours. The recorder was started about 10 minutes before the voyage and synchronized with a stop-watch. This stop-watch was later used to record events. In addition to collecting heart-beat, the recorder could record 16 different events by means of a IR transmitter-receiver combination, similar to the remote control of a television. This enabled easy and reliable synchronization of events with the heart beat data. The IR connection proved to be less successful on the bridge than in the laboratory because of the special low-reflective paint used on bridges. As a result were about 10% of the synchronization signals lost, but this had no serious consequences. There is no indication that the recordings were effected by transmitting equipment. The video could record 4 hours continuously on Hi-8 videotape.

The pilot was asked to fill in a TLX-questionnaire. This TLX questionnaire was an official Dutch translation of the NASA-TLX questionnaire. The six rating scales (mental demand, physical demand, temporal demand, performance, effort, frustration) with a 12-point scale were filled in for three stages of the voyage: sea, river or canal, and dock[1]. The sources–of–workload data, which provide the weights for the rating scales, were collected only

[1]See appendix B for a definition of the scales

once for each voyage. During 12 of the 25 voyages the full set of data for the TLX was collected from the pilot concerned. However, the use of this measure proved difficult for practical reasons. It was often not convenient to fill in this questionnaire during or very shortly after the voyage. The ship needs constant attention during the voyage, and as soon as the ship is secure work starts on board, and the pilot has to leave for the next ship. This resulted in 13 missing sets of TLX-data.

7.2.1 Data treatment

The data was treated in general as in the DURESS experiment. The ECG recordings were corrected for recording errors, and transformed into power density signals using Carspan 1.99 (Mulder, 1988). The power in the 0.1 Hz-band was calculated using a Discrete Fourier Transform based on the Integral Pulse Frequency Modulator model (Rompelman, 1986). The time window of 50(s) was considered to be acceptable for both a sufficient spectral resolution and a stationary signal.

The time–to–boundary was calculated using the position recording and the headings. From this the positions were interpolated using a map of the sailing area. This was done at regular intervals to obtain a TTB-curve that could be compared to the HRV for the fairway-related workload. For the traffic-related workload the position with respect to the other vessel was used as the independent variable.

The videos were transformed into activity lists which were synchronized with the heart-beat recordings. Eleven of the 25 voyages were fully transcribed: only daytime-videos were used because all pilot activity was recorded. The results were such that there was no need to transcribe nighttime videos. The following categories were recognized:

- order: a navigational order
- asking: asking for navigational information
- respond: being asked for navigational information
- social: social interaction
- radio: radio calls not addressed to the ship
- radio with ship call: radio call with the ship's call sign in it
- walking: physical movements, more than one step
- eating, drinking, etc.

The first three events were recorded into two classes: the pilot's part and the crew's part. Each of these events could be marked with video-frame accuracy. It was decided that 3 s accuracy was sufficient. This information was later added to the heart-beat recordings.

The TLX-ratings were converted into workload ratings according to the NASA-TLX paper and pencil package version 1.0 (NASA, 1986).

7.3 Individual differences

Until now, individual differences between navigators such as differences in goals and individual objectives have not been taken into account in the descriptions and discussions. It may be that some pilots strive to maximize the distance to obstacles, whereas others prefer to 'walk the boundary'. From my interviews and observation it was clear that some pilots feel more comfortable as far away from boundaries as possible, thereby maximizing their 'manoeuvrability margin'. Others prefer to stay near the boundaries for reasons of sensitivity (being able to "feel the edge"). Near the boundary changes in the operating points are more apparent, a clear reference is at hand, and when the process values change, it is immediately clear in what direction corrections have to be made: away from the boundary. These pilots will have developed skills which enable them to do so.

It will be shown that despite these differences between pilots the resulting workload profiles are very comparable when related to the TTB for that pilot. The reason for this may be that the workload can be higher or lower in absolute terms, and the differences between the lowest and highest workload may be larger, but the shape of the workload profile remains largely the same for each different pilot with their individual characteristics.

As for the risk-perception, uncertainty, knowledge, and skills, and their effect on workload, it is clear that these are largely constant for one voyage of each individual pilot and therefore have no effect on the variations of workload during that voyage. At this stage in developing the theory and testing it, it is therefore safe to ignore individual differences.

7.4 Data analysis

For this study various types of data were collected. From all subjects the heartbeat was recorded. As in the DURESS experiment all process variables were recorded continuously. The ship's position was recorded at regular intervals and all orders and confirmations were recorded. None of these data could readily be used directly to calculate the correlation functions: first the heartbeat data had to converted into HRV data using Fourier transform and the MTB in the harbour had to be reconstructed from the sailing plans.

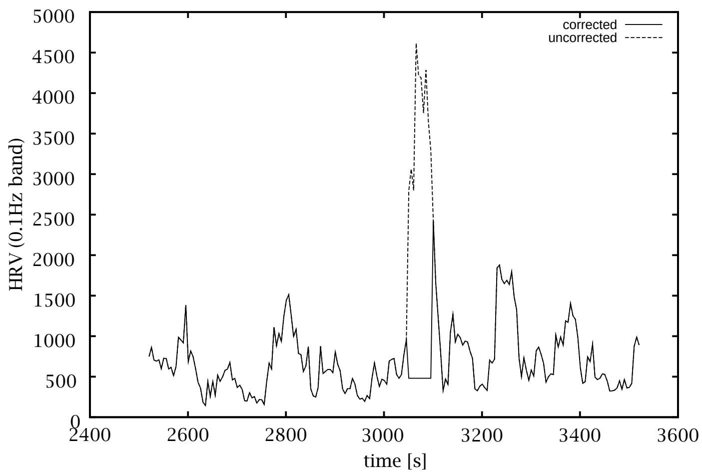

Figure 7.1: *HRV profile with artefact.*

7.4.1 Heart-rate variability

All heart-rate data were converted into heart-rate variability (HRV) data using Carspan 1.99[2]. The time window for the FFT and artefact detection was 50s and the maximum deviation of the RR-interval was 4.0 standard deviations (default). The frequency bands were 0.02–0.05Hz (low), 0.07–0.14Hz (middle), 0.15–0.50Hz (high) (all default). Harcor diagnosed about 2% of the heartbeats of the pilot data as artefacts and these were largely concentrated in the docking-phase of the voyage. About 75% of the artefacts was corrected by Harcor.

Two types of disturbances showed up in the pilot-HRV data: heart arrhythmia (random disturbances, artefacts), and the effects of other traffic (systematic disturbances). The artefacts resulted in outliers in the HRV signal. These outliers were corrected with an automated procedure using a window. The mean and standard deviation were calculated in a 300s window. If in the 5s interval right of the window the HRV data was outside the $\mu \pm 4\sigma$ interval it was replaced by the mean. After this the window was shifted 5s. An example of an artefact correction is shown in figure 7.1.

The effect of passing ships had to be treated completely different. In studying the effects of the fairway on the pilots workload they are unwanted disturbances and it is therefore preferred that they are removed for that analysis. However, when analysing the effects of traffic these events become the centre of focus (see also section 7.5.3). For this reason all passing ships were localized and the effects of these passages were removed by hand in the fairway analysis.

[2]Carspan is an environment for processing cardiovascular time-series

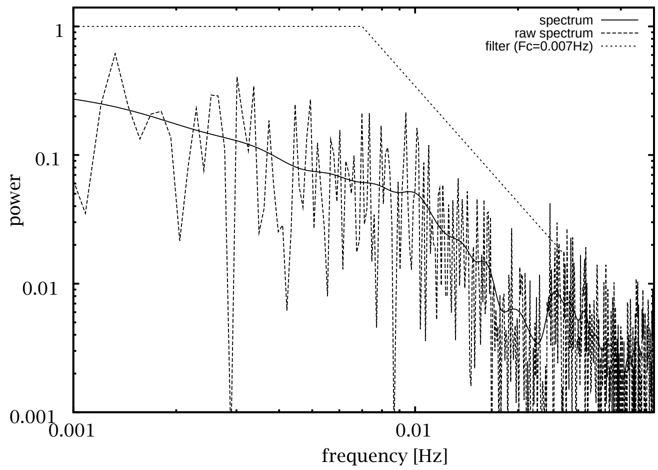

Figure 7.2: *HRV spectrum of the 0.1Hz signal of a pilot during work.*

Subsequently the spectrum of the HRV and of the MTB was calculated[3]. The HRV spectrum is much broader than that of the MTB. The highest frequency in the spectrum of the MTB of the pilot experiment is 0.007Hz. Therefore the HRV signal can be filtered in order to obtain signals that have the same characteristics. All HRV data with $f > 0.02$Hz has no meaning and can be treated as noise. This is the result of the 50s window used for the Fourier transform.

All HRV data of the pilots were filtered using a low-pass filter with $F_c = 0.007$Hz (6dB/oct, using FFT). With $F_c = 0.007$Hz about 30% of the spectrum below 0.02Hz was removed. The spectrum of a 0.1Hz HRV-signal of one pilot's voyage is shown in figure 7.2. The resulting HRV signal in shown in figure 7.3.

The HRV data is in the time-domain, but the MTB data is in the place domain. It was chosen to bring all data to the place domain before calculating the correlation functions. In order to be able to calculate confidence intervals for the correlation functions, a number of locations had to be chosen that provided information about several voyages. Six such locations were selected. These locations, or trials, were 3—10km in length, taking about 500—2000s. The HRV data on these trials were copied from the complete HRV recordings. For the MTB an 'average' course was constructed for each location using the headings ordered, the confirmations by the helmsmen, and the location of the ship as recorded during the voyages. These headings and locations were plotted on a map and every 150m the distance to the boundary was measured. The border of the sailable area according to the charts was taken as the boundary. These data, the HRV data and the MTB,

[3]All calculations were made using Octave, a high-level language mostly compatible with Matlab.

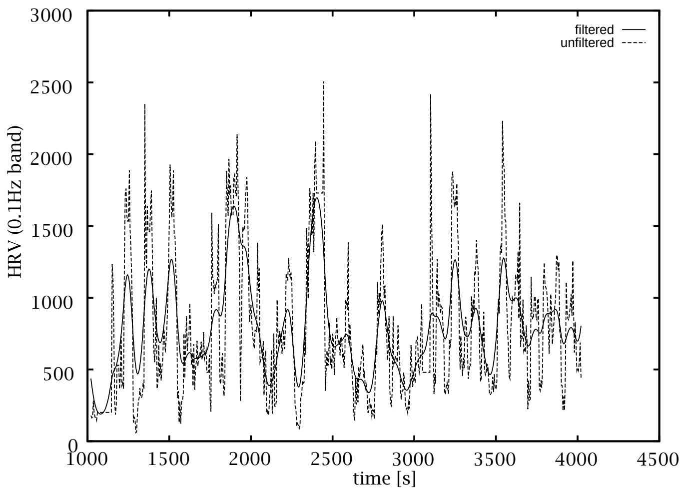

Figure 7.3: *An example of the 0.1Hz signal raw and filtered with* $F_c = 0.007Hz$.

were interpolated using a spline and re-sampled on 2048 points. These re-sampled signals were then used to calculate the correlation functions. This resulted in a correlation function for each trip, each with a similar shape but with different maxima and different displacements. All trips for one location were taken together by calculating the average displacement, shifting the individual correlation function to this average displacement, and calculating the average correlation function. All displacements within one section were all within 300m. The average displacement is later referred to as 'time shift'. One trial was removed: in the Beerkanaal one ship was passed on the port side instead of the normal starboard side, giving a deviant HRV pattern.

7.4.2 HRV-measurement complications

In laboratory studies it has been shown that speaking for longer periods (>10 s) will show up as an increased HRV in the 0.1 Hz band. Under the conditions on board this effect was not found to be a problem, mostly because most speaking is done in short orders. Also, standing up after resting has a strong effect on the heart-rate. This did not apply in this experiment as the pilots normally remain standing.

The physical activity of the pilot is very limited, with the exception of boarding and leaving the ship. Normally he stands on the bridge, although new bridges are often equipped with consoles that require sitting down to operate them comfortably. In three voyages this situation was present, but the pilot remained seated in the analysed areas. His usual place is in the middle of the bridge, near the compass and the rudder indicator, and a few steps away from radio equipment and a radar console. Most of the information is collected visually by regularly scanning the outside scenery, and little information is obtained from the radio and radar. Orders for heading, speed, rudder, and thrust are given verbally and are confirmed and reported back by the crew as soon as the order is implemented. In order to see what is beside and behind the ship the pilot will slowly walk to the side of the bridge every five to twenty minutes, depending on the traffic and manoeuvring. Social conversation is limited to avoid distraction from the task. Only when the traffic is very quiet does it pick up. The pilot is always offered coffee and cold beverages[4]. On longer voyages when it is about time to eat, or when the pilot asks for it, meals are offered to the pilot. This happened in five voyages, but this eating period fell largely outside the analysed areas and no corrections were made for any of these possible disturbances.

The heart-rate and HRV showed no significant reaction to the physical exercise as long as the pilot remained standing. The heart-rate remained about the same when sitting or standing, but the transition gave a significant short-term effect. Giving orders (the physical part, not the decision-making process) had no visible effect on the HRV. Speaking had no large effect on the HRV and speaking periods could be included uncorrected. This is probably because the conversation could be labeled 'small talk' with short sentences and long breaks in between. There was one event in the activities that showed a significant effect and that was when the radio mentioned the ship's name. Pilots are very sensitive to this, and on hearing 'their' name showed a sudden increase in mental workload. Three such disturbances on the HRV signal, which lasted for about 30s, were removed by hand.

7.4.3 Time–to–boundary standard

For the time–to–boundary a standard sailing plan was drawn based on the maps available, the recorded orders during the voyages, and the informal interviews with the pilots. The orders recorded on video largely confirmed the information obtained through interviews. The initial plan was based on the information from interviews. This plan was then refined based on the transcriptions of the videos where rudder orders were given. Orders which were the result of the traffic situation were excluded. This series of way-points provided the frame for interpolation of a course-line with a 1 cable[5] interval. From each of these interpolated points the time–to–boundary was calculated based on extrapolation of the heading and the boundary of the fairway. These distances could be interpreted as times because the speed of a ship can be largely considered constant. To become a full TTB-curve this set of times (or distances) was interpolated using a spline.

[4]It was agreed to restrict the coffee intake to two cups per shift during the experiment.

[5]1 cable = 185m

On the river this procedure was relatively easy to use, but near sea rather large differences between voyages and pilots emerged. Reasons for this are the tidal streams, the traffic that is more complicated in this area, and a less dominant river bed that leaves the pilot more freedom in choosing a plan. At sea this technique to construct a standard navigation-plan could not be used due to the large differences in the sailing conditions, and hence the data for these parts of the voyages could not be used further in this study.

A special problem was posed by the communication with the VTS, the radar operators who provide information to the pilots. At specific points in the river pilots exchange information with these stations and this results in increased levels of mental workload. To deal with this complication these points are always treated as locations with a short TTB. In the analysis there are three of these points.

Interviews

After a voyage the events were discussed with the pilot. Valuable information was obtained by interviewing afterwards. In the resting period between shifts the data were roughly analysed by comparing the voyage, the events, and the HRV profile. Some characteristic points in the HRV profile were related back to an event and/or location, and the pilot was asked about what had happened there, without telling him about the HRV-characteristics. In order to disguise the significance of these questions the pilots were also asked about other events that had occurred during the voyage. This provided some very useful information about currents and wind problems, events that passed unnoticed during the voyage. These events were later confirmed in a similar way by the other pilots. Although clarifying interviews proved valuable, the time for a thorough 'walk through – talk through' was not available. The time between ships is often very short, although long waiting periods did occur. The double irregular working hours (working hours change every shift) does not make it convenient to plan long interviews. This made it the more important to have the general knowledge about the work of the pilot as collected and described in chapter 6 in order to interpret the data in the first place.

7.5 Results

7.5.1 Heart-rate

With respect to the heart-rate the hypothesis is that the pilot will show an increased level of heart-rate due to very high levels of mental workload. This is expected to happen when the pilot is forced to switch from heading control to rudder control and when he manoeuvres the ship using tugs. The heart-rate will also show an increased heart-rate due to the physical

Table 7.1: *The pilot's heart-rate. This table summarizes the average heart-rate of the pilots during rest, normal navigation, entering a harbour, bridge, or dock, and climbing to the bridge, together with a 95% confidence interval. N=24*

resting	navigating	docking	climbing
69	82	108	134
67-72	80-85	103-113	127-140

effort needed to come on the bridge. This has no meaning for the mental workload.

The heart-rate can be directly calculated from the heartbeat recordings. The resulting data was a beat to beat heart-rate that varies continuously during the voyages. Special filtering was not required. The heart-rate of three phases for all voyages are summarized in table 7.1. Together with the heart-rates under the three conditions the heart-rate during rest on board is given. These phases are navigating, docking, and physical exertion.

The expectation with respect to physical exercise was very clearly found. Bridges can be as high as 30m (12 stories), pilot ladders can be up to 9m high, and not all big ships have elevators. The physical exercise that results from this will bring up the heart-rate to as much as 156 beats per minute, close to the maximum for the age of the pilot. This can be seen clearly in figure 7.4, top graph, at the first marker at T=750s. This figure is an example of a typical voyage. In this situation a rather small ship was navigated from sea to a dock near the city of Rotterdam and finally moored on a buoy. The top graph shows heart-rate during the voyage, and the bottom graph the heart-rate variability. The high but very narrow peaks in the HRV have virtually no meaning. Please note that high mental workload is associated with low HRV. The markers in this figure are placed at periods of high workload based of the HRV profile (except for the first marker).

A second phenomenon is an increase in heart-rate during the voyage. The normal rate of the pilot on the pilot-ship or ashore is about 70 beats per minute, averaged across all voyages recorded. This went up to around 82 beats per minute when he arrived on the ship that he had to navigate. This increase cannot be the result of the physical conditions on board because the situation on board the pilot station is often less favourable than the situation on board the ships. This is attributed to arousal.

An increase in heart-rate was expected during periods of high workload, such as entering a dock, bridge, or just before berthing. Such an event can be seen in figure 7.4 at the rightmost marker, at T=8000s. The heart-rate will often rise to around 110 beats per minute. This was not rated as stress by the pilot on the TLX-scale, which is discussed in a later section. However such increases were found in all voyages during docking (see average in table 7.1). A similar increase in heart-rate can be seen when the ship does

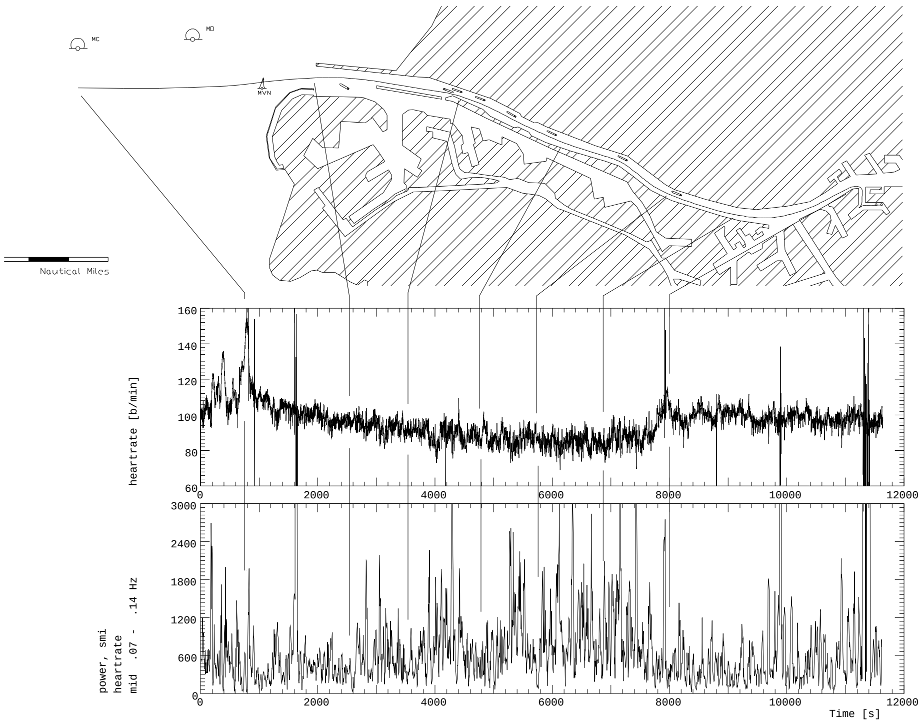

Figure 7.4: *Physiological data during piloting.*

not respond as expected, and the situation becomes critical. This can be the result of malfunctioning equipment, but also unanticipated situations may lead to such an increase. This effect was attributed to arousal required for being extra alert in a critical phase during the voyage, which is directly related to hormonal effects (Gellatly & Meyer, 1992).

Another example of high workload related to docking is to be found in figure 7.5. One can see the effect of a large ship (198m) being brought through a bridge and being moored backwards with the use of two tugs. From top to bottom one can see the voyage, the heart-rate, and the HRV. The mark at T=9150s marks the point where the ship touches the quay. The voyage starts as normal, with increasing heart-rate when the pilot climbs to the bridge of the pilot-station for a briefing (T=150s), an increase when the pilot boards the speed boat (T=500s), and when the pilot comes on board and climbs to the bridge (marked at T=700s). It can also be seen clearly that the base heart-rate for the pilot increases during navigation compared to the situation on the pilot station: 73 beats/minute at T=300s compared to 85 beats/minute at T=1500-5000s. This profile is very similar to figure 7.4 until the bridge at the entrance of the dock is approached (T=5600s). At this point there is a sharp rise in the heart-rate reaching a maximum in the opening of the bridge. After that a large number of arrhythmia in the ECG can be seen. This is a physiological phenomenon that was found with many recorded voyages during the manoeuvring phase in the dock, especially after such a high level of arousal. An explanation of this phenomenon is outside the scope of this study. The reduced HRV during the periods of increased heart-rate is normal, because HRV is inversely related to the mean heart-rate. The absolute values of HRV cannot be compared directly for this reason. The very high peaks in the HRV resulting from the arrhythmia have no significance.

7.5.2 Correlation functions

The hypothesis for the correlation functions is that there is causal relationship between the TTB pattern of the voyage and the pilot's HRV measured during that voyage. In addition must the correlation function show that the HRV pattern leads the TTB pattern by a significant distance in order to compensate for the large inertia of ships. For this the correlation must be sufficiently high and the correlation function must have one clear maximum at $\Delta x < 0$.

For this analysis six locations were selected that were predicted to show clear workload peaks at well defined places, and where there were a number of comparable passages. The minimal number used was four passages. At such points it should be possible to check whether the workload is predicted by the MTB parameter from the alarm-clock model. The locations are (see figure 7.6):

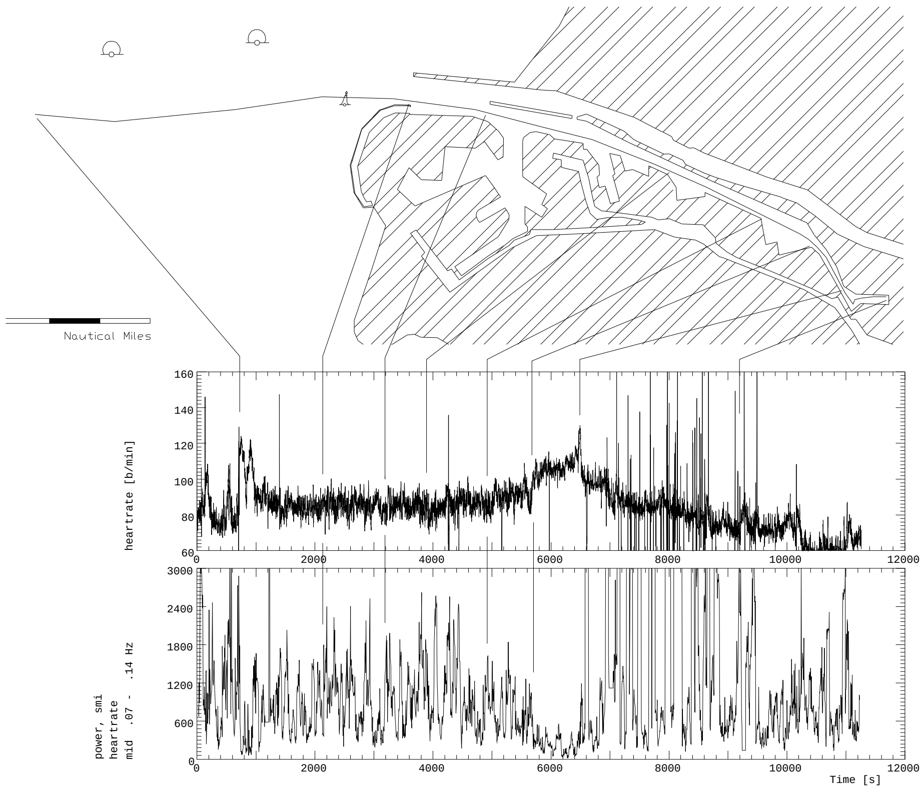

Figure 7.5: *HRV-profiles of another voyage.*

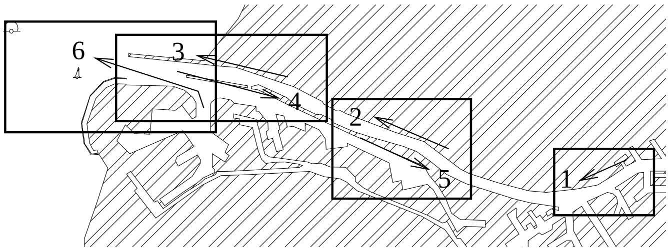

Figure 7.6: *Map with the locations of the analysed areas.*

1. Vulkaanhaven – Oude Maas
2. Maassluis – BMK[6]
3. BMK – Maasmond
4. Maasmond – BMK
5. BMK – Maassluis
6. Beerkanaal – Maasmond

The first three section are going downstream from Rotterdam to sea. In the next two sections they are going upstream with section 4 matching 3 and 5 matching 2. Finally a short downstream section from the Europort is analysed. In all the figures which follow in this section of the thesis three figures or graphs are shown:

1. A map of the section with some characteristic points marked. The dotted line is an indication of the course used for calculating the MTB.
2. The MTB graph for that section, showing the relationship between the location and the distance to the boundary of the navigable area considering the current heading. This is also the expected HRV profile, not considering the expected shift. A low MBT coincides with a high workload. The horizontal scale shows the location on the river in km[7].
3. The average correlation function calculated from the individual correlation functions between the MTB shown in the graph above and the HRV profiles measured. The horizontal scale is the shift between the HRV and the MBT given in this study in distance, since all data are analysed in the place and not time dimensions (see also section 7.4.1). When the maximum value of the correlation function is found on the left of the origin the HRV signal leads the MBT signal; this was the hypothesis. Together with the average correlation function a 95% confidence interval for this average is calculated.

[6]The BMK is an abbreviation for the high-water barrier in the Nieuwe Waterweg.
[7]One has to add 1000km to obtain the official location on the river

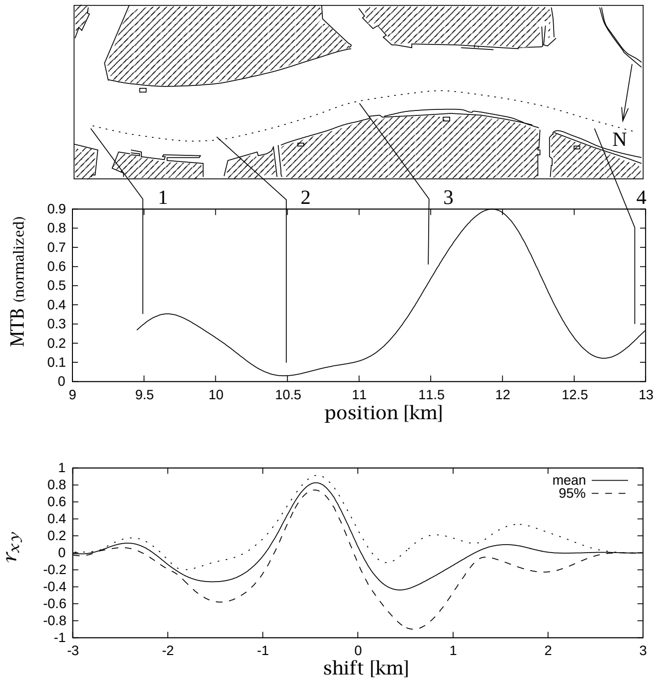

Figure 7.7: *Analysis from Vulkaanhaven to Oude Maas.*

1. Vulkaanhaven – Oude Maas

This section is a rather small, about 3km, near Vlaardingen (figure 7.7), and close to the city of Rotterdam. Large ships do not go more than 3km further upstream. When considering the fairway, after a short straight stretch of fairway the river first bends rather sharply to port (1,2) and then back to starboard (3). In the middle of this bend is a sector border for the VTS. At the end of the bend is the inlet of the river Oude Maas(4). The MTB's decrease due to the bend and at the inlet of the river Oude Maas. This part of the voyage takes almost 15 minutes, and is analysed for four voyages. The deep dip at mark 2 is near the border of the radar sectors, and the high workload remain during the entire bend. The final predicted workload peak corresponding to a MTB dipappears when approaching the river Oude Maas(4). The maximum correlation between MTB and HRV is $r_{xy} = 0.83$ at $\Delta x = -0.55$.

2. Maassluis – BMK

A few kilometers downstream is the second section analysed, from the city of Maassluis to the high-water barrier BMK (figure 7.8). In this section there are also two bends but they are much more spread and they are therefore negotiated very differently. The first bend is to starboard (1) and the second

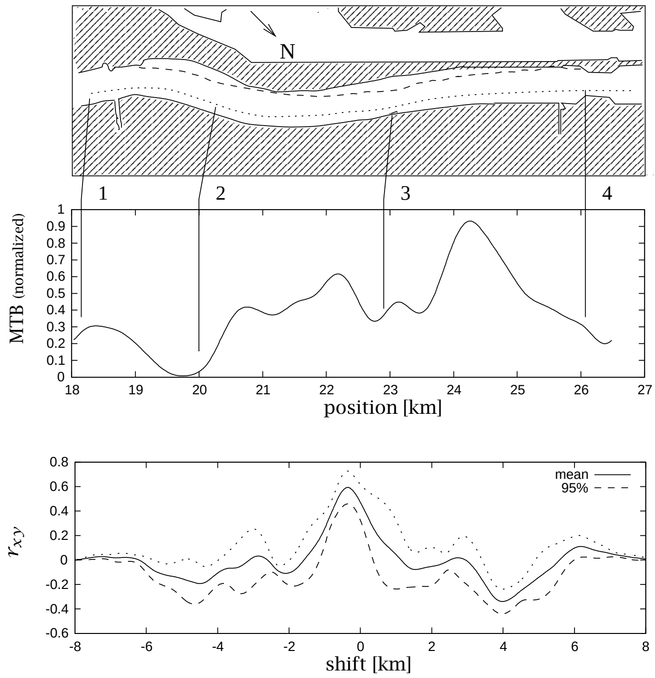

Figure 7.8: *Analysis from Maassluis to BMK.*

to port (2), and then there is a straight section of river to the BMK. About three kilometers before the BMK is the radar sector change (3), and here the pilots are told what the situation at the BMK is, and what traffic is expected. If two ships cannot pass at the same time it must be decided who must wait. This depends on current, wind, and manoeuvrability, and is decided by the pilot and the captain. The MTB dip due to the BMK(4) is very clear. The maximum correlation between MTB and HRV is $r_{xy} = 0.63$ at $\Delta x = -0.46$.

3. BMK – Maasmond

The last part of the voyage is from the BMK to sea (figure 7.9). From the BMK(1) the ship has to pass the Breeddiep, report at the VTS-radar, pass the lower light of the 112°light-line, pass the entrance to open sea(4), and navigate the vessel until disembarkment, about 3 n mile at sea. In the MTB profile, shown in figure 7.9, the dip of the BMK can be seen very clearly. The analysis starts on entering the BMK. An MTB minimum can be found at the bend near the Breeddiep(2). Shortly after that the ship has to go to port in two major stages (km 30 and 31). At the lower light (3) and shortly after two minor course corrections are required. When arriving at open water the MTB decreases due to the current and waves taking effect. The maximum correlation between MTB and HRV is $r_{xy} = 0.65$ at $\Delta x = -0.58$.

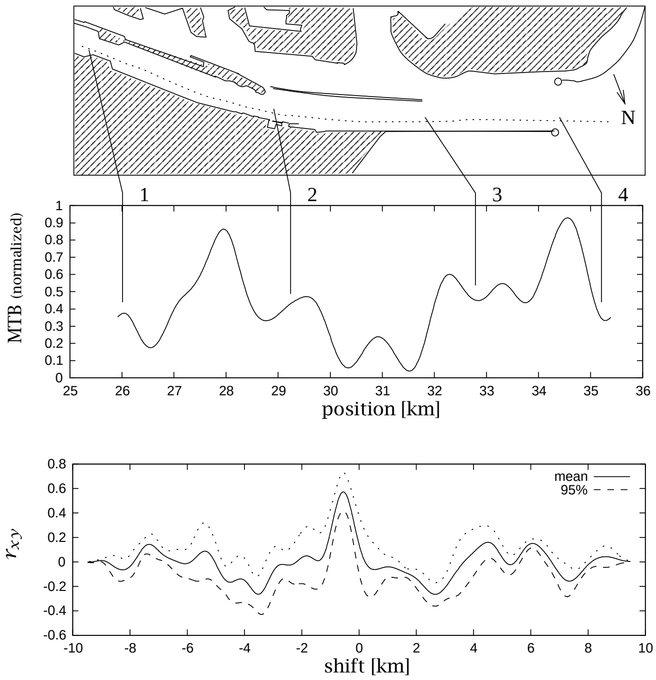

Figure 7.9: *Analysis from BMK to sea.*

4. Maasmond – BMK

Maasmond is the entrance of the Nieuwe Waterweg(1). Ships come from the south of the fairway and must cross to the north side of the fairway to go to Rotterdam. Near the Berghaven, at the Breeddiep(3), there is a bend in the river, and from there it is a straight line to the BMK(4). The voyage takes about half an hour. At the time of this project the river-bed was being prepared to accommodate the high-water barrier, and because of this the southern half of the river was closed. This resulted in a fairly narrow opening of about 150 m, compared to the normal width of the river of about 350 m at this point. Because of this layout five major MTB minima were found: one at entering the river (1), one shortly after entering the river near the lower light of the 112°light-line (2), one near the Breediep (3), a dip predicting a peak in workload at about five minutes before entering the BMK(km 28), and one near the BMK (4). The maximum correlation between MTB and HRV is $r_{xy} = 0.54$ at $\Delta x = -0.49$.

5. BMK – Maassluis

The section between the BMK(1) and Maassluis(5) is generally considered to be quiet by the pilots (figure 7.11). The river is wide and there is relatively little traffic. About 1.5 n mile upstream from the BMK is the 'Kruitsteiger'(2)

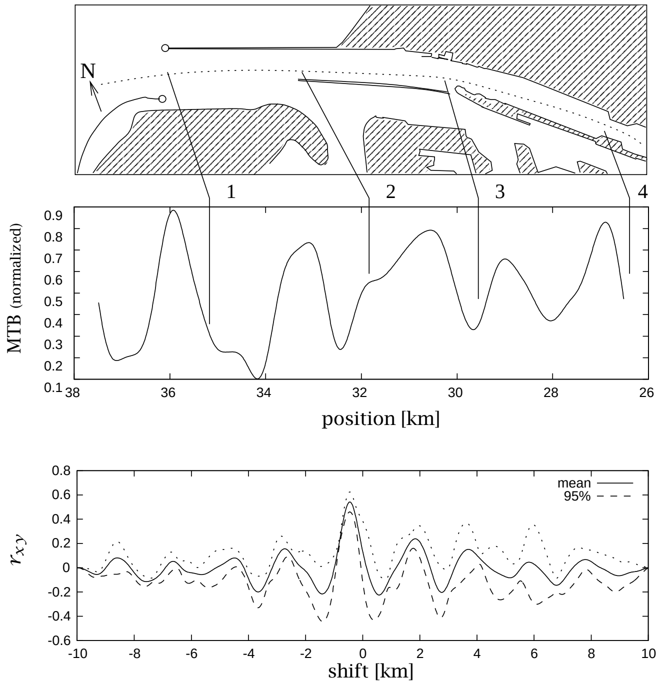

Figure 7.10: *Analysis from Maasmond to BMK.*

which is the border between two radar sectors. Further upstream the river first bends slightly to starboard (3), and bends to port to arrive at Maassluis (4,5). This section of the voyage also takes about half an hour. The curves in the river shape the MTB curve. The current was a complicating element in this section. The tide has a large influence on the current over here. When the current is at its peak it will push the ship to the outside of the river, requiring the pilot to pay attention. The maximum correlation between MTB and HRV is $r_{xy} = 0.52$ at $\Delta x = -0.50$.

6. Beerkanaal – Maasmond

The profiles above are all made for the river. There is also a rather new harbour area in Rotterdam, Europort, that is fit to accommodate large ships. This area is also close to sea, requiring much shorter pilot times than needed for the river. Only one section provided enough comparable voyages to convert into a profile, and that is ships coming from the Europahaven and the Mississippihaven (figure 7.12). They can be followed from the Beerkanaal, into the Calandkanaal, Maasmond, to sea. The analysis starts at the buoys B3 – B4(1). The second mark is placed at the lower light of the 112°light-line(2). Here is also the VTS-radar reporting point. The third mark is the entrance of the river(3). The MTB minima are located shortly after the B3

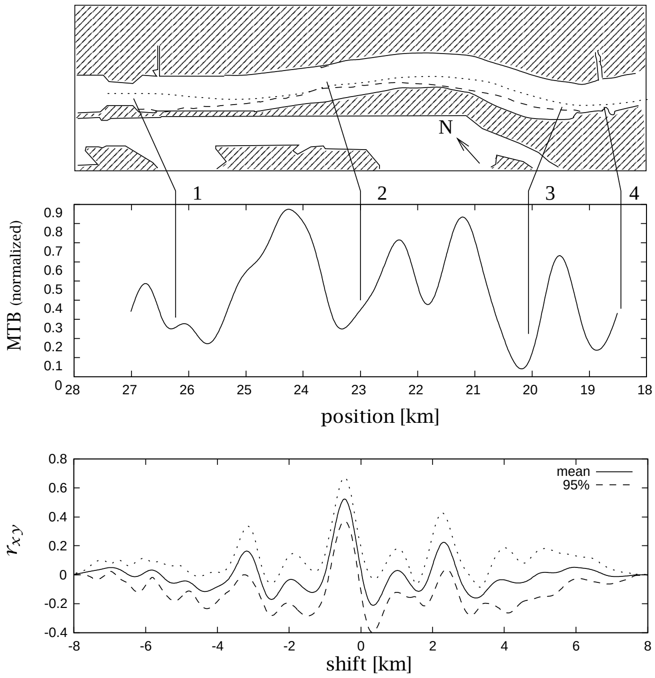

Figure 7.11: *Analysis from BMK to Maassluis.*

– B4 line, before the 'lower light', shortly before the sea where they have to change course. The maximum correlation between MTB and HRV is $r_{xy} = 0.73$ at $\Delta x = -0.41$.

7.5.3 Traffic-related workload

In this section some effects of traffic on the pilot's mental workload are analysed. For this the encounters with other ships were extracted from the HRV data. These HRV signals are then compared to each other. The hypothesis is that the workload will increase sharply when passing another ship as a result of a reduced TTB. A TTB criterion could not be reconstructed because the equipment did not allow accurate recording of the entire situation. Therefore these HRV signals are not analysed using correlation functions but they are inspected visually to discover whether clear indications can be found in the graphs of a peak in workload at a predicted moment in the encounter.

Head-on

Head-on encounters happen relatively frequent. With an average of four ships an hour entering and leaving the harbour head-on encounters are almost unavoidable. Each ship is compelled to keep to the starboard side of

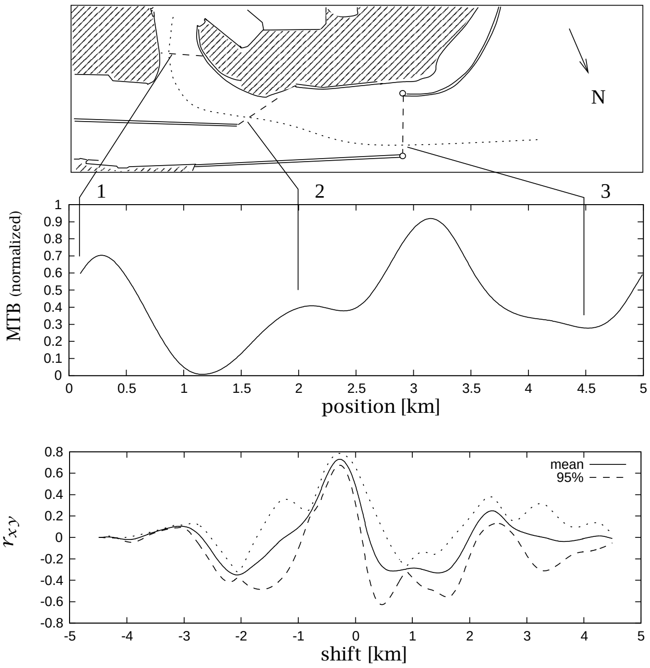

Figure 7.12: *Analysis from Beerkanaal to sea.*

the fairway, so conflicts are not very likely. Because the fairway is generally very wide compared to the beam of the ship, interaction will be weak, i.e. the chance of a collision is low. A total of 46 head-on encounters involving three pilots were used for the analysis. Three encounters were rejected because there was a strong interaction with the fairway layout which required control actions. The encounters were marked at the point at which the bow of the other ship was abeam. The hypothesis is that the TTB (and HRV) gradually decreases and reaches a minimum, i.e. the workload is maximum, shortly before the bows meet. After the bows passed each other there will be no relationship between the HRV and the position of the other ship. Figure 7.13 shows the graphs for the three pilots, each averaged over all their encounters. In all three there is a clear dip in the HRV just before the bows meet at $t = 0$. The peaks outside a small area around $t = 0$ have no relation to the encounter.

Overtaking

In overtaking two different situations may be distinguished: overtaking and being overtaken. The overtaking ship must stay free of the ship that is being overtaken. The ship being overtaken must keep course and let the overtaking take place quickly. During the experiment six overtaking situations occurred, four in which the ship was overtaken and two where the

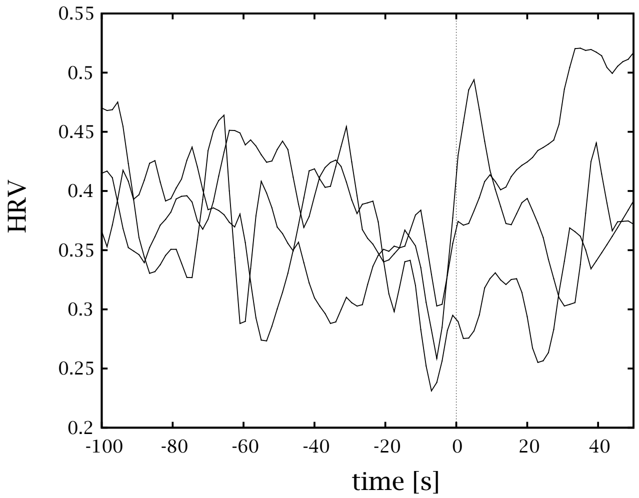

Figure 7.13: *These are the HRV profiles from head-on encounters for three pilots. The mark is placed where the bow was abeam. The analysis is done from 100 s before until 50 s after the encounter (n=18, 25, 3).*

ship overtook. The manoeuvres were marked at the point the two ships were directly alongside each other. Information on speed was not available. The hypothesis is that the TTB and HRV reach a minimum when the other ship is almost in front of the own ship. That is when the movements of the overtaking ship are critical and hydrodynamic forces become effective and hence the workload is high. In figure 7.14 the resulting HRV for four ships can be seen. Because the number of observations is very limited and only one point is marked, no averages were calculated. The picture from the graph is not very clear, but the first increase of workload appears to occur when the overtaking ship arrives more or less in front of the other ship, visible in the graph at around $t = 100$s. A second but much less clear increase in workload is when the overtaking ship is ready to come to its original position in the fairway, visible in the graph at around $t = 200$s.

When the ship on which the pilot was being studied carried out the overtaking manoeuvre itself the situation is rather different. In this situation two minima in the HRV are expected. The first minimum is when the other ship is just in front of the own ship. A second minimum is expected when the own ship turns in order to return to the original course-line. Overtaking happened only twice during 25 voyages.

The results are shown in figure 7.15. The first expected minimum is found at $t = -25$s. The second HRV minimum for returning to its original course can be seen in the HRV-pattern at $t = 100$s.

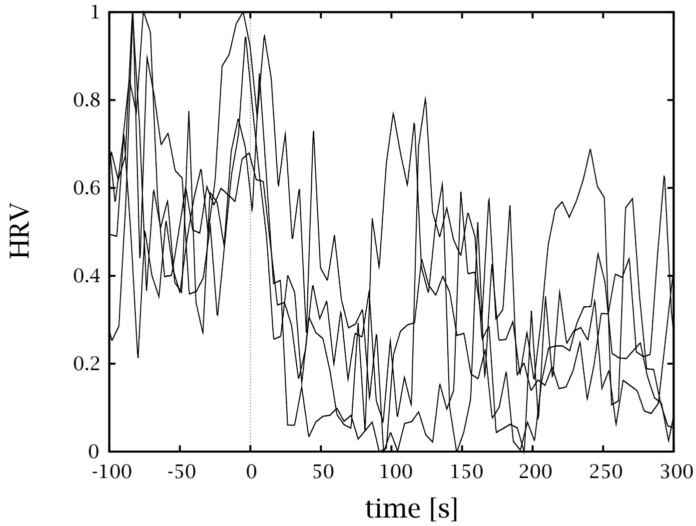

Figure 7.14: *The resulting HRV profiles of four ships being overtaken by another ship. The graph starts 100 s before the two ships are aside, and carries on until 300 s after that point. All HRV data is normalized.*

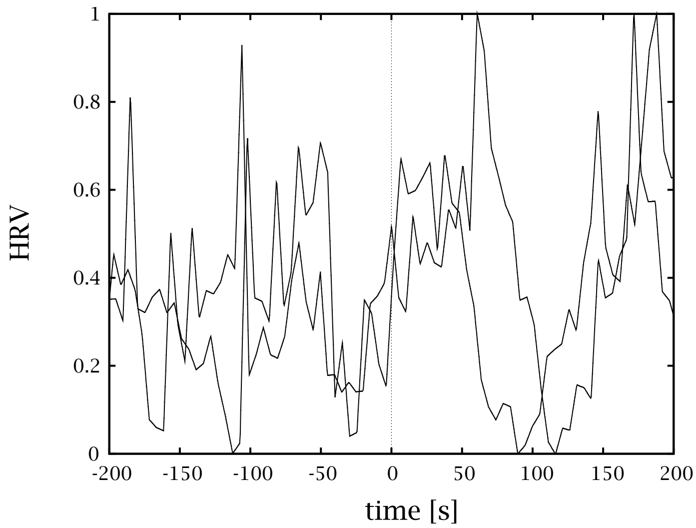

Figure 7.15: *The resulting HRV profile of a ship overtaking another ship. The graph starts 200 s seconds before the ships are side by side, and carries on until 200 s after that point. The two HRV profiles are normalized.*

Crossing

Crossing is a rather rare event on canals and rivers. This is an inherent result of the properties of such waters. During the 25 voyages a crossing situation occurred twice, leading to a temporarily high workload. These events happened at distances of about 1 n mile, making a clear analysis difficult. A much more common situation is a ‘green on green’ passage, where at least one of the ships sails on the port side of the fairway. This is common practice on rivers, when one sails with limited speed compared to the current and there are rather strong bends. In the harbour area ‘green on green’ is only accepted in special situations, such as just before entering a dock. This manoeuvre generally results in high workload. Because crossing is much more complex than head-on or overtaking and therefore marks cannot be placed easily, no graphs are produced.

7.5.4 Control-related workload

The normal mode of operation is that the pilot orders headings for the helmsman: this is known as heading control. Only when headings are not convenient will the pilot use rudder orders: this is known as rudder control. Even when using rudder control the pilot will use part of the helmsman’s skills by letting him stabilize the course (‘steady’). In situations where rudder-control is appropriate the heart-rate increases significantly as can be seen in figure 7.4 at $t = 8000$s, in figure 7.5 at $t = 6000$s (table 7.1). The average increase was from 82 to 108 beats per minute. The HRV values have no real meaning in these situations. Arousal is high due to the high task-demand load, and so is the mental workload.

7.5.5 View analysis

The pilot uses the visual scenery to observe the ship’s course and heading and any changes in these. The visual scenery is more accurate and sensitive than instruments such as a compass. It is expected that the pilot looks at the environment at regular intervals and that these intervals are related to the smallest time-constant of the ship with respect to manoeuvring and hence the required accuracy. This sampling is not so much related to prediction as to control. The fastest control on the ship is the rudder which takes about 10-20s to come to a full angle. It is therefore expected that the visual sampling period will be in this order of magnitude. The view time must be long enough to determine a rate of turn, and is expected to be a few seconds.

From the recorded voyages nine were found to be useful for ‘view-analysis’ where the light conditions were good and the pilot was in full view of the camera most of the time. It was recorded when and for how long the pilot observed the environment in front of the ship. The starting time and end of each observation period was logged using a computer-based logging-

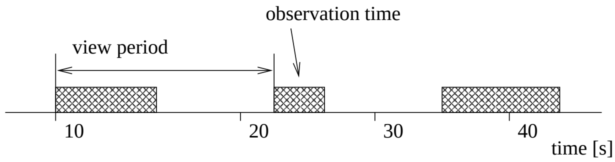

Figure 7.16: *View period and view fraction.*

Table 7.2: *The view period (in seconds) by three pilots on 9 voyages. For each pilot the mean and the 10% and 90% percentiles are determined.*

pilot	mean	10%	90%
a	9.4	2.6	18.4
b	10.6	2.8	21.2
c	9.6	2.6	19.8

system. The accuracy depended on the reaction time of the observer, and was found to be better than 0.3 seconds. The view fraction and view period was determined (see figure 7.16). The view period is the time between consecutive starts of an observation. These view periods have been calculated and the results are shown in table 7.2.

The view fraction is the time relative to the view period and is shown in table 7.3. The distribution is a flattened normal-distribution (S.E. kurtosis $\simeq 0.15$), slightly skewed to the left (S.E. skewness $\simeq 0.08$).

The results are consistent with the hypothesis. With a TTB in the order of a few minutes and a rudder time in the order of 20s the view period can be attributed to the heading control of the ship.

7.5.6 TLX-ratings

The TLX-rating were recorded to see if the subjective measures reflected the physiological measures. With respect to the different phases of the voyage (sea, river, and dock) it was expected that there would be significant differences between successive TLX-ratings within one voyage with dock being

Table 7.3: *The view fraction by three pilots on 9 voyages. For each pilot the mean and the 10% and 90% percentiles were determined.*

pilot	mean	10%	90%
a	0.48	0.13	0.84
b	0.39	0.07	0.79
c	0.42	0.10	0.79

clearly higher than river or sea. It is expected that each pilot has a different attitude towards their mental workload and that because of these differences there will be a significant difference between pilots in their average TLX-rating. Finally, the differences between ships should be reflected in their associated subjective mental workload, i.e. a significant spread within one pilot.

The TLX-ratings were obtained on twelve voyages. Before the experiment it was decided that the distinct sections of a voyage would be sea, river – canal, and dock. The pilots were instructed on how to use the TLX-scales, and what the scales meant. The explanation of the scales was printed on the back of the rating scales. The results for the three pilots can be seen in table 7.4.[8] The heart-rate and HRV measures show a very significant increase of arousal and workload in the harbour situation. This increase in subjective workload present in the TLX rating across voyages is small but significant: $F_{(36,2)} = 5.51, p < 0.02$, giving support to the hypothesis that TLX data can give a coarse measure of workload.. The differences of TLX ratings between three pilots the across voyages is very significant: $F_{(36,2)} = 103.6, p < 0.001$. However, the differences of the TLX ratings between voyages is marginal: $F_{(36,9)} = 2.54, p < 0.05$.

From the remarks made by the pilots when marking the scales it was concluded that the rating they gave was largely an indication of the workload experienced on that trip compared to the workload range they had experienced in their career as a pilot in that situation. This means that when high workloads are common in a situation, and nothing special happens, the subjective workload is rated as low. This means that the TLX data provides different type of information from the physiological data.

Part of the TLX rating is a ranking of the six sub-scales, providing information about how important the different elements are for that specific task. The order of the scales in this navigation task was (with their average weight): performance (4.7), mental demand (4.1), effort (2.2), temporal demand (1.7), physical demand (1.4), and frustration (0.8). This confirms the descriptions to be found in the introduction on pilots that navigation is largely a mental task, that the pilot depends largely on skills, which loads on the performance scale, and that the pilot is autonomous in his work, which is shown by the low weight on the temporal scale and the frustration scale (see also appendix B).

7.6 Discussion

HRV and workload

This experiment provided a series of results. The element that was considered most important was the relationship between the TTB that was the

[8] Different letters from the previous tables were chosen to be able to guarantee anonymity.

Table 7.4: *The results of the TLX rating provided by the pilots on a 12 point scale. The workload was rated for three parts of the voyage: sea, river – canal, and dock.*

pilot	sea	river	dock
α	2.0	3.0	5.5
	4.0	3.3	*[a]
	2.5	4.5	4
	2.4	3.4	4.2
	3.6	3.9	3.9
β	1.0	1.0	1.0
	2.3	2.8	2.1
	1.0	1.1	3.3
	1.0	1.1	1.7
	1.0	1.3	1.7
γ	7.1	6.6	7.5
	7.5	7.1	7.6

[a] No docking phase

result of the fairway layout and the mental workload. The correlation functions between HRV and MTB were largely as predicted by the theory. The correlation functions show one clear maximum with a significant positive correlation in each of the six sections covered in the different parts of the voyages. The maximum correlation levels found were: 0.83, 0.63, 0.65, 0.54, 0.52, and 0.73. The correlation functions of the voyages going up the river are not as well defined as those going towards sea. The reason for this is unknown. One would expect them to be very similar. It may be that the current has a significant effect or that the planning of the arrival disturbs the correlation function somewhat.

These correlation functions had a maximum on the negative side of $\Delta x = 0$ for all location, meaning that the minimum level of HRV was reached before the TTB had reached its minimum, i.e. the workload precedes the critical moment. The displacements at which the HRV and TTB signal reached their maximum correlations were: -0.55, -0.46, -0.58, -0.49, -0.50, and -0.41km. This suggests that the pilot looks ahead of the current situation for a distance of some 500m and initiates control actions well before they become a problem. There was not enough data to establish a relationship between the individual pilot and this 'looking-ahead' shift, between the physical location and shift, or between ship type and shift. This could be the subject of further study.

The mental workload at sea was higher than expected. The reason might be that the effect of currents is stronger than expected and the quality of the position fix is rather poor. At sea an increment of workload can be seen where the currents and waves take effect. The location where different currents meet is clearly visible in the real situation for the experienced eye. This point showed a strongly increased workload, although the ship's movements were entirely compensated by the helmsman.

The relationship between mental workload and the traffic situation was very hard to analyse quantitatively. Three types of encounters were recognized: head-on, overtaking, and crossing. The HRV signals were inspected visually and they show a pattern which was predicted by the theory. The head-on situation is the most common of the three. When all goes normally the two ships do not interfere with each other in any way because the lateral distance is large compared to the breadth of the ship and to the accuracy of the heading control. The results do show a well located decrease of the HRV signal shortly before the bows of the two ships meet. This peak in workload is located at about 7s before the encounter, which is associated with approx. 70m[9], much less then the 500m associated with the fairway-TTB.

When the ship is being overtaken the situation is rather different. The distance between the ships is small compared to the breadth of the ships and the interaction between the ships is therefore considerable. This interaction is most significant when the other ships stern (with the propeller) comes near the bow of the ship. The ship starts turning because of this force and steering corrections are often required. The effect of this interaction was visible in the HRV signal. A second element in this manoeuvre is when the overtaking ship returns to her original course-line. This gave a second minimum in the HRV signal, although not as clear as the initial one.

The situation of ships crossing is the rarest of the three, and the most complex to analyse. This event requires that the pilot revises his plan. The event led to a significant increase in the mental workload but this could not be attributed to a reduced TTB: the two ships were almost a nautical mile apart when this happened. The reaction is likely to be attributed to the sudden need for alternative plans, i.e. arousal.

All of this data is consistent with the theory on monitoring and control in this thesis which therefore stands the test of this practical situation as a predictor of workload.

Heart-rate data

The heart-rate may not be a very sensitive measure for mental workload, it did provide information not available in the heart-rate variability. It was expected that in a number of situations the mental workload would increase greatly and that this would have an effect on the heart-rate. These situations were: the change from heading control to rudder control, and when tugs are used. These situations, when combined with the process of docking, do indeed show a significant increase in workload from 82 to 108 beats

[9]This is based on a speed of 10 knots, which is an average speed

per minute. Individual events where rudder control was used during navigation are also visible as heart-rate peaks. This increase was attributed to the increased number of control variables that must be kept under control, particularly in docking, the limited knowledge about the manoeuvring characteristics of the ship with large rudder angles and a high rate–of–turn, and the complexity introduced by the non linearity of a ship makes.

An effect that was not expected was the increase from normal rest to the navigation situation, in which the heart-rate increased from 69 to 82 beats per minute in a physically identical situation (standing on the bridge of a ship in a comfortably warm environment). This suggests that the pilot has a rather high level of arousal during the entire voyage and that the normal level of mental workload during the navigation task is rather high. The increase of heart-rate that resulted from climbing was normal.

TLX data

The TLX scale proved to be rather difficult to use in this experiment. The pilot had no time to mark the scales in various stages of the voyage. Only 12 of the 25 voyages were rated. The voyages that were rated were not very different from the other voyages with respect to type and size of ship, heart-rate, HRV, time of day, or other factors. The results show that pilots do recognize the difference between different stages of the voyage. The increase on heart-rate also resulted in a higher TLX-rating, but not as significant as the difference in heart-rate. The differences between pilots in TLX score was very significant, but the difference between ships was not very significant. The remarks made by the pilots indicated that ratings were largely based on their expectations and not so much on the actual perceived workload. If this is normal with highly skilled professionals who are not trained in this type of rating the outcome is very different from the objectives.

General conclusions

The navigator model consists of three levels: planning, prediction, and control. The assumption was that the control level would have little effect on workload, planning only once, after which the plan was only updated, and the majority of the workload could be attributed to prediction. This was named the alarm-clock model. This was confirmed in large part. The control level only has an effect where the pilot takes direct control of the ship by changing from heading control to rudder control. It is assumed that the skill of the pilot is so extensive that the workload that results from the heading orders is negligible compared to the two higher levels. The HRV profiles show a direct relationship between the time–to–boundary and the workload. This means that the workload is largely the result of the prediction process of the pilot. It is at this level that the decisions are taken on a minute–to–minute base. The problems are dealt with one by one. The planning was hypothesized to produce a workload only once, but this turned out not to be entirely true. The pilot has a preliminary plan when he arrives on board, and this plan has to be adjusted as soon as he arrives on the bridge. But after this he is informed by the VTS radar station about the situation in the sector ahead of him. This event showed up clearly in the HRV profiles. Whether this increased workload results from reporting or from the new information received and its incorporation into the plan is unknown. Another effect on planning might the information received by radio. The pilot was very sensitive to hearing his ship's name on the radio, and on hearing this name the workload increased significantly. It is unknown if this must be attributed to arousal, an effect that is known from all sorts of research, or to incorporating the information to the plan.

Some cautionary statements about the adequacy of the data used in this part of the study are in order. The number of four pilots is very low for statistical purposes. In addition to this the voyages were very different in many respects, making the comparison between ships and pilots almost impossible, On the other hand, the purpose was not to differentiate between pilots or between ships, but to build a navigator model and to validate this using a physiological mental workload measure. For this the relationship between place and workload is of importance, and for this the 21 voyages which could be analysed gave sufficient data and a clear common structure.

CHAPTER 8

Summary and conclusions

The goal of this study was to arrive at and test a monitoring model appropriate to complex control situations in the laboratory, but particularly in real life. For this a theory was developed based on a group of monitoring theories known as conditional sampling models. Important criteria for the development of the new theory were that it was based on real operator behaviour involved in a control task, it could deal with multiple variables, it could explain the effects of informational 'overload' and 'underload', and that it could be easily applied in a real situation. Based on these requirements the alarm–clock–model was developed, based on the time–to–boundary concept.

The time–to–boundary theory is a conditional sampling theory, based on the idea that sampling is largely based on the time left for intervention at any time during the control process. When the time left for intervention is short, or the confidence about the state of the process is low, sampling rates are high and these sampling rates decrease with the increase of time left for intervention or with increased confidence. Sampling forms the main workload component in such tasks when they are highly practiced, i.e. they are skill based. The model has four components: prediction, planning, resampling, and implementation of control actions.

The theory includes a number of characteristics. The sampling interval depends fully on the quality of the prediction and the estimated time left until a system variable leaves the boundaries planned (time–to–boundary, TTB). Both poor predictions and short TTB's lead to high sampling frequencies. The model includes the assumption that the operator can only pay attention to one (complex) variable at a time. Variables are sampled and dealt with one after another, i.e. the minimum of the TTB's determines the selection. As a consequence performance will degrade when several variables need intervention at the same time. In such situations it is likely that one or more of these variables is ignored temporarily. The operator predicts the future state of all variables under control based on his knowledge of these variables. The assumption made in this study was that these predictions are simple extrapolations of the current state.

Monitoring behaviour was studied in two situations: the maritime pilots and the DURESS process-simulator. Because recording visual sampling behaviour is difficult, especially in the maritime situation, a recording technique was chosen that did not depend on recording sampling behaviour directly. Instead, heart-rate variability measures were used as a measure for mental workload which in turn was used as a measure for monitoring and prediction.

Results

Taking the results of the DURESS experiment and the maritime-pilot experiment together gives very strong support to the theory on the relationship between monitoring and time–to–boundary in a complex control task. This relationship was inferred from the relationship between heart-rate variability (HRV) and the minimum time–to–boundary (MTB). The correlation functions between HRV and MTB are very satisfactory and strongly suggest a causal relationship. In the DURESS experiment the maximum mean correlation between HRV and MTB was 0.72 (the maximum of the correlation function $r_{xy}(\tau)$, averaged over all trials). In the pilot experiment, with a much more complex task, the mean correlations for the different locations used in the predictions were: 0.83, 0.63, 0.65, 0.54, 0.52, and 0.73 (the maxima of the correlation functions $r_{xy}(\Delta x)$, averaged over voyages). The inferred causality was based on the fact that all correlation functions have a well defined maximum. This maximum averaged around $\tau = 0$ for the DURESS experiment and around $\Delta x = -500$m for the pilot experiment. This shift of 500m was interpreted as the look-ahead distance necessary for reliable ship control. The average around $\tau = 0$ for DURESS appeared to come from two different strategies by two groups of operators, one showing proactive control, similar to the pilots, one showing delayed, reactive control. It may be significant that the DURESS operators showing proactive control had the deepest knowledge of the simulator, making them arguably more similar to pilots.

In the process of achieving the operator goals each operator applies different tactics and changes these over time, thereby applying different skills. With any change in tactic the data must be treated differently. This was very clear in the DURESS experiment where three situations were recognized: startup, control, and finish. In the startup phase the activity rate was high, the rate of change of all process variables was high, and the mental workload was low. This was interpreted as showing that in such situations control is based on patterns of well trained actions (skills). In this phase the time–to–boundary theory did not apply. In the control phase the TTB-theory did apply. However, when the TTB went above a threshold (and therefore the sampling rate dropped below a threshold) the relationship between the process state and mental workload or sampling was lost because operators started to focus on other elements than the process. The time at which this happened was about 200s in this experiment. This phenomenon of changing focus to elements outside the task at hand is not uncommon in most real tasks and can be seen as an ‘underloading’ phenomenon. The last phase is the finish, where control was stopped as soon as the operator decided that no further control was needed to finish the trial correctly. Since there is then effectively no boundary within which to stay, and the goal has been reached no relation with these is to be expected.

The maritime pilots showed two modes of operation: heading control and rudder-thrust control. With heading control it was possible to predict HRV based on MTB. This confirms the TTB-theory. In the situation of rudder-thrust control the HRV could not be adequately predicted. In this situation the HRV was low (i.e. workload was high), the heart-rate was increased significantly, and the TTB’s were much lower than in the situations where heading control was applied.

Navigation

In modelling the task of the pilot three levels of ship control were recognized: control, prediction, and planning. Each of these levels is required for navigation but they are not of equal importance for understanding the navigators work. The control level is well described in literature, but this level is not considered the most important part of the pilot’s model when sailing at sea or in the river. The planning is a type of task that every navigator learns at school and can largely be done behind a desk, notwithstanding the fact that changes in the plan must be made regularly due to new information that becomes available. It was found that prediction is the key task of the pilot that connects all the other task elements and the one which required the most work, both for the pilot and for the researcher, to develop an appropriate explanatory model.

Selecting an adequate boundary within which to exercise control is not simple for the operator in all situations. For boundaries that are simply laid on the operator through instructions or physical laws, the situation is not that complicated. However, in the pilot situation determining the boundary is not as simple. Ships normally sail on the starboard side of the fairway, and that boundary can be expected to be close to the embankment. On the port side however the situation is much more complicated. In general a pilot prefers to stay on the starboard side of the fairway because of other traffic (and regulations). Current and wind can, however, require that the pilot choose to sail on the port side of the fairway. How ‘hard’ the fairway boundary on the port side is depends, however, on the presence of other traffic, current and wind. In the calculation made for the workload profiles the boundary on the port side was set at 2/3 of the fairway, just over the centre-line. It is unknown how good this choice is in all circumstances in practice. Finding out more about what sets the boundaries in the mental models used by the pilots can be a point for future research. Such a study could look at any potential problems or conflicts arising out of the definitions used.

The rise of 10-15 beats per minute of the heart-rate in the navigation phase compared to rest indicates that navigation poses a considerable workload on the pilot. All recorded voyages were made under reasonable weather conditions: wind was low to moderate and visibility was good, occasionally with some mild rain. The interviews showed that the workload would increase dramatically when visibility is seriously deteriorated, such as with fog, heavy rain, or snow. Given the fact that pilots depend largely on the view from the bridge it can be expected that heart-rates will increase even more under those conditions. During docking and other complicated manoeuvres the heart-rate rose to about 110 beats per minute, indicating that such manoeuvres are associated with a high workload. It was expected that workload would be high in these situations, but this result was much higher than expected. The high rates of change combined with the narrow margins and the rather unfamiliar ship put a high strain on the pilot with respect to monitoring and control. Whereas during normal navigation the pilot samples the environment approximately every 10s and can divide his attention between navigation and other tasks, during docking he has no time left for anything but manoeuvring.

The TLX measures did not provide much additional information with respect to pilot control behaviour. Where the heart-rate recordings showed large differences between normal navigation in the river and manoeuvring the TLX measures showed marginal differences. Only when these outcomes were combined with the remarks made by the maritime pilots when rating the scales was it clear that these outcomes might reflect the perceived mental workload relative to the expectations of the pilot. This would mean that this type of task, a task which depends very much on skills and expertise built up over more that ten years, is not suitable for the TLX or that users of the TLX scale must be trained in using them in order to let the outcomes reflect workload measures that are obtained independent of the TLX. Despite these shortcomings the coarse level of difference in physiological measures was reflected in TLX ratings.

Mental workload

An important part of the results are estimates of mental workload based on the heart-rate variability. The reason that researchers are often interested in mental workload is that there might be an increased level of human failure in situations of high load. However, the relationship between high level of mental workload and human failure is not clear. Situations of high workload bear two types of risk. One is that if a situation of high workload exists for a long time humans become error prone due to exhaustion. If such a situation of high workload exists only for a short period of time there is no reason to assume that this increases the risk of a wrong decision due to exhaustion. However, there is a second type of risk related to the complexity and time pressure of that situation. As a result, because many decisions must be made in a short period of time, there is an increased risk of decision errors, and in such situations an error can very quickly lead to an accident due to the short time period between the decision and the potential accident. Such situations of high workload therefore require special attention, and this thesis is largely concerned with this risk. Based on theoretical considerations one can say that high levels of workload bear the risk that not all variables of importance remain under control. Research on how this type of problem can be dealt with can be a point for future research. The results of this research would suggest the value of an aid for the operator to predict the near future with a reference to the boundaries, i.e. a coupling between related variables to reduce the number of critical variables. Some of these idea can be found in the "perspective tunnel" displays of the latest cockpit designs in aircraft (Grunwald et al., 1981).

At the other end of the workload spectrum we have very low levels of workload. The DURESS experiment showed the dissociation between workload as measured by HRV and performance when the minimum–boundary–time (MTB) exceeded approximately three minutes. Interviews showed that the operator started to focus on other elements than the process variables. In itself it is normal that the operator focuses on something else of interest when the process needs no immediate attention. The problem is how to regain the attention of the operator reliably in such situations. The design of displays that show the operator very clearly what time–to–boundary he has left can be a good start towards an optimal operator interface. Important in this respect might be the rate at which the TTB is dropping. Operators need time to recognize the state of the system and if the rate of change is much higher than the time needed by the operator to build up his situation awareness, the risk of failure may become high. It is expected that this problem is present in many process-control situations. For pilots in harbours it is expected that this phenomenon might show up slightly different. The pilot should only familiarize himself with the navigational equipment during a phase of low workload. The captain should independently ensure an adequate task assignment and a clear task-structure. The task assignment should be such that instrument interaction by the pilot is kept to a minimum. This should make sure that the pilot only has to focus on navigation and has adequate time to obtain the required level of situation awareness.

Heart-rate variability

In this study mental workload was measured using heart-rate variability (HRV). HRV measures proved to be effective and practicable in measuring mental workload for the purpose of this thesis. At low and intermediate levels of workload they provide a sensitive measure of mental workload. At high levels of workload when heart-rate increases and various hormonal systems are activated, the HRV-method is of very little use since the high heart-rate surpresses the HRV. During changes of the heart-rate the HRV also has no meaning. An increased heart rate is only an indication that a situation of high arousal is present. Other types of analysis are needed to learn more about such situations.

It is known that many physiological heart measures, such as heart-rate variability, reflect readiness. The results can also be interpreted based on this readiness theory. The question is then: did we measure readiness or mental workload? It is my opinion that in this type of task it is impossible to distinguish between them. The shorter the time left for intervention, the higher the required sampling frequency, the higher the mental workload, and the higher the level of bodily readiness. A more difficult question is how these elements are all related in the situation where mental workload is induced by tasks such as arithmetics.

Implications for the pilots and VTS

The pilot is added to the crew of a ship in situations where the navigation is difficult and the risk of grounding and collision is high. Pilots are officers trained in sailing in confined waterways close to other ships and they are trained in manoeuvring ships at very low speeds with the help of tugs. The crews of ships, who spend weeks or even months at sea and visit harbours all over the world, often do not possess the detailed knowledge and skills to perform this task. Therefore a local expert is brought on board in order to minimize the risk of grounding or collision. The work of the pilot is to avoid situations that might be of risk. He does this by having an adequate long-term plan, and by selecting such a short term plan that the ship is always well within the 'safety margins'. For this he needs to have extensive knowledge of the changing navigable area with all the possible disturbances and to select a course and speed that matches the manoeuvring characteristics of the ship.

Pilots can be shown to be rather successful in safe navigation. Data collected by the port of Rotterdam on accidents[1] in the period 1987–1991 showed that in the research area (sea/river/canal) 7 accidents per year on average happened where at least one pilot was involved. If the docks are included the accident rate is 15 per year. The number of voyages is about 150.000 per year (inbound + outbound)(van Kuilenburg, 1995). Almost all of these accidents result in minor damage.

This study has shown that there is a direct relationship between the state of the ship (i.e. location, speed, and heading) and the workload of the pilot. Knowing the sailing plan of the ship it is even possible to predict the level of workload as a function of the location of the ship. The pilot's workload-data

[1] An accident is an event where damages must be paid

also showed that the pilot is well aware of the situation ahead: i.e. changes in current showed in the HRV before the ship reached such positions.

It is the pilot who collects all information and decides on heading and speed, or rudder setting and engine thrust. We can conclude that the pilot is fully integrated in the control of the ship and not an external advisor of the captain as defined in the law. The reason for this integration is that the control decisions are based on many time-variant variables retrieved from the environment by skill. For his information the pilot depends on the information-rich visual environment of the ship, not on instruments: instruments are not as dependable as direct observation when this is available. Each instrument shows only one or two variables, and these variables have no direct relationship to the environment. It is this relationship that is the basis for control decisions. The pilot adapts his control strategy to the characteristics of the ship and the crew to compensate also for the large differences that can be found between ships.

In this thesis it was argued that the cognitive task of the pilot is dominated by the state of the ship in relation to the environment. It is not the physical activity nor the type of control settings that dominate mental activity. Because of the large time constants in shipping, small changes in the state of the ship can have large effects. It is therefore necessary that the pilot can detect these small changes. The visual view from the bridge is the best information source for the pilot. Almost half the time the pilot inspects the fairway in front of the ship. The rest of the time is spent evenly distributed looking in all other directions, on board and around the ship. The instruments on board are used very sparsely. In periods of high workload up to 90% of the time is used to observe the fairway in front of the ship.

In this research only the voyages at sea and in the river were analysed. A very important part of the pilot's work is docking. There are two reasons that this phase received minimal attention in this study. One is that the heart-rate increases greatly during docking, making the use of heart-rate variability as a workload measure impossible. Here one can see a clear limitation of the use of HRV as a measure of mental workload. Other techniques must be used to measure workload in these situations of high arousal. A second reason is that a very detailed state logging (of the ship's position relative to the dock, the tugs, and other traffic) is needed in order to derive TTB's. The equipment was not available to do this.

The high arousal during docking is however a good indication of the high workload in this phase of the voyage. Docking is different from the task at sea and in the river. By using propeller thrust and rudder angle the ship must manoeuvre between quays and ships at a very low speed, while the wind puts a large force on the ship. The models that describe the behaviour of the ship during normal sailing are not valid in this situation, so other skills than those for normal sailing are required to control the ship during docking. During docking tugs are often used, requiring even more skills and knowledge on manoeuvring a ship in such situations. From the accident data provided earlier it can be seen that sailing in the harbour and dock accounts for half the accidents although the sailing distance is very small compared to the entire voyage. I would predict that the TTB approach used in this study could be used in studying the docking task, but a bridge simulator may be an essential tool in order to control and log all the system-state variables.

Remote pilotage

One of the initial questions that led to this study was the issue of how shore-based radar can provide better information in order to increase the safety level in the estuary and the harbour. In addition to this it is expected that remote pilotage could start to play an increasingly important role in the future because of economic reasons, if the technology could make remote pilotage possible at an acceptable level of safety and provides a better functionality than the current system. It is important, therefore, to indicate how this study contributes to an answer to these questions of safety and functionality.

Remote sensing techniques such as radar make new types of control possible. The navigator is then provided with information other than that available in the visual environment, i.e. with those variables which are measured and which are considered essential. One such type of control is remote pilotage. During remote pilotage a pilot who is ashore gives navigational instructions to the crew on board. The information about the ship's position is obtained using radar, and radio transmitters are used to communicate with the crew. Remote pilotage is allowed by law for pilotage at sea if the pilot cannot be brought on board due to poor weather conditions, if the quality of communication between the pilot and the crew is good, and if the crew is expected to be skilled enough. Considering the experimental results of this study it can be concluded that remote pilotage is very different from normal pilotage. My results have shown that the workload of the pilot is the direct result of short-term planning, and this short-term planning directly depends on the actual movements of the ship particularly in relation to the visual environment. Radar does not provide such information, therefore the short-term planning-task must be done by someone on board. The pilot behind the radar can only inform the crew about the long-term plan that they must follow, because he has only limited information which a radar provides about the ship's state (at this moment only position), because of the inaccuracy of the radar, and time delays in the scanning system. All these elements make accurate short-term planning impossible with the technology available at present. Remote pilotage therefore requires the development of a completely new system of pilotage both ashore and on board. In such a system the navigator on board must have the knowledge, information, and skills to do the short-term planning.

Radar at present may not be useful for short-term planning, but it can be very useful in long-term planning. One goal of the navigator is to avoid bringing the ship into a situation that might be a safety risk. In order to control a ship that is in a transient state the navigator must deploy attention over many variables. In such a complex situation there is often more than one critical variable that must be monitored. It is this multitude of controls that imposes the highest risk of losing full control over the situation because there is ‘not enough time’. In order to ensure safe navigation this aspect needs careful consideration. Supporting the navigator in complex situations can be realized by making sure that the ‘problems’ come one after another and not simultaneously and that they are known about as far as possible in the future —preferably whilst they are still “on the horizon”—. This is where shore-based radar-control can be of help. The radar operator can, with the use of many radar scanners and computer generated images, monitor a large area of shipping traffic. By informing the crews (and pilot) on board about the upcoming situation he enables them to take appropriate measures to avoid multiple-problem situations, i.e. update their plan. By informing the navigator on board about the potential conflict developing, the navigator can decide on what to do to avoid such a conflict. However, this radar service must take care not to interfere with the pilot’s work in situations of high workload because in those situations the pilot does not have ‘time’ to pay attention to long term information.

Other variables

In this study the relationship between the state of the ship and the pilot’s workload was studied. There are two important elements in this relationship that were not studied: pilot experience and ship manoeuvrability. Only very experienced pilots participated in this study, senior pilots. The assumption was that they possess all the necessary skills and therefore the workload is indeed a reflection of the ship’s state, not of learning and problem solving behaviour. This is not entirely correct. Each and every ship has her own specific characteristics that makes her different from all other ships, and therefore the pilot will always show some knowledge-based behaviour; not to mention that needed if there is a system failure on the ship or equipment failure on the bridge. There are also junior pilots, pilots who learn every day. The assumption is that the workload is much higher with junior pilots than with senior pilots when the ship’s state is the same, and that the workload of junior pilots is less dependent on the state of the ship. Both predictions require testing in future research.

The second element excluded in this study is the ship’s manoeuvrability. One can recognize this from the course stability and the ship’s size. Pilot’s share experiences about the handling of the ship, and they guess the handling from the appearance of the ship. Ships are categorized from “very heavy"[2] to "being light on the helm"[3]. The second dimension in manoeuvrability is the ship's size. More experienced pilots are licensed for bigger ships. Bigger ship have a smaller navigable area, have a larger inertia, are less responsive on the engine and the helm, and have more equipment and a bigger crew. The assumption is therefore that bigger ships require better predictions about the future state that can only be made by more experienced pilots, and because of the smaller safety margins and larger inertia impose a higher workload on pilots. These are subjects for future research.

[2] In Dutch: "zware bak"
[3] "stuurt lekker licht"

Appendix A

Recorded voyages

This is a list of all recorded voyages with the four Rotterdam pilots. The list has ten entries:

1. number: sequential archive number
2. date: date of the start of the voyage
3. from: the starting point of the voyage, either the name of a harbour in Rotterdam, or sea (with the exception of voyage 4, that started at the Waterweg)
4. to: similar to from
5. ship: the ship's name
6. type: type of ship, e.g. chemical tanker, bulk carrier, container ship
7. length: the lengths over all (m)
8. BRT: the size of the ship
9. draught: the draught in meters, aft and bow if different, absolute or relative
10. time: local time of the start of the voyage and disembarkment

n.	date	from	to	ship	type	length	BRT	draught	time
1	94/8/30	sea	2ePet. h.	Crystal Pearl	chem	112		7.6	12:40-15:00
2	94/8/30	Frisoh.	sea	Pacific Prosperity	ferry	140		6.6 −0.35	17:30-20:00
3	94/9/2	Chemieh.	sea	Cadmus	bulk	150		6.4	10:45-13:15
4	94/9/2	Maas2	Merweh.	Pine King	bulk	74.5		5.3 −1	15:00-16:53
5	94/9/6	NW-Vulkaanh.	sea	Stolt Kingfisher	bulk	88	1500	5.7	11:30-13:46
6/7	94/9/6	sea	Bolnes	Energy Express	supplier	59		4.7	14:40-18:40
8	94/9/7	Merweh.	sea	EWL Suriname	cont	127	5969	6.5	09:15-12:00
9	94/9/7	sea	Alexanderh.	Pol Europe	cont	165.3	15000	7.25 −0.55	14:10-16:50
10	94/9/8	Dordrecht	sea	Harstraum	chem	115		5.2	18:20-21:00
11	94/9/8	sea	Merweh.	American Reefer	fruit	144.7		6.9	22:45-01:15
12	94/9/19	Europah.	sea	Thor	cont	87	2139	4.3 +0.3	02:35-03:20
13	94/9/19	sea	Europah.	Norasia Susan	cont	189	21887	8.8	18:00-20:50
14	94/9/20	Europah.	sea	Marstal Mearsk	cont	294	49779	12.1	01:15-02:15
15	94/9/21	5ePet. h.	sea	Super Lady	chem	274.3	150000	9.9 6.5	12:20-13:50
16	94/9/21	Brittannieh.	sea	Marwan	cont	167	14865	9.1 8.8	
17	94/9/21	sea	Beneluxh.	Commodore	bulk	110	3749	5.2	19:05-20:20
18	94/9/30	Mississippieh.	sea	Lowlands Grace	bulk	270		8.58 6.52	07:35-09:40
19	94/9/30	Europah.	sea	Providence Bay	cont	292.2	50350	10.37 7.06	13:50-15:20
20	94/10/4	Margrieth.	Europah.	York Orlando	cont	99.9	3632	5.2	05:45-08:30
21	94/10/4	Europah.	sea	Rhein Merchant	cont	93	1000	4	09:30-10:30
22	94/10/5	1ePet. h.	sea	Prince Henri	chem	110	2522	4.9	04:00-6:00
23	94/10/5	sea	Europah.	Nedlloyd Holland	cont	291.2	45678	12	08:20-09:40
24	94/10/6	Beatrixh.	sea	Canopus 1	cont	98	2860	3.7	02:55-05:50
25	94/10/6	sea	Dordrecht	Nordstraum	chem	85	2898	6.8	06:35-9:25
26	94/10/8	sea	Brittannieh.	Isolde	car	198	18636	7.5 +0.1	18:50-21:20

APPENDIX B

NASA-TLX

The NASA Task Load Index is a multi dimensional rating procedure that provides an overall workload score based on a weighted average of the ratings on six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration [1]. The definitions are provided below.

The NASA Task Load Index is a two-part evaluation procedure consisting of both weights and ratings. The first requirement is for each rater to evaluate the contribution of each factor (its weight) to the workload of a specific task. These weights account for two potential sources of between-rater variability: differences in workload definition between raters within a task and differences in the sources of workload between tasks. In addition, the weights themselves provide diagnostic information about the nature of the workload imposed by the task.

There are 15 possible pair-wise comparisons of the six scales. Each pair is presented on a card. Subjects circle the member of each pair that contributed more to the workload of that task. The number of times that each factor is selected is tallied. The tallies can range from 0 (not relevant) to 5 (more important than any other factor). A different set of weights is obtained for each distinctly different task or task element upon its completion.

The second requirement is to obtain numerical ratings for each scale that reflect the magnitude of that factor in a given task. The scales are presented on a rating sheet. Subjects respond by marking each scale at the desired location. Each scale is presented as 6-cm line divided into 12 equal intervals anchored by bipolar descriptors (e.g.. High/Low).

In the situation of the Rotterdam pilots the average weights of the scales were: performance (4.7), mental demand (4.1), effort (2.2), temporal demand (1.7), physical demand (1.4), and frustration (0.8). This confirms the introduction on pilots that navigation is largely a mental task, that the pilot depends largely on skills, and that the pilot is autonomous in his work.

[1] Most of the information in this appendix is taken from the NASA TLX paper and pencil package (NASA, 1986).

Table B.1: *The NASA TLX scales*

Title	Endpoints	Descriptions
mental demand	low/high	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
physical demand	low/high	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
temporal demand	low/high	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
performance	perfect/failure	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
effort	low/high	How hard did you have to work (mentally and physically) to accomplish your level of performance?
frustration level	low/high	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Appendix C

Transcription

This is a part of one transcription of a voyage. The transcription starts when the schip is in the river and a tug is being connected to the ship. After that the ship enters its harbour of destignation for docking. The transcription stops at the point the ship is almost in place. The ship's name is American Reefer, length 144m, draught 6.9m. The tug is the Zweden. This transcript was chosen because it was a characteristic voyage and the pilot was in a 'talking mood'.

Table C.1: *Pilot transcription*

time (s)	type	text
6.536,28	ORDER	i, 090
6.609,92	SOCIAL	zit de sleper nu nog niet vast? ach dat gaat hier niet zo snel
6.613,20	TERUG	i, 090
6.619,76	ORDER	i, 085
6.623,08	TERUG	85
6.627,80	SOCIAL	bij hollanders had 'ie al drie keer vast gezeten. Hollanders halen hem op het handje naar binnen. ...Die gasten doen alles rustig op de winch. Die zullen ook nooit een hernia krijgen
6.652,32	TERUG	i, tug is vast
6.656,72	ORDER	s, Zweden, de reefer, je zit vast hoor
6.660,88	TERUG	s, ja loods, bedankt, nog even wachten en dan kunnen we aan het werk
6.683,44	ORDER	w, sector eemhaven, de american reefer, die gaat wat vroegtijdig over naar de waalhaven, die gaat een beetje aanpassen voor de merwehaven, goede wacht
6.691,00	TERUG	w, american reefer, sector eemhaven, goede wacht, uit
6.695,56	ORDER	w, sector waalhaven, een goedemorgen de american reefer, dwars van de eemhaven, midwaters, gaat aanpassen voor de merwehaven, fruitterminal
6.705,92	TERUG	w, american reefer, bij de waalhaven, afkomende binnenvaart ...(niet te verstaan)

6.729,40	ORDER	i ,make it 075
6.730,80	TERUG	i, 075
6.742,56	TERUG	s, zweden is helemaal ingespannen, is er helemaal klaar voor
6.745,52	ORDER	s, zweden is er helemaal klaar voor, bedankt
6.783,20	ORDER	i, stop engine
6.785,36	TERUG	i, stop
6.789,24	Special	zoemer telegraaf stopt
6.798,88	TERUG	i, engine stopped
6.808,52	ORDER	i, and make it 070
6.809,52	TERUG	i, 070
6.811,64	ORDER	i, if no more steering, you tell me
6.847,52	Radio	sector waalhaven, de reefer sun, die is aan het opkorten voor vertrek
6.856,32	Radio	reefer sun, sector waalhaven, de american reefer is aan het oversteken voor de merwehaven, fruitterminal, paal 3, stuurboord
6.865,72	Radio	ok, dan wacht ik nog wel even
6.907,40	ORDER	i, make it dead slow ahead
6.910,72	TERUG	i, dead slow
6.913,80	Special	zoemer stopt
6.915,76	TERUG	i, steering is bad
6.972,88	Radio	
6.976,40	ORDER	i, stop engine
6.979,68	TERUG	i, stop engine
6.980,24	Special	zoemer telegraaf stopt
6.986,20	SOCIAL	zie je, er staat zo'n 5 mijl stroom
7.006,04	SOCIAL	die andere reefer ligt daar achterin he. Je ziet z'n groene boordlichtje al
7.034,72	SOCIAL	leuk, als je nu niets doet ga je dadelijk ineens bakboord uit
7.041,64	SOCIAL	die sleepboot begint namelijk na te denken
7.047,52	VRAAGK	i, the other ship, is it sailing?
7.050,80	REAKTIE	i, no, it is singled up, and it is waiting for us captain
7.060,92	REAKTIE	i, (na controle met kijker) still alongside
7.078,04	SOCIAL	kijk, daar gaat'ie bakboord uit. heeel zachtjes. die sleepboot denkt: "nu moet het maar eens gaan gebeuren"
7.091,00	ORDER	i, port 20
7.094,44	TERUG	i, port 20
7.097,96	TERUG	i, port 20, sir
7.100,68	ORDER	i, dead slow ahead
7.105,00	TERUG	i, dead slow
7.108,44	Special	zoemer stopt
7.114,80	TERUG	i, stop
7.115,40	ORDER	i, stop engine
7.119,44	ORDER	i, midships
7.119,68	Special	zoemer stopt
7.123,36	TERUG	i, midships sir
7.192,52	ORDER	i, port 20
7.192,92	TERUG	i, port 20
7.196,88	TERUG	i, port 20, sir
7.214,28	SOCIAL	kijk, daar gaat de sleepboot naar de verkeerde kant, ...dat moet ie niet doen
7.225,40	ORDER	s, zweden, we gaan stuurboord voor de kant hoor

7.229,20	TERUG	s, oh, sorry loods, ik dacht dat we bakboord moesten
7.244,48	ORDER	i, hard to port
7.247,76	TERUG	i, hard to port
7.248,20	TERUG	i, hard to port, sir
7.255,60	SOCIAL	die heeft ook niet op z'n briefje gekeken
7.271,64	ORDER	i, midships
7.272,08	TERUG	i, midships
7.278,12	TERUG	i, midships sir
7.305,96	ORDER	i, starboard 20
7.309,16	TERUG	i, starboard 20
7.312,64	TERUG	i, starboard 20, sir
7.326,40	ORDER	i, dead slow ahead
7.327,40	TERUG	i, dead slow ahead
7.332,48	Special	zoemer stopt
7.340,68	ORDER	i, stop engine
7.340,84	TERUG	i, stop engine
7.344,96	Special	zoemer stopt
7.348,16	TERUG	i, engine stopped
7.364,36	ORDER	s, zweden, trek dat koppie maar een beetje naar de kant hoor
7.368,36	TERUG	s, zweden trekt hem een beetje naar de kant, ja hoor
7.371,36	ORDER	i, slow astern
7.372,08	TERUG	i, slow astern
7.376,76	Special	zoemer stopt
7.379,32	ORDER	s, ja, en vaar het koppie maar weer een beetje op, een beetje naar bakboord komen zweden
7.383,36	TERUG	i, slow astern
7.384,72	ORDER	i, half astern
7.385,40	TERUG	i, half astern
7.387,40	TERUG	s, (onverstaanbaar) ik kom een beetje naar bakboord
7.390,20	ORDER	s, geef hem maar een goeie ruk naar bakboord zweden

APPENDIX D

Task analysis

This task-analysis has been done for a small-sized ship, some 4000 tons, sailing short to medium distances of approximately one to four days between harbours. The crew consists of five officers and five crew-members. The officers are the captain, chief engineer, chief mate, second engineer and a second mate. The engine room is unmanned during the night, and the two mates are on a 6-on, 6-off watch; the chief mate doing the 6-12 watch and the second mate doing the 12-6 watch.

"The officer of the watch is the master's representative, and his primary responsibility at all times is the safety of the ship. He is responsible for ensuring that the planned passage is properly carried out during his watch. He must at all times comply with the 1972 International Regulations for Preventing Collisions at Sea (COLREGS) and Regulation II/1 (Basic Principles to be Observed in Keeping a Navigational Watch) of the International Convention on Standards of Training, Certification and Watchkeeping for Seafarers, 1978 (STCW Convention)." (IMO, 1990)

On approaching a harbour five phases can be distinguished: 1. open sea 2. approaching the pilot 3. arrival of the pilot 4. sea-pilot on board 5. change of pilot (not in all harbours) 6. harbour-pilot on board 7. docking

Table D.1: *Task analysis*

time (s)	type	text	
1	**Open sea**		
			All companies provide "standing orders". These orders are to be followed by anyone on watch and tell him what his duties on the bridge are. In open sea the mate-on-watch does his normal tasks. The regular
1	1	monitoring	
		1	monitor all navigational equipment and the deck
		2	monitor the ships movements
		3	monitor the distress frequencies

4 monitor the sea for distress signals
2 navigation
1 survey other ships and make decisions for avoidance
2 determine position and fix this on the chart
3 determine necessary course changes if the ship is off course
4 set and monitor the engine revolutions
5 set and monitor the course on the autopilot
6 check for approaching waypoints
7 check the weather and sea conditions
8 warn other ships when there is need
3 administration
1 log course and speed, changes in course and speed and the distance sailed
2 enter standard log entries
4 warning
1 at night, with restricted visibility and in dense traffic a look-out is ordered to the bridge
2 warn the captain in case of any irregularity, or according to his orders

Tasks of the look-out

During the night and with limited vision there will be a look-out on the bridge

5 1 survey the sea for other ships, obstacles and signals, and report

2 Approaching the pilot

Some predetermined time before the sea-pilot comes, 0.5 to 1 hour, the ship is prepared to take the pilot on board. The extra task are

1 1 call the pilot by VHF and give an accurate ETA (expected time of arrival), sailing direction and draught
2 call VTS when in range of the radar
3 alert the captain
4 order the engine-room to prepare for manoeuvring
5 order a mate to prepare the pilot ladder
6 order a mate to prepare the anchor
7 slack down when approaching the pilot boat
8 give over the watch to the captain
9 take the pilot on board

Tasks of the captain

2 Supervise all activity
1 take the command from the watchkeeper
2 keep the ship on course
3 monitor the crew on deck, the pilot coming on board, and the pilot boat
4 monitor the ship’s status

3 Arrival of the pilot

When the sea-pilot comes on board several task have to be done

Tasks of the pilot

1 Begin to navigate
1 the pilot is informed about draught, destination, speed, propulsion, the crew and the state of the ship and its equipment
2 the pilot requests a radio channel on the VHF

3 the pilot informs radar assistance about the destination and draught, and is informed about problems he may encounter
4 fill in the paperwork and hand it over to the captain

Tasks of the captain

2 Supervise all tasks
1 inform the pilot
2 order a task allocation
3 the autopilot is switched off when the waterway becomes too congested, curved or narrow and the captain or a helmsman is steering manually
4 the second steering machine is switched on
5 fill in the paperwork for the pilot

4 Sea/river-pilot on board

The function of the pilot is to assist the captain in navigation

Tasks of the sea-pilot

1 Navigate the ship
1 survey other ships and make decisions for avoidance
2 keep watch for buoys and other navigational marks
3 make decisions for course changes
4 monitor ship's movements
5 monitor the VHF
6 give commands for speed and course changes
7 give commands or advises for anchoring
8 report VTS abeam selected marks

Tasks of the captain

2 Supervise all tasks
1 monitor all consoles and the deck
2 determine the ship's position
3 survey other ships and make decision for avoidance
4 monitor ship's movements
5 monitor the VHF
6 set the engine power
7 set course on the autopilot (when active)
8 monitor the weather and sea conditions
9 give commands for anchoring
10 assign a helmsman when needed
11 monitor the whole operation

Tasks of the mate-on-watch

3 Navigational administration and assistance
1 monitor all consoles and the deck
2 monitor the weather and sea conditions
3 survey other ships and report
4 monitor ship's movements
5 logging speed, course and location
6 log time abeam navigational marks
7 on stand-by for anchoring
8 implements orders from the captain

Tasks of the helmsman

4 Control the helm

A helmsman is recruited when needed. This depends on the steering characteristics of the ship, steering position on the bridge, controllability of all the ships functions, traffic density, vision etc. The helmsman can be the mate-on-watch, the lookout or a extra helmsman

1 steer the ordered course
2 look out for obstacles in the fairway of any kind and report

5 Change of pilot

At the entrance of the harbour the pilot-ladder is prepared again, the harbour-pilot comes on board, is informed by the sea-pilot and the sea-pilot leaves the ship

Tasks of the pilots

1 Take over navigation
1 the pilot is informed on draught, destination, maximum speed and manoeuvring speed, propulsion (including the type and direction of the propellor(s), and bow- and stern-thruster), crew and the state of the ship and its equipment. There are no procedures for this
2 the pilot requests a (new) radio channel on the VHF
3 the pilot informs radar assistance on the position, destination and draught, and is informed about problems he may encounter
4 when tugs are available, tuglines are taken and the tugs are informed by the pilot

Task of the captain

2 1 the captain orders the engine-room to make bow- and stern-thruster available

6 Harbour-pilot on board

Tasks of the harbour-pilot

1 Navigate the ship
1 monitor other traffic and the distance to piers, buoys and other objects and make decisions for avoidance
2 monitor the ship's movements and manoeuvring characteristics
3 monitor the VHF
4 give commands for speed
5 give commands for course changes or rudder-angle
6 give order to tugs (when available)
7 report VTS abeam selected marks

Tasks of the captain

2 Supervise all activity
1 monitor all consoles and the deck
2 monitor other traffic and the distance to piers, buoys and other objects and make decisions to avoid collisions
3 monitor the ship's movements
4 set course or rudder-angle (no helmsman)
5 set the engine power
6 assign a helmsman when needed

Tasks of the mate-on-watch

3 Supervice the crew on the deck
1 prepare the crew for docking
2 prepare the ship on deck for docking
3 go to the bow
4 look out for obstacles and report

Tasks of the helmsman (when assigned)

4 Control the helm
1 steer the ordered course or rudder-angle

2 look out for obstacles in the fairway of any kind and report

7 Docking

Tasks of the harbour-pilot

1 Navigate the ship
- 1 monitor the distance to pier, jetties and other objects
- 2 monitor the VHF
- 3 monitor the ship's movements
- 4 give orders to the tugs (when available)
- 5 give orders to the boatmen
- 6 give commands for speed
- 7 give commands for rudder-angle
- 8 give commands for mooring
- 9 inform VTS when done

Tasks for the captain

2 Supervise all activity
- 1 monitor the distance to pier, jetties and other objects
- 2 monitor the VHF
- 3 monitor the ship's movements
- 4 set rudder-angle (no helmsman)
- 5 set the engine power
- 6 give orders to the bow and stern for mooring, including number and type of lines
- 7 order bow and stern for adjusting lines
- 8 monitor the whole operation

Tasks of the helmsman (when assigned)

3 Control the helm
- 1 set the ordered rudder-angle

Tasks for the mate-on-watch (on the bow)

4 Supervise all activity on the deck
- 1 order the crew
- 1 give distances and speeds upon request
- 2 give out the lines and control the winches
- 3 monitor the docking operation

Tasks of the boatswain (at the stern)

5 Deck assistent (stern)
- 1 give distances and speeds upon request
- 2 give out the lines and control the winches
- 3 monitor the docking operation

8 After docking the captain has the following tasks

1
- 1 inform engine-room when finished
- 2 log standard entries such as time mooring, position etc.
- 3 order the mate-on-watch to do harbour watch

Summary

People have proven to be flexible and reliable in many control tasks, such as driving and ship navigation. Much effort has been invested in automating these tasks, but so far the results have been disappointing and the problems legion. Other tasks such as plant control - where complicated systems are tightly coupled in order to produce large volumes of high quality products given very stringent production requirements - show a much higher level of automation. This automation makes possible control of these complicated systems, relieving the controller of many tasks, improving production quality, and reducing the operator's workload. In both cases, in order to match the work to the operator and to make sure that his capacity is used to the maximum, that system safety is optimal and that working conditions meet human long-term needs, extensive knowledge of operator abilities and limitations is required.

This study examines the relationship between process characteristics and monitoring behaviour in order to learn more about operator control behaviour. This was done in two situations: a relatively complex process simulator in which the operator had to perform a rather complex and realistic control task, and in the real situation of maritime pilots on board sea ships. The operator's monitoring activity was measured using mental workload measures. The hypothesis was that workload is a linear function of the time-to-contact or time-to-boundary of each of the process variables. An assumption was that the mental workload is the result of sampling and decision making and is proportional to the frequency of this cycle (sampling and decision making).

To record mental workload, recordings of human heart rate were obtained, with special emphasis on heart-rate variability. The heart rate is not a constant, but fluctuates at about 10% around the mean heart rate. It is known from the literature that an increase in mental workload coincides with a decrease of the heart-rate variability (HRV), and this decrease of HRV was used as an indication of increased workload, which in turn was an indication for the operator's monitoring behaviour.

The hypothesis was tested using a simulated process (DURESS): a simulation of a hot-water production-plant. The subjects had to produce water of a certain temperature and quantity by carefully adjusting the controls of the simulator. During the experiment the subjects' heart rate was recorded, together with all their control activities and production performance. The

results show that the time-to-boundary (TTB) approach is successful in explaining a large part of the operator's monitoring behaviour: the TTB measure correlated well with the HRV during the control phase, which confirms the theory on monitoring behaviour. This means that there is a direct relationship between the time left for the operator to intervene and his sampling frequency.

The second experiment was a similar study but now with a real task: maritime pilots doing their normal work. Four Rotterdam pilots participated in an experiment in which they were recorded on video, their heartbeat was recorded, and their voyage was logged on maps, all during their normal work. Twenty-five voyages were recorded, with a large range of types and sizes of ships and destinations.

This experiment provided a series of results. The element that was considered most important was the relationship between the TTB that was the result of the fairway layout and the mental workload. The correlation functions between HRV and MTB were largely as predicted by the theory. These correlation functions also showed that the minimum level of HRV was reached about 0.5 km before the TTB had reached its minimum or, in other words, the pilot makes a decision about 0.5 km before the situation becomes critical.

Taken together, the DURESS experiment and the maritime-pilot experiment lend very strong support to the theory on the relationship between monitoring and time-to-boundary in a complex control task. This relationship was inferred from the relationship between heart-rate variability (HRV) and the minimum time-to-boundary (MTB).

In addition to these main results, various conclusions are drawn with respect to the recording techniques used, pilotage, and shore-based radar support.

Samenvatting

De loods aan het werk: evaluatie en gebruik van een tijd–tot–contact model van mentale werkbelasting in mens–machine systemen

Mensen zijn flexibel en betrouwbaar in vele besturingstaken, zoals het besturen van auto's en het navigeren van schepen. Er is veel onderzoek gedaan naar de mogelijkheden deze taken te automatiseren, maar tot nu toe zonder veel succes door de vele problemen. Andere taken zoals in de procesindustrie, waar gecompliceerde systemen met elkaar zijn verbonden om grote volumes van hoogwaardige producten te kunnen produceren, hebben een veel hogere graad van automatisering. Deze automatisering maakt het mogelijk om de processen onder controle te houden, de kwaliteit van de productie te verhogen, en de operator te ontlasten. In beide gevallen is een uitgebreide kennis van de menselijke mogelijkheden beperkingen nodig om de taak aan de mens aan te kunnen passen. Hierdoor kan een optimale prestatie geleverd worden, en veiligheid en gezondheid gewaarborgd zijn.

In deze studie wordt het het regelgedrag van operators bestudeerd door het verband te bepalen tussen procesvariabelen en het waarneemgedrag. Dit is gedaan in twee situaties: een relatief complexe en realistische processimulator en aan boord van schepen bij het uitvoeren van de loodstaak. Het waarneemgedrag van de loods is gemeten met behulp van mentale werkbelastingsmaten. De hypothese is dat mentale werkbelasting omgekeerd evenredig is met de tijd die resteert tot het proces buiten zijn grenzen treedt (tijd tot contact, TTC). De aanname hierbij is dat de mentale werkbelasting het gevolg is van observeren en beslissen en evenredig is met de frequentie waarmee deze cyclus van observeren en beslissen wordt doorlopen.

De mentale werkbelasting is gemeten met behalp van de hartslag, in het bijzonder de hartslag variabiliteit. De hartslag is niet constant, maar fluctueert ongeveer 10% rond de gemiddelde hartslag. Uit de literatuur is bekend dat een toename van de mentale werkbelasting samengaat met een afname van deze hartslag variabiliteit, en deze afname is gebruikt als een indicatie voor een toename in de mentale belasting, die op zijn beurt een indicatie voor het waarneemgedrag van de operator is.

De hypothese is getest met behulp van de proces-simulator DURESS, een simulatie van de productie van heet water. De proefpersonen hadden als opdracht water te produceren in een bepaalde hoeveelheid en van een bepaalde temperatuur door zorgvuldig de regelaars in te stellen. Gedurende dit experiment werd de hartslag van de proefpersoon opgenomen, werden al zijn handelingen vastgelegd, en werd zijn productie-prestatie bepaald. De resultaten tonen aan dat de tijd die resteert tot het proces buiten zijn grenzen treedt een belangrijke voorspeller is voor het observatiegedrag van de operator: deze tijd correleert goed met de hartslagvariabiliteit gedurende het productieproces. Dit betekent dat er een relatie is tussen de tijd die de operator nog heeft om in te grijpen en zijn observatiegedrag.

Het tweede experiment komt voor een deel overeen met het eerste maar nu met een echte taak: het beloodsen van schepen. Vier Rotterdamse loodsen namen deel aan dit experiment waarbij zij werden opgenomen op video, hun hartslag werd vastgelegd, en de reis werd ingetekend op de kaart. Vijentwintig reizen zijn opgenomen, met een grote varieteit in type, grootte en bestemming van de schepen.

Dit experiment levert een reeks resultaten. Het belangrijkste resultaat is de relatie tussen de tijd die resteerde tot de vaargrens (TTC), die het directe gevolg is van de vorm van de waterweg, en de mentale werkbelasting. De correlatie-functies tussen de TTC en de hartslag varibiliteit zijn grotendeels als voorspeld door de theorie. Deze correlatie-functies hadden een maximum op ongeveer een halve kilometer voor er contact zou zijn, wat aangaf dat de maximale werkbelasting ruim voor het kritieke moment lag, of anders gezegd, de loods maakt op ongeveer een halve kilometer voor het kritieke punt een beslissing.

Als de resultaten van beide experimenten, DURESS en de loodsen, worden samengenomen zijn deze een sterke indicatie voor de ondersteuning van de theorie over de relatie tussen observatiegedrag en de mentale werkbelasting. De relatie was afgeleid van de relatie tussen hartslagvariabiliteit en de tijd tot contact (TTC).

Naast dit resultaat zijn er verschillende conclusies getrokken met betrekking tot de gebruikte technieken, beloodsing, en beloodsen op afstand.

Fulko van Westrenen was born in 1961 in Zeist, the Netherlands. After completing the HAVO in 1979 he went to the Polytechnic School (HTS) in Utrecht, department of Electrical Engineering. He graduated in 1983 and continued studying Electrical Engineering at Delft University of Technology from 1983 to 1985. In 1986 he went to Leiden University to read psychology. In 1990 he was awarded his Masters degree in Experimental Psychology from Prof. W.A. Wagenaar. Returning to Delft University in 1991, he began work on his Ph.D. research with the Safety Science Group, specialising in man-machine interaction in real-world tasks, with an emphasis on mental workload. A large part of the study was done in a maritime environment.