Worker Fatigue. An Overview of Subjective and Objective Methods of Measurement

Marcin Kuliński, Joanna Koszela-Kulińska and Katarzyna Jach

Faculty of Computer Science and Management
Wrocław University of Technology
27 Wybrzeże Wyspiańskiego, 50-370 Wrocław, Poland

ABSTRACT

Fatigue, as a psychophysiological state, is a normal human condition experienced as a result of mental or physical activity that consumes one's biological resources. In order to relieve its symptoms and to restore human's ability to continue work, a return to a fresh, relaxed state is required. It is usually obtained by an adequate amount of rest and/or sleep, but under the presence of some disadvantageous factors, such as too hot or too cold microclimate, high levels of noise at a workplace, work overload, psychological stress, lack of motivation, and others, e.g., family or health-related problems, this needed amount of rest simply may not fit into a normal 24-hour cycle of activity, resulting in the fatigue built-up. A fatigued worker becomes prone to performance deterioration, loss of concentration, misjudgment of his/her environmental safety, hazardous behaviors, anxiety and sleepiness. Therefore, worker's fatigue management is essential for improving productivity and quality of work, as well as for building worker's satisfaction, motivation and overall wellbeing. To supervise fatigue, appropriate tools for its measurement are required. This paper presents an overview of existing methods of fatigue measurement, including psychological (self-reports and questionnaires), physiological (i.e., changes in various body functions), and performance related (e.g., reaction time, sensory abilities) ones.

Keywords: Human Performance, Fatigue Metric, Fatigue Assessment, Fatigue Management System

INTRODUCTION

A fatigued worker becomes prone to performance deterioration, loss of concentration, misjudgment of his/her environmental safety, hazardous behaviors, anxiety and sleepiness, all resulting in an elevated risk of accidents. Nowadays, high demands in the manufacturing industry, transportation, and public services inevitably lead to situations where adequate dose of sleep, being an essential and single most important fatigue mitigating factor, turn to be a luxury rather than a basic need and a necessity. Therefore, fatigue detection and management mechanisms are vital for improving work safety, labor productivity and quality, as well as for building worker's satisfaction, motivation and overall wellbeing. In this paper selected methods of fatigue assessment are briefly presented with respect to both subjective and objective perspectives. A special attention was put into picking these techniques and means that have a potential of being suitable for an integrated approach, i.e. a fatigue management system.

SUBJECTIVE METHODS OF FATIGUE EVALUATION

Fatigue questionnaires originate from health care, medical and therapeutic territories, where quality of life monitoring plays a significant role in treatment of chronic diseases and other conditions of prolonged physical or mental impairment. Being a psychometric instrument, a fatigue scale or questionnaire must fulfill strict requirements regarding its reliability (an ability to perform measure in a consistent, unbiased by situation, subject or questioner, way) and validity (i.e., measuring what was intended to measure). Two concurrent concepts of psychometric fatigue measurements exist. Unidimensional scales handle fatigue as an atomic, indivisible construct and try to measure its intensity only, while multidimensional ones assume that fatigue has a complex structure and attempt to judge its quality and amplitude across two or more “dimensions”, using relevant subscales.

One of the most frequently used and best validated unidimensional self-report scales is the Fatigue Severity Scale (Krupp et al., 1989). It consists of 9 statements regarding the impact of fatigue on specific types of activity and uses a 7-point Likert scale to assess one's levels of agreement on those. Another one is the Fatigue Assessment Scale (FAS) proposed by Michielsen et al. (2003), which includes 10 statements that refer to personal feelings regarding usual (mental and physical) activities and tiredness perception. As for the multidimensional approach, several reliable and well documented tools exist, including:

- Chalder Fatigue Questionnaire (CFQ, FQ), with 11 items divided between physical and mental symptoms (Chalder et al., 1993)
- Checklist Individual Strength (CIS-20) consisting of 4 subscales with a total of 20 items (Vercoulen et al., 1994)
- Multidimensional Fatigue Inventory (MFI), comprising 20 items divided among 5 subscales (Smets et al., 1995)
- Revised Piper Fatigue Self-report Scale (PFS), with 22 items grouped in 4 dimensions/factors (Piper et al., 1998).

All of above mentioned tools allow for assessing one’s perceived fatigue in less than 5 minutes. A comprehensive review and comparison of existing fatigue questionnaires was done by, e.g., De Vries et al. (2003) and Dittner et al. (2004).

MEASURING PHYSIOLOGICAL CHANGES IN RESPONSE TO FATIGUE

Fatigue affects various basic body functions, such as brain electrical activity, muscles activity, blood pressure, heart rate, or numerous hormones excretion, e.g., melatonin and cortisol (Gartner and Murphy, 1976; Åhsberg, 1998). Moreover, manifestations of these physiological states can be registered at the behavioral level, for example using oculographic (Dinges et al., 1998; Zhang and Zhang, 2010), posturographic (Avni et al., 2006; Karita et al., 2006), facial expressions reading (Vural et al., 2007; Saradadevi and Bajaj, 2008), or voice analysis techniques (Greeley et al., 2006; Krajewski et al., 2008). Selected measures that have a potential to be used as a self-administered fatigue assessment tool or by means of an automated fatigue monitoring system are described below.

Critical flicker fusion rate

An intermittent, flickering light can be perceived as a steady, continuous one when its flickering frequency is higher than the flicker fusion rate (or threshold) of the human vision system. This phenomenon was adopted and is being successfully exploited by the cinematography, where an illusion of movement is obtained via projecting series of still images that switch rapidly, as well as by electric light appliances for home and industry, where energy-efficient gas-discharge lamps are widely used. However, there is no one universal critical flicker fusion threshold, as it fluctuates from subject to subject and from one body state to another. For instance, CFF rate was found to be lower in case of fatigued individuals compared to rested ones (Krugman, 1947).

Nowadays, CFF measures are used frequently by researchers exploring fields of human performance, workload and fatigue. Saito and colleagues (1994) found a relationship between CFF rates and subjectively reported eye fatigue symptoms due to sessions of work with a visual display terminal. A statistically significant difference between pre- and post-work CFF values was found among high-elevation construction workers by Hsu et al. (2008). Various aspects of scientific feasibility of an inexpensive, personal CFF meter (called the Portable Fatigue Meter) were studied by Hosokawa et al. (1997).

PERCLOS

PERCLOS (abbrev. for percentage of eye closure) is a measure of fatigue introduced by Wierwille and colleagues (1994). Their three-year study on vehicle driver performance monitoring led to a conclusion that specific eyelid activity can be a strong predictor of driver's drowsiness, as manifested by slow eye closures, rather than usual blinks, and eyes that stay only partially open for prolonged periods of time. Specifically, they proposed a cut-off value of 80% eyelid closure (where 0% represents eyes wide open and 100% is associated with eyes fully closed) for calculations of the proportion of time when eyes remain fully or almost fully closed, which constitutes the metric.

Initially, percentages of eye closure were judged and quantified (with about 5% accuracy) from a pre-recorded video material by human scorers trained with simulated and actual video footage for several hours (see Dinges et al., 1998). Since then, numerous real-time implementations of PERCLOS have been proposed. Ji et al. (2004) have successfully integrated PERCLOS with several other measures computed from head- and gaze-tracking data and from output of their facial expression analysis algorithm, as well as positively validated the proposed composite fatigue score. Zhang and Zhang (2010) used PERCLOS in concert with a sophisticated eye-tracking method in another real-time fatigue monitoring system. A combined, computer vision based assessment method that incorporates PERCLOS and mouth opening (yawning) recognition, which yielded an increased accuracy in discriminating fatigued and non-fatigued car drivers compared to using percentage of eye closure exclusively, was proposed by Sacco and Farrugia (2012).

PERFORMANCE RELATED METHODS OF FATIGUE ASSESSMENT

Fatigue can be assessed by measuring outcomes of human occupational activity, for example quality and quantity of work performed (Grandjean, 1980). On a more basic level, this work can be scaled down to some simple, yet representative tasks executed more or less routinely during job activities, regardless the type and specificity of one's occupation. Such tasks may involve sensory and psychomotor components, which are fairly easy to measure and quantify. According to Welford (1968), performance deterioration resulting from a fatigue build-up manifests in a form of impaired sensory/perceptual functions, as well as slowed down psychomotor reactions. Several performance tests most widely used in the field of fatigue research are briefly presented and discussed below.

Psychomotor vigilance task

Known also as the psychomotor vigilance test, it was proposed by Dinges and Powell (1985) and since then became a universally accepted tool used in research dealing with behavioral alertness, sleepiness, sleep deprivation and disorders, and fatigue. Moreover, PVT outcome metrics are often employed as validation criteria for other methods and techniques of fatigue assessment, for instance, eye closure and head position monitoring (Dinges et al., 1998), posturographic analysis of body balance (Avni et al., 2006), driving task using an interactive simulator (Baulk et al., 2008), heart rate measures (Heinze et al., 2009), or speech analysis (Greeley et al., 2013). Originally presented as a standalone, hand-held device, it can also be used in a form of software running on a PC (Basner and Dinges, 2011) or on a palm-sized computer, as a PDA, smartphone, etc. (Lamond et al., 2005; Ferguson et al., 2011).

PVT works by measuring simple reaction time, that is, a single type motoric response to a single type stimulus, but in a sophisticated manner. The procedure relies on generating series of stimuli of a chosen type (visual or auditory) and on recording corresponding response times, typically as button presses done by index finger or the thumb, over a predefined period of time. This period is usually 10 minutes long, but other variants of the administration regimen do exist, namely 5-min. and 20-min. ones (see Lamond et al., 2005, and Dinges et al., 1998, respectively). Furthermore, stimuli are presented at random intervals, typically from 2 to 10 seconds, which guarantees that response times are independent from one's experience with the test and frequency of its previous usage.

Besides the most straightforward outcome metric, i.e., mean (or median) response time (RT), numerous scores may be computed from sampled data. Examples of suggested ones include (after Basner and Dinges, 2011): number of lapses (RTs equal to or longer than 500ms are considered as "lapses"), number of false starts (RTs shorter than 100ms are considered as "false starts"), lapse probability (defined as the number of lapses divided by the number of valid RTs), mean response speed (mean 1/RT), fastest 10% of RT, slowest 10% of 1/RT, performance score (computed as 1 minus the number of lapses and false starts divided by the number of valid RTs, including false starts).

Choice reaction time

Choice reaction time differs from simple RT in that it relies upon multiple classes of stimuli and, consequently, several possible ways of responding to them. At its simplest form, presenting one of two different colors (numbers, letters, pictograms, sounds, etc.) will require pressing one of two differently placed/labeled buttons. In general, choice reaction time is a function of the number of possible stimulus/response pairs, as expressed by the Hick-Hyman Law (Welford, 1988), and becomes gradually slower while this number increases.

There is no accepted standard for measuring choice reaction time, therefore its implementations, apparatus, and administration procedures differ greatly amongst researchers. A fundamental difference is the number of stimuli, for instance, 2-choice reaction time was used by Schroeder et al. (1998) in a study involving fatigue assessment for 8- and 10-hour work shifts, as well as by Sallinen et al. (1998) in a research investigating short naps as a countermeasure for performance degradation during night shift work, while 3-choice reaction time was chosen routinely by Takeyama and his colleagues for various studies about night shift work and fatigue (Takeyama et al., 2002; 2005; 2009). Test variations with more stimulus/response pairs have been also employed, as 4-CRT (Lieberman et al., 2002; Davranche et al., 2006), or even 10-CRT (Wesensten et al., 2002). A significant amount of diversity is seen regarding methods of stimuli presentation, as the use of words "true" and "false" (2-CRT), digits from a chosen range (1-3 for 3-CRT, 0-9 for 10-CRT), or solid boxes or circles of uniform color discriminated spatially using a single row or a two-dimensional grid. Similarly, responses may be registered using a regular computer keyboard, purpose-built input boxes, joysticks, etc. Nevertheless, one significant consistency worth noting is that visual stimuli are universally preferred over auditory ones, probably as a consequence of difficulties with creating a larger set of sounds mutually discriminative, which can be categorized in a short time, i.e., these produced only by means of fundamental frequencies (pure tones) and harmonics (complex sounds), but without any other form of temporal modulation.

Compensatory tracking task

A compensatory tracking task is a test suitable for assessing both vigilance and eye-hand coordination. It can be performed either as a one-dimensional or two-dimensional simulation, the latter one being the most frequently used. A computer implementation, the COMPTRACK, was proposed by Makeig and Jolley (1995). The aim is to continually maintain the position of a moving indicator presented in the form of a circular disk and to place it as close as possible to a fixed position reference point (also called a "bullseye") rendered as a thick ring. To obtain this, frequent corrections countermeasuring certain virtual forces that try to move the disk away from the ring are required. One of these forces is of a quasi-random, unpredictable nature, causing a drift that constantly changes its direction and magnitude, while the other one can be predicted, as it acts like the gravity force trying to pull the disk down of an invisible, slippery, bell-like shaped surface, which top is aligned with the ring displayed. Corrections of the disk position are applied using a trackball or a mouse, where both movement speed and direction inputs are constantly used to calculate the countermeasuring force.

Described above form of CTT was used by Van Orden et al. (2000) in a study of fatigue-related changes in various oculographic measures, while other researchers used it, along with EEG monitoring, during sleep deprivation experiments (e.g., Dinges et al., 1998; Makeig et al., 2000). Schenka et al. (2010) applied various pattern recognition and classification algorithms to CTT data obtained from partially sleep deprived subjects and concluded that the test is able to quantify performance decrements due to fatigue utilizing relatively short, 6- or even 4-min. administration protocols.

CONCLUSIONS

Probably the most promising direction in fatigue measurement concepts is expressed by tools that incorporate selected objective fatigue metrics into a computer system, which then can assess one's fatigue levels in an unobtrusive and fully automated manner. A possibility of quantifying and evaluating fatigue in real-time, e.g., from face expressions of a car driver, voice and speech characteristics of air traffic personnel, or from body balance and gait dynamics of an industrial worker is of indispensable value in the context of a fatigue management system and its continuous monitoring requirements. Nevertheless, subjective metrics can not be omitted and excluded from the equation, as such systems need to be widely accepted by supervised humans and perceived as adequately sensitive and specific (in terms of minimized false indication ratios). In other words, assessments generated by a fatigue monitor must correspond with subjective feelings of its users and beneficiaries, and this coherence can be achieved by incorporation of reliable and valid psychometric measurements.

REFERENCES

Avni, N., Avni, I., Barenboim, E., Azaria, B., Zadok, D., Kohen-Raz, R., Morad, Y. (2006). Brief posturographic test as an indicator of fatigue. Psychiatry and Clinical Neurosciences, 60, 340–346

Åhsberg, E. (1998). Perceived fatigue related to work. Arbete och Hälsa 1998:19

Basner, M., Dinges, D.F. (2011). Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. SLEEP, 34(5), 581-591

Baulk, S.D., Biggs, S.N., Reid, K.J., van den Heuvel, C.J., Dawson, D. (2008). Chasing the silver bullet: Measuring driver fatigue using simple and complex tasks. Accident Analysis and Prevention, 40, 396–402

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., Wallace, E.P. (1993). Development of a Fatigue Scale. Journal of Psychosomatic Research, 37(2), 147–153

Davranche, K., Audiffren, M., Denjean, A. (2006). A distributional analysis of the effect of physical exercise on a choice reaction time task. Journal of Sports Sciences, 24(3), 323-329

De Vries, J., Michielsen, H.J., Van Heck, G.L. (2003). Assessment of fatigue among working people: a comparison of six questionnaires. Occupational and Environmental Medicine, 60(suppl I), i10-i15

Dinges, D.F., Powell, J.W. (1985). Microcomputer analysis of performance on a portable, simple visual RT task sustained operations. Behavioral Research Methods, Instrumentation, and Computers, 17, 652–655

Dinges, D.F., Mallis, M.M., Maislin, G., Powell, J.W. (1998). Final report: Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, Report No. DOT HS 808 762

Dittner, A.J., Wessely, S.C., Brown, R.G. (2004). The assessment of fatigue: a practical guide for clinicians and researchers. Journal of Psychosomatic Research, 56(2), 157-170

Ferguson, S.A., Paech, G.M., Dorrian, J., Roach, G.D., Jay, S.M. (2011). Performance on a simple response time task: Is sleep or work more important for miners? Applied Ergonomics, 42, 210-213

Gartner, W.B., Murphy, M.R. (1976). Pilot workload and fatigue: A critical survey of concepts and assessment techniques. National Aeronautic and Space Administration, Washington, DC, Report No. NASA TN D-8365

Grandjean, E. (1980). Fitting the task to the man. An ergonomic approach. Taylor & Francis Ltd, London

Greeley, H.P., Friets, E., Wilson, J.P., Raghavan, S., Picone, J., Berg, J. (2006). Detecting fatigue from voice using speech recognition. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 567-571

Greeley, H.P., Roma, P.G. , Mallis, M.M., Hursh, S.R., Mead, A.M. (2013). Field Study Evaluation of Cepstrum Coefficient Speech Analysis for Fatigue in Aviation Cabin Crew. Office of Aerospace Medicine, Federal Aviation Administration, Washington, DC, Report No. DOT/FAA/AM-13/19

Heinze, C., Trutschel, U., Schnupp, T., Sommer, D., Schenka, A., Krajewski, J., Golz, M. (2009). Operator Fatigue Estimation Using Heart Rate Measures. World Congress of International Federation for Medical and Biological Engineering Proceedings, 25/IV, 930-933

Hosokawa, T., Mikami, K., Saito, K. (1997). Basic study of the portable fatigue meter: effects of illumination, distance from eyes and age. Ergonomics, 40(9), 887-894

Hsu, D.J., Sun, Y.M., Chuang, K.H., Juang, Y.J., Chang, F.L. (2008). Effect of elevation change on work fatigue and physiological symptoms for high-rise building construction workers. Safety Science, 46(5), 833-843

Ji, Q., Zhu, Z., Lan, P. (2004). Real-time nonintrusive monitoring and prediction of driver fatigue. IEEE Transactions on Vehicular Technology, 53(4), 1052-1068

Karita, K., Nakao, M., Nishikitani, M., Iwata, T., Murata, K., Yano, E. (2006). Effect of overtime work and insufficient sleep on postural sway in information-technology workers. Journal of Occupational Health, 48(1), 65-68

Krajewski, J., Wieland, R., Batliner, A. (2008). An acoustic framework for detecting fatigue in speech based Human-Computer-Interaction, in: Computers Helping People with Special Needs, Springer Berlin Heidelberg, 54-61

Krugman, H.E. (1947). Flicker fusion frequency as a function of anxiety reaction; an exploratory study. Psychosomatic Medicine, 9(4), 269-272

Krupp, L.B., LaRocca, N.G., Muir-Nash, J., Steinberg, A.D. (1989). The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Archives of Neurology, 46, 1121–1123

Lamond, N., Dawson, D., Roach, G.D. (2005). Fatigue assessment in the field: validation of a hand-held electronic psychomotor vigilance task. Aviation, Space, and Environmental Medicine, 76, 486–489

Lieberman, H.R., Tharion, W.J., Shukitt-Hale, B., Speckman, K.L., Tulley, R. (2002). Effects of caffeine, sleep loss, and stress on cognitive performance and mood during US Navy SEAL training. Psychopharmacology, 164(3), 250-261

Makeig, S., Jolley, K. (1995). COMPTRACK: A Compensatory Tracking Task for Monitoring Alertness. Naval Health Research Center, San Diego, California, Report No. NHRC-TD-96-30

Makeig, S., Jung, T.P., Sejnowski, T.J. (2000). Awareness during drowsiness: dynamics and electrophysiological correlates. Canadian Journal of Experimental Psychology, 54(4), 266

Michielsen, H.J., De Vries, J., Van Heck, G.L. (2003). Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale. Journal of Psychosomatic Research, 54(4), 345-352

Piper, B.F., Dibble, S.L., Dodd, M.J., Weiss, M.C., Slaughter, R.E., Paul, S.M. (1998). The Revised Piper Fatigue Scale: psychometric evaluation in women with breast cancer. Oncology Nursing Forum, 25, 677–684

Sacco, M., Farrugia, R.A. (2012). Driver fatigue monitoring system using support vector machines. IEEE 5th International Symposium on Communications Control and Signal Processing (ISCCSP), 1-5

Saito, S., Sotoyama, M., Saito, S., Taptagaporn, S. (1994). Physiological indices of visual fatigue due to VDT operation: pupillary reflexes and accommodative responses. Industrial Health, 32(2), 57-66

Sallinen, M., Härmä, M., Åkerstedt, T., Rosa, R., Lillqvist, O. (1998). Promoting alertness with a short nap during a night shift. Journal of Sleep Research, 7, 240–247

Saradadevi, M., Bajaj, P. (2008). Driver fatigue detection using mouth and yawning analysis. IJCSNS International Journal of Computer Science and Network Security, 8(6), 183-188

Schenka, C., Schnupp, T., Heinze, C., Krajewski, J., Golz, M. (2010). The compensatory tracking task: a pattern recognition based approach for classifying vigilance. Proceedings of Measuring Behavior, Eindhoven, The Netherlands, August 24-27, 2010, Vol. 7, 470-472

Schroeder, D.J., Rosa, R.R., Witt, L.A. (1998). Some effects of 8- vs. 10-hour work schedules on the test performance/alertness of air traffic control specialists. International Journal of Industrial Ergonomics, 21, 307-321

Smets, E.M.A., Garssen, B., Bonke, B., De Haes, J.C.J.M. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. Journal of Psychosomatic Research, 39(3), 315-325

Takeyama, H., Itani, T., Tachi, N., Sakamura, O., Suzumura, H. (2002). Psycho-physiological effects of naps during night shifts on morning types and evening types. Journal of Occupational Health, 44(2), 89-98

Takeyama, H., Itani, T., Tachi, N., Sakamura, O., Murata, K., Inoue, T., Takanishi, T., Suzumura, H., Niwa, S. (2005). Effects of shift schedules on fatigue and physiological functions among firefighters during night duty. Ergonomics, 48(1), 1-11

Takeyama, H., Itani, T., Tachi, N., Sakamura, O., Murata, K., Inoue, T., Takanishi, T., Suzumura, H., Niwa, S. (2009). Effects of a modified ambulance night shift system on fatigue and physiological function among ambulance paramedics. Journal of Occupational Health, 51(3), 204-209

Van Orden, K.F., Jung, T.P., Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. Biological Psychology, 52(3), 221-240

Vercoulen, J., Swanink, C., Fennis, J., Galama, J., van der Meer, J., Bleijenberg, G., (1994). Dimensional assessment of chronic fatigue syndrome. Journal of Psychosomatic Research, 38, 383–392

Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., Movellan, J. (2007). Drowsy driver detection through facial movement analysis, in: Human–Computer Interaction, Springer Berlin Heidelberg, 6-18

Welford, A.T. (1968). Fundamentals of skill. Barnes & Noble, New York

Welford, A.T. (1988). Reaction time, speed of performance, and age, in: Central determinants of age-related declines in motor function, Joseph, J. A. (Ed.), The New York Academy of Sciences, New York, 1-17

Wesensten, N., Belenky, G., Kautz, M.A., Thorne, D.R., Reichardt, R.M., Balkin, T.J. (2002). Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. Psychopharmacology, 159(3), 238-247

Wierwille, W.W., Wreggit, S.S., Kirn, C.L., Ellsworth, L.A., Fairbanks, R.J. (1994). Research on Vehicle-Based Driver Status Performance Monitoring; Development, Validation, and Refinement of Algorithms For Detection of Driver Drowsiness. Final Report. Office of Crash Avoidance Research, National Highway Traffic Safety Administration, Washington, DC, Report No. DOT HS 808 247

Zhang, Z., Zhang, J. (2010). A new real-time eye tracking based on nonlinear unscented Kalman filter for monitoring driver fatigue. Journal of Control Theory and Applications, 8(2), 181-188