Measuring Cognitive Load: Heart-rate Variability and Pupillometry Assessment

Nerea Urrestilla Anguiozar
École de Technologie Supérieure
Montréal, Canada
nerea.urrestilla-anguiozar.1@ens.etsmtl.ca

David St-Onge
École de Technologie Supérieure
Montréal, Canada
david.st-onge@etsmtl.ca

ABSTRACT

Cognitive load covers a wide field of study that triggers the interest of many disciplines, such as neuroscience, psychology and computer science since decades. With the growing impact of human factor in robotics, many more are diving into the topic, looking, namely, for a way to adapt the control of an autonomous system to the cognitive load of its operator. Theoretically, this can be achieved from heart-rate variability measurements, brain waves monitoring, pupillometry or even skin conductivity. This work introduces some recent algorithms to analyze the data from the first two and assess some of their limitations.

KEYWORDS

cognitive load, pupillometry, heart-rate variability

ACM Reference Format:

Nerea Urrestilla Anguiozar and David St-Onge. 2020. Measuring Cognitive Load: Heart-rate Variability and Pupillometry Assessment. In *Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3395035.3425203

1 INTRODUCTION

Robotic systems range over application scenarios that are quickly expanding. In some cases due to the lack of manpower, or to achieve tasks too tedious or dangerous for humans, or for entertainment, robots must understand new challenging environments. In many cases, the most robust deployment option is to rely, at least partially, on a human operator to monitor and control the robotic system.

Research at the INIT Robots laboratory[1] focuses on robot group control [23, 24], adding multiple layers to the complexity of the control to the operator. The operator needs to manage the coordination between the robots, the level of information exchanged between them, and he may be required to contribute to collision avoidance and task allocation. Furthermore, understanding the behavior of a group can be more confusing [17]. In order to design a robotic system centered on the user, the developers are required to assess the influence of these aspects: it can lead to high physical and mental stress for the operator. No stranger to the cognitive load theory applied in human-computer interaction [12], interfaces for multirobot systems aim at increasing their usability: allow novice users to get comfortable quickly and optimize the level of effort a trained user must invest to complete a task.

One of the main effects at play is that operators usually suffer from attentional selectivity [21]; meaning they cannot be fully invested on a single task (or robot-task) when they deal with other duties at the same time. The impact of lacking attention can rise up to high risk for tasks where small mistakes may lead to big fatalities, such as pilots of large vehicles. Many interaction modalities and control strategies have been proposed to help cope with that, but at the core of it stands the need to assess the cognitive load of the operator.

Computing the cognitive load related to a single controlled task is already difficult as it can be influenced by many physiological and psychological factors foreign to the task itself, such as personal concerns to the user and socio-cultural background.

There are many devices that measure the cognitive effort, but each with severe limitations. An electroencephalogram, for instance, can be invasive and uncomfortable while heart-related measures are sensitive to movements. In order to achieve cognitive load measurement in the wild, i.e. outdoor while manipulating the robots in a realistic context, the selected equipment must be portable, easy to deploy, flexible and robust.

The two most promising cognitive load metrics, which also fulfill the criteria of being non-invasive, wearable, reliable and inexpensive, are pupillometry and heart-rate variability. We tested the first in an outdoor setup discussed in [24] and quickly noticed the challenges of a non-controlled environment.

The heart-rate variability is a well-studied metric and was shown to work well in certain scenarios [22], but it is sensitive to variance in the electromagnetic field and breathing cycle [7]. Recent high frequency micro-camera grant the possibility to use pupil information that acts as a proxy of mental effort [25], but wearable (glasses) are noisy and sensitive to the environment [3].

In this paper, we present a methodology to deploy pupillometry and heart-rate variability for cognitive load estimation, while performing an *n*-back task procedure. We implement and test common algorithms for both metrics and discuss their performance. Our main goal is to see if both measures are equivalent and if they align with the NASA TLX results. We will use the data collected to assess their robustness to the lightning conditions.

This paper is organized as follows: Section 2 covers the main related works to our experiment, the methods to analyze the data

[1]https://initrobots.ca/en

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-8002-7/20/10…$15.00
https://doi.org/10.1145/3395035.3425203

will be explained in Section 3, followed by the experiment description and the results from a small sample in Section 4. We conclude with a discussion on the obtained results and future works.

2 RELATED WORKS

We focus this work on pupil size and heart-rate measurement, two metrics covered by a handful of major contributions on cognitive load assessment. For more information on cognitive load, we refer the reader to the work of *Heard et al.* [10]. They reviewed state-of-the-art algorithms for assessing human mental effort level, some of which can adapt to individual differences.

2.1 Pupillometry

The spontaneous variation of the pupil diameter, also known as pupillometry, is used to estimate (interpolate) the pupil dilation. The pupil measure is a sensitive physiological factor that indicates the activity of the Autonomic Nervous System (ANS) [19], which is a control system that acts unconsciously to regulate body functions such as breathing, pupil response and heart activity. The ANS also keeps the body alert to respond to any external stimulus; to defend itself. Thus, pupil variation is also linked to other physiological elements: the pupils dilate with inspirations, decrease with expiration, and pulse can be detected from it.

Pupillometry analysis is certainly one of the newest additions to the quantitative cognitive metrics, but already has undergone important research. The pioneer work of *Murata et al.* aimed at investigating if the fluctuation rhythm of pupil diameter can be used to evaluate mental load [19]. They concluded that it can be as good as the well-known heart-rate variability for that purpose. Another of these contributions, by *Peysakhovich et al.* [21], evaluates the influence of the cognitive load by recording the pupil diameter while performing a visual task, when also subject to auditory distractions.

We now know that the pupil diameter changes according to various characteristic of the user state, namely amusement, surprise, his/her current task complexity, but also following environmental factors such as light. In order to measure only task-related influence, one must first get rid of the influence of all other factors, or otherwise control them. To cope with the complex nature of the pupil data and generate adaptive models, *Buettner et al.* [4] suggest a Random Forest learning algorithm to infer the user's effort and adapt live the device he/she is manipulating. However, the trained algorithm is far from universal and requires new datasets for each user-task pair.

A novel frequency based analysis from the pupil oscillation done by *Duchowski et al.* is meant to be an indicator of cognitive load [6]. The main motivation for their new indicator is that the pupil diameter is susceptible to light and to camera angles, leading to misunderstandings on the cognitive load result. In addition, it is presented as an alternative to baseline-related measures.

According to the state-of-the-art on pupillometry cognitive load detection, pupil diameter suffers big influence on light, so newer algorithms are being explored with the use of wavelet transforms and in the frequency domain.

2.2 Heart-Rate Variability

Heart-rate variability (HRV) was deeply studied as a cognitive load indicator [22]. The autonomic nervous system contains the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS), which give respond to the body on activity and rest situations respectively. SNS and PNS are reflected on the frequency analysis of the HRV as low-frequency (LF) and high-frequency (HF). And the ratio between LF and HF is commonly used to get the sympathovagal balance, related to the cognitive activity of a user. The HRV is the physiological parameter of the variation in the time interval between consecutive heartbeats in milliseconds. This value is reduced during stress phases, so that the SNS can keep up to the body demand.

When HRV is used in sensitive medical application, it was shown that the influence of respiration can lead to wrong diagnostics and so should be removed from the HRV analysis [16]. In [1], they provide several mechanisms to cope with the respiration effect on the data.

The relation between HRV and mental workload has been a field of study for many years. In [20], they searched for cognitive load evidence on HRV and respiratory rate, with the use of three different tasks, one of them being the *n*-back test. They conclude that both parameters suffer the effect of load. Similar study was followed in [11], where they searched for mental workload in ECG during driving task and *n*-back test.

In order to improve the fact that the common LF/HF ratio provides only a single degree of freedom, *von Rosenberg et al.* propose a joint treatment of the LF and HF powers in HRV within a two-dimensional representation framework [27]. In the frequency domain, contributions of SNS and PNS are manifested in separate frequency bands, enabling a more detailed examination of the HRV. They state that their system can distinguish between different cognitive load levels.

In this paper, the joint treatment was not considered and simple LF/HF was used instead, as only a binary load level is defined, and because of the small sample of subjects in the experiment.

3 METHODS

Before detailing the experiment and its results, this section will cover how to manage the selected measurements: pupil diameter and Electrocardiogram (ECG) data, to compute cognitive load indexes.

3.1 Pupillometry

The pupillometry is gathered from Pupil Labs Pupil Core glasses. It is an eye-tracking platform that includes an open-source software suite and a wearable eye tracking headset (see Fig. 1). The software allows to stream all raw data as well as the pupil gaze and diameter. A customized version of the software allows us to record the data in the Robotic Operating System (ROS) environment and thus to synchronize the data with other sources. This synchronization is, however, not mandatory for this work.

The recorded data had to follow a filtering process. Data was normalized in order to have 60 seconds length for all the condition participants. Then outliers were removed, considering only values between 2 and 9 mm, interpolating the ones deleted. Values further

Figure 1: Left: A team member adjusting the eye camera. Right: Pupil Core glasses from Pupil Labs.

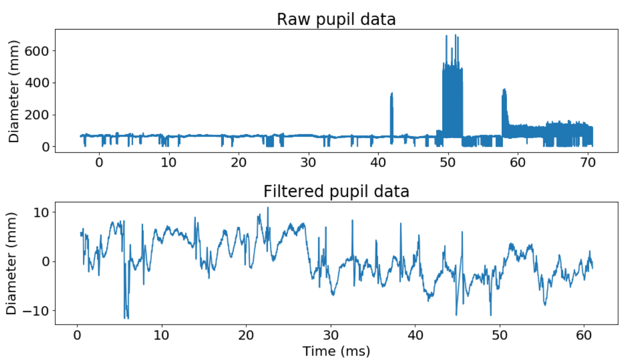

Figure 2: Top: Raw pupil data from the Pupil Core system captured at 200Hz. Down: Filtered data with a low-pass filter, outliers and median removed and cropped at 60s.

than 2 standard deviations were also extracted. The remaining signal was filtered with a butter low pass filter of order 3 and 10 Hz of cutoff frequency [14]. Figure 2 shows the pupil data, with the top plot of the raw output from the Pupil Core system, and the one below showing the filtered signal (cropped, outlier removal, filtered, removed median).

And finally, in order to help with comparing the pupil values between candidates, the median was removed. This last step enables the transfer and comparison between users pupil diameters, as each one has different baseline and reactive diameters [15].

Once all the data is filtered, the average for each condition is calculated, as the main target is to study the pupil diameter reaction to our conditions. From each experiment completed a single average diameter value is obtained. Higher diameter values denote higher cognitive load, and lower values lower load [15].

3.2 Heart-rate variability

This measurement was made from the Biopac MP35. The device captures the analog signals coming from the electrodes placed on the body and outputs digital signal to be processed by their proprietary *AcQKnowledge* software.

In order to measure the heart-rate variability, a single analog channel was used with three electrodes placed on the body as illustrated in Fig. 3: right arm (white), right leg (black) and left leg

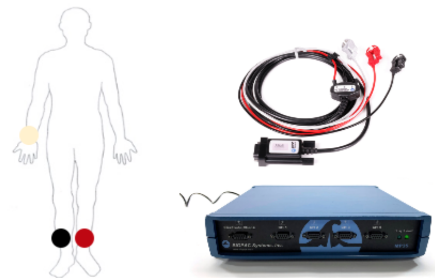

Figure 3: Left: Electrodes localization in the body. Right: The MP35 device with the electrodes.

(red). This configuration enables a triangulation to calculate the electrocardiogram (ECG).

The Biopac software can manage a lot of the filtering and processing of the data, but we only benefit from three channels: the raw ECG, the filtered ECG and the RR peaks intervals. The latter defines the time elapsed between two successive R peaks of the ECG. All are acquired at a sampling rate of 1000 Hz.

The Biopac wired electrodes are used as a proof of concept, to assess how HRV permits to distinguish cognitive load states and relate it to the pupillometry results. For further developments portable ECG measurements will be used, such as Polar and TEA Ergo chest bands, enabling a more realistic setup for in-the-wild recordings but also incorporating measure-related inaccuracies [2].

In this study the filtered ECG was used and then, with the help of the Python *pyHRV* library, we computed the R peaks interval. In Fig. 4 the filtered ECG is at the top and the RR peaks from *pyHRV* below. When using the RR intervals directly from BioPac, one must re-sample the 1000 Hz signal to 4 Hz or 10 Hz [5], to be able to use the most common HRV algorithms and Python libraries (pyHRV, hrvanalysis, Biosppy, etc.). However, resampling has large effect on the frequency response of the signal and we preferred to ensure no artifact were added to the signal. We thus use directly the ECG signal from BioPac and feed it to the *pyHRV* functions.

The sampling frequency of the incoming data is 1000 Hz. And after obtaining the RR peaks, the Python *pyHRV* library [9] is used to analyze the data on the frequency domain and get the LF/HF ratio. This ratio is already known to illustrate the cognitive activity. As the mental effort increases, this ratio will decrease.

3.3 Hypothesis

The hypothesis is that the light should have an effect on the pupil diameter value (as stated in [6]) and we expect a difference between rest and load states (as shown in [21]).

From the heart-rate variability, we expect light not to have any effect and cognitive load difference should be visible (similar to [22]).

4 EXPERIMENTS

With the current challenging pandemic context, the study was kept small and limited to a close circle of people surrounding the main researcher. In the end, the experiments were performed on

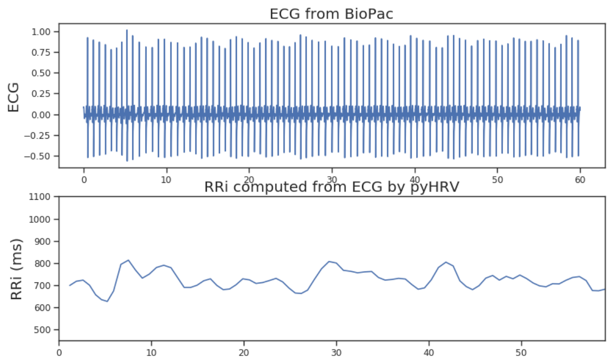

Figure 4: Top: ECG Filtered signal from Biopac captured at 1000Hz. Down: RR intervals computed by *pyHRV* Python package.

10 students involved in this research program. This homogeneous group of participants presents severe limitation to the applicability of our results, but we focus on the methodology of the experiment.

Each participant was asked to perform two sessions, one resting on a chair, without any mental task to achieve and another following the *n*-back test. This method is commonly used as an assessment in physiology and cognitive neuroscience to estimate the working memory capacity [18]. Our implementation relies on the researcher enumerating out loud numbers to the user for one minute, while the user needs to speak up when a number was already mentioned before. This forces the user to keep an active mental workload in order to remember the previous numbers.

The *n* selected was 3, meaning that the participant needs to check if the number currently indicated was repeated 3 mentions before [13]. We selected the *n*-back task as it was previously used in several cognitive load tests with many variations [26], and can now be considered a baseline.

On top of comparing two distinct task loads, we also compared two different lightning conditions: a bright room lightning (neon) and a full dark one. Bringing up 4 different measures from each participant.

According to the conditions above, we did four recordings for each participant: as light condition and mental workload needed to be compared. First light with rest and light with workload were performed. And then the same was repeated without light. The same order was kept through the recording process. Following the recommendation in [1], in order to increase the likelihood that respiration rates will not differ between conditions, the participants were asked to stay sat in the same position.

A detailed description of the structure of the experiment was explained to each participant, and there was not a prior test of the system before the recordings.

4.1 Results

To analyze the results obtained from the experiments above, we first tried a two-way analysis of variance, but several of the test assumptions were not met. So, we had to rely on within-subjects nonparametric tests instead and so we run Wilcoxon tests that enables comparisons between two paired groups. The Wilcoxon test essentially computes the difference between sets of pairs, and studies the differences to determine if both groups are statistically different one from the other. The common null hypothesis is that both groups are the same. We selected a threshold of 5% to reject the null hypothesis.

Table 1 shows, the results of cognitive load for each measure in load and rest conditions. On the pupillometry analysis, the p-value between light and no light conditions with load is 0.241: the diameter does not suffer any effect from luminance because both conditions (light and no light) are equal. This opposes to most of the experiments done in the reference works. However, the boxplot of the pupil (Fig. 5 to the right) shows more variance for no light conditions, which confirms that the measurement is indeed affected by ambient lightning conditions. The small sample size, may lead to not enough statistical power to firmly accept or reject the Null hypothesis, which will explain this incongruity in the results regarding the luminance effect.

The pupillometry analysis (see Table 1) for load and rest conditions with light leads to a p-value of 0.046, which indicates a significant difference between load and rest.

As for the HRV analysis, the results are similar to the pupil (see Table 1). The p-value between light and no light on load case is 0.284, showing no effect of the light. However, for this measure, all variances are similar. Then the p-value between rest and load conditions with light is 0.036, meaning that HRV measurement can distinguish between load and rest situations. This result matches most of the previous works found in this field, as light is not necessarily a directly influencing parameter on controlled HRV measures.

In relation to the hypothesis stated in the section 3.3, it can be proved that according to the pupillometry not everything is fulfilled. Light does not seem to have significant effect on the measurements but does influence the variance of the results. The low number of participants of our study might prevent us from getting normally distributed data, thus not providing a real statement. More importantly, the cognitive conditions were successfully differentiated, both for pupillometry and HRV, confirming the hypothesis.

Finally, as a baseline, each participant was asked to fill a NASA TLX survey for each task. It is a well-known subjective workload assessment method [8], where the user answers questions regarding the performed experiment, and leads to evaluate the task load. A Wilcoxon test comparing the perceived mental workload between rest and load states gave a p-value of 0.0074. The average for rest is 17.77 and for load 48.88, which means the participants felt significantly more mental effort for the *n*-back task than for the rest condition.

CONCLUSION

In this paper the cognitive load detection was analyzed, recording the physiological parameters of pupil diameter and heart-rate variability, with the target of seeing if two workload conditions (load and rest) can be distinguished. Besides the load, we also looked into lightning effect on the measures. The study was conducted for two load conditions, each under different lightning conditions: for the

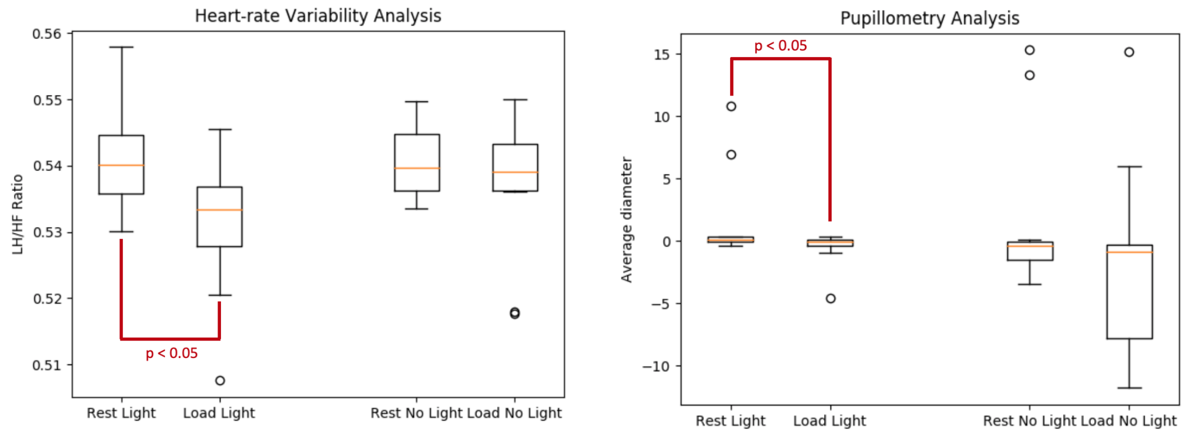

Figure 5: Cognitive load results from the experiment. Left: The boxplot for the HRV dataset. Right: The boxplot for the pupillometry analysis.

Table 1: Results from the three measurement sources used, for rest and load conditions with light.

Measure	CL Value - **Rest**	CL Value - **Load**	p-value
Pupillometry	$\mu = 1.89, \sigma^2 = 26.52$	$\mu = -1.01, \sigma^2 = 30.44$	0.046
HRV	$\mu = 0.54, \sigma^2 = 4.41$	$\mu = 0.53, \sigma^2 = 0.00011$	0.036
NASA TLX	$\mu = 17.77, \sigma^2 = 761.72$	$\mu = 48.88, \sigma^2 = 532.09$	0.0074

rest condition participants were asked to relax on a chair, while for the high load one, the n-back test was used.

The pupil diameter analysis did not show any significant effect of light (Wilcoxon), as opposed to several previous works, but it was highly likely influenced by our low sample size. Moreover, pupillometry showed significant difference between load and rest states. As for the LF/HF Ratio computed from the HRV, light did not had any effect and cognitive load could be distinguished. Furthermore, as expected, NASA TLX results show that users felt more mentally demanding during the load task than in the rest case.

These preliminary results are encouraging and allow to demonstrate a methodology for testing specific parameters influence. The lightning influence needs more testing with a larger group to confirm our hypothesis. We are on the way of testing more advanced algorithms that could get rid of this effect such as the Index of Pupillary Activity (IPA) [6], a wavelet-based method that uses the Discrete Wavelet Transform to decompose and analyze the signal at multiple levels of resolution.

As soon as we confirm the potential of pupillometry, combined with HRV for more robustness, we will be able to conduct our cognitive ergonomy study in the wild. And another point to look at is the effect of Respiratory Sinus Arrhythmia (RSA), that should be assessed with respect to HRV readings.

REFERENCES

[1] J. J. B. Allen, A. S. Chambers, and D. N. Towers. 2007. The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology* 74 (2007), 243–262. https://doi.org/10.1016/j.biopsycho.2006.08.005

[2] M.M. Baig, H. Gholamhosseini, and M.J. Connolly. 2013. *A comprehensive survey of wearable and wireless ECG monitoring systems for older adults.* https://doi.org/10.1007/s11517-012-1021-6

[3] Matthias Beggiato. 2018. Using Smartbands, Pupillometry and Body Motion to Detect Discomfort in Automated Driving. *Front. Hum. Neurosci* 12: 338. (2018). https://doi.org/10.3389/fnhum.2018.00338

[4] Ricardo Buettner. 2018. Real-time Prediction of User Performance based on Pupillary Assessment via Eye Tracking. *AIS Transactions on Human-Computer Interaction* 10(1), 26-56 (2018). https://doi.org/10.17705/1thci.00103

[5] S C Saxena D Singh, K Vinod. 2004. Sampling frequency of the RR interval time series for spectral analysis of heart rate variability. *Journal of medical engineering and technology* 28, 263-72 (2004). https://doi.org/10.1080/03091900410001662350

[6] Andrew T. Duchowski. 2018. The Index of Pupillary Activity. Measuring Cognitive Load vis-à-vis Task Difficulty with Pupil Oscillation. *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–13 (2018). https://doi.org/10.1145/3173574.3173856

[7] J. P. Ginsberg Fred Shaffer. 2017. An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health. 2017 Sep 28* 5:258 (2017). https://doi.org/10.3389/fpubh.2017.00258

[8] Sandra G.Hart. 2008. *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.* Vol. 52. Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

[9] Pedro Gomes. 2019. Python toolbox for Heart Rate Variability. https://pypi.org/project/pyhrv/

[10] Jamison Heard. 2018. A Survey of Workload Assessment Algorithms. *IEEE Transactions on Human-Machine Systems* 48, no. 5, pp. 434-451 (2018). https://doi.org/10.1109/THMS.2017.2782483

[11] Tobias Heine. 2017. *Electrocardiographic features for the measurement of drivers' mental workload.* Vol. 61. Elsevier. 31–43 pages. https://doi.org/10.1016/j.apergo.2016.12.015

[12] Nina Hollender, Cristian Hofmann, Michael Deneke, and Bernhard Schmitz. 2010. Integrating cognitive load theory and concepts of human-computer interaction. *Computers in Human Behavior* 26, 6 (2010), 1278–1288. https://doi.org/10.1016/j.chb.2010.05.031

[13] Susanne M. Jaeggi. 2009. The concurrent validity of the N-back task as a working memory measure. *Memory (Hove, England)* 18 (2009), 394–412. https://doi.org/10.1080/09658211003702171

[14] Jeff Klingner, Rakshit Kumar, and Pat Hanrahan. 2008. Measuring the task-evoked pupillary response with a remote eye tracker. *Eye Tracking Research and Applications Symposium (ETRA)* 1, 212 (2008), 69–72. https://doi.org/10.1145/1344471.1344489

[15] Mariska E. Kret and Elio E. Sjak-Shie. 2019. Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods* 51, 3 (2019), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

[16] Gustavo Lenis, Michael Kircher, Jesús Lázaro, Raquel Bailón, Eduardo Gil, and Olaf Doessel. 2017. Separating the effect of respiration on the heart rate variability using Granger's causality and linear filtering. *Biomedical Signal Processing and Control* 31 (2017), 272–287. https://doi.org/10.1016/j.bspc.2016.07.014

[17] F. Levillain, D. St-Onge, G. Beltrame, and E. Zibetti. 2019. Towards situational awareness from robotic group motion. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–6.

[18] K.M. Miller. 2009. Is the N-Back Task a Valid Neuropsychological Measure for Assessing Working Memory? *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists* 24 (2009), 711–7. https://doi.org/10.1093/arclin/acp063

[19] Atsuo Murata. 1998. Evaluation of Mental Workload by Fluctuation Analysis of Pupil Area. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 20 (1998). https://doi.org/10.1109/IEMBS.1998.746146

[20] Taishi Nagasawa. 2016. Workload induces changes in hemodynamics, respiratory rate and heart rate variability. (2016), 176–181. https://doi.org/10.1109/BIBE.2016.27

[21] Vsevolod Peysakhovich. 2015. Pupil Diameter as a Measure of Cognitive Load during Auditory-Visual Interference in a Simple Piloting Task. *Procedia Manufacturing* 3 (2015). https://doi.org/10.1016/j.promfg.2015.07.583

[22] Soroosh Solhjoo, Mark C. Haigney, Elexis McBee, Jeroen J.G. van Merrienboer, Lambert Schuwirth, Anthony R. Artino, Alexis Battista, Temple A. Ratcliffe, Howard D. Lee, and Steven J. Durning. 2019. Heart Rate and Heart Rate Variability Correlate with Clinical Reasoning Performance and Self-Reported Measures of Cognitive Load. *Scientific Reports* 9, 1 (2019), 1–9. https://doi.org/10.1038/s41598-019-50280-3

[23] David St-Onge, Ulysse Côté-Allard, Kyrre Glette, Benoit Gosselin, and Giovanni Beltrame. 2019. Engaging with Robotic Swarms: Commands from Expressive Motion. *J. Hum.-Robot Interact.* 8, 2, Article 11 (June 2019), 26 pages. https://doi.org/10.1145/3323213

[24] D. St-Onge, M. Kaufmann, J. Panerati, B. Ramtoula, Y. Cao, E. B. J. Coffey, and G. Beltrame. 2020. Planetary Exploration With Robot Teams: Implementing Higher Autonomy With Swarm Intelligence. *IEEE Robotics Automation Magazine* 27, 2 (2020), 159–168.

[25] P. Van der Wel and H van Steenbergen. 2018. A Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychon Bull Rev. 2018 De* 25(6) (2018). https://doi.org/10.3758/s13423-018-1432-y

[26] Valeria Villani. 2020. Wearable devices for the assessment of cognitive effort for human-robot interaction. *IEEE Sensors Journal* (2020). https://doi.org/10.1109/JSEN.2020.3001635

[27] Wilhelm von Rosenberg. 2017. Resolving Ambiguities in the LF/HF Ratio: LF-HF Scatter Plots for the Categorization of Mental and Physical Stress from HRV. *Front Physiol. 2017 Jun 14* 8:360. (2017). https://doi.org/10.3389/fphys.2017.00360