M. Grootjen, M.A. Neerincx, J.C.M. van Weert, "Task Based Interpretation of Operator State Information for Adaptive Support", ACI/HFES 2006, San Francisco.

Task Based Interpretation of Operator State Information for Adaptive Support

M. Grootjen[a,b]
Marc@Grootjen.nl

M.A. Neerincx[b,c]
Mark.Neerincx@tno.nl

J.C.M. van Weert[b]
J.C.M.v.Weert@student.tue.nl

[a] Defense Materiel Organization, Directorate Materiel Royal Netherlands Navy, Dep. of Naval Architecture and Marine Engineering, P.O. Box 20702, 2500 ES The Hague, The Netherlands
[b] Technical University of Delft, P.O. Box 5031, 2628 CD Delft, the Netherlands
[c] TNO Human Factors, Kampweg 5, P.O. Box 23, 3769 ZG Soesterberg, the Netherlands

Abstract

In the last 2 decades, the application of new technologies in process control caused a radical change in the role of the operator. To cope with these changes, the operator needs personalized support that can vary over time. To realize this support, we present the design and first testing of a framework for adaptive support. First an analysis was done which data is necessary for the framework. After selection, data was collected during an experiment. Based on this data a software tool was designed to visualize task load of operators over time. Evaluation of the tool shows that combination of cognitive task load, effort and performance gives insight into the task load of an operator over time. Furthermore, it enables generation of critical and optimal task load work area's. Further implementation of the framework needs to be done to generate a real time adaptation plan.

1 INTRODUCTION

In the last 2 decades, major changes in information technology have taken place. In process control, the ongoing automation and the application of new technologies caused a radical change in the position of the operator. This position shifted from monitoring and control to supervision. During multiple seminars and projects we identified 6 main topics:

1. Changing of information type and volume. The operator has to handle large amounts of digital information, varying between sensor values and 'higher level' information, often used together. Highly dynamic systems cause enormous fluctuations in information volume.
2. Task integration. The integration of tasks from different domains (e.g. navigation and propulsion) increases.
3. Increasing autonomy. Autonomous systems may cause 'out of the loop' problems for an operator.
4. Increasing complexity. Systems and their dependencies become more and more complex.
5. Decreasing personnel and training budgets. The pressure to work with the lowest possible costs and the highest possible efficiency asks for an optimal distribution of personnel and intuitive systems.
6. Increasing legislative constraints. The increasing number of laws (Teltzrow & Kobsa, 2004) puts extra constraints on the system design.

To deal with the mentioned problems, the operator needs specifically personalized support which can differ in time (Neerincx et al., 2005). This paper describes the design (section 2) and first testing (section 3) of a high level framework for adaptive support. Finally, section 4 presents conclusion and discussion.

2 FRAMEWORK FOR ADAPTIVE SUPPORT

Our goal is to develop a user friendly, efficient and effective system which adapts to the situation and the operator. It should accommodate the user with the right task support at the right time. Consequently we have to know how and when this system should be adjusted. For this we need information about user's momentary processing capacities and a mechanism that transforms this information into a plan for employing these capacities effectively and efficiently. This plan should describe when and how to adapt the system, and to dynamically allocate tasks between the actors. We define this plan here as the work plan, the mechanism that generates this work plan the dynamic task allocator or shortly allocator, the sources of information are models and the mechanisms that produce the models by filtering and collecting information from the environment are modeling components. For generation of this work plan, we designed a framework. First step in the design of this framework is the identification of the models we need. Literature shows a large variety of models, for example Schneider-Hufschmidt et. al (1993) describe a structural model of the interface composed of about 13 elementary models which focus on the various functional aspects relevant to adaptive interaction. Of course, if the amount of models is increased, the information will be more precise. However, the complexity of the system will also increase enormously. If we want to keep the system controllable, we should minimize the number of models. Based on our goal, we need at least the following models (figure 1):

1. operator model (or user model)
2. task model
3. system model
4. context model (or domain model)

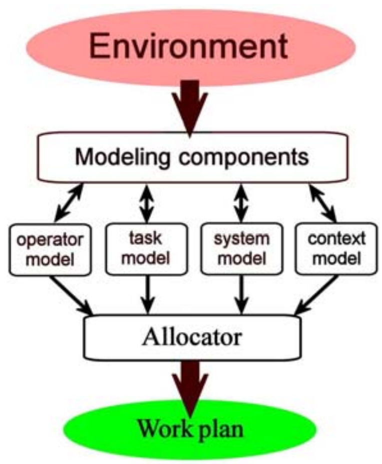

Figure 1. Information from the environment is used to produce four models. Accordingly, the allocator generates an effective and efficient work plan.

Operator model. An essential source of information for the allocator is the operator model. In order to adapt the system to the characteristics of the operator, the operator model can contain a large variety of information:

- To collect task data, experts performed a video analysis afterwards using the Observer XT (Noldus, 2006). At the moment we are testing the use of logfiles and gaze tracking to collect task information.
- Real-time taskload measures (see allocator).
- Profile parameters; routine on the job (Benyon & Murray, 1988), spatial abilities, memory span (Neerincx et al., 2005), education & characteristic knowledge, preferences, generic role in environment.
- Physiological measures.
- Performance; reaction time, time pressure, correct handling.
- Subjective effort;
- Behavioral variables; eye tracker, gaze tracker.

Task model. The task model should contain information about the possible tasks for an operator. It should be noted that this information represents task demands that affect human operator performance and effort (i.e. it is not a definition of the operator cognitive state). We see the task model as some kind of static representation expressing all possible tasks that can be executed (Neerincx, 2003). The effects these tasks have on the operator belong to the operator model.

System model. In order to support a user, a system needs to have information about itself available. Consequently, the system model contains technical information about the different system components (e.g. layout, software applications and dependencies). Special cases of system models are for example application models and dialogue (interaction) models (Brown et al., 1990).

Context model. The context model (or domain model) contains relatively static, high level information of the operational unit and its environment (e.g. the readiness state of a combat unit).

Allocator. Neerincx (2003) developed a model for Cognitive Task Load (CTL) which is very useful for the allocator. The CTL model (figure 2) distinguishes three load factors that have a substantial effect on task performance and mental effort. The first load factor is the classical factor percentage time occupied (%TO). The second load factor is the level of information processing (LIP) (cf. Rasmussen, 1986). To address the demands of attention shifts, the cognitive load model distinguishes task-set switching (TSS) as a third load factor.

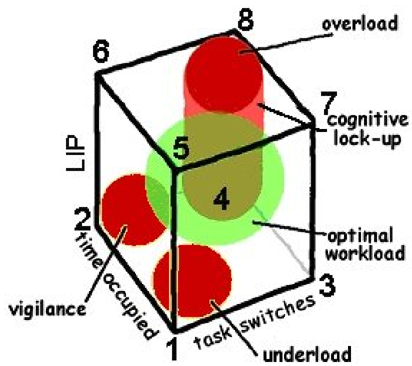

Figure 2. Dimensions of the CTL model (Neerincx, 2003) (see text for explanation).

Figure 2 presents a 3-dimensional 'load' space in which human activities can be projected with regions indicating the cognitive demands that the activity imposes on the operator. In the middle area, CTL matches the operator's mental capacity. At angular point 8 CTL is high and an overload situation occurs. Angular point 1 represents the area in which CTL is not optimal due to underload. When TO is high, and LIP and TSS are low, vigilance problems can appear (angular point 2) (Parasuraman, 1986). When TO and TSS are high, lock-up can appear (Boehne & Paese, 2000).

With information of the different models, the allocator should be able to generate 'CTL space', determine the current and future working point for a specific operator and generate a work plan. The next section will focus on the selection of data for the operator and task model and the generation of CTL space.

3. SELECTING, MEASURING AND PROCESSING OF DATA

Based on a literature review and empirical studies (table 1 and Grootjen et al., in press; Grootjen et al., 2006), we selected different types of data that should feed into the task and operator models. In recent experiments, we subsequently investigated and applied different technologies for recording this data and analyses of these data. For analysis, two different tools were used: the Observer XT (Noldus, 2006) and SOWAT (Ship Operator Workload Assessment Tool). SOWAT should enable an offline generation of CTL space and an online generation of the operators path through this space. Explanation of the Observer XT software is out of the scope of this paper.

3.1 Selecting Effort Parameters

During the first experiments we concentrated on the collection of performance, effort and task data. With a combination of this data we should be able to determine the position of the mentioned regions in CTL space. Effort data can be collected with use of subjective or physiological measurements. Research has shown that it is very hard to find an unique measure that can be directly interpreted as a effort indicator (Veltman, et al., 1996; Wilson, 2001; Wilson et al., 2003; Miyake, 2000). Moreover, different measures provide insight in different aspects of effort, e.g. eye blink rate variations can correspond specifically with visual demands (Wilson, 2001) (Wilson et al., 2003). We decided to use a mixture of subjective and physiological measurements for effort. To determine what measurements are suitable for us, we rated available measurements on three criteria:

1. Sensitivity to workload changes.
2. Obtrusiveness for the operator.
3. Availability of the equipment.

Table 1 shows the rating results. Used literature is shown in the in the last column of the table.

Table 1. Overview of effort parameters with rating from -2 (bad/low) to 2 (good/high)

Measurement	*Sensitivity*	*Obtrusiveness*	*Available*	*Literature*
Heart measures	2	0	yes	Veltman & Gaillard, 1996; Veltman & Gaillard, 1998; Wilson, 2001; Wilson & Russel, 2003; Miyake, 2000
Respiratory	2	0	yes	Veltman et al., 1998; Wientjes, 1992
Blinking, EOG	1	2	yes	Veltman et al., 1998; Wilson et al., 2003
Gaze tracking	0	2	yes	Marshall et al., 2003; Zhai et al., 1999
Pupil dilation	2	2	yes	Marshall, 2002
Voice stress	0	2	no	Rothkrantz, et al., 2004
Facial expression	-1	2	no	Kuilenburg, et al., 2005
Face warmth	0	1	yes	Veltman & Vos, 2005
Brain activity (EEG)	2	-2	yes	Wilson, 2001; Prinzel et al., 1999
Blood pressure	0	1	yes	Veltman, et al., 1996; Miyake, 2000
Skin conductivity	0	0	yes	Haag et al., 2004
Gesture recognition	-2	2	no	Ehlert,2003
Muscles (Lundberg, U. et al. 1994)	1	0	yes	Lundberg et al., 1994
Cortisol (saliva)	-1	-2	no	Veltman et al., 1996; Lundberg et al., 1994
NASA-TLX	2	-2	yes	Hart & Staveland, 1988
Self Assessment Manikin (SAM)	-1	-1	yes	Bradley & Lang, 1994
Workload Watch	2	0	yes	Boer, 1997; Likert, 1932
Rating Scale Mental Effort (RSME)	2	0	yes	Zijlstra, 1993
Continuous memory task (CMT)	2	-2	yes	Veltman et al., 1996

3.2 Measuring Data

Data was collected during an experiment called OLA 2 (Operator Load Assessment 2). OLA 2 was done to evaluate the ship control centre of the air defense and command frigate. Three scenarios, designed using the CTL method (Neerincx, 2003), were executed by a total of 12 teams divided over three different frigates. The experiment took place at sea, the problems were simulated but all actions had to be operated on the actual systems. During the experiment physiological data was recorded using Mobi8 (heart and respiratory measures, TMS International, 2005) and Tobii (eyeblinks and pupil dilation, Tobii Technology AB, 2006). During video analysis, subjective effort and complexity was measured using a workload watch (Boer, 1997) (i.e. a device that beeps on predefined moments and asks for subjective input). Performance and task data was available after video analysis of experts.

3.3 Visualization of Data

The previous sections described which information we selected for the models from figure 1. Now we should define how this information can be used by the allocator to create a work plan. To visualize the information we developed SOWAT (figure 3). Each numbered area will be explained below:

1. Selection area. In this area a start and stop time of the data that has to be visualized can be selected. Below the end time, the amount of cube points determines the number of CTL points (i.e. the interval time for each CTL calculation). The LIP variable can be calculated in three ways, which can be selected at the drop down menu ‘LIP source’. First option is the use of subjective ratings, where the subjective complexity ratings from the workload watch are used as LIP values. Second option is the use of expert LIP ratings, which were defined in advance for each action by experts, using the decision tree for LIP (Neerincx & Lindenberg, 1999). Third option is to use an average of those two ratings. For the operator

skill level the user can choose ‘beginner’, ‘intermediate’ and ‘expert’, which is of influence on the expert LIP ratings.

2. Presentation area. After data selection and pressing the plot button, all figures will be shown in the presentation area. First the CTL cube and projections of the three sides of the cube are plotted. On the right side of the area, additional plots from all recorded variables against time are shown using tab pages. Possible variables are: TO, LIP, TSS, WLW, subjective effort and complexity, expert rating, heart rate, upper/lower respiratory values, pupil size.
3. CTL table area. In the lower left corner, a table is displayed containing the values of all variables for each cube point. Selecting a point in this table will highlight the accompanying point in each figure in the presentation area.
4. Configuration area. At the left part of this area the coloring of points can be configured. Next to this, the coloring thresholds (subjective effort and expert ratings) can be set. There are three colors, green (low effort or high performance), orange (medium effort or performance) and red (high effort or low performance).
5. Export area. After pressing one of the export buttons, all plotted data including a large variety of other data (e.g. average LIP and total TSS over all data) will be saved in text and jpg files.

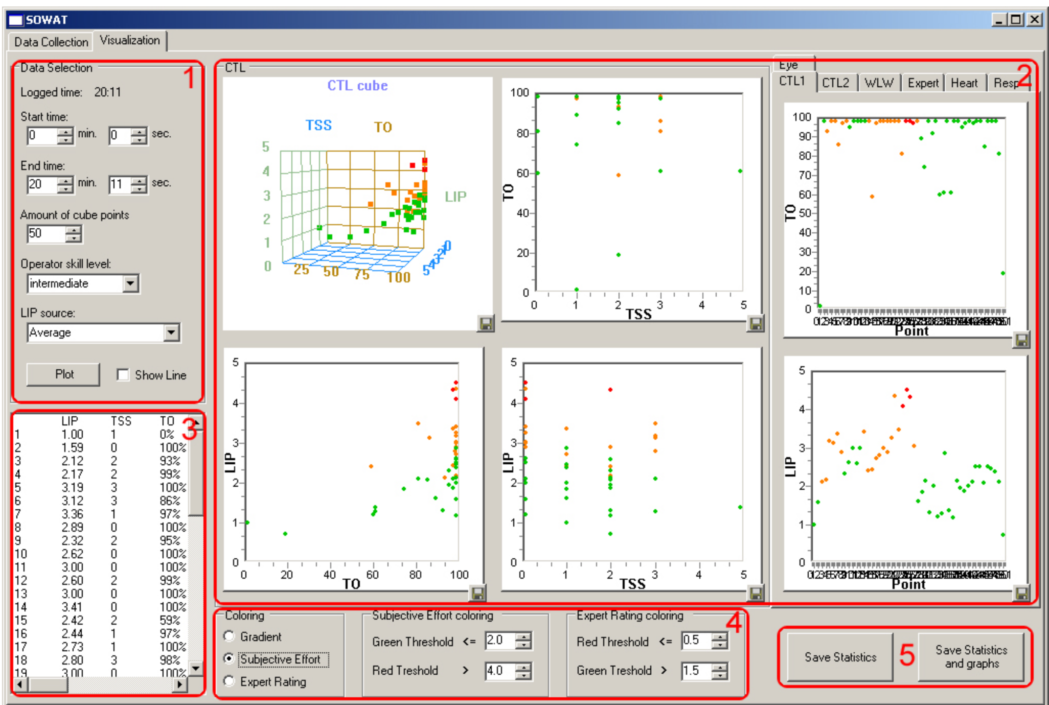

Figure 3. Ship Operator Workload Assessment Tool (SOWAT). A first construction of CTL area’s is made.

3.4 Interpretation of Results

The results visualized with SOWAT can be used in two ways. The first way is an ‘online’ interpretation of an operators CTL in time. A line between the CTL points can be turned on, presenting the chronological path of an operator in CTL space. Also the gradient coloring, which can be turned on in the configuration area, indicates an operators path in time through the cube. It is preferred for the overview to keep the amount of CTL points low. The second way is an ‘offline’ interpretation of data to construct the CTL area’s of figure 2. Using a higher number of CTL points directly shows the critical area’s in CTL space for an operator, for example in figure 3. Depending on operator profile parameters (operator model, figure 1), critical area’s can be determined for specific groups of operators.

4 CONCLUSION AND DISCUSSION

The increase in automation in process control emerged a strong need for a user friendly, efficient and effective system which accommodates the user with the right task support at the right time. This paper shows the design and first application of a framework for adaptive support which results in the following conclusion:
Combination of CTL, effort and performance information gives great insight in the state of the operator. With use of SOWAT, data can be processed and visualized using video observations, subjective user input and objective physiological and performance data. Specifically, two main applications can be distinguished:

1. An offline analysis of data with SOWAT enables a generation of optimal and critical area's in CTL space
2. An online analysis by SOWAT enables generation of the operators path through CTL space

Future research will focus on the usage of offline and online analysis to generate a work plan. Main topics during this research will be:

- Specification of system and context model.
- Implementation of operator classes using profile parameters.
- Further specification of the operator model. This paper presented a literature review of possible information for the operator model. However, an evaluation of chosen measurements should still be done. Depending on this evaluation other measurements could be added.
- Integration of effort and performance data into one or two variables. Current visualization shows all variables in separate figures, however, combination into more generic variables makes it easier to interpret the data.
- To collect task data, experts performed a video analysis afterwards using the Observer XT. At the moment we are testing the use of logfiles and gaze tracking to collect task information.
- In the development of an adaptive support system, we need an real time framework. The work plan needs to be generated using real time information which can be used to keep the operator out of critical area's. At the moment analyzing is only possible afterwards.
- When a work plan can be generated another important question needs to be answered. Who should make the decision for dynamic task allocation? The decision can be made by automation, the operator or by a mix of both. Even another operator (e.g. a superior) can be involved and make the decision.

5. REFERENCES

Benyon, D. & Murray, D. 1988. Experience with adaptive interfaces. *The Computer Journal*, 31.

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

Boer, L. C. Workload-watch as an element of human engineering testing and evaluation.

Bradley, M. M. and Lang, P. J. Measuring emotion: The self-assessment manikin and the sematic differential. *Journal of behavioural Therapy & Experimental Psychiatry* 25[1], 49-59. 1994. Elsevier Science Ltd.

Brown, D., Totterdell, P., & Nonnan, M. 1990. *Adaptive user interfaces*. London: Academic Press.

Ehlert, P. 2003. Intelligent User Interfaces. Report Technical University Delft.

Grootjen, M, Bierman, E. P. B., and Neerincx, M. A. 2006. Optimizing cognitive task load in naval ship control centres: Design of an adaptive interface. IEA 16th World Congress on Ergonomics.

Grootjen, M, Neerincx, M. A., and Veltman, J. A. in press. Cognitive Task Load in a Naval Ship Control Centre. *Ergonomics*.

Haag, A., Goronzy, S., Schaich, P., and Williams, J. 2004. Emotion Recognition Using Bio-Sensors: First Steps Towards an Automatic System.

Hart, S. G. & Staveland, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Human Mental Workload*.

Kuilenburg, H. van, Wiering, M., and Uyl, M. J. den. 2005. A Model Based Method for Automatic Facial Expression Recognition. 16th European Conference on Machine Learning.

Likert, R. 1932. A technique for the measurement of attitudes. *Archive of Psychology.*

Lundberg, U., Kadefors, R., Melin, B., Palmerud, G., Hassmen, P., Engstrom, M., & Dohns, I. 1994. Psychophysiological stress and EMG activity of the Trapezius muscle. *International Journal of Behavioral Medicine*.

Marshall, S. P. 2002. The index of Cognitive Activity: Measuring Cognitive Workload..
Marshall, S. P, Pleydell-Pearce, C. W., and Dickson, B. T. 2003. Integrating PsychoPhysiological Measures of Cognitive Workload and Eye Movements to Detect Strategy Shifts.
Miyake, S. 2000. Multivariate workload evaluation combining physiological and subjective measures. *Internation Journal of Psychophysiology*, 40: 233-238.
Neerincx, M. A. 2003. Cognitive task load design: model, methods and examples. In Hollnagel, E. (Ed.), *Handbook of Cognitive Task Design*: 283-305. Mahwah, NJ: Lawrence Erlbaum Associates.
Neerincx, M. A. and Lindenberg, J. 1999. DISCII: User requirements for support of task-set switching and integration. TM-99-D001. TNO.
Neerincx, M. A., Lindenberg, J., and Grootjen, M. 2005. Accessibility on the Job: Cognitive Capacity Driven Personalization. HCII 2005.
Neerincx, M. A., Ruijsendaal, M, and Wollf, M. 2002. Usability Engineering Guide for Integrated Operation Support in Space Station Payloads. *International Journal of Cognitive Ergonomics*.
Noldus. Observer XT. [6.0]. 2006. www.noldus.com
Parasuraman, R. 1986. Vigilance, Monitoring, and Search. *Handbook of Perception and Human Performance*: 1-39. New York: John Wily.
Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. 1999. A Closed-loop system for examining Psychophysiological Measures for Adaptive Task allocation. *International Journal of Aviation Psychology*, 10(4): 393-410.
Rasmussen, J. 1986. Information processing and human-machine interaction: An approach to cognitive engineering. Amsterdam: Elsevier.
Rothkrantz, L. J. M., Wiggers, P., Wees, J. van, and Vark, R. J. van. 2004. Voice stress analysis.
Schneider-Hufschmidt, M., Kühme, T., & Malinowski, U. 1993. *Adaptive User Interfaces - Principles and Practice*. Amsterdam: Elsevier.
Teltzrow, M. and Kobsa, A. 2004. Impact of User Privacy Preferences on Personalized Systems - a Comparative Study.
Karat, C. M., Blom, J., and Karat, J. 2004. Designing Personalized User Experiences for eCommerce. Dordrecht, Kluwer Academic Publishers.
TMS International. Mobi 8. 2005. TMS International, Oldenzaal, the Netherlands. www.tmsi.com
Tobii Technology AB. Tobii. 2006. www.tobii.com
Veltman, J. A. and Gaillard, A. W. K. 1996. Pilot workload evaluated with subjective and physiological measures.
K.Brookhuis, C. Weikert J. Moraal & D. d. Waard 1996 (Eds.) *Aging and Human Factors*, 107-128.
Veltman, J. A. & Gaillard, A. W. K. 1998. Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5): 656-669.
Veltman, J. A. and Vos, W. K. 2005. Facial temperature as a measure of mental workload. 597-601. International Symposium on Aviation Psychology.
Wientjes, C. J. 1992. Respiration in psychophysiology: methods and applications. *Biological Psychology*.
Wilson, G. F. 2001. An analysis of Mental Workload in Pilots during flight using multiple psychophysiological measures. *International Journal of Aviation Psychology*, 12: 3-18.
Wilson, G. F. & Russel, C. A. 2003. Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, 45(4): 635-643.
Zhai, S, Morimoto, C., and Ihde, S. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing.
Zijlstra, F. R. H. 1993. *Efficiency in work behaviour*. Delft university press.