International Journal of Industrial Ergonomics 54 (2016) 93–102

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: www.elsevier.com/locate/ergon

Quantitative relationship model between workload and time pressure under different flight operation tasks

Lijing Wang*, Xueli He [1], Yingchun Chen [2]

Fundamental Science on Ergonomics and Environment Control Laboratory, School of Aeronautic Science and Engineering, Beihang University, Beijing, 100191, China

ARTICLE INFO

Article history:
Received 23 June 2015
Received in revised form 16 May 2016
Accepted 17 May 2016

Keywords:
Subjective evaluation
Accuracy
Pupil diameter
Time pressure
Redline
Workload

ABSTRACT

The goal of this study was to establish a quantitative relationship model between workload and task demand under different tasks, when time pressure was set as the main influential factor to the task demand, with three workload measurement parameters. The workload "redline" was also analyzed and determined with the relationship models between the workload measurement parameters and time pressure. The experiment was designed with three different tasks under different time pressures. Three workload measurement parameters (subjective evaluation workload, accuracy and pupil diameter) and the subjective feeling threshold of time pressure were measured experimentally and then used in a comprehensive analysis for the relationship model. The data analysis result showed significant differences in workload under different time pressures, but workload was not affected by the task type. With a time pressure of 0.8, participants felt a sense of time urgency and the accuracy decreased by approximately 85%. The results demonstrate that the subjective evaluation workload, accuracy and pupil diameter can be used as the measurement parameters for the workload under different time pressures and for different tasks. Thus, for a time pressure of 0.8, an accuracy of 80%–85% was determined as the workload "redline". Linear relationships were found between subjective evaluation workload, and pupil diameter and time pressure, and a quadratic curve relationship was found between accuracy and time pressure. Workload prediction can thus be performed using these relationship models between workload and time pressure.

1. Introduction

Mental workload is often conceptualized as the interaction between the capacity and the operator (Hilburn and Jorna, 2001). The ratio between task demands and capacity therefore determines the level of workload. The workload acting on operators includes time load, mental effort load and psychological stress load (Reid and Nygren, 1988). Capacity is determined by the skills and training of the operator, but may also be influenced by stressors such as fatigue, noise, etc. Task demands are determined by the number of tasks to be performed, the amount of attention needed, and the time available (Hancock and Meshkati, 1988). Cillie (1992) defined the task variables as task criticality, task structure (High versus Low), task novelty, time response, amount of task, task complexity, and task rate. The time available for task demand can lead to conversion of time load and psychological stress load to workload. Engineers typically emphasize operational definitions based on time available to perform a task (Krüger, 2008). It is crucial to assess workload by the time available, which is one of the main factors influencing task demand. Thus, the relationship between time available and workload is an important issue for the workload prediction and assessment.

Time available, as an input variable referring to task demand (Gaillard, 2001), was defined by the task loading index and used in workload assessment with time pressure. The time pressure can be defined as the ratio of time required to complete a task to the time available (Siegel and Wolf, 1969). The time pressure is defined as the ratio of time required for tasks to time available for tasks (Mioch et al., 2010). The ratio of time required to time available (Time Pressure, TP) is widely used in workload prediction and assessment. The time line approach treats the workload as a function of the time

* Corresponding author.
E-mail addresses: wanglijing505@126.com (L. Wang), 15210988283@163.com (X. He), chenyingchun@comac.cc (Y. Chen).
[1] AVIC China Aero-polytechnology Establishment, China.
[2] Shanghai Aircraft Design and Research Institute, China.

http://dx.doi.org/10.1016/j.ergon.2016.05.008
0169-8141/

available to perform a task (Beevis, 1994), and calculates workload with TP, which is used to identify workload encountered in the overall aircraft operations, and then used to develop workload prediction models, such as the Workload Assessment Model (WAM), Time Based Analysis of Significant Coordinated Operations (TASCO) (Roberts and Crites, 1985), the Workload Index (W/INDEX)(North, 1986), etc. The proportion of time required to time available is also taken as an objective indication of workload levels for ATM operators (Cullen, 1999), in which 100% occupation is taken to indicate high workload (and is plotted as a workload value of 5), and 0% occupation is taken to indicate low workload (and is plotted as a workload value of 0).

The time pressure is a critical factor in terms of pilot workload, especially in the flight scenario of emergency operations. The flight operation procedure was designed and fixed after the design of the human-machine interface in the cockpit. The standard normal operating procedure is fixed at each flight phase. In an emergency situation, the number of operation tasks in a unit time increases so that the time pressure becomes the main factor influencing workload. There are various types of flight operation tasks, including visual tasks, such as altimeter monitoring and airspeed indicator monitoring, and operational tasks, such as manipulating the stick (disk), operating the throttle lever, etc. Under these different types of task units, are the relationship between time pressure and workload constant? Understanding the linear relationship between workload and time pressure under different types of tasks can improve the accuracy of workload prediction.

The method of using time pressure (the proportion of time required to time available) to calculate the workload only considers temporal relationships in a fixed period of time without considering task type. However, there are other factors that influence workload, such as task difficulty, but there is no quantitative corresponding relationship between time pressure and workload for more precise workload prediction that considers these additional factors.

1.1. Task demand, performance and workload

The earliest relation model between task demand and performance has been described by Meister (Meister, 1976). Meister defined three regions: performance remains unchanged with the increase in task demand, performance decreases with the increase in task demand, performance remains at this minimum level with further increase in demand. Other studies (Meijman & O'Hanlon, 1984) described the relationship among task demand, performance and workload with a reverse “U” model, including the underload and overload regions. There are six regions in the relation model among workload, performance and task demand (De Waard, 1996). The relationship among task demand, performance and workload was detailed in these models. However, there is no quantitative analysis on their relationship because the task demands are determined by task difficulty and time pressure. A quantitative relationship model between workload and task demand can be established when the task demand mainly comes from one factor, task difficulty or time pressure.

1.2. Workload measurement

There are numerous methods to measure workload, such as subjective measurement, physiological measurement (pupil dilation, heart rate, respiratory, sinus, arrhythmia, etc.) (Berka et al., 2007; Glenn et al., 1994; Vigo et al., 2012; Miriam Reiner and Gelfeld, 2013; Faure et al., 2016; Mansikka et al., 2016), primary task performance (accuracy, reaction time, etc.) and secondary task performance (Brookhuis and De Waard, 2010; Mazur et al., 2013; Reiner and Gelfeld, 2013). Under different methods, there are many parameters that can be used to indicate workload. To qualify the relationship of workload and time pressure, the parameters used to indicate the workload should also be defined with qualification. Among all the parameters, performance accuracy is the most direct indicator of the workload (Mazur et al., 2013; Krüger, 2008). Subjective measurement is the most widely used method in measuring workload (Fairclough et al., 2005; Prichard et al., 2011; Stuiver et al., 2012; Fallahi et al., 2016; García-Mas et al., 2016). Meanwhile, the correlation of pupil size in physiological measurement with mental workload has long been supported (Juris and Velden, 1977). Pupil size is the most promising single measure of mental workload because it does not disrupt a user's ongoing activities, and provides real-time information about the user's mental workload (Kahneman, 1973). Startle eyeblink was attenuated during both tasks, and the attenuation was greater during the multiple-task condition than during the single-task condition (Neumann, 2010). Although the difference in pupil size between different emergency operation procedures (EOPs) was not significant, there was a tendency for pupil size in the high complexity EOP to be larger than that in the low complexity EOP (Gao et al., 2013). When participants drove a simulator and performed the n-back task, the initial results show their pupil size (PS) measurements have the expected trends, but significant differences between n-back levels found in the PS data suggest that PS may be more sensitive to differences in workload (Gable et al., 2015).

It is well known from a variety of studies that an observer's pupils dilate with increasing cognitive workload being imposed (Klingner et al., 2008). All of the parameters are not indicated to individually influence workload, but rather combining physiological information with subjective and performance information leads to a more pronounced insight into workload. We chose the workload measurement parameters of subjective evaluation, accuracy and pupil diameter to represent workload for the comprehensive analysis on the relationship model between task demand and workload.

When using pupil diameter as the measurement indicator to workload, the factors influencing pupil diameter require consideration. Winn et al. (1994) reported that pupil size decreased linearly as a function of age at all illuminance levels. Pupil diameter changes under different tasks (Batmaz and Ozturk, 2008). Pupil dilation is known to quickly respond to changes in the brightness in the visual field and a person's cognitive workload while performing a visual task (Pomplun and Sunkara, 2003). Pupil diameter as a workload measure should be introduced and discussed in light of some recent research demonstrating, for instance, an effect of blink duration but not pupil size with workload (Gao et al., 2013). Also Schulz et al. (2013) showed that pupil diameter increases were associated with a shift of visual attention from monitoring towards manual tasks. Gao et al. (2007) reported that pupil diameter was affected by stress. Thus, pupil diameter can be affected by various factors, such as task, brightness, stress, etc.

1.3. Workload “redline”

Estimates of workload can determine whether specified functions and tasks allocated to human operators are feasible in terms of time and capability requirements (Krüger, 2008). Therefore, it is important to determine the workload “redline”. The value of the workload is called “redline”. Workload “redline” is always presented as another parameter, for instance, performance has been used as the presentation of workload “redline” (Colle et al., 1988).

Therefore, it may be more useful to place a workload redline at the transition from the optimal performance region to the taskrelated effort region (Krüger, 2008). Reid et al. used SWAT scale as the representative of the workload "redline", and identified 40 ± 10 as the redline range for the Subjective Workload Assessment Technique (SWAT) (Reid and Colle, 1988). Colle et al. (Colle and Reid, 2005) also estimated a SWAT based redline with SWAT scores of approximately 40. The value achieved 85%–95% accuracy in predicting a "redline" workload level in training data, and on completely new data, the accuracy was in the 70–75% range (Schvaneveldt, 1997). The redline value was adopted from earlier work (Reid and Colle, 1988) showing that at that value of workload, performance measures begin to show effects of workload. If the workload level is calculated at 85%, then it could be expected that, in the normal performance of the tasks, the operator would shed some activities to bring the workload below 80% occupation. In the timeline task analysis, the situations of concern are those which cause the operator to approach the edges of the performance envelope, i.e., as TP approaches 1.0., simulation for the Workload Assessment and Manning (SIMWAM).

Although there are many studies on workload redline, they determined the workload redline with one single parameter. Is the "redline" the same when using multiple parameters under comprehensive consideration?

The purposes of this study were, therefore to 1) analyze the relationship between time pressure and workload with measurement parameters, and test the impact of individual differences and task type on measurement results; 2) establish a quantitative relationship between time pressure and workload, using three workload parameters, subjective evaluation, accuracy and pupil diameter; 3) determine the workload "redline" under different time pressures, with comprehensive multi-parameter considerations.

2. Methods

2.1. Participants

The participants were 29 healthy right-handed males between 19 and 26 years old who reported no orthopedic or neurological disease, with a CVA (corrected visual acuity) of 1.5. Ethical approval was requested and received from the Ethics committee in the Technion for human experiments. The participants signed consent forms.

2.2. Apparatus

Experimental stimuli were displayed using pictures on the computer screen. The content of the experiment was edited and shown in the software (Experiment Buide, EB) of Eyelink II. The participants sat approximately 70 cm from the bottom of the stimulus. Data were collected using data analysis software of Eyelink II. Eyelink II is a head-mounted video-based eye tracker, consisting of a video camera and infrared light source that were pointed at a participant's eye, and a device that tracked the location and size of the pupil using these tools. Pupil size can be recorded at 250 Hz (pupil) or 500 Hz (pupil with cornea). In the experiments, the pupil size was recorded at 250 Hz because the eyes of most participants cannot be identified well with cornea.

2.3. Experimental content

According to the human information processing stage model and the MRT (Multiple Resource Theory) proposed by Wickens (Wickens and McCarley, 2007), the resource processing stages include perception, cognition, and response. Perception included both visual and auditory perception; cognition included both spatial and verbal cognition; response included both manual and vocal response.

Then, according to the flight operation procedure of civil aircraft A320, the flight operation tasks include perception tasks (e.g., monitoring altimeter, scanning navigation information and receiving warning sounds), cognition tasks (e.g., evaluation, calculation and selection), and responding tasks (e.g., operating toggle switch, having a conversation). A whole operation task commonly includes perception, cognition and response in sequence.

In the analysis of flight operation tasks, the experiments were designed including three types of flight operation tasks: Identify the status of the toggle (Task 1), i.e., perception (visual) −cognition (spatial, simple) − response (manual) as a whole task; Identify the warning information with warning light (Task 2), i.e., perception (visual) − cognition (spatial, complex) − response (manual); Recognize the notification aural (Task 3), i.e., perception (auditory) − cognition (verbal) − responding (manual).

The experiments are described in detail as follows.

Task 1: Identify the status of the toggle: Identify the position status of the toggle "ANTI COLLISION". When it is on "ON", press button "N"; when it is on "OFF", press button "F".

Task 2: Identify the warning information with a warning light: Press different buttons according to the warning light in different positions. When the warning light glows in area "ENG", identify the danger posed by the engine, then press button "E"; when it glows in area "HYD", identify the danger posed by the hydraulic pressure system, and press button "H"; when it glows in area "AIR", identify the danger posed by the air condition system, and press button "A"; when it glows in area "BLEED", identify the danger posed by the bleed air system, and press button "B".

Task 3: Recognize the notification aural: Press different buttons according to different notification notes that participants hear. When the aural "Ding" occurs, press button "A", which stands for "Attention"; when the aural "short buzz" sounds, press button "W", which stands for "Wait"; when the aural "long buzz" sounds, press button "E", which stands for "Emergency"; when the aural "fire alarm sound" occurs, press the button "F", which stands for "Fire alarm".

Siegel and Wolf (Siegel and Wolf, 1969) proposed that the time pressure (TP) of the task operation can be calculated by the equation $TP = Tr/Ta$, in which Ta represents the time available to perform a task and Tr represents the time required to perform the task. In the experiment, Tr was measured in the pre-test before the official experiment, and then Ta can be obtained based on the time pressure calculation equation with the value of Tr and TP. The experimental time was set by Ta.

The required time (Tr) of each participant to perform the task was recorded without any time limitations in the experiment, operating as quickly as possible while ensuring accuracy. TP was set at the level of 0.6, 0.7, 0.8, 0.9, 1.0 and 1.2. Then, based on Tr and the equation of $TP = Tr/Ta$, the available operation time Ta can be calculated, which was taken as the experiment time. Each experimental task under different TPs was performed thirty times, so that the participants fully felt the workload of the operation tasks. It has been said that such psychophysical experiments might cause habituation error and expectation error, in which case the results might be affected by the non-susceptibility factors. Therefore, the experiment was designed with the method of minimal change smallest variation to allow habituation error and expectation error to cancel each other out to eliminate the effect of the errors (Yang, 1988). The time pressures of the experimental tasks are set from small to big, then to small: 0.6, 0.7, 0.8, 0.9, 1.0, 1.2,1.2, 1.0 0.9, 0.7, 0.8 and 0.6.

2.4. Parameters

Three workload parameters, including subjective evaluation workload, pupil diameter and accuracy, from twenty-nine participants in the experiment were measured.

2.4.1. Subjective evaluation for workload

A five-point Likert scale with instructions was used for subjective evaluation, of which one point indicates that the task can be easily completed with no difficulty; two points indicates that the task can be easily completed with little difficulty; three points indicates that the task can be completed with little difficulty; four points indicates that the task is difficult to complete; five points indicates that the task is too difficult to complete.

2.4.2. Accuracy

Each task required the participant to operate the corresponding keyboard in the experiment, which are described in the experimental contents. The correct keyboard operation was defined as the accuracy of each task in the experiment. The accuracy was recorded in EB.

2.4.3. Pupil diameter

In each experiment, under different tasks and after completion of each task, the pupil diameter of each participant was recorded with eye tracker for further analysis.

2.4.4. Threshold value of time pressure

In each experiment, under different tasks and after the completion of each TP task, the participant was required to provide a subjective feeling rating. If the participant felt the time pressure of the task operation showed a sign of positive reaction, they marked the operation as "+"; if they did felt the time pressure of task operation showed a sign of negative reaction, they marked the operation as "-".

2.5. Procedure

In order to eliminate the effect of fatigue on the experimental results, it is necessary to strictly control the time of the experiment. Therefore, before the formal experiment, there are a lot of pre-experiments, to determine the time of experimental executing, ensuring that the workload changes can be measured without any fatigue.

Before the experiment started, twenty minutes were allocated to each participant in order to become familiar with the experimental contents. Then, the participants were asked to execute the experimental tasks as quickly as they could with accuracy as the premise. This group of experiments contained eleven tasks with TP levels of 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.2, 1.0 0.9, 0.7, 0.8 and 0.6, respectively. The absolute feeling threshold of time pressure is the boundary at which the participants change from one reaction to the other. Each task was presented to the participants for ten minutes, and the reaction times of the participants were recorded. If the participant felt the time pressure of task operation showed a sign of positive reaction, they marked the operation as "+"; if they felt the time pressure of task operation showed a sign of negative reaction, they marked the operation as "-". Then, the subject was equipped with the eye tracker and subjective evaluation methods and standards were instructed to the participants, and the official experiments began. Each TP level of the experiment lasted 3 min, and the subjective evaluation was performed after each level. When the subjective evaluation finished, the subjects pushed the "task" button to enter the next level. The whole experiment lasted approximately 90 min. The accuracy of operation, the subjective evaluation of the workload and the pupil diameter were obtained.

2.6. Data analysis methods

It is expected that three workload measurement parameters, i.e., subjective evaluation, accuracy and pupil diameter, can show the same change orientation and degree as the change of time pressure. The time pressure was taken as the only variable in each experimental task, the effect of individual diversity was taken as the fixed factor. ANOVA was used to test the different significant of the workload under different time pressure.

Then, the data of all participants under the same experiment were averaged to establish a relationship model between workload and time pressure. UNIANOVA was used to test the effect of task type and time pressure on the experimental results. Pearson correlation coefficients (γ) were used to test the simple correlation between workload measurement parameters and time pressure, which fit the normal distribution, at a confidence level of 0.05.

All analyses were completed using SPSS 18.0 (SPSS Inc., Chicago, USA). A significance level of 0.05 was used throughout the manuscript.

3. Results

The experiments were performed with TP levels of 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.0, 0.9, 0.8, 0.7 and 0.6 in sequence. All the data of workload measurement parameters were distributed in the form of normality ($p > 0.05$ for all). The same TP tasks were first averaged, resulting in six sets of data, i.e., TP under 0.6, 0.7, 0.8, 0.9, 1.0, and 1.2. The workload parameters from twenty-nine participants showed that there is high correlation among each other with the mean value, with a significance level of 0.05 ($p < 0.05$).

3.1. Analysis of workload measurement

The test results of the ANOVA on the subjective evaluation, pupil diameter, and accuracy were recorded and analyzed as follows.

(1) Subjective evaluation workload.

The impact variance analysis of the TP on workload was shown by analyzing the workload subjective evaluation of the same task under different time pressures. During the analysis, the TP was set as the fixed factor and the subjective evaluation workload was set as the dependent variable.

- Task 1: Identify the status of the toggle

Significant difference in terms of subjective evaluation workload was observed under different time pressures, $F(11, 335) = 24.92$, $p < 0.05$.

- Task 2: Identify the warning information with a warning light

Significant difference in terms of subjective evaluation workload was observed under different time pressures, $F(11, 335) = 19.85$, $p < 0.05$.

- Task 3: Recognize the notification aural

Significant difference were noted for subjective evaluation workload was observed under different time pressures, $F(11, 335) = 16.58$, $p < 0.05$.

The above results suggest that for the three different tasks, the difference of the subjective evaluation workload under different time pressures is significant.

(2) Accuracy.

The impact variance analysis of TP in terms of accuracy was studied by analyzing the accuracy of the same task. During the analysis, the TP was set as the fixed factor, and the accuracy was set as the dependent variable.

- Task 1: Identify the status of the toggle

Significant difference was noted in terms of accuracy under different time pressures, $F\,(11, 335) = 77.81$, $p < 0.05$.

- Task 2: Identify the warning information with the warning light

Significant difference was noted for accuracy under different time pressures, $F\,(11, 335) = 96.56$, $p < 0.05$.

- Task 3: Recognizing the aural notification

Significant difference in terms of accuracy was observed under different time pressures, $F\,(11, 335) = 28.68$, $p < 0.05$.

The above results show that for the three different tasks, the difference in accuracy under different time pressures is significant.

(3) Pupil diameter.

The impact of the TP over pupil diameter was studied by analyzing the pupil diameter under different time pressures of the same task. During the analysis, the TP was set as the fixed factor, and pupil diameter was set as the dependent variable.

- Task 1: Identify the status of the toggle

There is a significant difference in pupil diameter under different time pressures, $F\,(11, 335) = 6.98$, $p < 0.05$.

- Task 2: Identify the warning information with the warning light

The pupil diameter under different time pressures is significant, $F\,(11, 335) = 8.69$, $p = 0.025 < 0.05$.

- Task 3: Recognizing the aural notification

The pupil diameter under different time pressure is significant, $F\,(11, 335) = 5.48$, $p = 0.\,04 < 0.05$.

The above results show that for the three different tasks, the difference in pupil diameter under different time pressures is significant.

The accuracy, subjective evaluation, and pupil diameter can be used for workload measurement study because the three parameters show great differences under different TPs among all the tasks.

3.2. Relationship of TP and workload parameters with TP

All the data of the participants within the same experiment were averaged to test the effect of task type and TP with UNIANOVA.

(1) Subjective evaluation.

Task type and TP were set to be fixed factors and the subjective evaluation workload was set as the dependent variable. The subjective evaluation workload under different tasks shows no significant difference: $F\,(2, 15) = 2.105$, $p = 0.173 > 0.05$. The pupil diameter under different time pressures is significantly different: $F\,(5, 13) = 266.277$, $p < 0.05$.

$\gamma = -0.037$ ($p = 0.885 > 0.05$), between the subjective ratings and task types demonstrates a low correlation between the two parameters. However, $\gamma = 0.981$($p < 0.05$), between the subjective ratings and TP, shows strong correlation.

The analysis diagram of subjective evaluation workload of different tasks under different time pressures is shown in Fig. 1. From the figure, we observe that the subjective evaluation workload increases with the time pressure in a linear relationship.

(2) Accuracy.

Task type and TP were set to be fixed factors and the accuracy was set as the dependent variable. The accuracy under different tasks shows no significant difference: $F\,(2, 15) = 2.70$, $p = 0.11 > 0.05$. The accuracy under different time pressures shows a significant difference: $F\,(5, 13) = 261.31$, $p < 0.05$.

$\gamma = 0.05$ ($p = 0.843 > 0.05$), between the accuracy and task types demonstrates a low correlation between the two parameters. However, $\gamma = -0.931$ ($p < 0.05$) between the accuracy and TP shows a strong correlation.

The analysis diagram of pupil accuracy under different time pressures is shown in Fig. 2. From the figure, we observe that the accuracy decreases with increasing time pressure.

(3) Pupil diameter.

Task type and TP were set to be fixed factors and the pupil diameter was set as the dependent variable. The pupil diameter under different tasks shows no significant difference: $F\,(2, 15) = 2.70$, $p = 0.11 > 0.05$. The diameter under different time pressures shows a significant difference: $F(5, 13) = 261.31$, $p < 0.05$.

$\gamma = 0.008$($p = 0.975 > 0.05$), between the pupil diameter and task types demonstrates a low correlation between the two parameters. However, $\gamma = 0.854$ ($p = 0.049 < 0.05$) between the pupil diameter and TP shows strong correlation.

The analysis diagram of pupil diameter under different time pressures is shown in Fig. 3. From the figure, we can observe that the pupil diameter increases with increasing time pressure.

3.3. Determine the workload "redline"

(1) The sensory threshold of TP.

In this paper, the sensory threshold of TP indicates the lowest time pressure value when the participants feel time pressure. TPs were set to levels of 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.0, 0.9, 0.8, 0.7 and 0.6, respectively, for different tasks. The subjective evaluation values of the sensory threshold, i.e., the time pressure of participants under

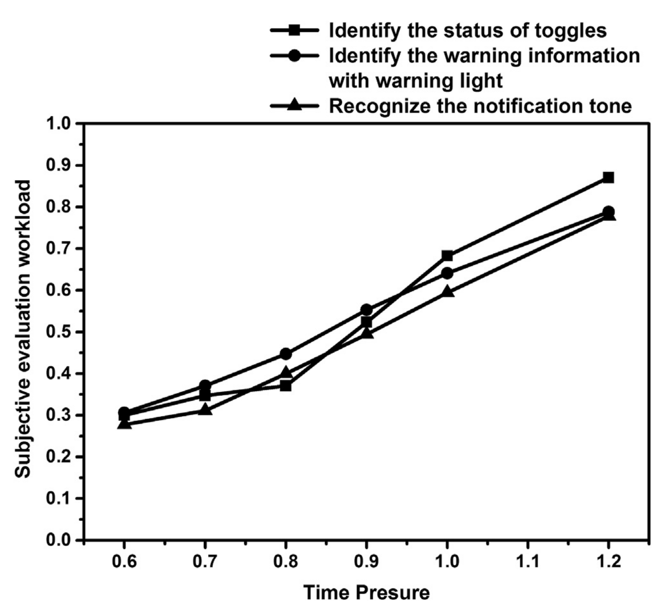

Fig. 1. The change in subjective evaluation workload with time pressure for three different tasks.

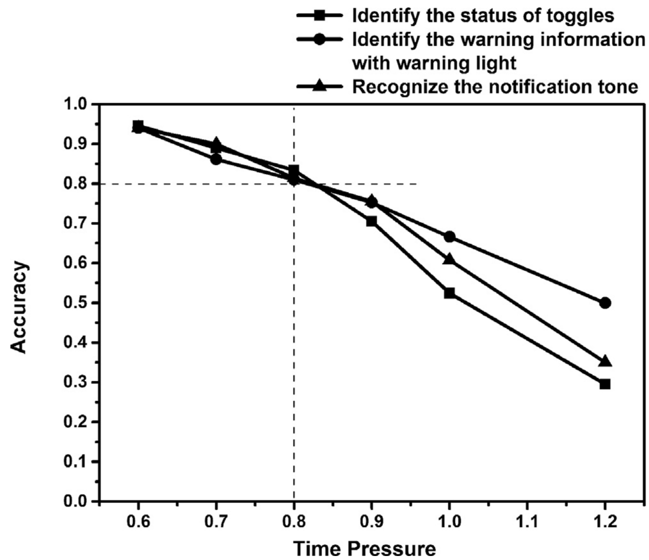

Fig. 2. The change in accuracy with time pressure for three different tasks.

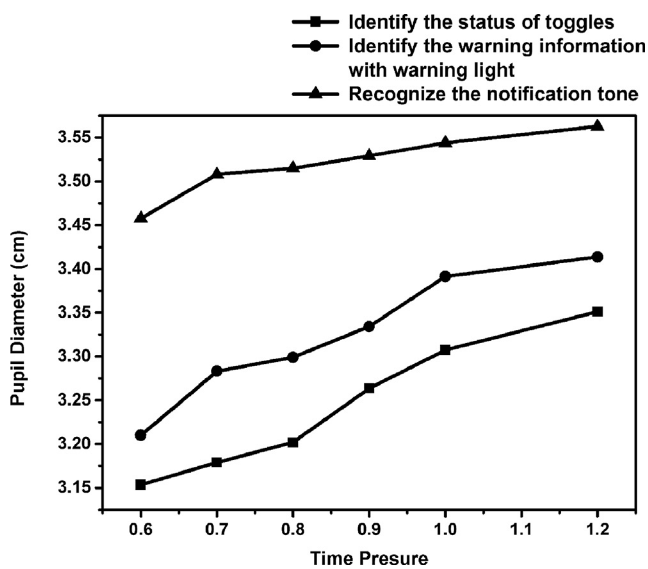

Fig. 3. The change in pupil diameter with time pressure for three different tasks.

the same TP, were averaged. Due to the fact of the limited the number of participants, the 50*th* percentile of the sensory threshold was chosen as the absolute TP threshold. The absolute TP threshold in different experiments was shown in Fig. 4. The analysis of the data could be seen in Table 1. It was demonstrated from Fig. 4 and Table 1 that TP at 0.8 was the sensory threshold for different tasks, i.e., the workload exceeds the human capacity when the ratio of time required to time available is at 0.8.

(2) “Redline” analysis.

The sensory threshold of TP showed that participants felt the workload was influenced obviously when TP reached 0.8. The subjective evaluation workload was normalized and converted to the new standard with ten points as the highest points. The analysis diagrams of subjective evaluation workload, accuracy, and pupil diameter in the same task, are shown in Figs. 5–7. The analysis of the relationship of workload, pupil diameter and accuracy indicates that 1) the accuracy dropped into the range from 85% to 80% when the TP increased to 0.8; and 2) the changes of pupil diameter

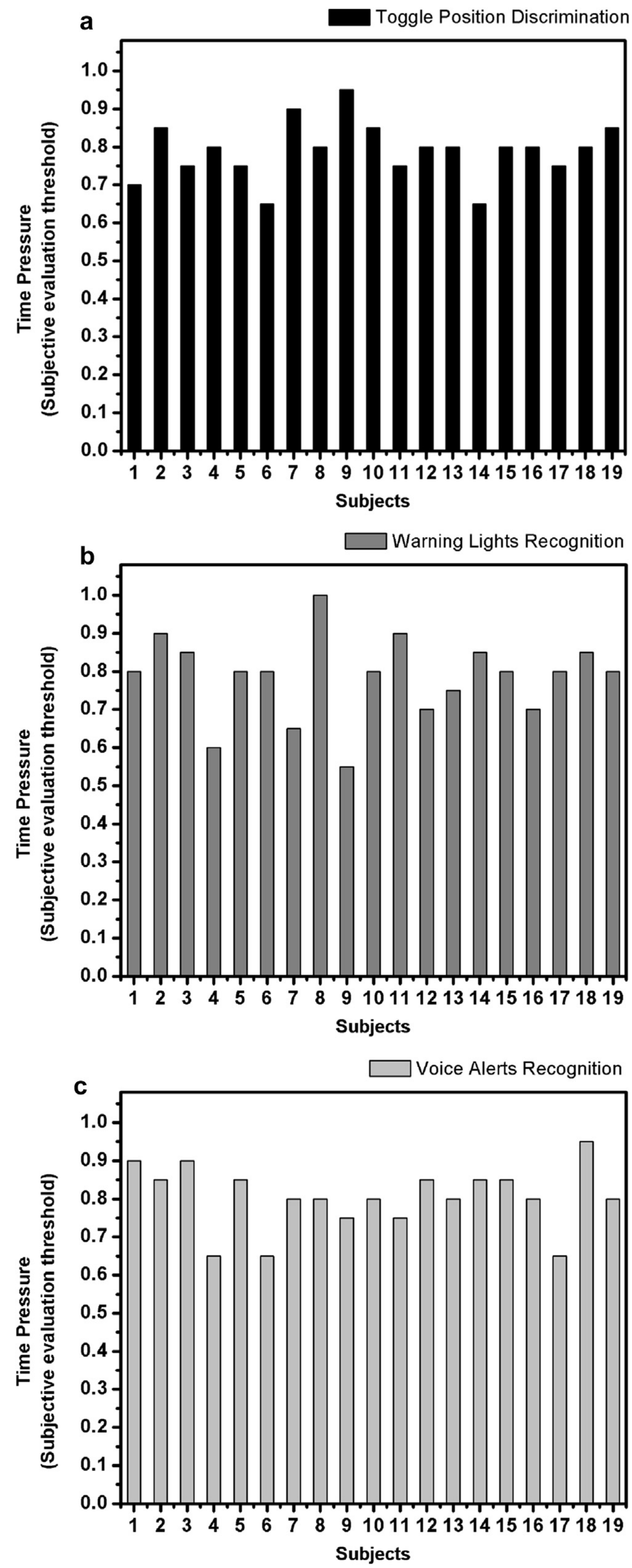

Fig. 4. The diagram of subjective feeling threshold in terms of time pressure.

started to show linear trends when TP reached 0.8. Therefore, for different cognitive tasks, there are no differences in the relationship between time pressure and accuracy, pupil diameter and workload.

Table 1
The subjective feeling threshold of TP over different experiments.

	Identify the status of toggles	Identify the warning information with warning light	Recognize the notification tone
Average	0.79	0.78	0.80
Medium value	0.80	0.80	0.80
Standard deviation	0.076	0.11	0.08
Variance	0.01	0.01	0.01
50*th* percentile	0.80	0.80	0.80

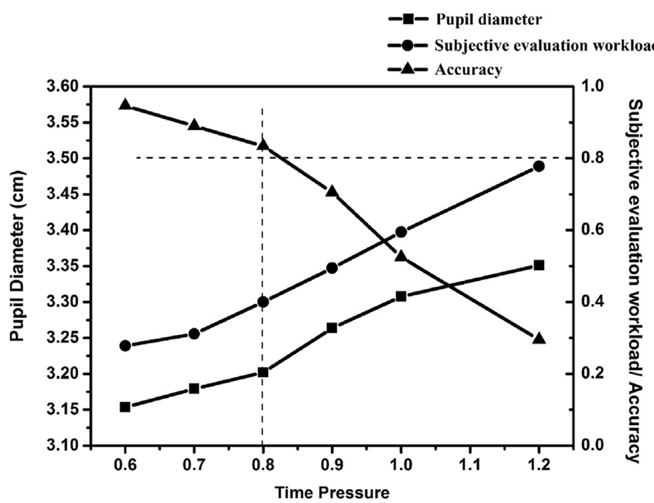

Fig. 5. The workload measurement parameters in Task 1 for defining workload "Redline".

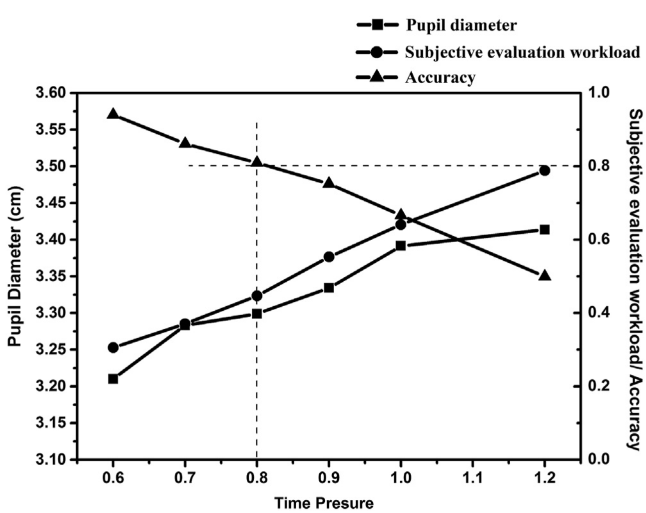

Fig. 6. The workload measurement parameters in Task 2 for defining workload "Redline".

Thus, TP 0.8 could be set as the workload "redline", with an accuracy of 80%–85% and a subjective evaluation workload of approximately 5.

3.4. Relationship models

(1) Subjective evaluation workload and time pressure.

The subjective evaluation workload and time pressure showed a significant linear relationship: $y = -0.279 + 0.903x$, $0.6 \le TP \le 1.2$, where y indicates subjective workload ratings and x indicates time pressure, as shown in Fig. 8.

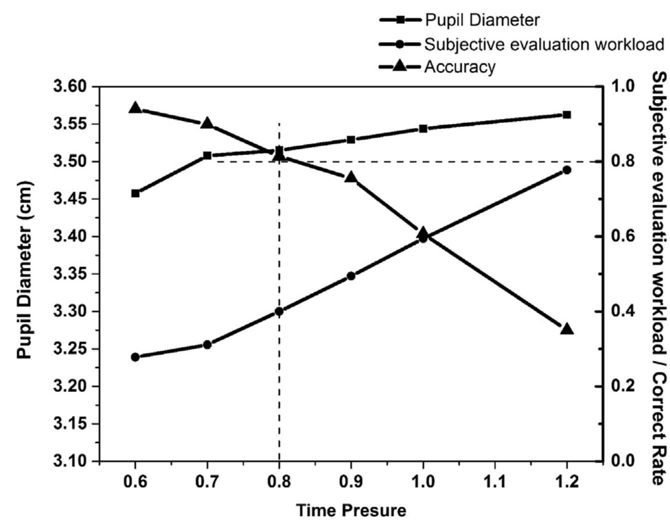

Fig. 7. The workload measurement parameters in Task 3 for defining workload "Redline".

(2) Accuracy and time pressure.

The accuracy and time pressure showed a significant curve linear relationship: $y = 1.012 + 0.309x - 0.699x^2$, $0.6 \le TP \le 1.2$, where y indicates accuracy and x indicates time pressure, as shown in Fig. 9.

(3) Pupil diameter and time pressure.

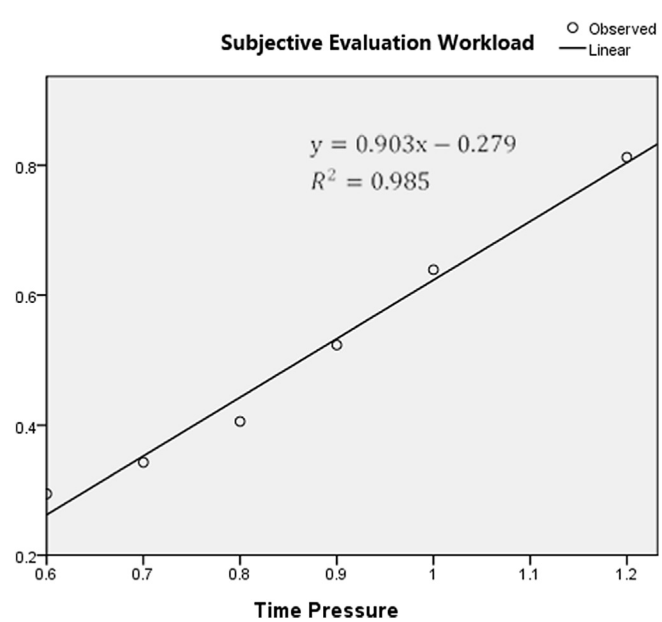

Fig. 8. Relationship model between subjective evaluation workload and time pressure.

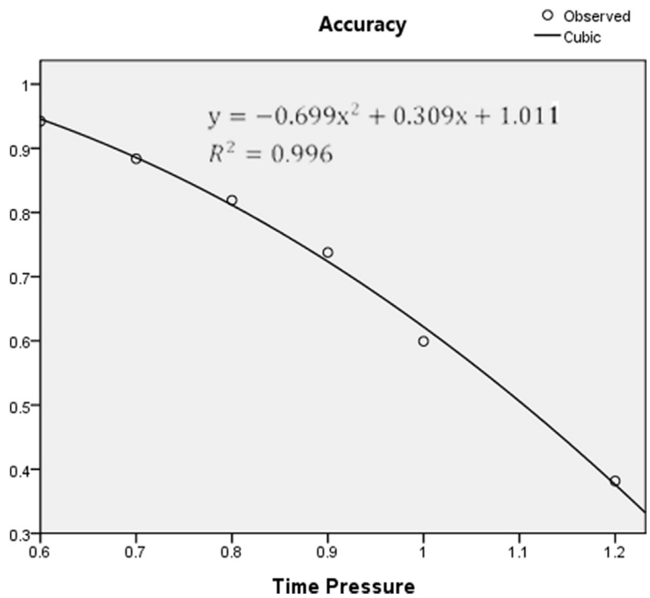

Fig. 9. Relationship model between accuracy and time pressure.

The pupil diameter and time pressure showed a significant linear relationship: $y = 3.117 + 0282x$, $0.6 \le \mathrm{TP} \le 1.2$, where y indicates pupil diameter and x indicates time pressure, as shown in Fig. 10.

4. Discussion

The first research result of this study is the quantitative relationship model between time pressure and workload. When TP is in a small special range, i.e., 0.6 to 1.2, time pressure and subjective evaluation workload show a linear trend in the relationship model; time pressure and accuracy show a quadratic curve trend; time pressure and pupil diameter also show a linear trend. Although many researchers predict workload with time pressure, i.e., the ratio of operating time required to time available (Cullen, 1999), there are no quantitative relationship models that exist for time pressure and workload. Through a detailed study on the definitions of task demand and workload, and their relationships, we found that time pressure is one of the important elements influencing task demand. When the task type remains unchanged, the change in time pressure can lead to different workloads. Thus, the quantitative relationship model between time pressure and workload can be established under this situation.

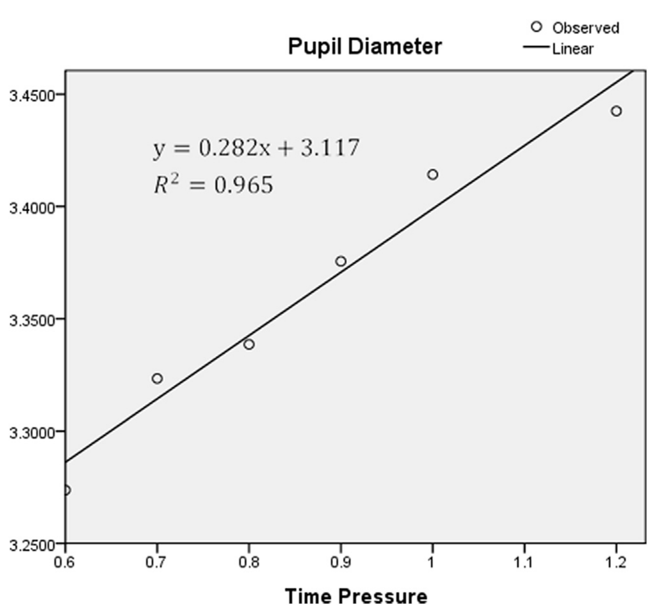

Fig. 10. Relationship model between pupil diameter and time pressure.

The second research result of this study is that there is no difference in the relationship model between time pressure and workload under three different tasks, i.e., there is no effect of task type on the relationship models. Because the task load is not high in the experiment, time pressure is the only factor affecting task demand, thereby increasing the workload. Thus, when the task load is not high, the relationship between time pressure and workload is not affected by task type. When the task load is high, the effect analysis of task type on the relationship model is worthy for further study.

The third achievement of this study is that according to the comprehensive analysis of subjective and objective measurement of workload, the time pressure is the only factor influencing the task demand, and workload "redline" for different tasks all appear to be TP 0.8, with an accuracy of 80%–85%. Schvaneveldt et al. (1997) showed that the value achieved 85–95% accuracy in predicting a "redline" workload level in training data, and for completely new data, the accuracy was in the 70–75% range. Beevis et al. (1994) showed that human reach overload caused the time pressure to be more than 70–80%. In this experiment, the training time is approximately ten minutes, which is between that of the training data and new data, and the measurement accuracy is also between them. However, the redline defined with time pressure is similar to the research results from Beevis. Thus, the result is the same as the study results of workload redline in previous studies that used signal measurement indices.

In this study, the time pressure is an objective parameter, defined as the ratio of time required to complete a task to the time available (Siegel and Wolf, 1969), which is different than in other studies. Monod and Kapitaniak (1999) proposed that the time pressure would activate an emotional component and would thus have an indirect effect on cognitive load. Time pressure involves a conflict between the imposed completion time for a task and the time it actually takes to perform the task, and leads to highly emotional reactions. Koslowsky et al. (1995) demonstrated that time pressure is one of the most common stressors in work environments in which time may be part of a mediating process that influences perception of control. Monod et al. not only defined time pressure as the ratio of time required to complete a task to the time available but also regarded it as a production result with emotion. Koslowsky et al. considered time pressure from the perspective of human feelings.

Accuracy and subjective evaluation workload showed no significant difference over different tasks under the same time pressure. However, pupil diameter showed significant differences with different tasks under different time pressures. However, the average processing can eliminate a certain amount of individual differences in the same task. We know that the pupil diameter can be influenced by many factors, such as lighting, distance, brightness, etc. Thus, in the experiment, we control all the factors influencing the pupil, such as lighting, and distance. Based on Fig. 3 and Figs. 5–7, we found that pupil diameter increases can be very sensitive to the changes in time pressure and task type. Relative to the accuracy and subjective evaluation, the pupil diameter may be more sensitive to the changes of workload, even to the slight changes in task types, so this part of the study will be continued in our research.

Subjective evaluation workload is approximately 5 on the point of workload "redline" determined with multiple parameters, and approximately 8 on TP 1.2 in Figs. 5–7. Workload is usually evaluated from multiple dimensions: time load, effort load, psychological stress load in SWAT (Subjective Workload Assessment Technique, Reid and Eggemeier, 1982), mental demand, physical demand, temporal demand, effort, performance, and frustration level in NASA-TLX (NASA Task Load Index, Hart and Staveland, 1988). In this experiment, workload is derived only from the factor of time pressure, with low task load. Thus, the participants still felt the workload was low even when the performance was declining. In this situation, subjective evaluation workload cannot be used as the reference parameter in workload redline determination.

In the study of psychophysics, the absolute threshold is defined as the stimuli that causes positive reaction of 50% in the experiment (Yang, 1988). Because this experiment could not last too long because of participant fatigue, there are only a few experiments in the process of determining the threshold. The method in the study of psychophysics could not be used. Therefore, the *50th* percentage of the subjective evaluation value of threshold in time pressure was set as the threshold in this article, as shown in Table 1.

5. Conclusions

When time pressure is the main influential factor of the task demand, the quantitative relationship model between workload and task demand under different tasks can be established with the workload measurement parameters. When TP is in a small special range, 0.6 to 1.2, time pressure and subjective evaluation workload shows a linear trend relationship; time pressure and the accuracy shows a quadratic curve trend; time pressure and pupil diameter shows a linear trend.

When the time pressure is the only factor to the task demand, workload "redline" for different task types all appear to be TP 0.8 (i.e., the ratio of time required to complete a task to the time available is 0.8), with an accuracy 80%–85%.

Acknowledgments

This research was supported by the strong support of the 973 Project of National Technological Department (2010CB734103).

References

Beevis, D., Bost, R., Döring, B., Nordø, E., Oberman, F., Papin, J., et al., 1994. Analysis Techniques for Manmachine System Design. AC/243(Panel 8) TR/7 Vol, 2.
Batmaz, I., Ozturk, M., 2008. Using pupil diameter changes for measuring mental workload under mental processing. J. Appl. Sci. 8 (1), 68–76.
Berka, C., Levendowski, D.J., Lumicao, M.N., 2007. EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviat. Space Environ. Med. 78 (Suppl. 1), B231–B244.
Brookhuis, K.A., De Waard, D., 2010. Monitoring drivers' mental workload in driving simulators using physiological measures. Accid. Anal. Prev. 42 (3), 898–903.
Cilliers, M., 1992. An Empirical Investigation into the Measurement of Fixed Wing Fighter Pilot Workload. Doctoral thesis. University of Stellenbosch.
Colle, H.A., Amell, J.R., Ewry, M.E., 1988. Capacity equivalence curves: a double trade-off curve method for equating task performance. Hum. Factors J. Hum. Factors Ergonomics Soc. 30 (5), 645–656.
Colle, H.A., Reid, G.B., 2005. Estimating a mental workload redline in a simulated air-to-ground combat mission. Int. J. Aviat. Psychol. 15 (4), 303–319.
Cullen, L., 1999. Validation of a Methodology for Predicting Performance and Workload. Eurocontrol Experimental Centre, pp. 28–29.
De Waard, D., 1996. The Measurement of Drivers' Mental Workload. Groningen University, Traffic Research Center.
Fairclough, S.H., Venables, L., Tattersall, A., 2005. The influence of task demand and learning on the psychophysiological response. Int. J. Psychophysiol. 56 (2), 171–184.
Fallahi, M., Motamedzade, M., Heidarimoghadam, R., Soltanian, A.R., Miyake, S., 2016. Effects of mental workload on physiological and subjective responses during traffic density monitoring: a field study. Appl. Ergon. 52, 95–103.
Faure, V., Lobjois, R., Benguigui, N., 2016. The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. Transp. Res. Part F Psychol. Behav. 40, 78–90.
Gable, T.M., Kun, A.L., Walker, B.N., et al., 2015. Comparing heart rate and pupil size as objective measures of workload in the driving context: initial look. In: Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. ACM, pp. 20–25.
Gao, Q., Wang, Y., Song, F., Li, Z.Z., Dong, X.L., 2013. Mental workload measurement for emergency operating procedures in digital nuclear power plants. Ergonomics 56 (7), 1070–1085.
Gao, Y., Barreto, A., Zhai, J., et al., 2007. Digital filtering of pupil diameter variations for the detection of stress in computer users. In: Proc 11th World Multi-conference on Systemics. Cybernetics and Informatics, Orlando, pp. 30–35.
Gaillard, A.W., 2001. Stress, workload, and fatigue as three biobehavioral states: a general overview.
García-Mas, A., Ortega, E., Ponseti, J., de Teresa, C., Cárdenas, D., 2016. Workload and cortisol levels in helicopter combat pilots during simulated flights. Rev. Andal Med. Deporte. 9 (1), 7–11.
Glenn, F.W., Penny, F., Iris, D., 1994. Evoked potential, cardiac, blink, and respiration measures of pilot workload in air-to-ground missions. Aviat. Space Environ. Med. 100–105.
Hancock, P.A., Meshkati, N.E., 1988. Human mental workload. North-Holland.
Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. Elsevier Science, Amsterdam, pp. 139–183.
Hilburn, B., Jorna, P.G.A.M., 2001. Workload and air traffic control.Stress, Workload and Fatigue: Theory, Research and Practice. Erlbaum, Hillsdale, New Jersey, USA.
Juris, M., Velden, M., 1977. The pupillary response to mental overload. Physiol. Psychol. 5 (4), 421–424.
Kahneman, D., 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, NJ.
Klingner, J., Kumar, R., Hanrahan, P., 2008. Measuring the task-evoked pupillary response with a remote eye tracker. In: Proceedings of the 2008 Symposium on Eye Tracking Research & Applications. ACM, pp. 69–72.
Koslowsky, M., Kluger, A.N., Reich, M., 1995. Commuting Stress: Causes, Effects, and Methods of Coping. Plenum, New York.
Krüger, A., 2008. A Systems Approach to the Assessment of Mental Workload in a Safety-critical Environment. University of Pretoria.
Mansikka, H., Virtanen, K., Harris, D., Simola, P., 2016. Fighter pilots' heart rate, heart rate variation and performance during an instrument flight rules proficiency test. Appl. Ergon. 56, 213–219.
Mazur, L.M., Mosaly, P.R., Hoyle, L.M., 2013. Subjective and objective quantification of physician's workload and performance during radiation therapy planning tasks. Pract. Radiat. Oncol. 3 (4), 171–177.
Meijman, T.F., O'Hanlon, J.F., 1984. Workload: an introduction to psychological theories and measurement methods. Handbook of Work and Organizational Psychology 1, 257–288.
Meister, D., 1976. Behavioral Foundations of System Development. Wiley, New York.
Mioch, T., Looije, R., Neerincx, M.A., 2010. Developing a model of cognitive lockup for user interface engineering. In: Proceedings of the 10th International Conference on Cognitive Modeling, pp. 157–162.
Monod, H., Kapitaniak, B., 1999. Ergonomie. Masson, Paris.
Neumann, D.L., 2010. Effect of varying levels of mental workload on startle eyeblink modulation. Ergonomics 45 (8), 583–602.
Pomplun, M., Sunkara, S., 2003. Pupil dilation as an indicator of cognitive workload in human-computer interaction. Int. Conf. Humancomputer Interact. 542–546.
Prichard, J.S., Bizo, L.A., Stratford, R.J., 2011. Evaluating the effects of team-skills training on subjective workload. Learn. Instr. 21 (3), 429–440.
Reiner, M., Gelfeld, T.M., 2013. Estimating mental workload through event-related fluctuations of pupil area during a task in a virtual world. Int. J. Psychophysiol 93 (1), 38–44.
Reid, G.B., Colle, H.A., 1988. Critical SWAT values for predicting operator overload. In: Proceedings of the human factors and ergonomics society annual meeting, 32(19). SAGE Publications, pp. 1414–1418.
Reid, G.B., Eggemeier, F.T., Shingledecker, C.A., 1982. Subjective workload assessment technique. AIR FORCE FLIGHT TEST CENTER EDWARDS AFB CA.
Reid, G.B., Nygren, T.E., 1988. The subjective workload assessment technique: a scaling procedure for measuring mental workload. In: Hancock, P.A., Mehtaki, N. (Eds.), Human Mental Workload. North Holland, Amsterdam, pp. 185–218.
Roberts, B.B., Crites, C.D., 1985. Computer-aided cockpit workload analysis for all weather, multirole tactical aircraft.
Schulz, C.M., Schneider, E., Kohlbecher, S., Hapfelmeier, A., Heuser, F., Wagner, K.J., et al., 2013. The influence of anaesthetists' experience on workload, performance and visual attention during simulated critical incidents. J. Clin. Monit. Comput. 28, 475–480.
Schvaneveldt, R.W., Reid, G.B., Gomez, R.L., Rice, S., 1997. Modeling mental workload. New Mexico State Univ Las Cruces Comput.ing Research Lab.
Siegel, A.I., Wolf, J.J., 1969. Man-machine Simulation Models: Psychosocial and Performance Interaction. Wiley-Interscience.
Stuiver, A., De Waard, D., Brookhuis, K.A., 2012. Short-term cardiovascular responses to changing task demands. Int. J. Psychophysiol. 85 (2), 153–160.
Vigo, D.E., Ogrinz, B., Wan, L., 2012. Sleep-wake differences in heart rate variability

during a 105-day simulated mission to Mars. Aviat. Space Environ. Med. 83 (2), 125–130.
Wickens, C.D., McCarley, J.S., 2007. Applied Attention Theory. Taylor & Francis.
Winn, B., Whitaker, D., Elliott, D.B., Phillips, N.J., 1994. Factors affecting light-adapted pupil size in normal human subjects. Invest. Ophthalmol. Vis. Sci. 35 (3), 1132–1137.
Yang, Z.L., 1988. Psychophysics. GanSu People's Publishing House, pp. 15–19.