Association of sleep deprivation with speech volume and pitch

Alfred. L.C ROELEN[1,2] and Rutger STUUT[1]

[1] *Aviation Academy, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands,* [2]*Air Operations Safety Institute, Netherlands Aerospace Centre NLR, Amsterdam, the Netherlands*

Abstract. Research was conducted to determine if alterations in the acoustical characteristics of voice occur after moderate cumulative sleep deprivation. Eight subjects participated in the study. Sleep deprivation was obtained by prescribing four nights of reduced sleep (6 hrs instead of 8). Speech data were obtained with sociometric badges, cognitive and subjective fatigue data were also collected. Speech volume and pitch were found to be significantly different when subjects were sleep deprived. Secondary circadian effects were not observed. The results support the proposition that speech can be used to measure the fatigue state of individuals.

Keywords. Sleep deprivation, circadian rhythm, speech, sociometric badges.

1. Introduction

1.1 Background

Fatigue is defined by the International Civil Aviation Organisation (ICAO) as a physiological state of reduced mental or physical performance capability resulting from sleep loss or extended wakefulness, circadian phase, or workload (IATA/ICAO/IFALPA, 2011). Operators' fatigue is a major safety concern in transportation (NTSB, 1999); an estimated 21% of fatal motor vehicle crashes in the United States involved a drowsy driver (Tefft, 2014). Traditionally, work scheduling is used to manage operator's fatigue; however, duty time limitations imposed by regulations tend to be very rigid and they might limit operational flexibility and efficiency (Hellerström et al., 2010).
Increasing scientific knowledge on human fatigue in combination with safety management principles resulted in the concept of Fatigue Risk Management (FRM) that is being introduced in aviation (IATA/ICAO/IFALPA, 2011). Under an FRM program, duty time limitations are identified by each organisation through its own FRM processes, specific to a defined operational context, and are continually evaluated and updated in response to its own risk assessments and the data the organisation collects. The availability of objective fatigue related data is essential for FRM. However, there is no accepted practical way to obtain this in the operational context. There is a need for a method to allow fatigue measurement in a fast, easy and non-invasive manner (Whitehead, 2009).

1.2 Measuring fatigue in an operational setting

Many efforts for measuring fatigue have been reported in the scientific literature. The most often used physiological measurements in an operational setting are ocular and cardiographic indices. Percentage of eyelid closure (PERCLOS) is considered to be a reliable indicator of the onset of sleep (Wierwille at al., 1994). Saccadic velocity, pupil size and pupil constriction are not under voluntary control (Di Stasi et al., 2013) and are therefore robust against compensation effects. Measuring these eye metrics is minimally intrusive. However, they do not always function reliably and accurately in all illumination conditions and do not always accommodate corrective eyeglasses and sunglasses (Barr et al., 2009). Among the existing cardiographic indices, the heart rate variability (HRV) is cited as a good indicator of sleepiness but HRV is also influenced

by other factors such as exercise and digestion (Fogt et al., 2011).
One of the most widely used neurobehavioral tests in studies of sleep and circadian rhythm is the Psychomotor Vigilance Task (PVT), which is a simple visual reaction time task that tests sustained attention. (Dinges and Powell, 1985; Lamond et al., 2005; Drummond et al., 2005; Basner et al., 2011). The PVT has proven to be a sensitive measure of sleep loss (Dinges et al., 1997) but requires attention of the subject and is therefore not practical in an operational environment.
Subjective information on fatigue can be obtained from subjective rating scales such as the Stanford Sleepiness Scale (SSS) (Hoddes et al., 1973), the Karolinska Sleepiness Scale (KSS) (Åkerstedt and Gillberg, 1990) and the Samn-Perelli Scale (SPS) (Samn and Perelli, 1982). However, operational pressure may result in subjects adjusting their scores.
The results of limited empirical research suggest that fatigue affects speech. Differences between alert and fatigued speech in fundamental frequency, loudness, formants, and duration features were found, (Krajewski et al., 2009), as well as effects on speech rhythm, tone (pitch) and clarity of speech (Morris et al., 1960). Significant differences in speech timing were observed when participants were fatigued (Vogel et al., 2010). A significant reduction was found in the use of voice intonation after sleep deprivation, resulted in more flattened or monotonic voices (Harrison and Horne, 1997).

1.3 Research objective

The objective of the research is to determine if fatigue resulting from moderate cumulative sleep deprivation and circadian influences is associated with speech audio parameters.

2. Methods

Subjects for this research study included 8 students (5 male and 3 female) of the Amsterdam University of Applied Sciences with ages varying from 21 to 27 years (Mean ± SD = 22.4 ± 1.9 years). All subjects were requested to sleep 8 hours (i.e. the recommended amount of sleep according to the National Sleep Foundation; Hirshkowitz et al., 2015) for seven consecutive nights, after which they were subjected to a morning test, starting at 10 AM, and an afternoon test, starting at 2 PM. Due to the circadian rhythms, the morning tests were expected to correspond with a secondary peak in alertness while the afternoon test corresponds with a secondary dip in alertness (Hursh and Van Dongen, 2010). On the following 4 nights the subjects were requested to limit the amount of sleep to 6 hours per night and they were then again tested in the morning and in the afternoon (Figure 1).

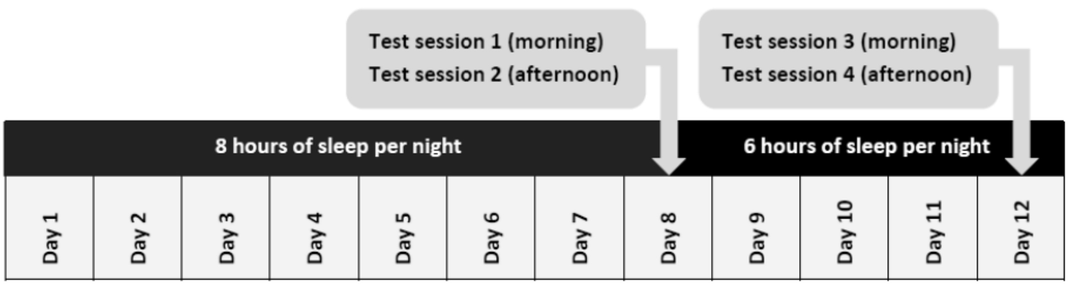

Figure 1: Timing of the experiment

During each test session, subjects performed a PVT, rated their fatigue according to the KSS and held a 20 min. conversation with a researcher to obtain speech samples.

Subjects tracked actual sleeping behaviour during the 12 days of the experiment with a sleep diary. Other than the daily amount of sleep, subjects were asked not to change their normal activities during the 12 days of the full research period. The test sessions took place in a meeting room at the Amsterdam University of Applied Sciences.

2.1 Psychomotor Vigilance Task (PVT)

The reaction time (RT) of the subjects was recorded with a PVT test on a MacBook Air that runs Boot Camp Windows 7. Version 1.1.0 of the PC-PVT software from the Biotechnology HPC Software Applications Institute (BHSAI) was used (Khitrov et al., 2014). The subjects performed three consecutive PVT tests of 180 seconds each during all four test sessions.

2.2 Subjective fatigue

Participants were asked to rate their subjective fatigue according to the KSS during each test session just before performing the PVT test and immediately after the conversation with the researcher.

2.3 Speech data

In order to obtain spontaneous speech samples from the subjects a semi-scripted conversation with one of the researchers was set up. The conversation is based on ‘would you rather’ dilemmas from Rrrather.com; a collection of user-submitted controversial or thought-provoking questions (Rrrather, 2015). These questions start with the words "Would you rather" and always have two possible answers. An example is “Would you rather have more time or more money?” The questions were asked by the researcher and the subjects answered. For each test session a list of 30 ‘would you rather’ dilemmas was prepared resulting in a conversation of approximately 20 minutes.

Speech data (speech volume and pitch) was recorded with wearable electronic badges, called sociometric badges (Olguín Olguín et al., 2009), which were developed by Sociometric Solutions and operate with the 3.1.2063M version of the firmware. Speech volume is defined as the volume of the speech audio sampled by the microphone. It ranges from 0 (silent) to 1 (loud). Pitch is defined as the average frequency of the speech audio sampled by the microphone. The audio signal is recorded with two microphones (one at the front, the other at the back of the badge) with a high-pass filtering cut-off frequency of 85 Hz and a low-pass filtering cut-off frequency of 4000 Hz. The audio signal is captured at 8000Hz. Parameters are averaged over 64 samples (8ms) to ensure that the content of the conversion or the identity of the speaker cannot be determined from the data. Data were further averaged to obtain 1 sample per minute.

3. Results

3.1 Sleep diary

The sleep diaries indicate that subjects slept between 6 and 9 hours (Mean ± SD = 7.9 ± 0.40 hours) in the seven nights days prior to the first day of testing and between 5 and 6.5 hours (Mean ± SD = 6.0 ± 0.23 hours) in the four nights prior to the second day of testing. This indicates that the subjects were indeed sleep deprived at the second day of testing.

3.2 PVT

A Wilcoxon signed-rank test shows that the RTs of the non-sleep deprived and the sleep deprived condition are statistically significant ($Z = -2.379$, $p = 0.017$). Subjects scored significantly poorer when they were sleep deprived. Comparing the RTs between the morning and the afternoon, the Wilcoxon signed-rank test shows that these are not statistically significant ($Z = -0.052$, $p = 0.959$).

3.3 KSS

A Wilcoxon signed-rank test comparing the results of the first and second test day shows that the KSS ratings of the non-sleep deprived and the sleep deprived condition are statistically significant ($Z = -2.582$, $p = 0.010$). Subjects gave a significantly higher fatigue rating when they were sleep deprived. Comparing the KSS ratings between the morning and the afternoon, the Wilcoxon signed-rank test showed that the KSS ratings were not statistically significant ($Z = -0.517$, $p = 0.605$) based on negative ranks (KSS afternoon < KSS morning).

3.4 Speech parameters

A Wilcoxon signed-rank test shows that the speech parameter values of the non-sleep deprived and the sleep deprived condition are statistically significant, see Table 1.

Table 1: Comparing speech parameters for the sleep deprived with the non-sleep deprived condition; results of Wilcoxon signed rank test.

	Front microphone	Back microphone
Pitch	$Z = -2.379$ $p = 0.017$	$Z = -2.482$ $p = 0.013$
Speech volume	$Z = -3.464$ $p = 0.001$	$Z = -3.516$ $p = 0.000$

4. Discussion and Conclusion

The results of the sleep diaries, KSS ratings and PVT scores indicate correspondingly that subjects were sleep deprived during the second day of testing, felt more fatigued and performed less well than during the first day of testing. Speech volume and pitch were significantly different for the sleep deprived condition with a reduction of the speech volume and lower pitch compared to the non-sleep deprived condition. The reduction of speech volume with increased fatigue was also found in previous research for fatigue induced by sustained wakefulness (Krajewski et al., 2009). Using speech analysis for measuring fatigue potentially offers several advantages because it is non-obtrusive, relatively inexpensive and can be used in a wide variety of operational settings (Krajewski et al., 2009).

The results from the KSS ratings and PVT score did not indicate circadian effects and these were also not identified in speech volume and pitch. Core body temperature is the best known indicator of circadian phase (Kräuchi, 2002) but was not measured in this experiment. Therefore it is not known if the time of testing (morning and afternoon) indeed corresponded with a secondary circadian peak and low. Previous research found the circadian cycle in fundamental frequency, but only the primary peak and low and not the secondary (Whitmore and Fisher, 1996). It is therefore recommended to conduct additional measurements to identify if a correlation exists between core body temperature and speech audio.

Operator speech samples can easily be routine collected, for instance pilot speech data is already collected with the cockpit voice recorder. To practically apply this information in the context of fatigue risk measurement, the issues of inter- and intra-individual variation must be better understood. It is therefore recommended to conduct further research on these topics.

References

Åkerstedt, T., Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52: 29–37.

Barr, L., Popkin, S., Howarth, H. (2009). An evaluation of emerging driver fatigue detection measures, Final Report, FMCSA-RRR-09-005, U.S. Department of

Transportation, Federal Motor Carrier Safety Administration, Washington D.C., USA.

Basner, M., Mollicone, D., Dinges, D.F. (2011). Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. Acta Astronautica, 69: 949-959.

Dinges, D.F., Powell, J.W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behavior Research Methods, Instruments, & Computers, 17(6): 652-655.

Dinges, D.F., Pack, F., Williams, K., Gillen, K.A., Powell, J.W., Ott, G.E., Aptowicz, C., Pack, A.I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep, 20(4): 267-277.

Di Stasi, L.L., McCamy, M.B., Macknik, S.L., Mankin, J.A., Hooft, N., Catena, A., Martinez-Conde, S. (2013). Saccadic eye movement metrics reflect surgical resident's fatigue, Annals of Surgery, 00: 1-6.

Drummond, S.P., Grethe, A.B., Dinges, D.F., Ayalon, L., Mednick, S.C., Meloy, M.J. (2005). The neural basis of the psychomotor vigilance task. Sleep, 28(9): 1059-1068.

Fogt, D.L., Cooke, W.H., Kalns, J.E., Michael, D.J. (2011). Linear mixed-effects modelling of the relationship between heart rate variability and fatigue arising from sleep deprivation. Aviation, Space, and Environmental Medicine, 82(12): 1104-1109.

Harrison, Y., Horne, J. (1997). Sleep Deprivation Affects Speech. Sleep 20(10): 871-877.

Hellerström, D., Eriksson, E., Romig, E., Klemets, T. (2010). Flight Time Limitations and Fatigue Risk Management: A comparison of three regulatory approaches. Presented at the European Aviation Safety Seminar 2010, Lisbon, Portugal.

Hirshkowitz, M. et al. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary, Sleep Health 1:40–43.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., Dement, W.C. (1973). Quantification of sleepiness: a new approach. Psychophysiology, 10(4): 431–436.

Hursh, S.R., Van Dongen, H.P.A. (2010). Fatigue and performance modeling. In M. H. Kryger, T. Roth & W. C. Dement (Eds.), Principles and Practice of Sleep Medicine, Fifth Edition (pp. 745-752). Philadelphia: Elsevier Saunders.

IATA/ICAO/IFALPA. (2011). Fatigue Risk Management System (FRMS) - Implementation guide for operators (First edition). International Air Transport Association (IATA), International Civil Aviation Organization (ICAO) and International Federation of Air Line Pilots' Associations (IFALPA).

Khitrov, M.Y., Laxminarayan, S., Thorsley, D., Ramakrishnan, S., Rajaramen, s., Wesensten, N.J., Reifman, J. (2014). PC-PVT: A platform for psychomotor vigilance task testing, analysis, and prediction. Behavior Research Methods, 46: 140-147.

Krajewski, J., Trutschel, U., Golz, M., Sommer, D., Edwards, D. (2009). Estimating fatigue from predetermined speech samples transmitted by operator communication systems. Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, p 468-474.

Kräuchi, K. (2002). How is the circadian rhythm of core body temperature regulated? Clinical Autonomic Research,12: 147–149.

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.

Morris, G.O., Williams, H.L., Lubin, A. (1960). Misperception and disorientation during sleep deprivation. Archives of General Psychiatry, 2: 247-54.

NTSB. (1999). Evaluation of U.S. Department of Transportation Efforts in the 1990s to Address Operator Fatigue, NTSB/SR-99/01, National Transportation Safety Board, Washington D.C., USA.

Olguín Olguín, D., Waber, B.N., Kim, T, Mohan, A., Ara, K., Pentland, A. (2009). Sensible Organisations: Technology and Methodology for automatically measuring organisational behavior. IEEE Transactions on Systems, Man, and Cybernetics, 39(1): 43-55.

rrrather. (2015). What would you rather? Retrieved 17 March 2015 from http://www.rrrather.com.

Samn, S.W., Perelli, L.P. (1982). Estimating aircrew fatigue: a technique with implications to airlift operations. Technical Report No. SAM-TR-82- 21. USAF School of Aerospace Medicine, Brooks AFB, TX. USA.

Tefft, B.C. (2014). Prevalence of motor vehicle crashes involving drowsy drivers, United States 2009-2013, AAA Foundation for Traffic Safety, Washington D.C., USA.

Vogel, A., Fletcher, J., Maruff, P. (2010). Acoustic analysis of the effects of sustained wakefulness on sleep, Journal of the Acoustical Society of America, 128(6):3747-56.

Whitehead, L. (2009). The measurement of fatigue in chronic illness: a systematic review of unidimensional and multidimensional fatigue measures. Journal of Pain and Symptom Management, 37(1):107-128.

Whitmore, J., Fisher, S. (1996). Speech during sustained operations. Speech Communication, 20:55-70.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L. (1994). Research on vehicle-based driver status/performance monitoring; development, validation and refinement of algorithms for detection of driver drowsiness. DOT HS 808 247, U.S. Department of Transportation, National Highway Traffic Safety Administration, Office of Crash Avoidance Research, Washington D.C., USA.