ISSN 1009-3095 Journal of Zhejiang University(SCIENCE) V.1, No.3, P.337-341, July-Sep., 2000
http://www.chinainfo.gov.cn/periodical; http://www.zju.edu.cn;
http://lib.zju.edu.cn/English; http://www.zjupress.com

MULTIPLE MEASURES OF MONITORING MENTAL WORKLOAD*

ZHANG Zhijun(张智君)

(*School of Psychology and Behavior Science, Zhejiang University, Hangzhou, 310028, China*)

Received May.21, 1999; revision accepted Dec.6, 1999

Abstract: This study explored the possibility of developing a technique combining multiple measures of mental workload. Dual task technique was applied. Twenty subjects were asked to perform computer-simulated monitoring tasks with three different levels of difficulties. A reaction time performance with digital stimuli was added as subsidiary task. The four indexes, accurate reaction time of primary task, variation rate of accurate reaction time of subsidiary task, weighted subjective workload rating and variation rate of heart rate variance, were used and tested. The correlation matrix of the four indexes was examined with principal component analysis technique, and two mathematical models of comprehensive mental workload indexes (CMWIA, CMWIB) were then advanced. It is verified that these two comprehensive indexes were far more sensitive for monitoring mental workload as compared with any one of the preceding four indexes. It was concluded that it was useful to construct some comprehensive mental workload indexes for practical industrial purposes.

Key words: monitoring mental workload, multiple measures, mental overload
Document code: A **CLC number:** B849; TB18

INTRODUCTION

The problem of mental overload became increasingly serious with development of industrial automation. In many industrial stations, it was believed, operators were frequently under the pressure of high monitoring mental workload because many panels and other instruments must be monitored and controlled simultaneously and continuously. The heavy mental workload must make an adverse impact on operators, and it was also one of major factors which resulted in various accidents and losses. The assessment and prediction of monitoring mental workload is highly important for industrial engineers to design systems with adequate mental workload, which will in turn promote work efficiency, improve work environment and operator satisfaction, and finally, make performances much more reliable (Nachreiner, 1995; Svensson et al., 1997).

There have been quite a few studies on mental workload measures for dozens of years. Some important techniques have been suggested and applied, including primary task measures, subsidiary task measures, subjective workload ratings and physiological measures (Myrtek, 1994; Tattersali et al., 1996). However, it was found it is not easy to evaluate the mental workload reliably in some industrial conditions. In other words, almost all common procedures had specific weaknesses. For example, primary task measures were just sensitive to medium mental workload. The subsidiary task measures necessarily interrupted primary tasks. They also required the specific-resource performances as subsidiary tasks to get high sensitivity. Subjective rating techniques just reflected the state of general workload. It was not easy to diagnose the sources which caused mental overload with subjective rating techniques (Tsang et al., 1996). Physiological measures had to be supported by some special equipment and skills. The physiological responses of mental workload were sometime likely to get confused with physical stresses (Zhang et al., 1994). Therefore, multiple measures of mental workload were suggested and approved (Rouse et al., 1993; Tsang et al., 1997).

The usability and validity of multiple indexes in measuring monitoring mental workload was firstly examined in this study. Then, a technique of integrating these indexes into comprehensive mental workload indexes (CMWI) was further explored. It was expected that our results

* The project (395174) supported by Zhejiang Provincial Natural Science Foundation of China

will provide a basis for assessing practical monitoring of mental workload.

METHODS

1. Design of monitoring performance

A computerized-simulated monitoring task was designed and used. On the upper part of the computer screen, three stimulus frames were set, each with one active belt in it. These belts lengthened or shortened continuously, randomly, and independently of each other. The subjects were directed to monitor the changes of these belts and asked to respond with right hands if ' stimuli' emerged. In other words, when any one of these belts lengthened into the warning areas, which were set on the right side of the frames, a stimulus occurred, the subjects were to press on the keyboard selectively as accurately and rapidly as possible. The belt went out of the warning areas and the stimulus disappeared as soon as the subjects acted. Three response keys, arranged from left to right, were positioned corresponding to three stimulus frames. There were three levels of monitoring difficulties, which depended on the number of stimulus frames watched at the same time. The stimuli occurred at the rate of 50 items per three minutes. Seven other static panels and two dynamic targets were added on the screen to prompt the immersion of reality.

2. Method and indexes

Dual task technique was employed. The monitoring performance was defined as the primary task, and a simple reaction time performance was assigned as the subsidiary task. For the latter, the subjects were required to respond to digital stimuli (which appeared within the stimulus frame on the lower left part of the screen) with their left hands by pressing a specific key.

The following four indexes were selected and used:

Primary task measure: Accurate Response Time of the Primary task (ARTP)

Subsidiary task measure: Variation rate of Accurate Response Time of the Subsidiary task (VARTS)

Subjective workload rating: Weighted Subjective Workload Rating (WSWR)

Physiological measure: Variation rate of Heart Rate Variance (VHRV)

In which, Accurate Response Time (ART) equaled Reaction Time divided by Reaction Accuracy. ARTP was average reaction time of monitoring task divided by its accuracy, and VARTS was calculated according to the following formula:

$$\text{VARTS}(\%) = \frac{\text{ART of simple reaction time task in dual task conditions}}{\text{ART of simple reaction time task in single task conditions}}$$

WSWR was the weighted sum of four subjective workload rating values. The four subjective ratings were Task Complexity (TC), Task Pressure (TP), mental Effort and Fatigue (EF) and Strain Level (SL), respectively, which had been tested in our previous study (Zhang et al., 1995).

VHRV was the ratio of heart rate variance of subjects in single monitoring task conditions to that at rest. Heart rate variance was calculated according to the R – R intervals of ECG.

3. Subjects and procedure

The twenty undergraduate students (10 males and 10 females) participated in the study as subjects. They were 18 to 23 years old, had normal or corrected to normal vision, and all right-handed.

All subjects engaged in four 5-min practice sessions at first. They were asked to reach an accuracy of 95% or over after these practices.

After a break, the trials began. During the whole experiment, each subject had to perform seven trials, three for primary tasks with different difficulties, one for subsidiary task, and three for dual tasks with different monitoring difficulties. The sequences of trials were predetermined at random. Each trial lasted three minutes. There was a 30-min break between any two trials.

The ECG was recorded during each trial. Subjects were asked to give subjective ratings of TC, TP, EF and SL based on their feelings which were experienced in each trial. The subjective rating scale was a 100 mm line, which included two bi-polar adjectives and meant 'very, very low' or 'very, very high', respectively. When all trials ended, subjects were asked to arrange these subjective items according to their priority, which were used to calculate their weighted coefficients in WSWR.

RESULTS

1. Sensitivity of four indexes

The results showed that the reactions to monitoring targets were delayed with the increase in workload both in single task and dual task conditions. The ARTP was prolonged from 186.92 ms to 251.49 ms to 347.47 ms in single task conditions when the number of stimuli frames changed from 1 to 2 to 3. Likewise, the ARTP extended from 327.94 ms to 377.72 ms to 590.28 ms in dual task conditions. There were 0.01 – level significant differences among three workload levels both in single task ($F = 10.23$) and dual task ($F = 7.20$) conditions.

The average values of VARTS markedly increased from 184.90 to 236.21 to 332.99 with the changes of the numbers of stimulus frames of primary tasks from 1 to 2 to 3 ($F = 12.65$, $p < 0.01$). Because these increases possibly related to the change of primary task performances under single to dual task conditions, the covariance analysis of VARTS was made to exclude the effect of the primary task interference with subsidiary task performances. The results showed that average covariance modified values of VARTS were also significantly influenced by monitoring difficulties, but that there were no significant differences between VARTSs and their average covariance modified values ($p > 0.05$).

Likewise, the values of WSWR ($F = 11.86$, $p < 0.01$) and VHRV ($F = 5.56$, $p < 0.01$) both significantly increased with the increase of the number of stimulus frames. WSWR were 18.94, 33.38 and 45.21, and VHRV were 86.22, 77.09 and 66.41 respectively in three monitoring difficulties.

2. Correlation among four indexes

In order to examine individual and overlap contributions of four indexes to assessing monitoring workload, their correlation matrix was calculated (results in Table 1). The matrix showed significant correlation between ARTP and VARTS, between VARTS and WSWR, and between WSWR and VHRV.

This matrix was processed with principal component analysis technique. It was orthogonally rotated according to the principle of variance maximum. Finally, the matrix of factor loading was obtained, as shown in Table 2. It was found that these four indexes could be classified into two common factors. One mainly related to ARTP and VARTS, the other to WSWR and VHRV.

Table 1 The correlation matrix of four mental workload indexes

	ARTP	VARTS	WSWR
VARTS	0.66**		
WSWR	0.23	0.33**	
VHRV	−0.01	−0.12	−0.27*

** $p < 0.01$, * $p < 0.05$.

Table 2 The factor loading matrix of four mental workload indexes

	Factor 1	Factor 2	Common variance
ARTP	0.90613	0.00807	0.82113
VARTS	0.88117	−0.18271	0.80984
WSWR	0.36295	−0.67000	0.58063
VHRV	0.09533	0.87889	0.78154

3. Mathematical models of comprehensive mental workload index (CMWI)

The study used the values of factor loading displayed in Table 2 as weighted coefficients of four indexes to calculate their weighted sums. In other words, two mathematical models of comprehensive mental workload indexes (CMWI) were constructed based on following formulae:

$$\text{CMWIA} = a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4$$

$$\text{CMWIB} = b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4$$

in which, x_1, x_2, x_3 and x_4 were standard scores of ARTP, VARTS, WSWR and VHRV respectively. a_1, a_2, a_3 and a_4 equaled to 0.90613, 0.88117, 0.36295 and 0.09533, and b_1, b_2, b_3 and b_4 were 0.00807, −0.18271, −0.67000 and 0.87889 separately.

It was found that ARTP and VARTS were especially emphasized in CMWIA. On the contrary, WSWR and VHRV had major influences on CMWIB. Because VHRV decreased with the increase of monitoring workload, the CMWIB had a negative correlation with the monitoring difficulty.

The results of CMWIA and CMWIB are presented in Fig 1. The CMWIA were −1.2506, −0.1748 and 1.4255 respectively for three different monitoring difficulties from low to high, and there were very significant differences among them ($F = 35.21$, $p < 0.001$). Likewise, the CMWIB values of three different monitoring difficulties also changed significantly ($F = 32.5$, $p < 0.001$), and were 0.9787, 0.0176 and −0.9962 separately. Obviously, two CMWIs were far sensitive to monitoring mental workload as compared with any one of four indexes.

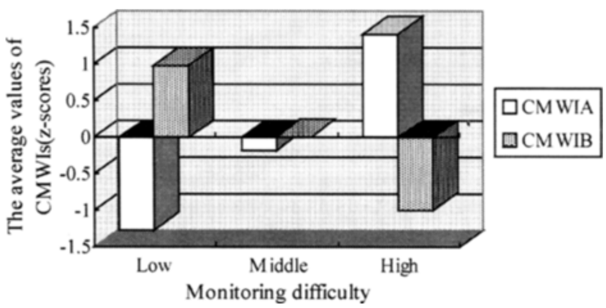

Fig. 1 The average values of CMWIs in three different monitoring difficulties

DISCUSSION

Many mental workload assessment techniques or indexes had been suggested, including primary task performances, subsidiary task performances, subjective ratings, heart rate, heart rate variance, pupil diameter, EEG, etc. But, these techniques or indexes generally had their own specific application limitations. Therefore, multiple measures were often recommended in practical situations (Myrtek, 1994; Casali et al., 1983).

When multiple measures were employed, a good tool was needed to combine different measures. There had been several important efforts in some previous studies. For example, SWAT, NASA − TLX and SWORD all applied multidimensional scale techniques (Reid et al., 1988; Hart et al., 1988; Vidulich et al., 1991). Sloan and Cooper used the factor analysis technique to analyze fifteen self-reported pilot performances and intended to find a synthetic subjective index (Sloan et al., 1986). However, those efforts were almost limited to the field of subjective workload rating.

The study also tried to develop a technique, by which ARTP, VARTS, WSWR and VHRV could be incorporated to construct some comprehensive workload indexes. This effort had been recognized as valuable because the sensitivity and validity of two comprehensive workload indexes (CMWIA, CMWIB) had been confirmed. Although the mathematical models deduced here would probably be inappropriate for other cases due to ubiquitous task, environmental, and individual variance, the study suggested a good operative technique for comprehensively assessing practical monitoring workload.

CONCLUSIONS

The study explored the possibility of developing a technique to combine multiple measures of monitoring mental workload. As a result, two mathematical models of comprehensive mental workload indexes (CMWIA, CMWIB) were constructed by analyzing the correlation matrix of four indexes, including accurate reaction time of primary task, variation rate of accurate reaction time of subsidiary task, weighted subjective workload rating and variation rate of heart rate variance, with principal component analysis technique. It was found that these two comprehensive indexes were far more sensitive for measuring monitoring workload as compared with any one of the four indexes. It was predicted that the technique of comprehensively assessing monitoring workload developed in the study would be available in practical industrial conditions.

ACKNOWLEDGMENTS

This paper was finished under the guidance of Professor Zuxiang Zhu, Zhejiang University.

References

Casali, J. G., Wierwille, W. W., 1983. A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human*

Factors, **25**(6): 623 – 642.

Hart, S. G., Staveland, L. E., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *In*: Human Mental Workload, P.A. Hancock, N. Meshkati (Eds.), North-Holland, Amsterdam, p. 139 – 183.

Myrtek, M., 1994. Physical, mental, emotional, and subjective workload components in train drivers. *Ergonomics*, **37**(7): 1195 – 1203.

Nachreiner, F., 1995. Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics*, **26**(4): 259 – 263.

Reid, G. B., Nygren, T. E., 1988. The subjective workload assessment technique: A scaling procedure for measuring mental workload. *In*: Human Mental Workload, P.A. Hancock, N. Meshkati (Eds.), North-Holland, Amsterdam, p. 185 – 218.

Rouse, W. B., Edwards, S. L., Hammer, J. M., 1993. Modeling the dynamics of mental workload and human performance in complex systems. *IEEE Transactions on System, Man, and Cybernetics*, **23**(6): 1662 – 1671.

Sloan, S. J., Copper, C. L., 1986. The measurement of self-reported pilot performance. *Aviation, Space, and Environmental Medicine*, **57**(5): 466 – 471.

Svensson, E., Angelborg-Thanderz, M., Sjoberg, L., et al., 1997. Information complexity-mental workload and performance in combat aircraft. *Ergonomics*, **40**(3): 362 – 380.

Tattersali, A. J., Foord, P. S., 1996. An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, **39**(5): 740 – 748.

Tsang, P. S., Velazquez, V. L., 1996. Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, **39**(3): 358 – 381.

Tsang, P.S., Wilson, G.F., 1997. Mental Workload. *In*: Handbook of Human Factors and Ergonomics, G. Salvendy (Ed.), John Wiley and Sons, New York, p. 417 – 449.

Vidulich, M.A., Ward, G.F., Schueren, J., 1991. Using the subjective workload dominance (SWORD) technique for projective workload assessment. *Human Factors*, **33**(6): 677 – 692.

Zhang, Z.J., Zhu, Z.X., 1994. A review on research of mental workload assessment methods. *Chinese Journal of Applied psychology*, **9**(4): 18 – 22 (in Chinese, with English abstract).

Zhang, Z. J., Zhu, Z. X., 1995. Experimental study on multidimensional subjective assessment for mental workload. *Chinese Ergonomics*, **1**(2): 4 – 7 (in Chinese, with English abstract).