Stationarity of a user's pupil size signal as a precondition of pupillary-based mental workload evaluation

Completed research

Ricardo Buettner, Ingo F. Scheuermann, Christian Koot, Manfred Rössle
Aalen University, Germany
ricardo.buettner@hs-aalen.de

Ingo J. Timm
Trier University, Germany

Abstract. We discuss the concept of stationarity as a precondition of pupillary-based assessments of a user's mental workload and report results from an experiment differentiating stationarity and non-stationarity pupillary size signals.

Keywords: NeuroIS, eye-tracking, mental workload, pupil diameter, stationarity, Augmented Dickey-Fuller test

1 Introduction

While a user's mental workload can be evaluated by pupillary-based eye-tracking [1-8], environmental (e.g. luminescence [9]) and factors other than workload related mental processes (e.g. emotional arousal [10]) also influence a user's pupil size.

In order to detect luminescence changes, which were reflected in sustainable shifts of the user's pupil size, we apply the concept of stationarity analysis. A stationary process is a stochastic process X_t whose probability distribution does not change when shifted in time t: $E(X_t) = E(X_{t-1}) = \mu$; $Var(X_t) = \gamma_0 < \infty$, $Cov(X_{t,} X_{t-1}) = \gamma_k$. While a stationary pupil size signal has been discussed as a precondition for assessing mental workload [11, 12], no pupillary signal related guideline exists.

2 Methodology

2.1 Applying the NeuroIS guidelines

In order to clearly contribute to NeuroIS research and show strong methodological rigor, we followed the NeuroIS guidelines established by vom Brocke et al. [13]. In particular, to assess prior research in the field of measuring mental workload as an important IS concept, a comprehensive literature review was conducted (cf. [14]). To base the experimental design adequately on solid research in related fields of neuroscience [15] we reviewed the fundamental anatomic mechanism of the pupillary dilation controlled by the vegetative nervous system and the key role of the Edinger-

Westphal nucleus that is inhibited by mental workload and directly leads to a pupillary dilation. The methodology uses eye-tracking-based pupillometry as a well-established approach in physiology and psychology for “widening the 'window' of data collection" [15, p. 93]. With this method, bio-data (i.e. pupil diameter) can be used to better understand mental workload as an IS construct (cf. guideline 4 of [14]). In comparison to other neuroscience tools, eye-tracking-based pupillometry is the contact-free and efficient method of choice [16]. I applied the guidelines and standards from Duchowski [17] and the Eyegaze Edge™ manual.

2.2 Measurements

To capture the pupillary diameter, eye-tracking was performed using the binocular double Eyegaze Edge™ System eye-tracker paired with a 19" LCD monitor (86 dpi) set at a resolution of 1280x1024, whereby the eye-tracker samples the pupillary diameter at a rate of 60 Hz for each eye separately.

2.3 Stimuli and test procedure

Following Beatty [18] and Hess & Polt [19] we manipulated mental workload using two well-documented experimental settings in psychology. In the first stage our participants had to memorize and reproduce numbers consisting of three to nine digits [18]. In this stage the luminescence changes on the computer screen were small, with only numbers on a bright background presented, which were interrupted by bright and light green break slides. In the second stage, we showed a fixed order of dark and bright screens [black (5s) → white (5s) → black (5s) → white (2s)] without any mental task. In the third stage the participants had to solve four arithmetic multiplication problems representing a high cognitive demand level as documented [19].

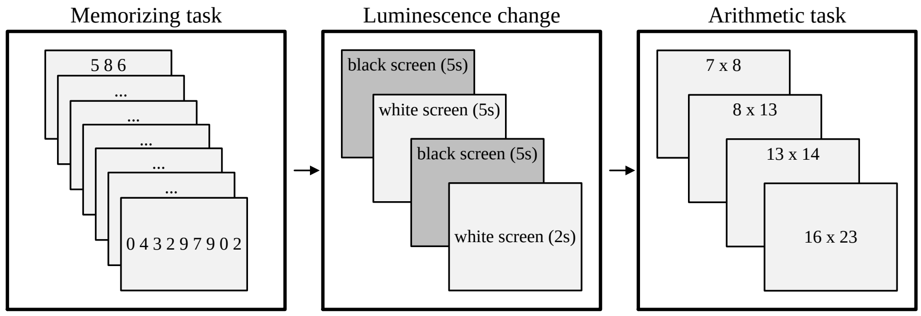

Fig. 1. Test procedure

Prior to all data collection, each test participant was welcomed by the experimenter (supervisor of the experiment). After that the participant was asked to fill out a consent form and also a questionnaire with demographics. After that, we took the necessary precautions for the experiment, in which we make use of the eye-tracking system. Hence, the eye-tracker was calibrated. In the next stage, the experiment began with the memorizing task, followed by the luminescence change stage without mental workload before the participants were invited to solve the last mental workload task (arithmetic) (Fig. 1).

2.4 Augmented Dicker-Fuller test

The Augmented Dickey-Fuller (ADF) test evaluates if a time series variable follows a unit-root process. The null hypothesis is that the variable contains a unit root, and the alternative is that the variable is generated by a stationary process.

To calculate the ADF statistics we used the tseries package within the R x64 3.3.3 environment [20].

3 Results

3.1 Sample characteristics

Twelve volunteers (six females) aged from 21 to 38 (M=26.2, SD=4.1) participated.

3.2 Stationarity test results

Table 1 shows the Augmented Dickey-Fuller t-statistic test results while small p-values suggest that the pupillary size time series is stationary. Vice versa, large p-values indicate non-stationarity.

Table 1: Stationarity test results

Parti-cipant	Mental workload task (memorizing)		Luminescence change (no mental workload)		Mental workload task (arithmetic)	
	left eye	right eye	left eye	right eye	left eye	right eye
1	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.48$)	Non-stationary ($p = 0.52$)	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)
2	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	NA (to less data points)	NA (to less data points)	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)
3	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.59$)	Non-stationary ($p = 0.33$)	Stationary ($p < 0.05$)	Stationary ($p < 0.1$)
4	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.50$)	Non-stationary ($p = 0.69$)	Stationary ($p < 0.01$)	Stationary ($p < 0.05$)
5	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.62$)	Non-stationary ($p = 0.47$)	Stationary ($p < 0.05$)	Stationary ($p < 0.1$)
6	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.39$)	Non-stationary ($p = 0.73$)	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)
7	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.54$)	Non-stationary ($p = 0.66$)	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)
8	Stationary ($p < 0.01$)	Stationary ($p < 0.01$)	Non-stationary ($p = 0.50$)	Non-stationary ($p = 0.74$)	Stationary ($p < 0.01$)	Stationary ($p < 0.05$)
9	Stationary	Stationary	Non-stationary	Non-stationary	Stationary	Stationary

	(p < 0.01)	(p < 0.01)	(p = 0.28)	(p = 0.43)	(p < 0.01)	(p < 0.01)
10	Stationary (p < 0.01)	Stationary (p < 0.01)	Non-stationary (p = 0.37)	Non-stationary (p = 0.31)	Stationary (p < 0.01)	Stationary (p < 0.01)
11	Stationary (p < 0.01)	Stationary (p < 0.01)	Non-stationary (p = 0.44)	Non-stationary (p = 0.42)	Stationary (p < 0.01)	Stationary (p < 0.01)
12	Stationary (p < 0.01)	Stationary (p < 0.01)	Non-stationary (p = 0.69)	Non-stationary (p = 0.80)	Stationary (p < 0.05)	Stationary (p < 0.1)

4 Discussion, limitations and future research

As shown in Table 1, the ADF test identified with 100 percent accuracy whether the related pupillary size time series is stationary or not.

The results can be used in NeuroIS research evaluating the stationarity of pupillary signals to exclude unidirectional luminescence changes, typically caused by sunrise, sunset or monitor brightness changes – which regularly occurred in non-laboratory environments [1, 21]. The method could be applied as a precondition of mental workload assessment excluding some biases in pupil size. However, it should be mentioned that short-time balanced pupil changes, for instance caused by a user's emotions [10] or the pupils' light reflexes [22] cannot be detected using this method. Spectral analysis based procedures such as the calculated Index of Cognitive Activity [22-24] could subsequently be applied to exclude the effect of pupils' light reflexes.

Since we only compared two extreme scenarios (either a mental workload task without a substantial luminescence change or a substantial luminescence change without a mental workload task) future work should verify the ADF results through a two times two factor design (mental workload task yes/no times substantial luminescence change yes/no).

References

1. R. Buettner, The relationship between visual website complexity and a user's mental workload: A NeuroIS Perspective. In Information Systems and Neuro Science, vol. 10 of LNISO, pp. 107-113. Gmunden, Austria (2016)
2. R. Buettner, Analyzing Mental Workload States on the Basis of the Pupillary Hippus, in NeuroIS '14 Proc., p. 52 (2014)
3. R. Buettner, S. Sauer, C. Maier, and A. Eckhardt, Towards ex ante Prediction of User Performance: A novel NeuroIS Methodology based on Real-Time Measurement of Mental Effort, in HICSS-48 Proc., pp. 533-542 (2015)
4. R. Buettner, Cognitive Workload of Humans Using Artificial Intelligence Systems: Towards Objective Measurement Applying Eye-Tracking Technology, in KI 2013 Proc., ser. LNAI, vol. 8077, pp. 37-48 (2013)
5. R. Buettner, B. Daxenberger, A. Eckhardt, and C. Maier, Cognitive Workload Induced by Information Systems: Introducing an Objective Way of Measuring based on Pupillary Diameter Responses, in Pre-ICIS HCI/MIS 2013 Proc., 2013, paper 20 (2013)

6. A. Dimoka, P. A. Pavlou, and F. D. Davis, NeuroIS: The Potential of Cognitive Neuroscience for Information Systems Research, ISR, vol. 22, no. 4, pp. 687-702 (2011)
7. A. Dimoka, R. D. Banker, I. Benbasat, F. D. Davis, A. R. Dennis, D. Gefen, A. Gupta, A. Ischebeck, P. H. Kenning, P. A. Pavlou, G. Müller-Putz, R. Riedl, J. vom Brocke, and B. Weber, On the Use of Neurophysiological Tools in IS Research: Developing a Research Agenda for NeuroIS, MISQ, vol. 36, no. 3, pp. 679-A19 (2012)
8. R. Riedl, R. D. Banker, I. Benbasat, F. D. Davis, A. R. Dennis, A. Dimoka, D. Gefen, A. Gupta, A. Ischebeck, P. Kenning, G. Müller-Putz, P. A. Pavlou, D. W. Straub, J. vom Brocke, and B. Weber, On the Foundations of NeuroIS: Reflections on the Gmunden Retreat 2009, CAIS, vol. 27, pp. 243-264 (2010)
9. S. R. Steinhauer, G. J. Siegle, R. Condray, and M. Pless, Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing, International Journal of Psychophysiology. vol. 52, no. 1, pp. 77-86 (2004)
10. J. L. Bradshaw, Pupil Size as a Measure of Arousal during Information Processing, Nature, vol. 216, no. 5114, pp. 515-516 (1967).
11. T. Knapen, J. W. de Gee, J. Brascamp, S. Nuiten, S. Hoppenbrouwers, and J. Theeuwes, Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance, PLoS One, vol. 11, no. 5, e0155574 (2016)
12. A. Czajka. Pupil dynamics for iris liveness detection, IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 726-735 (2015)
13. J. vom Brocke and T.-P. Liang, Guidelines for Neuroscience Studies in Information Systems Research, JMIS, vol. 30, no. 4, pp. 211-234 (2014)
14. J. vom Brocke, A. Simons, B. Niehaves, K. Riemer, R. Plattfaut, and A. Cleven, Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research, CAIS, vol. 37, no. 1, article 9 (2015)
15. S. D. Goldinger and M. H. Papesh, Pupil Dilation Reflects the Creation and Retrieval of Memories, Curr Dir Psychol Sci, vol. 21, no. 2, pp. 90-95 (2012)
16. B. Laeng, S. Sirois, and G. Gredebäck, Pupillometry: A Window to the Preconscious? Perspect Psychol Sci, vol. 7, no. 1, pp. 18-27 (2012)
17. A. T. Duchowski, Eye Tracking Methodology: Theory and Practice, 2nd ed. London, UK: Springer (2007)
18. J. Beatty, Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. Psychological Bulletin, vol. 91, no. 2, pp. 276-292 (1982)
19. E. H. Hess, and J. M. Polt, Pupil Size in Relation to Mental Activity during Simple Problem-Solving. Science, vol. 143, no. 3611, pp. 1190-1192 (1964)
20. R. Core Team, A Language and Environment for Statistical Computing. Austria: R Foundation for Statistical Computing Vienna (2017).
21. R. Buettner, A user's cognitive workload perspective in negotiation support systems: An eye-tracking experiment. In PACIS 2016 Proceedings (2016)
22. S. P. Marshall, The Index of Cognitive Activity: measuring cognitive workload, In Proceedings of the 7th IEEE Conference on Human Factors and Power Plants, Scottsdale Arizona, September 15-19, pp. 7–5-7–9 (2002)
23. S. P. Marshall, Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity, U.S. Patent 6,090,051 (2000)
24. S. P. Marshall, Identifying Cognitive State from Eye Metrics, Aviation, Space, and Environmental Medicine 78 (Supplement 1), pp. B165–B175 (2007)