2017

Online Measuring of Available Resources

Enrique Munoz-de-Escalona
University of Granada, enriquemef.delagado@ugr.es

José Juan Canas
University of Granada

Online Measuring of Available Resources

Enrique Muñoz-de-Escalona[1*] , José Juan Cañas[1],

[1] Mind, Brain and Behaviour Research Center, University of Granada, Granada, Spain
{enriquemef,delagado}@ugr.es

Abstract. This paper present a proposal for measuring available mental resources during the accomplishment of a task. Our proposal consists in measuring emotions provoked by perceived self-efficacy in the execution of the task. Self-efficacy is one of the most important factors that affect the resources that a person puts at the disposal of the execution of the task. When a person perceives that he/she is not being effective he/she will activate more resources to improve his performance. This self-efficacy will be reflected in the emotions that the person experiences. A good efficacy will provoke positive emotions and a bad efficacy negative emotions. The results of our study show that poor execution leads to negative emotions and psychophysiological activation as measured by pupil dilation. According to these results we propose that a possible method for measuring available resources during the execution of the task could be online measuring of emotions.

Keywords: Mental workload, available resources, online methods, emotions, self-efficacy.

1 Introduction

We define mental workload by relating task demanded mental resources and operators' available mental resources. Specifically, mental workload is the result of dividing demanded resources by available resources. Accordingly, if demanded resources are higher than the available, we say there exists an "overload" situation problem. On the contrary, if demanded resources are much lower than the available, we can assume there exist an "underload" situation problem. Although underload can be a problem because it can lead to loss of attention and drowsiness, with the consequent risk of incidents and accidents, the more important safety issues occur when there is an overload situation [1]. But, in any case, the important thing to consider is that human being will always try to maintain a balance between demanded and available resources in a way that does not endanger safety, health is not affected and satisfaction can be maintained at an acceptable level.

Demanded resources depend on the complexity of the task, in this sense, we say that in order to analyze and calculate demanded resources, we need to "measure" task complexity [1]. On the contrary, available resources depend on the level of activation, individual differences and time factors (time of the day, time on task, etc.) [2].

Therefore, available resources would be measured in different ways depending on intended purposed.

In any case, we can say that although we can measure the global mental workload that an operator is supporting, it would be interesting and possible to measure on the one hand the demanded resources and on the other hand the available resources, both of them separately. For example, if we are designing a new system, we could just measure human-system interaction demanded resources in order to make it less complex, without considering the system operators' available resources. Likewise, if we wanted to evaluate available resources when the operator changes work shifts, we could keep demanded resources constant while measuring the available.

For decades, a great effort has been made to measure the total mental workload and also task demanded resources. However, the available resources, the denominator of the equation, have received less attention from researchers. For this reason, we believe it necessary to develop methodologies that allow measuring available resources from a person, regardless from task demanded resources.

2 The Need for Measuring Available Resources Online

A main difference between task demanded resources and available resources is that the latter depend on both external and internal factors to the task. For example, sleep is an external factor that widely affects available resources: A well-rested employer will get to work with many available resources and vice versa. External factors are relatively easy to measure and have been investigated for decades. For example, there is an extensive literature of sleep deprivation effects on fatigue [3]. Those external factors are relatively easy to measure and manipulate.

However, it is more difficult to measure available resources variation during task performance, in order to do research on the factors that might affect them. In this sense, the difficulties for measuring them come from two facts:

1. We need to identify the factors that affect that variation during the task.
2. We have to decide what will be the dependent variable that we are going to measure, which correctly reflects available resources.

These two facts must be addressed in order to design reliable methods for measuring available resources during task performance.

2.1 Self Efficacy

One well-established factor, among others, that would affect available resources is the satisfaction that the operator has with how well she/he is doing in the task. In this sense, we can say that a "satisfied" person will be the one who perceives that she/he has done her/his work correctly from all points of view. Whereas, on the contrary, an "unsatisfied" person will be the one who perceives that she/he has done his work incorrectly or unsatisfactorily. In the latter case, unsatisfactorily means that she/he could have been done better in terms of saving resources (time, errors, etc.) or avoiding potential conflicts.

Satisfaction with performance in the task has been called Self-efficacy by Bandura [4] who defined as "... how well one can execute courses of action required to deal with prospective situations". This self-efficacy is an important determinant of how a person thinks, behaves and feels [5].

Self-efficacy consists of an executive control system of evaluation with two components:

1. A proactive targeting component
2. A negative feedback system that checks the results obtained with those expected and corrects the discrepancy

As for the first component, which proactively sets the objectives, it is worth noting that the term self-efficacy also includes what we would call "self-efficiency." In the ergonomic literature the word efficacy is used to refer to how well the task has been executed. However, the term efficiency refers to the amount of resources that have been used to perform the task correctly. It is precisely because of this efficiency component that we assume that self-efficacy will affect available resources.

The relation between self-efficacy and emotions is well established [5], as well as the relationship between self-efficacy and performance. In this regard, Alessandri et al. [6], found a positive correlation between self-efficacy, task engagement and performance. Thus, a bad performance on task development would trigger a negative self-efficacy perception. Nevertheless, little literature exists about the relationship between self-efficacy and mental workload.

We assume that a bad perceived self-efficacy would lead to a negative emotion and would increment available resources through activation of compensatory mechanisms [7]. According to the "Associative Network Theory" proposed by Bower and Forgas, an activated emotion nodule would extend its excitation to a series of connected indicators, such as facial expressions, physiological reactions and action trends [8]. Thus, a negative elicited emotion would trigger a measurable facial negative expression, as well as a variation in physiological activation.

2.2 Our Proposal: Online Measuring of Emotions

Emotions are the result of estimation processes during which individuals evaluate external stimuli or mental representations in terms of their relevance to their current needs and objectives, including aspects of their ability to cope with consequences [9].

Psychological research has been designed, among other things, to develop a model to identify what are the basic emotions, the dimensions that underlie them and the method to quantify them. The result of this research has been the establishment of a model where emotions are distributed in a two-dimensional space in which one dimension represents "Activation" or "Arousal" level and the second one represents "Valence", positive or negative, of that emotion (see Figure 1). Thus, for example, "fear" would be an emotion with a high level of activation and a negative valence, while "calm" would be a low-activation positive-valence emotion.

Thus, due to the fact that emotions are directly linked with physiological activation (arousal dimension) [10] and that it has been seen in the literature that physiological activation can be a good measure for available resources [7], we think that online measuring of emotions would be an interesting (advantageous) indicator of current available resources.

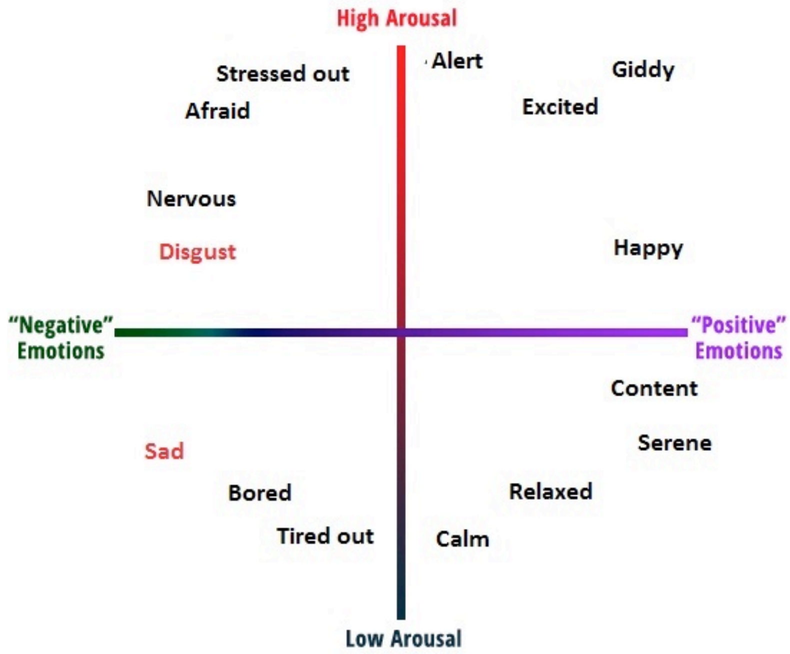

Fig. 1. The two dimensions of emotions. Valence (negative/positive) and arousal (low/high). Every single emotion can be placed on this two dimension graphic.

2.3 The Purpose of Our Study

In the empirical study presented below, we have asked a group of people to perform an air traffic control (ATCo) task on a simulation radar system. Participants' task goal is to avoid conflicts between aircraft. We have manipulated aircraft traffic density on the radar at different times to modify task difficulty and thus to affect the effectiveness avoiding conflicts. We expect that the perception of this efficacy will affect the psychophysiological activation and the emotions of the participants in the study.

We hypothesize that when complexity on task gets higher, performance would be negatively affected and so self-efficacy, that would lead, to a higher activation (higher pupil diameter) in order to successfully cope with the task which would also be reflected on emotions by an increase in a negative high arousal emotion ("disgusted").

3 Method

3.1 Instruments and Materials

Tobii T120 Eyetracker

To obtain the pupil diameter variable, an infrared-based eye tracker system, the Tobii T120 model marketed by Tobii Video System was used. This system is characterized by its high sampling frequency (120 Hz).

This equipment is completely non-intrusive, has no visible eye movement monitoring system and provides high precision and an excellent head compensatory movement mechanism, which ensures high-quality data collection. In addition, a calibration procedure is completed within seconds, and the freedom of movement it offers participants allows them to act naturally in front of the screen, as if it were an ordinary computer display.

Webcam Recorder

A full HD resolution webcam was used to record participants face during the experimental session. The camera was located in front of the participant, attached on the middle upper side of the computer screen, and would record participants face in order to obtain emotional data throughout the session development.

^{ATC}Lab-Advance Sofware

The software used for simulating ATCo tasks was an air traffic simulator called ^{ATC}Lab-Advance, which is available for free public download [11]. (See Figure 2).

The ^{ATC}Lab-Advance software provided a high level of realism (high similarity to real air-traffic control operational scenarios) as well as a simplified and easy handling, which allowed it to be used by all participants in several learning sessions. Additionally, it allowed strong experimental control of air traffic scenarios parameters since its XML code could be modified to develop scenarios consistent with research needs and objectives. For experimental scenario development, first, the static characteristics of the simulation environment (control sector size and possible pathways through which aircraft could travel) were defined. Next, aircraft quantity (density) and initial aircraft parameters (altitude, assigned altitude, speed, time of appearance on stage and planned route) were defined for each aircraft presented in the scenario. Once the structural and dynamic scenario parameters were established, a file that could be launched by the simulator was obtained. This file recorded a '.log' file with performance data for each participant during simulation.

Finally, we note that the [ATC]Lab-Advance simulator provided participants with every needed tool to carry out the ATCo task, such as the route (the aircrafts' fixed route was displayed), distance scale (which allows horizontal aircraft distance measurement) and altitude and speed change tools.

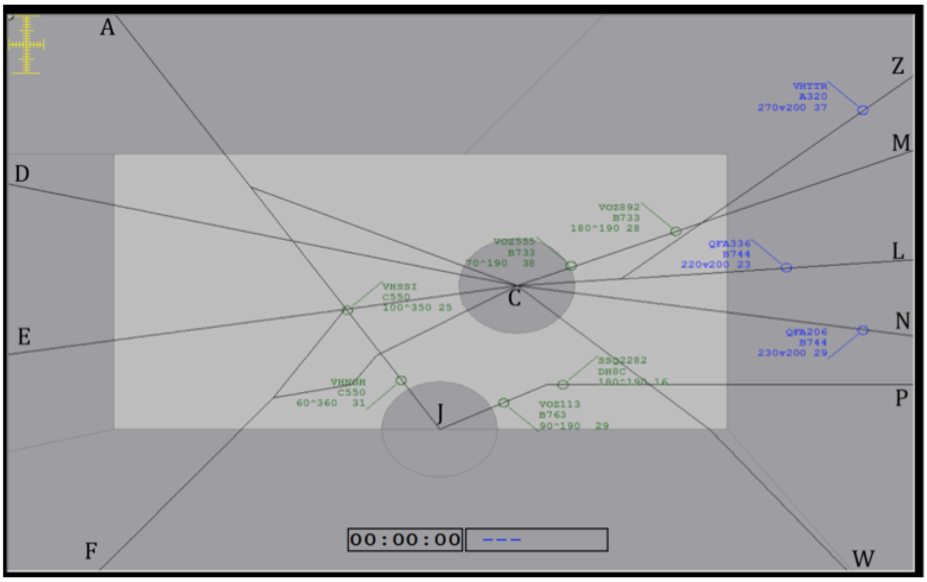

Fig. 2. ATCLab-Advance Software initial screenshot presented to participants.

Scenarios

Scenarios used in the study varied according to whether the participant was in the training or the experimental stage. Thus, during the training sessions, the standard scenarios provided by the software creator software were used, but a specific scenario was programmed by the experimenters to achieve experimental session goals.

The structural features of the experimental scenario was the following: 9 initial number of aircraft and 6 of them under participant control. For a better understanding, refer to Figure 2, which represents the initial simulator screen presented to participants; the capital black letters (starting route spots) do not appear on the radar screen. A total of 70 aircraft were presented. 50 came from external locations A,D,E,F,W,P,N,L,M and Z and 20 from inner locations. Specifically, 6 each came from A,D,E,F and W; 4 from P,N,L,M and Z; 12 from J (8 of them coming to P and 4 to A); and 8 from C (coming to A,D,E and F respectively).

Intraface: Emotion Software Recognition

IntraFace is a free research software for facial image analysis, which integrates algorithms for facial feature tracking, head pose estimation, facial attribute detection, synchrony detection, and facial expression analysis in a recorded/real-time video. It allows us to recognize participants' emotional status throughout the development of the ATCo tasks: The software detects the face within the video and monitor it, in order to recognize micro expressions through "face spots" relationship analysis, based on well stablished academic databases. That way, it offers a .log file with the emotional information in beta version 1.0.0: 4 emotions (happy, sad, disgusted, surprised) plus neutral face, experienced by participant through experimental session.

3.2 Participants

Thirty psychology students at the University of Granada participated in the study under the motivation of earning extra credit. Participants' ages ranged from 19 to 25, with an average of 22.45 and a median of 20. A total of 24 women (80%) and 6 men (20%) participated. It was essential that none of the participants would had any previous experience in ATCo tasks.

3.3 Procedure

Participants had to perform the ATCo tasks with the previously described ^{ATC}Lab-Advance software, so they had to learn how to use it before proceeding through the experimental stage in which the performance data were collected. Thus, we established 2 distinct stages:

1. **Training stage:** This took place for a total of 1.5 hours. The main objective of this first stage was for participants to familiarize themselves with the software so that they could handle it comfortably during the experimental stage. The training stage procedure was as follows: during the first day, once informed consent was given by the participants and their main task goal explained (maintaining air traffic security and preventing potential conflicts between aircraft), they started reading a short manual about the operation of the simulator for about 20 minutes and were asked to call the researcher once they had finished. Later, the participants sat in front of the running simulator while the researcher reviewed the manual in detail with the participants to ensure both correct understanding of the task and the assimilation of knowledge through content review. Participants then started using the simulator on their own while the researcher executed different air traffic control scenarios in order of difficulty (a total of six scenarios). The participants had free access to both the manual and researcher at all times in case of doubts or questions. The researcher also periodically checked the participants' performance to monitor their learning. Once the training period concluded, participants were ready for the experimental session, which took place the next day.
2. **Data collection stage:** The aim of the data collection stage, which lasted a total of 2 total hours, was to collect experimental data from participants while they

performed ATCo tasks. The participants were told the differences between the training and experimental stages, which were as follows: first, they would perform ATCo tasks in front of an eye-tracker system that had been previously calibrated. Secondly, participants were instructed to minimize head and body movements during the session. Lastly, they were told a webcam would record their face during task execution. Once the differences were explained, participants begin the task. At the end of the session, the participants were thanked and given extra study credits.

3.4 Experimental Room Conditions

Sessions were held in several different rooms, depending on whether the participant was in the training or data collection stage: During the training stage, participants could work in one of three different rooms equipped for training with the simulator, and no special attention to room conditions was needed. However, during the data collection stage, standardizing room conditions was essential. Thus, the testing rooms were temperature controlled to 21ºC, and lighting conditions (the main extraneous variable in pupil diameter measurement) were kept constant, with constant artificial lighting and no natural light in the rooms. Moreover, participants always sat into the same place, a comfortable chair spaced 60 cm from the eye-tracker system.

3.5 Variables

Independent Variable: Task Complexity

Complexity on the task is directly related with demanded mental resources: The more complex the task gets the higher mental resources requires. Thus, in order to successfully cope with their task, participants would have to increase available resources, which would be reflected in pupil diameter and emotions.

We have manipulated complexity on task by modifying aircraft traffic density (occupancy) presented through the 2 hour that lasted the experimental session, with two main complexity peaks at intervals 5 (higher complexity peak) and 10.

Although it is well known that complexity of ATCo scenarios are not only dependent by traffic density, "it seems from the literature that no single traffic characteristic has been as cited, studied and evaluated as has traffic density in terms of its influence on complexity and controller workload" [12, page 7]. Furthermore, it is very easy to measure and manipulate in experimental context. However, despite we find it appropriate only manipulating air traffic density in this first experimental approach, we do not rule out taking in account some other variables such as airway route structure to modify task complexity in future research.

Dependent Variables:

Performance

Performance is linked with self-efficacy and so with feelings: We assume that a bad performance would lead to a bad perceived self-efficacy, which would provoke negative feelings and would trigger, among other variables, compensatory mechanisms to increment available resources [7].

Even though performance related to ATCo task can include a wide range of indicators, including for example flown miles vs planned miles, “it has been noted that traffic density is not only an important driver of complexity, but also correlates well with conflict rate” [12, page 8]. In that sense, we found it appropriate to consider using conflict rate as our performance indicator, taking on account its well stablished correlation with our independent variable: Traffic density.

With regard to methodological aspects of our study, we operationalized the dependent variable ‘performance’ by dividing the number of conflicts by total number of aircraft present in the radar at a given time, since we thought that using only the number of conflicts as a performance measurement would not be an appropriate performance indicator, since it largely reflects air traffic density.

Pupil Dilatation

It is well established in literature that pupil diameter size positively correlates with the amount of resources used to perform certain task (available resources) [13-17]. Thus, we think we can use this variable as a solid indicator of available resources used by participants, in order to study its correlation with emotions.

Regarding to methodological issues about the variable, while our eye tracking system allows continuous sampling rate recording at 120Hz, we set a total of 12 intervals lasting 10 minutes each to facilitate subsequent analyses. However, given high inter-individual variability relative to average pupil size, it is not possible to analyze the variable without a process of standardization. To do this, we took the first interval (10 minutes) as a reference of standard individual average pupil size, which was then subtracted from the obtained value in each of the remaining 11 intervals, thereby giving a differential standardized value that allowed us to compare participants. Analyses were carried out both for the left and right pupils. A negative value meant that the pupil was contracting while a positive value meant that it was dilating.

Emotions

Emotions can be measured due to the fact that they are facially expressed. Therefore, in order to analyze facial expressions, Ekman and Friesen proposed the Facial Action Coding System (FACS) [18], which conform databases of combined based action units which describes different emotions.

Many emotion recognition software has been developed based on FACS, but normally people cannot afford using it due to its high price. However, we have decided to use Intraface software, which is a free emotion recognition software released for research purposes whose current beta version offers 4 of the main emotions plus neutral, as it has been said in previous lines.

3.6 Experimental Design

For testing the effect on performance the design was one-way within-subject design with interval (11 levels) as the only variable. For testing the psychophysiological activation the experimental design was a within-subjects factorial design with intervals (11 levels) and pupil (right pupil and left pupil) as the two dependent variables. Finally, for testing the effect of performance on emotions, the design was a within-subjects factorial design with interval (11 levels) and emotions (with 5 levels).

4 Results

Performance showed a significant effect of Intervals $F(10,290) = 10.74$, $MSe = .001$, $p < .001$. This effect is due mainly to a worse performance in interval 5. There was also a decrease in performance in in interval 10. Those intervals were when we inserted more aircrafts.

Pupil diameter showed a main effect of intervals, $F(1,29) < 1$, $MSe = .038$, $p > .05$, Interval $(10, 290) = 11.34$, $p < .01$, and a non significant interaction of Pupil by interval $(10,290) = 1.108$, $p > .05$. These results meant that in interval 6 there was an increase in activation, but that activation was equal for both right and left pupils.

There were significant main effect emotions $F(3, 87) = 3.67$, $MSe = .065$, $p < .05$, and Intervals $F(20,290) = 3.55$, $MSe = .001$. However, more important was the significant interaction of Emotions X Intervals $F(30,870) = 1.69$, $MSe = .002$, $p < .05$. As can be seen in Figure 2, the emotion disgusted showed a higher values during the first five intervals and increase in interval 6 that corresponded with the increase of pupil diameter. We can interpret that results by considering that disgust is an emotion that is negative and shows high activation. The poorer performance in interval 6 corresponded to a work auto-efficacy. This extra activation is shown also in the pupil diameter.

As happens with the pupil diameter, disgust decreased due to the better performance and better self-efficacy. However, the bad performance and the previous poor perceived self-efficacy let a negative emotional feeling as can be seen the increase in the sad emotion that has low arousal and negative valence.

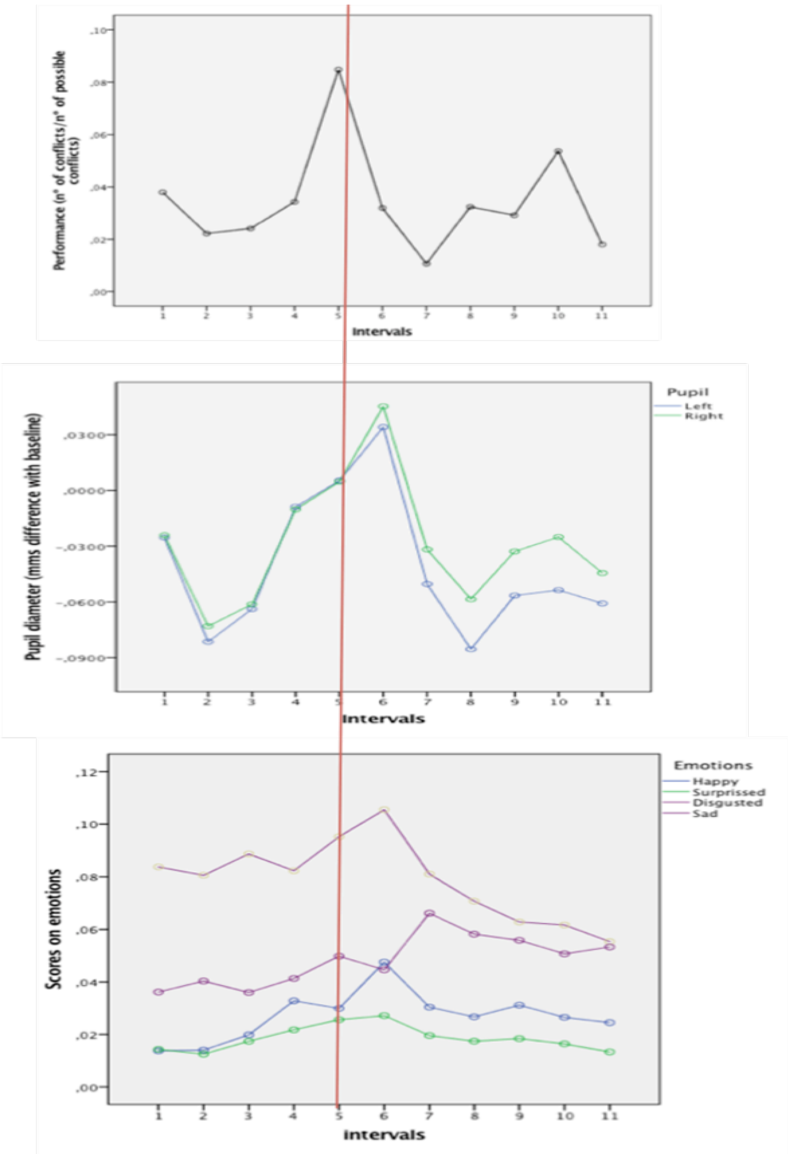

Fig. 3. Results graphics, from up to down: Performance, pupil dilatation and emotions.

5 Discussion

Nowadays it is clear and nobody discuss about the importance of mental workload, especially when it comes to talk about work efficiency and work-related accidents. As it has correctly been said in precedent lines, mental workload not only depends on demanded resources but in available resources, which we think is the key to improve our strategies to deal with workload related problems such as mental fatigue or mental underload problems. However there has been made many research on measuring demanded resources and "external" available resources factors, but few on measuring online available resources, despite its high importance.

We firmly believe that having an online measurement of available mental resources would strongly benefit workload related problems, since the denominator of the equation, indeed the most important component of the formula, could be monitored and controlled in most situations, preventing the outcome of workload faults. The actual issue is that nobody is using this method mainly because of 2 principal reasons: 1) lack of research in available resources and 2) lack of advantageous ways of measuring it.

Our purpose with this study was to begin to fill this literature gap with a first explorative and correlational approach to find out an advantageous possible way for online measurement of available resources, through emotion measuring, which our data seem to sustain. Nevertheless, our results must be interpreted with caution due to the fact of some methodological limitations such as a low number of non-ATCo participants, the use of limited independent and dependent variables and the correlational nature of the study itself.

As it can be seen in the results in 5^{th} interval we can notice a bad performance peak which would involve a higher resources demand in order to successfully confront the task. This increase in demanded resources would be reflected on the next interval (6^{th}) on both physiological and emotional variables, as an increase in pupil diameter and in "disgusted" emotion respectively (available resources).

Moreover, in 6^{th} interval performance improves again, resulting within the next intervals in a reduction in physiological activation levels (pupil size decrement) and in the negative felt emotion "disgusted". However, curiously it is also noticed an increment in "sad" emotion, which could be explained about the fact that the worse performance obtained previously would negatively have affected perceived self-efficacy, leading to a state of personal discomfort.

It should be taking in account in further research that it would be interesting to include qualitative users' subjective comments during or after performing the task in order to improve our knowledge about the findings. For example, it would be interesting to ask participants about their own emotional valence and arousal perception during task development.

6 Conclusions

Our result showed a promising methodology in using emotional measurement as an indicator of available resources: Taking in account our experimental data, it seems that the negative high arousal emotion "disgusted" can be proposed to be used as an indicator of online available mental resources, due to the fact that positively correlates with other well-stablished indicators such as pupil diameter. Nonetheless, it is worthy of note that this results should be taken with caution due to methodological limitations and should be validated in the real ATCo setting, as well as in other applied fields.

We are convinced that measuring emotions online would be an advantageous way of measuring available resources, considering its numerous benefits: a) it is cheap, just requires a simple webcam and the special software, some of them free; b) it is not intrusive, the webcam goes unnoticed and will not affect workers performance; c) it is easy to implement, attaching it on the computer screen and d) it is easy to use, with no special training requirements. And would be better than measuring it through other physiological activation measurements such as pupil dilatation, since this methods requires high-cost intrusive technology which requires special and limiting conditions to function properly (such as illumination control). Nevertheless, we must be cautious about the fact that emotions are not only influenced by self-efficacy, but also by other numerous factors. Thus, we believe further research should be done to validate the methodology and to find out ways of controlling extraneous variables that may affect emotions during task development.

In conclusion, we need to deepen the study of available resources and the way to measure it in order to improve efficiency and to prevent work-related problems with a new approach that seems to be next step in workload knowledge achievement.

References

1. Wickens, C.D.: Multiple resources and mental workload. Human Factors, 50: 449-455 (2008). doi:10.1518/001872008X288394
2. Hockey, R.: The psychology of fatigue: Work, effort and control. Cambridge University Press (2013).
3. Dawson, D., McCulloch, K.: Managing fatigue: it's about sleep. Sleep Med. Rev. 9 (5), 365-380 (2005). doi:10.1016/j.smrv.2005.03.002
4. Bandura, A.: Self-efficacy mechanism in human agency. American psychologist, 37(2), 122 (1982). doi: 10.1037/0003-066X.37.2.122
5. Bandura, A. Editorial. American Journal of Health Promotion, 12(1), 8-10 (1997).
6. Alessandri, G., Borgogni, L., Schaufeli, W. B., Caprara, G. V., & Consiglio, C. From positive orientation to job performance: The role of work engagement and self-efficacy beliefs. Journal of Happiness Studies, 16(3), 767–788. (2015). doi:10.1007/s10902-014-9533-4
7. Muñoz-de-Escalona, E., Cañas, J.J.: Subjective Mental Fatigue and the Compensatory Mechanism for Task Engagement. Submitted (2017).
8. G.H. Bower, J.P. Forgas. Afecto, memoria y cognición social. E. Eich, J. Kihlstrom, G. Gower, J. Forgas, P. Niedenthal (Eds.), Cognición y Emoción, Descleé De Brouwer, Bilbao, Spain, pp. 83–169 (2003).

9. Scherer, K. R. Psychological models of emotion. The neuropsychology of emotion, 137(3), 137-162 (2000).
10. Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology, 45(4), 602-607. doi:10.1111/j.1469-8986.2008.00654.x
11. Fothergill, S., Loft, S., & Neal, A.: ATC-labAdvanced: An air traffic control simulator with realism and control. Behavior Research Methods, 41(1), 118-127 (2009). doi:10.3758/BRM.41.1.118
12. Hilburn, B. Cognitive complexity in air traffic control: A literature review. EEC note, 4(04) (2004).
13. Kahneman, D., & Beatty, J.: Pupil diameter and load on memory. Science, 154, 1583-1585 (1966). doi: 10.1126/science.154.3756.1583
14. Beatty, J., & Lucero-Wagoner, B.: The pupillary system. Handbook of Psychophysiology, 2, 142-162 (2000).
15. LeDuc, P. A., Greig, J. L., & Dumond, S. L.: Involuntary eye responses as measures of fatigue in US army apache aviators. Aviation, Space, and Environmental Medicine, 76(Supplement 1), C91 (2005).
16. Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P.: Eye activity correlates of workload during a visuospatial memory task. Human Factors, 43(1), 111-121 (2001). doi: 10.1518/001872001775992570
17. Ahlstrom, U., & Friedman-Berg, F. J.: Using eye movement activity as a correlate of cognitive workload. International Journal of Industrial Ergonomics, 36(7), 623-636 (2006). doi: 10.1016/j.ergon.2006.04.002
18. P. Ekman, W.V. Friesen, J.C. Hager. Facial Action Coding System: Research Nexus, Network Research Information, Salt Lake City UT (2002).