Development of a navigator model by the use of mental workload measures

Fulko van Westrenen

Safety Science Group
Department of Technology and Society
Delft University of Technology

Abstract

This paper describes a navigator model clearing the approach into a harbour. The model is a three stage decision model, containing tracking, short-term planning and long-term planning behaviour. This model is verified by measuring the mental workload of pilots whose task is to bring a ship to berth. The mental workload is obtained by ECG recording, from which heart rate and heart rate variability (the 0.1 Hz power density) were calculated. It is shown that the workload varies with the difficulty of manoeuvring the ship, and the situation ahead.

Introduction

Due to technological innovations more information has become available for the navigator. This is especially important during the approach and departure of a harbour, when the risks of grounding and collision are highest. In most harbours around the world a pilot comes on board to assist the crew with the navigation task. Normally the pilot performs the entire navigation task. If better information is to be provided, we need to know what information is required at what time and place. Very few attempts have been made to do a thorough task analysis (Mara, 1968). Navigation is a difficult process to describe. Most information needed is obtained from the visual scene, few orders are given, and performance is hard to quantify. One of the few variables that can be quantified is mental workload. This provides an opportunity to measure the effort during decision making.

In process control all relevant parameters must be known. But when a pilot comes on board he has very little knowledge about the specific characteristics of the ship. A number of characteristics is available on a manoeuvring sheet, but since this information is only valid at full sea, the pilot will not even look at it. Personal observations and interviews have learned that the pilot only knows the basic dimensions of the ship, and he estimates characteristics by simple rules of the thumb and experience. Yet he will bring the ship to berth with a very high success rate.

Task analysis

In task analysis decomposition methods are best known. They start from the goals and narrows down towards detailed operations. From this an analysis can be done of all necessary information, controls, training, workload, allocation, etc. (Kirwan, 1992). A key factor in this approach is the set of operations. Unfortunately in shipping, and especially in piloting ships, there are very few operations. For operations all the system components must be known beforehand. In piloting this information is not available: all ships and crews are different and the environment is highly variable, making a unique operation for each instance necessary. All other decomposition methods suffer from the same drawback. A different approach is needed.

What is known, are the goals. Most other elements are not known beforehand, with the exception of the topography of the harbour with its infrastructure. To describe the task of the pilot, a goal directed approach is chosen to construct a pilot model. The centre of focus is what is what must be done, not how it is done.

Navigator Model

Many navigator models have been made over the years. Most of them are based on open-sea manoeuvring. Few of them are actually focused on inland navigation, and these often take a control theoretic approach (Papenhuijzen, 1994). The navigator model presented here is a decision model with three levels of control (see figure 1).

On the top level the overall planning of the voyage, long-term planning (LTP). Decisions on this level range from very simple decisions such as a free berth and the availability of tugs, to very complicated decisions about an overall plan to sail the ship to its berth. Most planning is done before the pilot boards the ship, and the pilot will check on board if everything is in order. Characteristic of long-term planning is that procedural and it is done once.

The middle level is short-term planning (STP). When a pilot brings in a ship, he will not plan the track as a whole in advance: he will sail the ship from waypoint to waypoint. Each waypoint will have its own objective: a new course, a curve, a change in current, an intersection. In STP the decision is made about a situation at hand. The controls which the pilot has available are very limited: rudder and engine. The decision horizon depends largely on the manoeuvrability of the ship in combination with the environment. Due to the inherent inertia of the ship, decisions often has the form of a point of no return, like in during the take-off sequence of an airplane. The choice made on this level is which track to follow. Short-term planning is based on knowledge about the forthcoming local situation, traffic, the training of the pilot, standard manoeuvres, and his personal style. This, and the manoeuvring characteristics of the ship, will be applied to plan a track. The manoeuvring characteristics and the required accuracy determine what type of control he will apply: speed-course (V,φ) or thrust-rudderangle (F,δ). Ships react slow to new settings, depending on the ship's size and type of machine control. Therefore, setpoints must be planned well in advance. Short-term planning is a discrete process, it has the form of methods, it is proactive, occurring at specific points, depending on the on-coming situation.

The lowest level of control is tracking behaviour. The track that has been chosen on the STP-level must be followed as well as possible. For this the navigator must observe the position and movements of the ship relative to the environment. When deviations from the desired track occur corrections must be made. Important factors are wind, current, ground-effect, and various elements in ship control, and the accuracy with which the position fixing can be done. Tracking is a continuous process, it is based on techniques, and it is reactive by nature. The most important characteristics are summarised in table 1.

Table 1. Most characteristic elements of the three levels of control.

level	character	type
long-term planning	procedures	once
short-term planning	methods	discrete
tracking	techniques	continuous

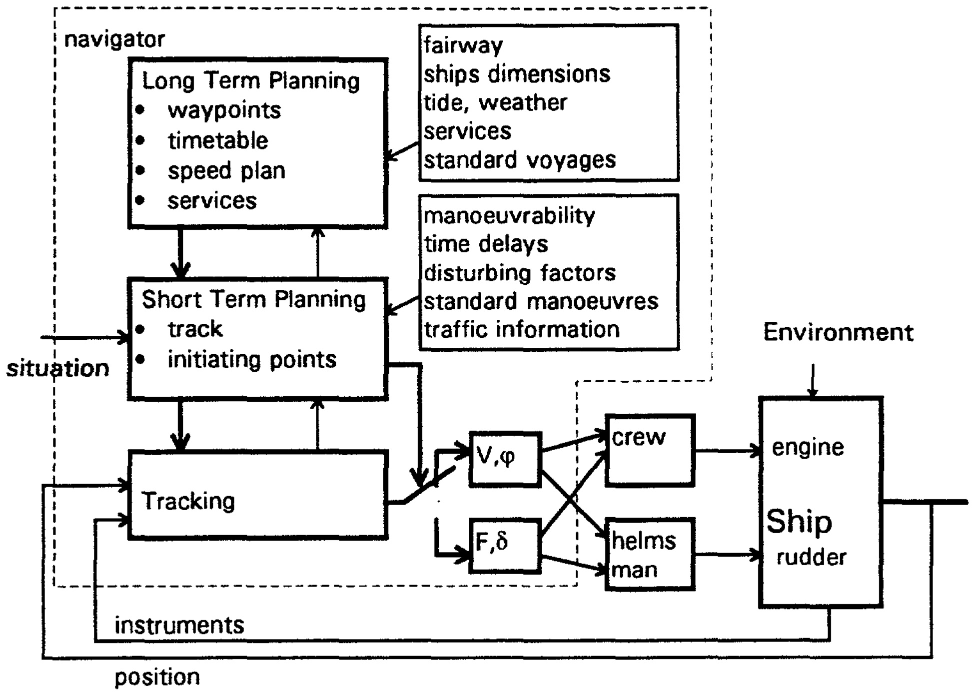

Figure 1: A navigator model.

Workload

In this paper, the term 'workload' always refers to mental workload. Physical workload resulting from boarding and leaving the ship, or irregular working hours is not considered. Performance shaping factors are also not considered.
Not all factors that may add to the workload are present at all times, and even if some factors are present, they may be unimportant. A voyage can be cut down into a number of sections, each characterised by a few elements. These elements determine the track to be chosen. From this track it follows what manoeuvres must be made, and how they are to be initiated.

Two elements are needed for tracking: observation and control. After a track is chosen it must be possible to observe the deviation from the desired track. The uncertainty of the position, related to the required accuracy, will add to the workload. During tracking different types of disturbances act of the ship. It is assumed that they are not always of importance. When the effect of the disturbance is below a certain threshold it has no effect on the task. Over the threshold it does have effect. One could model the influence of such disturbances as a linear combination adding to the workload. The use of speed-course control is considered to be easier than power-rudder control. This would also linearly add to the workload.. All these workload components are directly related to the situation present. Additionally, changes in the disturbances will add temporary to the workload.

Short-term planning will result in a workload present well before a situation occurs. It will be related to a specific situation and occur at a time that is related to the manoeuvrability of the ship. It will depend either on the fairway with its local problems, in which case it will be related to a specific location/situation, or it will depend on other ships, in which case it will be related to the traffic situation.

Long-term planning will lead to an increase of workload for a short time when arriving on board, and play a role later only if the planning has to be changed for reasons unknown at first.

Measurement

To verify the model a method must be applied which will provide workload information about the transitions. Generally, there are four ways to measure mental workload: task demand load in relation to performance, primary task measures, secondary task measures, and physiological measures. Much has been written about each of them and about their possibilities and limitations. (see for instance: Hancock & Meshkate, 1988; Sheridan & Stassen, 1979; Johanssen et al, 1979)
Performance measures are very hard to obtain in navigation tasks. The number of unknown factors is high, and personal preferences will lead to different outcomes.
Subjective workload estimates are not very good for measuring transients in workload. They provide an overall rating of the workload. Due to the specific working conditions, one may doubt the possibilities for self-rating by pilots. The TLX-ratings obtained provide information for wild speculation. Our idea is that pilots have learned to ignore their perception of workload.
To apply a secondary task, the primary task must be known exactly, which is impossible due to the very nature of this project. Additionally, a secondary task will increase the workload and is intrusive by its very nature, which is considered impermissible (Damos, 1991).
Despite the fact that the relation between mental workload and various physiological measures is not clear at all, some measures have shown high correlation between workload and the physiological reaction. Blood pressure, hormonal concentrations, and muscle tension are well known examples. Heart rate and heart rate variability (the power in the 0.1 Hz-band) were chosen. It is easy to record, sensitive, especially for central processing and visual perception, and non-intrusive. Although its reliability and validity have often been questioned, there is evidence that this measure is a reliable measure of workload. A few of these authors are Vincente et al (1987), Wastell (1989), Mulder, G (1983), and Mulder, L.J.M. (1988). The major disadvantage is its problematic interpretation, a drawback for all physiological techniques. It was assumed that the effects caused by a little physical exercise would be small compared to the differences caused by mental workload.

In a series of recordings, four experienced Rotterdam pilots participated. Each pilot did at least two inward bound and two outward bound voyages. The ships were selected for maximum size from the ships available. Ships for the river were preferred to ships for Europort. All pilots were videotaped, and marker points were plotted on a map for specific locations and traffic.

The ECG recordings were corrected for recording errors, and transformed into power density signals using Carspan 1.99 (Mulder, 1988). The power density was calculated using a Direct Fourier Transform based on the Integral Pulse Frequency Modulator model (Rompelman, 1986). The time window of 50(s) was found to be acceptable for both a sufficient spectral resolution and stationary signal. As a smoothing filter, a modified discounted-least square filter was applied with τ=30(s) (SWOV, 1994).

As was argued when you need to start, you need to know these goals. But goals, unlike actions, cannot be observed. Goals can only be obtained by interviewing, and sometimes by interpretation of the actions by an expert. Interviewing during the voyage may be a serious disturbance of the ongoing process, and was rejected for that reason. A compromise was found by asking the pilot afterwards what had happened during the voyage, what the plans were, and what the problems were in executing these plans. Over the time, some knowledge was obtained about the pilot's task, and this could be used in asking specific questions about the events that had taken place. Additionally, some knowledge about standard voyages was obtained by interviewing outside the recorded voyages.

Results

Because of the many differences between the voyages, no statistical analysis has been done over different pilots, ships, and situations. Voyages were analysed by comparing the voyage with the heart rate and power density. One voyage has been chosen as an example. In figure 2 one can see the track, and in figure 3 the resulting ECG-data. Apart from climbing on board, there is no physical exercise by the pilot. The heart rate of all pilots on board stabilised around 90 beats/min, about 10 beats higher than on the pilot-ship. Climbing on and off board can be seen very clear in all heart rate plots. Beforehand we distinguished between inward-bound and outward-bound voyages. In the heart rate plots this difference shows very clearly when entering the harbour. The heart rate increased significantly, to reach its peak at the start of the turn. This effect could not be found with outward-bound ships. The same could be found during critical manoeuvres and passing bridges. This effect was attributed to arousal, that is directly related to hormonal effects (Gellatly & Meyer, 1992).

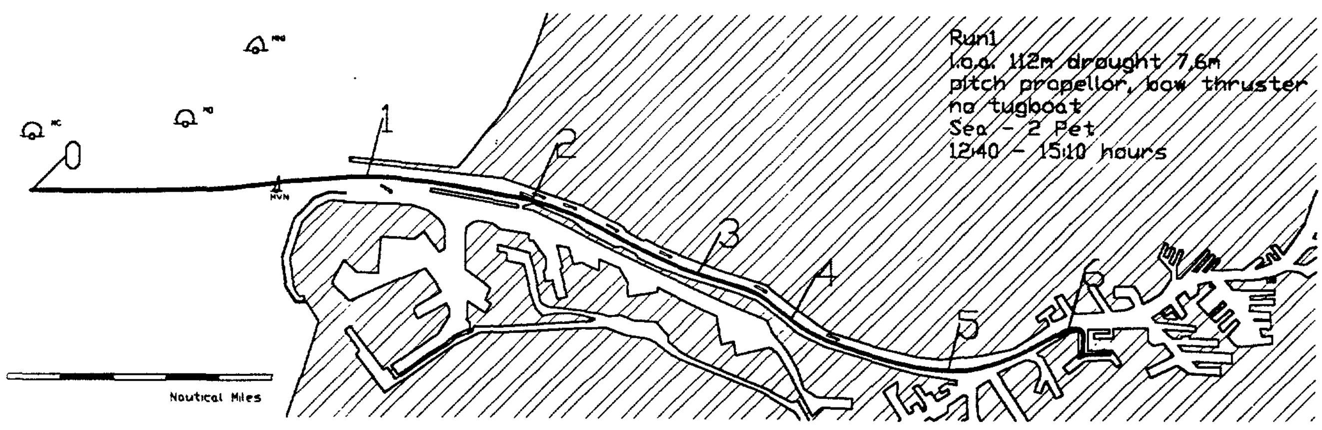

Figure 2. A voyage from sea to a Rotterdam harbour. The nmubers correspond with figure 3.

In computing the power density it was found that in the harbour many artefacts occurred. This is very troublesome because they act like Dirac-functions, masking all other data in that time segment. It is known that stress-like situations often result in extra-systolic contractions of the first ventricle, recording as an extra heartbeat, and disturbing the Fourier transforming process. A second outcome was that the power density is highly variable, which made it difficult to interpret. Moving average filters and frequency based filter techniques were found unsatisfying. The Modified Discounted Least Square filter was found to be most useful. This filter is an a-symmetrical exponential time-weighted filter.

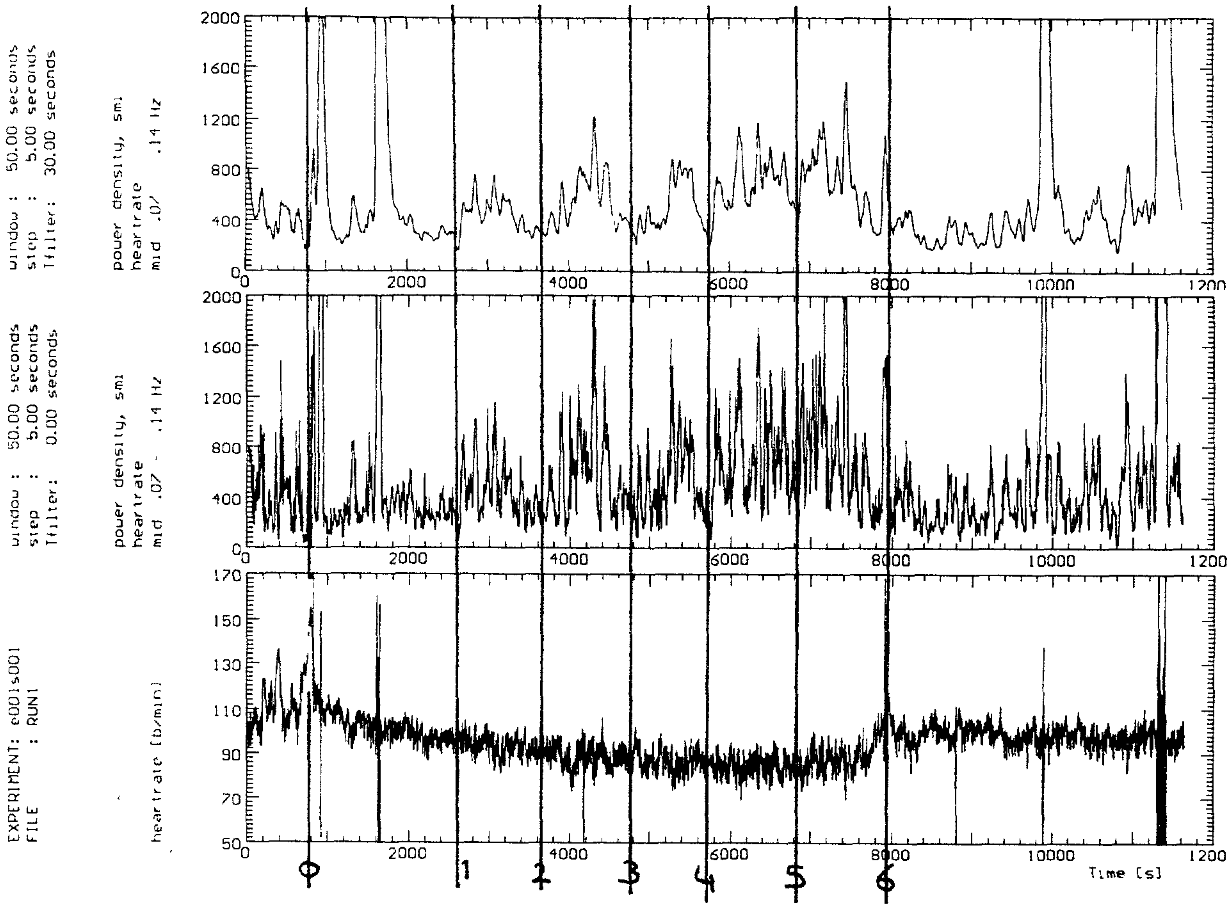

Figure 3: Physiological data from the voyage shown in figure 2. From bottom to top: the heart rate, the power density (unfiltered), the power density (filtered). The numbers on top of the vertical lines correspond with numbers in figure 2. Note that the mental workload is reversely related to the power density of the 0.1 Hz-band. The peaks in the power density result from various artefact in the heartrate.
The increase in heartrate is clearly visable during boarding at sea (0), and when entering the harbour (6). At marker 2 there is an increased workload due to traffic and construction work (between 2 and 3). Marker 3 is the result of a ship that came to the wrong side of the fairway. Marker 4 is located at a place where difficult currents exist. Marker 6 placed just before a location of usually intense traffic.

The overall profile of the power density gave a good indication of the workload (face value). The peaks in workload as indicated by the power density occurred at pre-defined places, such as before a significant change in the current, or when decisions had to be made about passing a ship coming from another fairway. Orders, passing of other ships and disturbances could often be well identified. As is usual with these techniques, lot of variation is left unaccounted for. In the first series, peaks in workload were found which later, during interviews with other pilots, were identified as specific problems.

Conclusions

The use of ECG-recording seems a valuable tool for developing a navigator model. Data obtained from interviews and the ECG-recordings provided valuable information about the pilot's task. For a more sophisticated interpretation a thorough talk-through afterwards is required, making the interpretation much easier. In field experiments such as these, this is not always easy to achieve. Often there is very little time between ships, due to high time-pressure, and the work is done around the clock. Only with full co-operation from the entire organisation time is made available.
The data obtained seem to confirm the model presented here. All major workload peaks which were found to be related to locations or situations, were confirmed in interviews. For a more detailed model, more detailed analysis is needed. We think it will be possible to identify most of the dominant factors in workload using this technique.
For the interpretation of the heartbeat data, the pilot actions, the goals, the characteristics of the ship, the location of the ship, and the situation it is in must be known. To do this, a lot of information must be recorded during and after the voyage. Unfortunately, due to many unknown factors, not all changes in power density may be explained.
For a good analysis of the heartbeat data a good reconstruction of that data is most valuable. Disturbances make the time-window in which it occurs useless, and when disturbances occur regularly it can make a large portion of the recording useless. It seems to us that a better reconstruction may be achieved than present available with Carspan 1.99.

Acknowledgements

We would like to thank the Rotterdam Pilots, the participating pilots, and the captains and crews of the ships for their co-operation and hospitality.

Literature

Damos, D.L., 1991, Dual task methodology: some common problems. In: D.L. Damos (ed), *Multiple task performance*. Taylor and Francis, London.

Gellatly, I.R. & J.P. Meyer, 1992, The effects of goal difficulty on physiological arausal, cognition, and task performance. *Journal of Applied Psychology*, vol 77, no 5, pp 694-704.

Hangcock P.A. & N. Meshkati (eds), 1988, *Human mental workload*. Advances in psychology 52, North Holland, Amsterdam.

Johanssen N.; N. Moray; R. Pew; J. Rasmussen; A. Sanders & C. Wickens, 1979, Final report of the experimental psychology group. In: N. Moray (ed), *Mental workload: it's theory and measurement*, p 101-114, Plenum Press, New York.

Kirwan B. & L.K. Ainsworth (eds), 1992, *A guide to task analysis*. Taylor & Francis, London.

Mara, T.D., 1968, *Human factors in ship control, Vol III: Function and information requirements of deck watch officers and other deck personnel*, Clearinghouse, Springfield, U.S.A..

Mulder, G., 1983, Attention, effort and sinusarrythmia: how far are we? In: J.F. Orlebeke, G. Mulder & L.J.P. van Doornen (eds), *Psychophysiology of cardiovascular control.* p 407-424, Plenum Press, New York.

Mulder, L.J.M.; H.J. Van Dellen; P. Van Der Meulen & B. Opheikens, 1988, Carspan: a spectral analysis program for cardiovascular time series. In: F.J. Maarse; L.J.M. Mulder; W. Sjouw & A. Akkerman (eds) *Computers in psychology: methods, instrumentation & psychodiagnostics.* pp 39-47. Swets & Zeitlinger, Lisse.

Mulder, L.J.M., 1988, *Assessment of cardio vascular reactivity by means of spectral analysis.* State University of Groningen, the Netherlands, Thesis.

Sheridan T.B. & H.G. Stassen, 1979, Definitions, models and measures of human workload. In: N. Moray (ed), *Mental workload, it's theory and measurement.* Plenum Press, New York, 1979.

Papenhuijzen, B., 1994, *Towards a human operator model of the navigator.* Delft University of Technology, the Netherlands, Thesis.

Rompelman, O., 1986, Tutorial review on processing the cardiac event series: a signal analysis approach. *Automedica*, vol 7, pp 191-212.

SWOV, 1994, Control strategies for a highway network; A joint research project of the SWOV, the Technical University Delft and the Institute for Perception TNO, sponsored by the Dutch Ministery of Transport and Watermanagement. Part I, II and III. R-94-34 I/II/III. SWOV, Leidschendam, the Netherlands.

Vincente K.J.; D.C. Thornton & N. Moray, 1987, Spectral analysis of sinus arrhythmenia: a measure of mental effort. *Human Factors*, vol 29, no 2, pp 171-182.

Wastell, D., 1989, Mental effort and task performance: towards a psychophysiology of human computer interaction. In: D. Diaper et al (eds), *Human-Computer Interaction - INTERACT '90.* North Holland: Elsevier Science Publishers B.V., Amsterdam.