Management and Production Engineering Review
Volume 11 • Number 3 • September 2020 • pp. 56–64
DOI: 10.24425/mper.2020.134932

HEART RATE VARIABILITY BASED ASSESSMENT OF COGNITIVE WORKLOAD IN SMART OPERATORS

Salvatore Digiesi[1], Vito Modesto Manghisi[1], Francesco Facchini[1], Elisa Maria Klose[2], Mario Massimo Foglia[1], Carlotta Mummolo[3]

[1] *Polytechnic University of Bari, Department of Mechanics, Mathematics and Management, Italy*
[2] *University of Kassel, FG Industrial and Organizational Psychology, Germany*
[3] *Department of Biomedical Engineering, New Jersey Institute of Technology, Newark NJ USA*

Corresponding author:
Francesco Facchini
Department of Mechanics, Mathematics and Management
Polytechnic University of Bari
Via Orabona 4, Bari, Italy
phone: +39 0805693612
e-mail: francesco.facchini@poliba.it

Received: 19 December 2019
Accepted: 25 July 2020

Abstract
The study on cognitive workload is a field of research of high interest in the digital society. The implementation of 'Industry 4.0' paradigm asks the smart operators in the digital factory to accomplish more 'cognitive-oriented' than 'physical-oriented' tasks. The Authors propose an analytical model in the information theory framework to estimate the cognitive workload of operators. In the model, subjective and physiological measures are adopted to measure the work load. The former refers to NASA-TLX test expressing subjective perceived work load. The latter adopts Heart Rate Variability (HRV) of individuals as an objective indirect measure of the work load. Subjective and physiological measures have been obtained by experiments on a sample subjects. Subjects were asked to accomplish standardized tasks with different cognitive loads according to the '*n*-back' test procedure defined in literature. Results obtained showed potentialities and limits of the analytical model proposed as well as of the experimental subjective and physiological measures adopted. Research findings pave the way for future developments.

Keywords
Smart operators, cognitive load, NASA-TLX, heart rate variability, information theory model.

Introduction

The Fourth Industrial Revolution, the so-called Industry 4.0, is born as a response to the need of improvement of current production systems, by providing a wide integration of digital technologies in manufacturing [1]. 'Smart Manufacturing Systems' are based on a digital network where the physical context is closely intertwined with artificial intelligence allowing to manage and monitor the production process at operational and procedural levels. Consistently with this approach, the human factors are at the heart of a virtuous closed-loop chain that leads the industries in the new era of manufacturing.

According to recent studies, the tasks typology as well as their complexity in smart factories will significantly change. An increased role of automation and artificial intelligence will lead to evolving workplaces in which the people will frequently interact with ever-smart machines [2]. The technological change led by the fourth industrial revolution has redefined the 'traditional' manufacturing process in a 'smart' manufacturing process where the physical context is closely intertwined with the corresponding cyber twin by means of IoT and Cloud Computing Infrastructures. In this perspective, a crucial role is played by operators who became "smart-operators", who are required to be highly flexible and to demonstrate adaptive capabilities in a very dynamic working environment [2]. Recently statistical studies showed that in industries the overall time spent by industrial operators on manual tasks is reducing over the years; as a consequence, the manual tasks assigned to machines, from 2018 to 2022, are averagely increased of 13%, in terms of tasks' number [3]. If on one hand, the manual tasks performed by 'smart operator' is reducing over time, on the other hand, it is possible to observe that, in the next years, the predicted time that each smart-operators will dedicate to cognitive-oriented tasks will significantly increase in the next years (Fig. 1) [3].

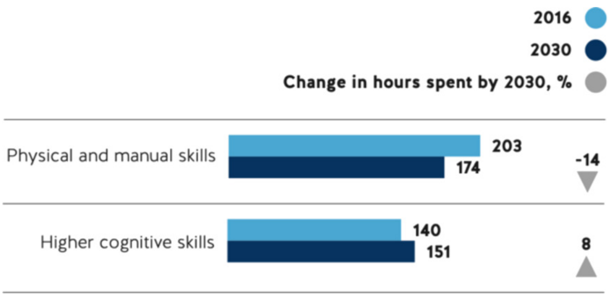

Fig. 1. Total hours worked in Europe and the United States, 2016 vs 2030 estimate, billion (adapted by www.worldmanufacturingforum.org/report-2019).

The experts consider this trend to be a relevant indicator of the changing and of the emerging role of the human factor in manufacturing systems. The requirement of new types of interactions between operators and machines [4]. These interactions will generate a new intelligent workforce and have significant effects on the work content. Therefore, the workforce demand will change significantly in the next years.

While the tasks of the new work environments are rapidly switching from manual- to cognitive-based, the average age of the workforce is significantly increasing. According to OECD Data (Fig. 2), the percentage of the working population in EU28 with an age between 55–64 years old is increased in the last fifteen years of 5%, and this trend will lead to a further increase of further 3% in the next ten years.

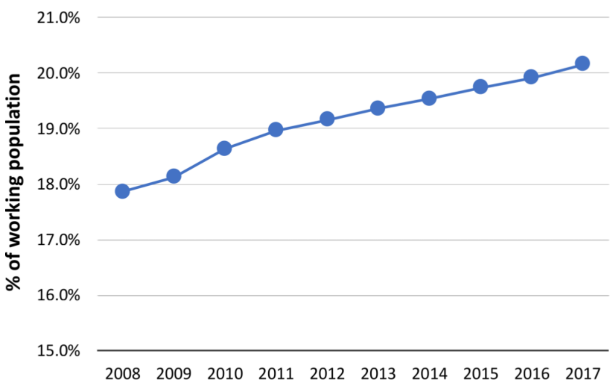

Fig. 2. Percentage of the working population in EU28 with age between 55–64 years old (source: OECD Data, available on URL http://www.oecd.org/).

The aging of the population directly affects the workforce performance, as shown in recent studies highlighting an age-related reduction of cognitive capacity [5].

In this context, the evaluation of Cognitive Work-Load (CWL) can be crucial to design, manage, and monitor the workloads in a smart factory. This strategy allows ensuring both a higher efficiency of the manufacturing system and the well-being of the workforces. Strong limitations on the evaluation of the CWL are highlighted by the available scientific literature on this topic. Although many physiological and subjective parameters are identified in order to measure the perceived CWL (e.g. Heart Rate Variability (HRV), electroencephalogram (EEG), pupil diameter, etc.), nowadays there is not a methodology allowing to quantify the information content of a specific cognitive task, in terms of bits, for example, and then to assess the corresponding CWL required to process it.

Further gaps in scientific literature showed that there are not 'conventional' assessment models to be adopted for evaluating the performance of the worker exposed to cognitive load. In most cases, the existing assessment models consider the aspects related to physical load at the workplace (e.g. OCRA, RULA, OWAS, etc.). Currently, the CWL evaluation methods are considered a hotspot in the scientific community. If, on the one hand, it is possible to claim that there are three different assessment approaches (i.e. subjective evaluation methods, work performance evaluation methods, and physiological evaluation methods), on the other hand for each of them the experts identify important advantages and disadvantages that do not allow an easy and flexible implementation in manufacturing systems [6].

Consistently with the observations mentioned above, to fully investigate the research problem, the following subsidiary research questions are raised:

- RQ1. Whether it is possible to identify a mathematical model that identifies the CWL on the basis of the information content of a corresponding cognitive task;
- RQ2. How the well-known physiological parameters detection (e.g. HRV, EEG, PUPIL DIAMETER) allows to dynamically estimate the perceived CWL;
- RQ3. How the higher mental load affects the workers' performance in terms of reliability and safety.

Therefore, the aim of the proposed study consists in investigating the relation between CWL, evaluated by means of an analytical model, and the corresponding perceived CWL, experimentally detected. In other words, the present work conducted allows identifying how the perceived CWL increases with the increasing of the information content to be processed. For this scope, an Analytical Model of Advanced Manufacturing Systems has been adopted to evaluate the human workload of specific tasks, characterized by different complexity. Consistently with the purpose of the work, the same tasks have been performed by a sample of workers in order to identify and evaluate the corresponding perceived CWL. In experiments conducted, the perceived CWL was estimated through the HRV analysis, concerning the tasks to be performed, the 'n-back' tasks test [7, 8] (at n-levels) was adopted for the workers' cognitive assessment. n-back tasks test, originally introduced by Kirchner [8], has become a standardized tool to simulate tasks with different cognitive complexities; it consists of standardized working memory and attention tasks with four incremental levels of difficulty.

The remainder of the paper is organized as follows: the literature review is introduced in the next section, the case study is introduced in Sec. 3; materials and methods of experiments conducted are in Sec. 4; finally, conclusions of this work are in Sec. 5.

Literature review

CWL is an important research topic in the field of human factors engineering. According to Fan et al. the higher amount of brain activity, the higher occupancy rate of the brain, as well as the higher psychological pressure will cause rapid fatigue, reduced flexibility, increased human errors and frustration, which will lead to errors in information acquisition, analysis, and decision making [6]. On the contrary, low CWL leads to the decline of job performance [9]. Therefore, nowadays the research on the evaluation method of CWL is considered one of the most important challenges for the manufacturing systems. According to available scientific literature, CWL can be evaluated adopting different well-known methodologies, so-called: subjective, physiological and performance, each of them including different evaluation techniques (Fig. 3).

The first methodology is based on a subjective evaluation of cognitive load led by the judgment of the same worker to be analyzed. In this case, a survey is conducted, by questionnaires administration, when the worker(s) has(have) completed the assigned task, in order to evaluate the cognitive effort perceived during the task execution. The output of the survey consists in a task load index returned by questionnaires evaluation. National Aeronautics and Space Administration – Task Load Index (Nasa-TLX) is one of the most common techniques included in the 'subjective' methodologies. It is based on a multidimensional measuring scale in which are provided six sub-scales representing different dimensions of cognitive effort: mental demand, physical demand, temporal demand, frustration level, effort, and performance [10]. Similar procedures are adopted for the other techniques, included in similar methodologies, although different dimensions of the sub-scales are considered.

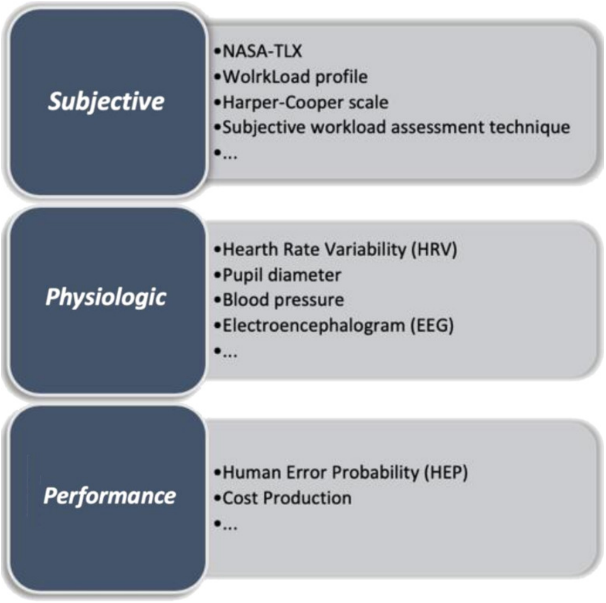

Fig. 3. Most common methodologies adopted for CWL evaluation.

The second methodologies are based on the monitoring of the human-body parameters, under the assumption that when the tasks complexity changes, physical changes are observed in the worker analyzed [11]. There are several techniques to evaluate the CWL by bodily responses such as ECG measures, EEG, Blood pressure, pupil diameter, etc.

The workers' performance evaluation (i.e. third methodologies) are based on the assumption that the increase of the CWL of the assigned task leads lessening the efficiency and increasing the error rate. Therefore, starting from performance parameters (e.g. the Human Error Probability – HEP, the cycle time, the number of accidents, etc.) it is possible to identify the corresponding CWL [12].

Considering the limitation of the single evaluation methodology, recent studies focused on the comprehensive adoption of more methods, so that the various methods complement each other and avoid their shortcomings [6].

Mansikka et al. evaluated the CWL of pilots in a flight simulation comparing the NASA-TLX scale and the modified Cooper-Harper scale (MCH). The study conducted showed a strong correlation between the results of the adopted techniques and the variation of the mean inter-beat-interval (IBI) parameter [13]. A demonstration of HRV's sensitivity to the cognitive demand is given by [14]. In this case, the CWL of the participants to the test has been evaluated accordingly to NASA-TLX and HRV techniques. Three different task kinds (i.e. psychomotor vigilance task, duration discrimination task, and n-back task) have been evaluated and, in every case, the authors showed a high sensitivity of the HRV to the change of task complexity.

Bommer and Fendley [12] developed a theoretical framework based on a systematic approach to measuring CWL through a computer simulation. The model is based on the joint adoption of the NASA-TLX, the Workload Profile, and the HEP. The authors tested the model on a real full-case study, showing its effectiveness to predict the operator's performance to variations of the CWL [12].

A set of case studies has been evaluated, adopting the physiological methods, in order to evaluate the change of CWL over time (before, during and after the execution of the task) for different tasks characterized by different complexity. All evaluations collected have been included in a simulation program, so-called IMPRINT, sponsored by the US weapons research laboratory. The software allows predicting the continuous workload profile of a specific task (included in a database), in order to suggest the cognitive load, required to the operator, for task performing [15]. The EEG technique was adopted for evaluating the CWL in assembly operations. The data collected on a sample of 46 workers, involved on a manufacturing line, showed significant changes in EEG measures by varying the task complexity. Three different levels of task-hardness (low, medium, high) have been considered [16].

Summing up, to the best of authors' knowledge, no studies investigate the relationship between human workload, evaluated in accordance with an existing analytical model, and the corresponded perceived CWL, experimentally detected.

The case study

As a case study, we analyzed the n-back task.

n-back tasks test, originally introduced by Kirchner [8] and by Mackworth [7], has become a standardized tool to simulate tasks with different cognitive complexities; it consists of standardized working memory and attention tasks with four incremental levels of difficulty.

A sequence of stimuli (letters on a computer screen) are shown to the tester (subject who runs the test). The tester is asked to digit the right shift button on a keyboard when the current stimulus matches the one observed n steps earlier in the sequence; in case of mismatching stimuli, the tester has to digit the left shift button. Increasing the number (n) of letters included between two target letters, increases the difficulty of the task and the mental effort required to accomplish it. We implemented the zero- and the two-back levels by exploiting the Psychology Experiment Building Language toolkit (PEBL) version 2.1 (freely available at http://pebl.sourceforge.net/download.html). In the 0-back task, testers responded to a predefined single target letter (i.e., "X"); while in the 2-back task, the targets were defined as any letter that was identical to the one presented two trials back.

We exploited the model of Salvendy in order to implement a custom model for evaluating the CWL associated with the execution of the two implemented "n-back" task levels [18]. The total task load (TL) is evaluated in terms of 'bit' on the basis of task arrival rate, task complexity, task uncertainty as well as task performance requirement. According to the authors, their model can be applied to all manufacturing systems independently on automation levels [19].

In the implementation of the model of Salvendiy, we used both subjective and objective measures.

We carried out a within-subjects experiment by asking participants to execute two n-back test sessions corresponding to the two implemented levels. As subjective measures, we used participants ratings collected by administering the NASA-TLX questionnaire. As objective measures, we used both participants' performance (i.e. reaction-time) and task parameters (such as inter-target time), and physiological parameters extracted by monitoring heart-activity (i.e. HRV analysis).

Material and methods

We recruited 13 participants (all volunteer, 7 male and 6 female), with mean age 26.9 years and deviation 2.98 years. They all had a Mechanical and Management Engineering background (8 master-degree students and 5 PhD. students). Before starting the experiment, they all filled in a preliminary questionnaire in order to check whether their life habits and physical conditions allowed their participation in the trial (participants were required to maintain a regular sleep-wake cycle for at least one day before the study and to abstain from stimulating beverages or intense physical activity).

The experimental procedure consisted in the execution of the two implemented n-back task levels. Before starting the experiment, each participant received written and verbal information explaining the experimental procedure and her/his task in the test. Then, in order to record the Electrocardiographic signal (ECG), an experimenter positioned three pre-gelled electrodes on the participant's chest. The ECG signal was acquired with a 1000 Hz sampling rate by using the BITalino®Plugged Kit BLE (https://bitalino.com/en/plugged-kit-ble), a low-cost multimodal platform for physiological signals acquisition [20].

Before each level execution, each participant carried out a training phase to get used to the correct procedure. The execution order of the two "n-back" task levels was counterbalanced among participants. The zero-back level had 100 prompts and the two-back 102 prompts, both levels had an inter-prompt period of 3000 milliseconds (i.e. 500 ms of stimulus presentation and 2500 ms of fixed delay). Each level execution lasted about 5 minutes and had 33 targets that were prompted randomly.

During the task execution, the ECG signal was recorded together with participants' performance (i.e. reaction time and error rate). At the end of each level each participant filled in the NASA-TLX questionnaire. The two "n-back" levels execution were interleaved with a resting phase of at least 5 minutes. During this phase, the ECG signal was recorded in order to obtain a baseline measure of the heart activity in rest conditions.

Analytical model evaluation

The analytical model adopted for CWL evaluation was proposed by Bi and Salvendy [19] it is based on 5 input parameters, one output and one, or more, physiologic and/or subjective workload measures. According to the authors, the basic task load can be divided into three categories: information input, information process, and information output [19]. For the task required by the "n-back" test, the information input consists of indicating when the current stimulus matches the one from n steps earlier in the sequence; the information output is provided through keystrokes activation (i.e. left 'shift' key if the stimulus coincides with the target, right 'shift' key if the stimulus not coincides with the target). The stimulus appears regularly at a fixed interval time (3 (s)). As a consequence, no uncertainty has to be considered in the task arrival. Moreover, in each run of the n-back test there are no sub-tasks.

Under the above-mentioned hypotheses, the average task load of the j-th subject for each k-th level of the n-back test ($TL_{k,j}$) is derived from the general model suggested in [18] and evaluated as:

$$TL_{k,j} = a + b\lambda_k + c\lambda_k \frac{Tc_k}{1 - P_{k,j}}, \tag{1}$$

where k is the level of n-back test: $k = 0$ for the "0-back test"; $k = 1$ for the "2-back test"; Tc_k is the task complexity of each run (in bit) identified by Shannon entropy measures for discrete equiprobable states. In case of $k = 0$ (low level), the task required to the operator is relatively easy: her/his work consists to identify whether the variable showed matches the fixed target-variable, hence only one bit is the content of information to be processed (yes/no). In case of $k = 1$ (high level), the operator is asked to identify a target variable on the basis of the sequence memorized (previous 3 letters displayed), therefore, before confirming whether the observed variable matches the target, she/he should remember the sequence. Since the target could be identified only after a sequence of 3 letters have been displayed, the subject has to compare the target appeared with the letters in the sequence memorized. In this case, the content of the information is quantified in six bits. λ_k ($i = 1, ..., \lambda_k$) is the interarrival rate of runs (100/300 (s^{-1}) for $k = 0$; 102/306 (s^{-1}) for $k = 1$); $P_{k,j} = Tr_{k,j}/Ta$ is the average schedule tightness of the j-th subject in accomplishing all runs of the test (100 for $k = 0$; 102 for $k = 2$), being: $Tr_{k,j}$ the average value of the time required by the j-th subject to accomplish a run; Ta is the available time for each task run (3 (s)); 'a' (bit), 'b' (bit · s), and 'c' (non-dimensional) are parameters obtained by regression analysis.

In this study, the normalized (0;1) values of the perceived cognitive load measured for each operator by means of the Raw NASA-TLX questionnaire (RTLX) and of HR measures were considered as a measure of subjective or physiological workloads, respectively. For both levels of the n-back test, regression parameters have been obtained. On the basis of Eq. (1), the average values (among subjects) of the TL_k based on the subjective and physiological measures carried out during the experiments are in Table 1.

Table 1
Normalized task load (TL_k) obtained by Eq. (1).

	Task load TL_k (bit)	
	0-back level	2-back level
RTLX data set	0.27	0.55
HR data set	0.74	0.76

As it is shown in Table 1, a meaningful increase in the average TL_k values between the two levels ($k = 0, 1$) has been observed when a subjective measure is adopted (RTLX). On the contrary, no significant increase has been observed in TL_k estimates when referring to a physiological measure (HR). Such a result obtained by the analytical model is in accordance with the statistical analysis provided in the following sections.

HRV assessment

Heart rate is the number of heartbeats per minute. Heart rate variability (HRV) is the fluctuation in the time intervals between adjacent heartbeats [21]. Healthy heart rhythm is not fixed. The oscillations of a healthy heart allow the cardiovascular system to rapidly adjust to sudden physical and psychological challenges to homeostasis. Many studies in the literature are exploiting HRV analysis as an indirect indicator of cognitive and physical workload [13, 14, 22, 23]. According to Shaffer and Ginsberg, HRV analysis can be carried out with respect to time, frequency, and non-linear measurements [24]. Among the factors influencing the reliability of HVR analysis, the recording period length plays a crucial role. In particular, the length ranges from 2 minutes (ultra shorth-term) to 24 hours (long-term). We analyzed 5 minutes (short-term) time recordings that guarantee a good tradeoff between recording length and HRV results reliability [24].

In order to evaluate how participants reacted to the different CWL associated with the implemented n-back task levels, we carried out the analysis in the time domain. Starting from the ECG signal we extracted the NN (i.e. RR) tachogram, by exploiting the HRV analysis software tool [25] and, from it, the HR indicators.

In the time domain, we assessed the Heart Rate (HR) and the square root of the mean squared differences of successive NN intervals (RMSSD).

We assessed all the indicators for the three recording conditions: the "zero-back", the "two-back", and the "baseline". In order to evaluate the effectiveness of such parameters as indicators of the CWL, we considered the ratio between the indicators in each one of the n-back levels and the "baseline" level.

In accordance with the literature, we observed that tasks involving higher CWL are associated with an increase in HR and a decrease of RMSSD (Figs 4 and 5, respectively).

The Wilcoxon signed-rank test shows that the HR ratio to baseline mean value for the Zero-back level is significantly lower than the one for the Two-back level (0.95 vs 1.03, $p = 0.0327$, Fig. 6).

The Wilcoxon signed-rank test shows that the RMSSD ratio to baseline mean value for the Zero-back level is significantly higher than the one for the Two-back level (1.16 vs 0.80, $p = 0.0327$, Fig. 7).

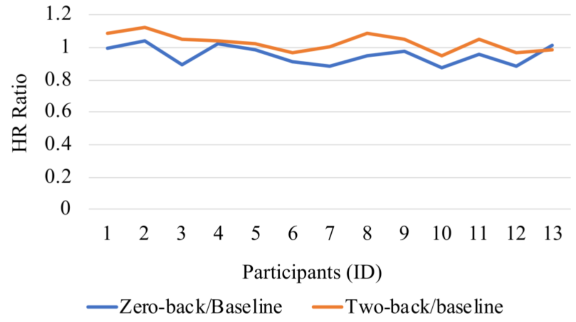

Fig. 4. Values of the HR ratio to the baseline for the 13 participants.

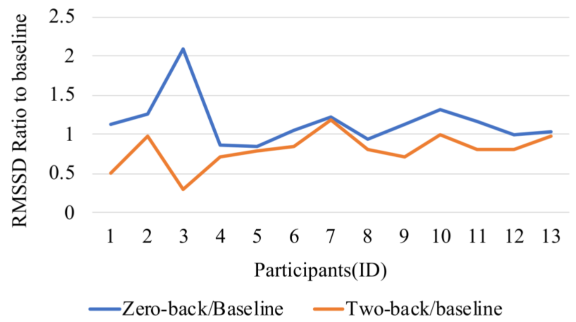

Fig. 5. Values of the RMSSD ratio to the baseline for the 13 participants.

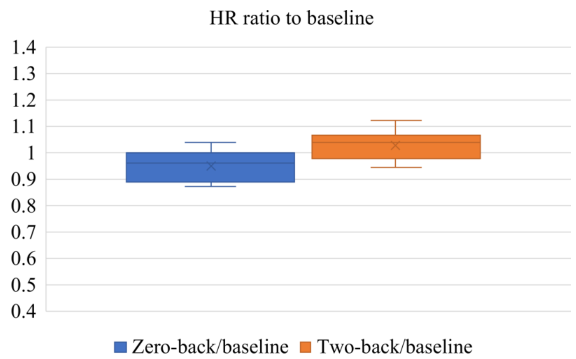

Fig. 6. Heart rate ratio to the baseline boxplots.

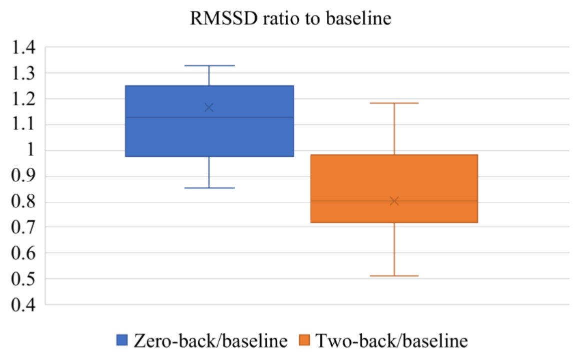

Fig. 7. RMSSD ratio to the baseline boxplots.

NASA-TLX assessment

The NASA-TLX provides an overall workload score (from 0 to 100 points) based on a weighted average of ratings on six dimensions: mental demands, physical demands, temporal demands, own performance, effort, and frustration. The NASA-TLX questionnaire sensitivity to mental-workload has been demonstrated to be useful in a variety of cognitively demanding tasks from aircraft piloting [26, 27], to surgery [28], or laboratory tasks context [29]. In the literature, the use of an unweighted or raw TLX (RTLX) is the most common, because high correlations have been shown between the weighted and unweighted scores [30, 31]. The assessment was carried out by administering post task questionnaires for both the "n-back" levels. Results on the overall RTLX score evidence the different perceived effort (Fig. 8).

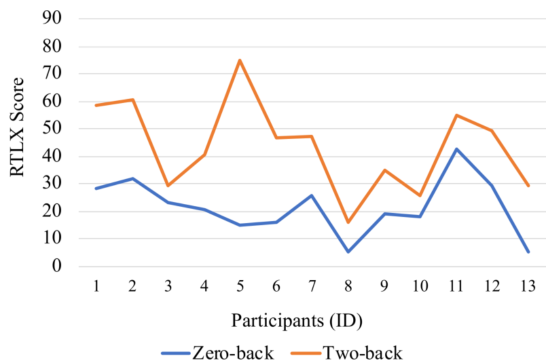

Fig. 8. Values of the RTLX score for the 13 participants.

The RTLX samples were positively tested for homoscedasticity and normality; the T-student test showed that the RTLX mean value for the Zero-back level is significantly different from the one for the Two-back level (21.53 vs 43.71, $p < 0.005$, Fig. 9).

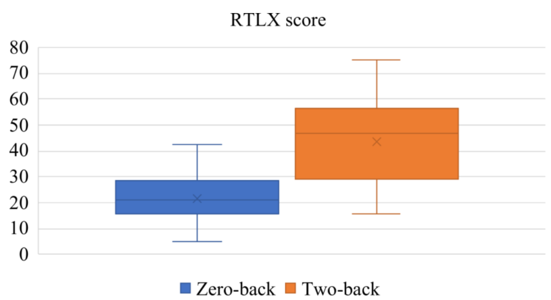

Fig. 9. NASA-TLX overall score (RTLX) box plots.

Conclusions and future research

This study investigated the relation between CWL, evaluated in accordance with an analytical model derived by [18], and the corresponding perceived CWL, estimated from experimental data on a sample of 13 operators. Both a physiologic (HRV) and a subjective (NASA-TLX) method has been adopted in order to estimate CWL. The assessment conducted showed good reliability of the adoption of both experimental methodologies. The application of the analytical model and the measures of the HRV and RTLX score on the field showed that is possible identifying two distinct subsets of data characterized by very different average values of the HRV and RTLX score for tasks of 1 bit (0-back) rather than 6 bits (2-back).

This study contributes in answering the first two research questions raised:

- RQ1. It is possible to identify a mathematical model that identifies the CWL on the basis of the information content of a corresponding cognitive task, under abovementioned limitations;
- RQ2. Detection of physiological parameters (i.e. HRV) estimated the perceived CWL in accordance with subjective tests (i.e., RTLX).

Regarding the third question (RQ3), no feedbacks were provided on the issue related to how the cognitive load affects the workers' performance in terms of reliability and safety. From this point of view, there are no pieces of evidence, on the tests conducted, on the possible relationship between CWL and the corresponding human error rate.

Limitations of the proposed study can be outlined. In order to obtain more generalized findings, further field research will be extended to further groups of subjects differently aged and sexed.

The analytical model is difficult to apply for complex tasks. Although the model can be generalized for most manufacturing activities, it requires a complex and time-consuming process that cannot ignore subjective and/or physiologic workload measures directly detected on the field: the HR and RTLX estimation could be hard to measure in many manufacturing systems. The HR is characterized by a signal that can be easily affected by noise due to frequent body movements; the RTLX requires the interruption of work, in order to allow the compilation of the test by the operator.

Future developments will address the highlighted limitations and provide a more effective and easy to use tool to evaluate the CWL and the corresponding effect on the human performances.

Future research will extend the investigation of human workload in manufacturing by addressing work tasks with both cognitive and motor components. To complete the estimation of the overall human workload, the contribution of the motor task to the work load will be addressed by breaking-down complex motor tasks in elementary motor sub-tasks. Shannon Entropy measures will be adopted to measure information in 'trajectory-based' motor tasks within the framework of the Fitts' law. A general 'trajectory-based' approach is proposed in [32] for Human Computer Interface (HCI) studies. The approach is meaningful as HCI research is very close to digital work environment of 'smart operators' and human-robot cooperation. In spite of the nature of motor tasks that will be investigated, which differ from the NASA TLX tasks, the operators are still required to process information which contribute to the cognitive workload. Difficulty index of differently shaped trajectories will be put into statistical relation with HR and RTLX estimates. In this way the influence of HR variability on workload of smart operators engaged to accomplish simple motor tasks in digital work environment will be investigated under the same information theory framework.

The research has been supported by the Italian Ministry of University and Research under the grant Research Project of National Interest (PRIN-2017) "Smart Operators 4.0 based on Simulation for Industry and Manufacturing Systems".

References

[1] Lu Y., *Industry 4.0: A survey on technologies, applications and open research issues*, Journal of Industrial Information Integration, 1, 1–10, 2017.

[2] Longo F., Nicoletti L., Padovano A., *Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context*, Comput. Ind. Eng., 113, 144–159, 2017.

[3] Bauer H., Baur C., Mohr D., Tschiesner A., Weskamp T., *Industry 4.0 after the initial hype – where manufacturers are finding value and how they can best capture it*, McKinsey Co., 2016.

[4] Romero D., Bernus P., Noran O., Stahre J., Berglund Å.F., *The operator 4.0: Human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems*, [in:] Nääs I. et al. [Eds], *Advances in Production Management Systems. Initiatives for a Sustainable World*, APMS 2016, IFIP Advances in Information and Communication Technology, Springer, Cham, vol. 488, 2016.

[5] Rose D.M., Gordon R., *Age-related cognitive changes and distributed leadership*, J. Manag. Dev., 34, 3, 330–339, 2015.

[6] Fan X., Zhao C., Hu H., Jiang Y., *Review of the evaluation methods of mental workload*, Springer International Publishing, vol. 967, 2020.

[7] Mackworth J.F., *Paced memorizing in a continuous task*, J. Exp. Psychol., 58, 3, 206–211, 1959.

[8] Kirchner W.K., *Age differences in short-term retention of rapidly changing information*, J. Exp. Psychol., 55, 4, 352–358, 1958.

[9] Zülch G. *et al.*, *Influence of mental workload on job performance*, Elektrotechnik und Informationstechnik, 131, 7, 207–211, 2015.

[10] Şeker A., *Using outputs of NASA-TLX for building a mental workload expert system*, Gazi Univ. J. Sci., 27, 4, 1131–1142, 2014.

[11] Charles R.L., Nixon J., *Measuring mental workload using physiological measures: A systematic review*, Applied Ergonomics, 74, 221–232, 2019.

[12] Bommer S.C., Fendley M., *A theoretical framework for evaluating mental workload resources in human systems design for manufacturing operations*, Int. J. Ind. Ergon., 63, 7–17, 2018.

[13] Mansikka H., Virtanen K., Harris D., *Comparison of NASA-TLX scale, modified Cooper-Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks*, Ergonomics, 62, 2, 246–254, 2019.

[14] Luque-Casado A., Perales J.C., Cárdenas D., Sanabria D., *Heart rate variability and cognitive processing: The autonomic response to task demands*, Biol. Psychol., 113, 83–90, 2016.

[15] Rusnock C.F., Borghetti B.J., *Workload profiles: A continuous measure of mental workload*, Int. J. Ind. Ergon., 63, 49–64, 2018.

[16] Morton J. *et al.*, *Identifying predictive EEG features for cognitive overload detection in assembly workers in Industry 4.0*, Proceedings of the 3rd Int. Symp. Hum. Ment. Workload Model. Appl. (H-WORKLOAD 2019), 2019.

[17] Coulacoglou C., Saklofske D.H., *Executive Function, Theory of Mind, and Adaptive Behavior*, [in:] Coulacoglou C., Saklofske D.H., Psychometrics and Psychological Assessment, Academic Press, Elsevier, UK, Chapter V, 89–130, 2017.

[18] Bi S., Salvendy G., *Analytical modeling and experimental study of human workload in scheduling of advanced manufacturing systems*, Int. J. Hum. factors Manuf., 4, 2, 205–234,1994.

[19] Bi S., Salvendy G., *A proposed methodology for the prediction of mental workload, based on engineering system parameters*, Work Stress, 8, 4, 355–371, 1994.

[20] Batista D., da Silva H.P., Fred A., Moreira C., Reis M., Ferreira H., *Benchmarking of the BITalino Biomedical Toolkit Against an Established Gold Standard*, IET Healthcare Technology Letters, 6, 2, 32–36, 2019.

[21] McCraty R., Shaffer F., *Heart rate variability: New perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk*, Global Advances In Health and Medicine, 4, 1, 46–61, 2015.

[22] Wilson G.F., *An analysis of mental workload in pilots during flight using multiple psychophysiological measures*, International Journal of Aviation Psychology, 12, 1, 3–18, 2002.

[23] Landi C.T., Villani V., Ferraguti F., Sabattini L., Secchi C., Fantuzzi C., *Relieving operators' workload: Towards affective robotics in industrial scenarios*, Mechatronics, 54, 144–154, 2018.

[24] Shaffer F., Ginsberg J.P, *An overview of heart rate variability metrics and norms*, Front. Public Heal., 2017.

[25] Pichot V., Roche F., Celle S., Barthélémy J.C., Chouchou F., *HRV analysis: A free software for analyzing cardiac autonomic activity*, Front. Physiol., 7, 557, 2016.

[26] Karavidas M.K., Lehrer P.M., Lu S.E., Vaschillo E., Vaschillo B., Cheng A., *The effects of workload on respiratory variables in simulated flight: A preliminary study*, Biol. Psychol., 84, 1, 157–160, 2010.

[27] Ma J. *et al.*, *Workload influence on fatigue related psychological and physiological performance changes of aviators*, PLoS One, 9, 2, e87121, 2014.

[28] Zheng B. *et al.*, *Workload assessment of surgeons: correlation between NASA TLX and blinks*, Surg. Endosc., 26, 10, 2746–2750, 2012.

[29] Muth E.R., Moss J.D., Rosopa P.J., Salley J.N., Walker A.D., *Respiratory sinus arrhythmia as a measure of cognitive workload*, Int. J. Psychophysiol., 83, 1, 96–101, 2012.

[30] Byers J.C., Bittner A.C., Hill S.G., *Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary?*, [in:] Mital A. [Ed.], Advances in industrial ergonomics & safety, Taylor & Francis, London, UK, 1, 481–485, 1989.

[31] Moroney W.F., Biers D.W., Eggemeier F.T., Mitchell J.A., *A comparison of two scoring procedures with the NASA Task Load Index in a simulated flight task*, Proceedings of the IEEE 1992 National Aerospace and Electronics Conference, NAECON 1992, 2, 734–740, 1992

[32] Accot J., Zhai S., *Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks*, proceedings of the ACM SIGCHI Conference on Human factors in computing systems, March 1997, 295–302, 1997.