A Mental Workload Predicator Model for the Design of Pre Alarm Systems

Sheue-Ling Hwang[1], Yi-Jan Yau[1], Yu-Ting Lin[1], Jun Hao Chen[1], Tsun-Hung Huang[1,3], Tzu-Chung Yenn[2], and Chong-Cheng Hsu[2]

[1] Institute of Industrial Engineering & Engineering Management
National Tsing Hua University, Hsinchu,
Taiwan, R.O.C. 300
[2] Institute of Nuclear Energy Research, Lungtan,
Taiwan, R.O.C. 325
[3] Department of Industrial Engineering and Management,
National Chin-Yi Institute of Technology, Taichung,
Taiwan, R.O.C. 411
slhwang@ie.nthu.edu.tw, d927814@oz.nthu.edu.tw,
d943847@oz.nthu.edu.tw, g943848@oz.nthu.edu.tw,
toby.hwang@gmail.com, tcyenn@iner.gov.tw,
worm@iner.gov.tw

Abstract. This study investigated the operator's mental workload of the fourth Nuclear Power Plant in Taiwan. An experiment including the primary and secondary tasks was designed to simulate the reactor shutdown procedure of the Nuclear Power Plant. The performance of secondary task, the subjective mental workload and seven physiological signals of participant were measured. The Group Method of Data Handling (GMDH) was applied to integrate these physiological signals to develop a mental workload predictive model. The relationship between subject mental workload and the performance of secondary task is highly correlated with Pearson correlation coefficient as 0.691. The validity of the proposed model is very high with R2=0.85. The proposed model is expected to provide supervisor a reference value of operator's performance by giving physiological signals. Besides nuclear power plant, the proposed model could be applied to other fields such as aviation, air transportation control, driving and radar vigilance, etc.

Keywords: Mental workload; Physiological signal; GMDH; Predictor.

1 Introduction

Modern complexity system could bring heavy mental workload to their operators. The high rate of information flow, complexity of the information, numerous hard decisions and task time stress could overwhelm the system operators. On the other hand, high level of automation could lead to low mental workload [1, 2]. For most of the operators, their performance would be degraded when their mental workload is either too high or too low. Only in an appropriate level of mental workload, the

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2007, LNAI 4562, pp. 316–323, 2007.

operators could perform as anticipated [3, 4]. Although both high and low mental workloads would cause operator's performance degradation, most of the previous studies paid much attention on the problems caused by high mental workload rather than by low mental workload. Hence, developing an early warning system to measure the operator's mental workload is a very important issue.

Performance measurement, subjective ratings and physiological measurement have been considered as the three most general mental workload measurements [5, 6]. Among these measurement technologies, physiological measurement could record the continuous data and be less intrusive on work activities and high sensitive to the cognitive requirements of a complex task. Hence, it is more suitable for measuring real time mental workload than others [7, 8]. However, Chen and Vertegaal [9] pointed out that there has been little use of real-time physiological measure to dynamically manage the operator's mental workload during system operation.

The physiological responses to mental tasks were different for each person and the physiological response patterns were also different from task to task [8, 10]. Thus, it is important to consider the situations while one try to develop a real time mental workload predictive model. One approach to solving these problems is to record several physiological indexes and integrate them into one synthesized index by individual difference, and analyze it immediately [9, 7].

Numerous researches have successfully used EEG signals to classify the operator's mental workload [2, 9, 11]. Their results indicated that the average rate of successful classification was over 80%. However, Farmer and Brownson [5] indicated that using EEG to classify mental workload was not practical since the collecting data was difficult to be analyzed and had high noise-to-signal ratio. Furthermore, the equipment required to calibrate to each individual and be operated by trained person. Therefore, it is not suitable for using in the real field.

The purpose of the present research is to develop a real-time, non-intrusive mental workload predicative model by using Group Method of Data Handling (GMDH) to integrate seven physiological indexes into a synthesized index. Comparing this index with a performance level set by supervisor, if the value of the synthesized index was lower than the setting performance level, the early warning system would be started to alert the operators to adjust their workload. The physiological indexes used in this study were parasympathetic/sympathetic ratio (LF/HF) (X1), heart rate variability (HRV) (X2), heart rate (X3), diastolic pressure (X4), systolic pressure (X5), eye blink frequency (X6), and blink duration (X7).

2 Method

2.1 Subjects

Fifteen paid NTHU graduate students participated in this experiment. All of them had normal eyesight and good health. The mean age of the participants was 24 years.

2.2 Apparatus

Meditech ABPM-04 ambulatory blood pressure measuring device was used to measure the heart rate and blood pressure (Systolic Pressure and Diastolic Pressure).

ANSWatch TS0411 was used to measure Heart Rate Variability (HRV) and Parasympathetic/Sympathetic Ratio (LF/HF). Face lab version 4.0 was used to measure the blink frequency and blink duration.

2.3 Procedure

Prior to the experiment, the subject took a 15-minute rest, and then wore the measurement apparatus and proceeded with Face Lab adjustment. The initial physiological indexes were taken as a base line before the experiment. After the adjustment and measurement, the experimenter illustrated the experimental task and instructed the participant how to operate the system. The participant took about 30 minutes to practice the control procedure and familiarize with location of information display. There were three phases simulated the mental workload of secondary task from heavy, median to low. The subject simultaneously executed the primary tasks and the secondary tasks. The primary task was to shut down the reactor following the procedure of “shut down cooling model” provided by the Institute of Nuclear Energy Research (INER) (see Fig. 1). The shut down procedure was to decrease the power by inserting the nuclear rods or adjusting the core flow rate by turning off the pump. The secondary tasks were to complete a series of mathematical comparisons at each phase. The different mental workload levels of the secondary task were designed by different mathematical calculation. The tasks were presented in four monitors except the Face-Lab monitor interface, and their layout was showed in Fig. 2 The physiological indexes (LF/HF, HRV, heart rate, diastolic/ systolic pressure, blink frequency and blink duration) were measured prior to experiment and during each phase, and the NASA-TLX questionnaire was conducted after each phase. Each phase lasted for about thirty minutes. The seven physiological indexes were transferred into individual difference and acted as input variables for the model development. The performance of secondary task was evaluated by the error rate of the number comparison. The subjective mental workload was rated by the NASA-TLX score after each phase.

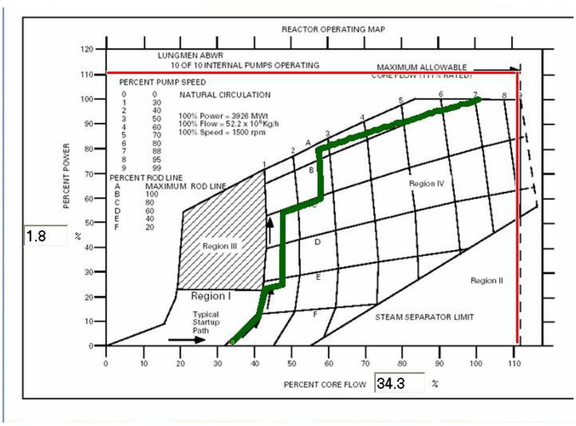

Fig. 1. The power and flow monitor interface

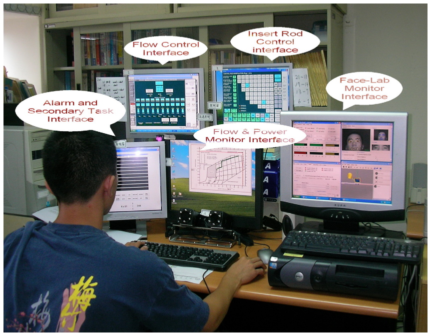

Fig. 2. The control interface screen layout

2.4 Analysis

The error rate of the series of mathematical comparisons at each phase was collected to compare with the corresponding scores of NASA-TLX questionnaire. The higher score of NASA-TLX questionnaire reflected higher level of mental workload during experiment. Correlation between the performance of secondary task (error rate) and the evaluation of subjective mental workload was analyzed by the Statistical Products and Services Solution (SPSS).

Group Method and Data Handling (GMDH), one of the well-know neural network methodologies, was applied to develop the predicative model. GMDH was developed in 1971 [12] and is one of the exact prediction methods in last decade. The GMDH algorithm has been widely used in various fields such as education [13], business, and vehicle factory [14]. This study investigated the relationship between seven physiological indexes and mental workload. These physiological indexes (X1~X7) were collected as input variables and then created a model to predict different degree of mental workload.

3 Results

All participants finished their primary task before time limitation and met the requirements set by INER; hence, differences in mental workload would be reflected by the performance of the secondary task.

3.1 Correlation Between Secondary Task Performance and Subjective Mental Workload

The relationship between the subjective mental workload and performance of secondary (error rate) is shown in Fig. 3. In general, the directions of two trend lines of error rate and subjective mental workload were definitely the same. Almost all subjects made fewer errors and rated lower score of NASA-TLX questionnaire as the secondary task becomes less complex.

The analysis of Pearson-product moment correlation was used to examine the relationship between secondary task performance and NASA-TLX subjective mental workload as shown in Table 1. It indicated that the error rate and the subjective mental workload were positively correlated with each other. The correlation coefficient of 0.691 was found to be statistically significant at $p < 0.01$.

As a result, an increasing number of the error of secondary task in the appearance implied that subjects felt heavier mental workload. In contrast, the better performance of secondary task meant that the subject felt lower degree of mental workload.

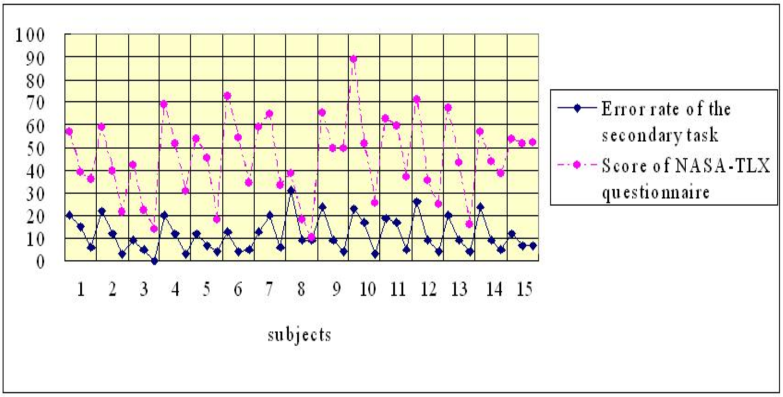

Fig. 3. The relationship between two dependent variables

Table 1. The analysis of Pearson-product moment correlation

Correlations		Error rate of the secondary task	Score of subjective mental workload
Error rate of the secondary task	Pearson Correlation	1	0.691 [a]
	Sig. (2-tailed)	—	0.000
	N	45	45
Score of subjective mental workload	Pearson Correlation	0.691 [a]	1
	Sig. (2-tailed)	0.000	—
	N	45	45

[a] Correlation is significant at the 0.01 level (2-tailed)

3.2 Model Establishment

Seven physiological indexes, including LF/HF (X1), HRV (X2), heart rate (X3), diastolic pressure (X4), systolic pressure (X5), blink frequency (X6), and blink duration (X7) were transferred to the forms of variability ranging from -1 to 1. The accuracy (Y) was to evaluate the performance of the secondary task.

The data X1~X7 were used as inputs in the network system to get an output Y where the output Y is a value from 0 to 100 (percentage of accuracy). Group Method of Data Handling (GMDH) was used to establish model of early warning system.

Thirty-nine data (thirteen subjects) were used to construct the model using NeuroShell software.

The result indicated that all factors of X1~X7 significantly affected the accuracy of the secondary task performance. Also, given the values for X1~X7, a model of performance was yielded and expressed by the following equation:

$$Y = 4.5X_6-33X_4+87-14X_3-0.52X_1+79X_7+590X_3^2-220X_4^2-120X_7^2-1900X_3^3+1100X_4^3 - 270X_3X_4+840X_3X_7+18X_2^2-34X_6^2-7.1X_2^3+6.1X_2X_6+130X_4X_7-39X_1X_5 \quad (1)$$

In equation (1), MSE of the model was 9.04 and R square of the model was 0.84. Such equation was expected to provide supervisors/operators a reference value of performance (Y) by giving inputs X1~X7.

For validation, 6 data (2 subjects) were taken into the model. The relational information was described in Table2. In this model, the six estimated values were very close to the real values and all fell in the 95% confidence interval. Therefore, the model was suitable and accurate to estimate the performance.

Table 2. Model validation

Subject Number	X_1	X_2	X_3	X_4	X_5	X_6	X_7	Estimative Value	Real Value	Low Bound of 95% C.I.	Upper Bound of 95% C.I.
14(High)	0.10	-0.61	0.03	0.05	-0.14	0.20	-0.07	83.71	81	77.491	89.929
(Median)	0.00	-0.45	0.02	0.19	-0.12	-0.01	0.02	85.19	83	78.971	91.409
(Low)	0.00	-0.52	-0.05	0.15	-0.17	-0.03	0.00	90.85	95	84.631	97.069
15(High)	-0.33	-0.30	0.00	0.23	-0.04	-0.03	-0.05	77.15	76	70.931	83.369
(Median)	-0.60	-0.22	-0.11	-0.10	-0.05	0.14	-0.03	95.86	91	89.641	102.079
(Low)	-0.67	0.14	-0.15	-0.15	0.01	0.25	0.04	96.44	95	90.221	102.659

4 Discussions

4.1 NASA–TLX Mental Workload and the Performance of Secondary Task

The NASA–TLX subjective mental workload assessment showed a significant correlation with the different level of mental workload of secondary task. For almost all subjects, the highest NASA–TLX scores occurred in the heavy mental workload whereas the lowest scores happened in the low mental workload. This result was not only consisted with the previous studies [15, 16] but also confirmed that the tasks used in this experiment could distinguish the different level of mental workload.

4.2 Physiological Indexes

From literature review, it could be found that some physiological indexes were significantly affected by the mental workload. The physiological indexes measured in this experiment included heart rate, heart rate variability, blood pressure (systolic pressure and diastolic pressure), parasympathetic/sympathetic ratio (LF/HF ratio), eyes blink frequency and eyes blink duration. The experimental result showed that most of the participants' heart rate and LF/HF components increased when the mental workload increased. On the contrary, the heart rate variety (HRV) decreased when the mental workload increased. These findings were consistent with some previous studies [6; 17]. Aside from these, many participants' blood pressure was not increased as the mental workload increases. This was not consistent with the previous studies [6; 18]. The possible reason could be that the mental workload of this experiment was not high enough to affect the blood pressure change. The experimental result also showed that most of the participants' eyes blink duration was shorter and eyes blink frequency was fewer during high mental workload than during low mental workload. These findings were also similar to many previous studies [17, 18, 19]. The reason could be the participants paid more attention on the interface during high mental workload phase than during low mental workload phase; hence, the activities of eyes (eye blinking frequency and duration) were lower during high mental workload phase.

5 Conclusions

This study developed an early warning model allowing the operators/supervisors to monitor operators' mental workload by physiological indicators. Comparing the synthesized index with the performance level set by supervisor, the early warning system could be commenced to alert the operator while the synthesized index was lower than the setting performance level. For further applications in different fields such as nuclear power plant, flight, aviation, air transportation control, driving and radar vigilance, etc, more data or physiological indicators are needed to train or add to this model.

Acknowledgements. This research has been supported by the National Science Council, Taiwan, Republic of China, Project No. NSC95-NU-7-007-003). The authors wish to thank for the help provided by the Institute of Nuclear Energy Research, Taiwan, Republic of China**.**

References

1. Billings, C.E.: Some questions about advanced automation. In: Mouloua, M., Koonce, J.M. (eds.) Human-automation interaction, pp. 314–320. Lawrence Erlbaum, Mahwah, NJ (1997)
2. Wilson, G.F., Russell, C.A.: Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. Human Factors 45(4), 635–643 (2003)
3. Moray, N.: Mental workload since 1979. International Review of Ergonomics 2, 123–150 (1988)
4. Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Linton, P.M., Wierwille, W.W., Zaklad, A.L., Bittner, A.C.J., Wherry Jr., R.J.: Operator workload: Comprehensive review and evaluation of workload methodologies. US Army Research Institute for the Behavioral and Social Sciences, Willow Grove (1989)

5. Farmer, E., Brownson, A.: Review of workload measurement, Analysis and interpretation methods. European air traffic management programme, CARE-Integra-TRS-130-02-WP2-1-0 (2003)
6. Veltman, J.A., Gaillard, A.W.K.: Physiological indices of workload in a simulated flight task. Biological Psychology 42, 323–342 (1996)
7. Guhe, M., Liao, W., Zhu, Z., Ji, Q., Gray, W.D., Schoelles, M.J.: Non-intrusive measurement of workload in real-time. In: Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society. Orlando, FL (2005)
8. Miyake, S.: Multivariate workload evaluation combining physiological and subjective measures. International Journal of Psychophysiology 40, 233–238 (2001)
9. Chen, D., Vertegaal, R.: Using Mental Load for Managing Interruptions in physiologically attentive user interfaces. In: Proceeding of Conference on Human Factors in Computing Systems, pp. 1513–1516. ACM, New York (2004)
10. Xie, B., Salvendy, G.: Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. Work. & Stress 14(1), 74–99 (2000)
11. Gevins, A., Smith, M.E.: Detecting transient cognitive impairment with EEG pattern recognition methods. Aviation, Space, and Environmental Medicine 70(10), 1018–1024 (1999)
12. Ivakhnanko, A.G.: Polynomial Theory of Complex Systems. IEEE Transactions on Systems, Man, and Cybernetics SMC 1(4), 363–378 (1971)
13. Pavel, N., Miroslav, S.: Modeling of student's quality by means of GMDH algorithms. Systems Analysis Modeling Simulation 43(10), 1415–1426 (2003)
14. Tomonori, N., Shizue, S.: Identification and forecasting in management systems using the GMDH method. Appl. Math. Modeling 6, 7–15 (1982)
15. Battiste, V., Bortolussi, M.: Transport pilot workload - A comparison of two subjective techniques in Human Factors Society. In: Annual Meeting 32nd. 1989, Human Factors Society, Anaheim, CA, USA (1989)
16. Jones, D.G., Endsley, M.R.: Can real-time probes provide a valid measure of situation awareness? In: Proceedings of the Human Performance, Situation Awareness and Automation (2000)
17. Wilson, G.F.: An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. The. International Journal of Aviation Psychology 12(1), 3–18 (2002(a))
18. Wastell, D.G., Newman, M.: Stress, control and computer system design: a psychophysiological field study. Behaviour & Information on Technology 15(3), 183–192 (1996)
19. Wilson, G.F.: Psychophysiological Test Methods and procedures. In: Charlton, S.G., O'Brien, T.G. (eds.) Handbook of human factors testing and evaluation, pp. 127–156. Lawrence Erlbraum Associates, Mahwah, NJ (2002(b))