Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P312-cd

Evaluation of workload for operators in the aeronautical sector

Sarah Francisca de Souza Borges

Aerospace Science and Technology, Aeronautics Institute of Technology, Brazil. E-mail: sarah@ita.br

Patrícia de Morais

Aerospace Science and Technology, Flight Testing and Research Institute, Brazil. E-mail: castilhods@msn.com

Mayara Gomes Bovo

Aerospace Science and Technology, Flight Testing and Research Institute, Brazil. E-mail: castilhods@msn.com

Diogo Silva Castilho

Aerospace Science and Technology, Flight Testing and Research Institute, Brazil. E-mail: castilhods@msn.com

Moacyr Machado Cardoso Junior

Aerospace Science and Technology, Aeronautics Institute of Technology, Brazil. E-mail: moacyr@ita.br

The workload affects the physical and mental aspects of the individual and consequently their performance. The concept comes from the interaction between task requirements and human achievement capacity. The objective of this study is to evaluate the mental and operational workload of operators in the aeronautical sector. This experiment involved the collaboration of four engineers (Structure, Stress and Manufacturing Designer; with 9 to 17 years of experience in the area) and four operators (between 18 and 35 years of experience in the area) who work in aircraft production. Developed by Hart and Staveland in 1988,the NASA TLX is a multidimensional rate procedure that provides a global score of the Workload based on a weighted average of evaluations in six subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration. This method was used to apply and evaluate four tasks with different levels of difficulty. For the group of engineers, the tasks were part design and nailing. For the group of production operators, the task was drilling and driving. As a result, in the first group, the mental demand was greater, as calculations and analyzes are required. In the second group, the effort demand was the highest, showing that physical and mental demand need to be applied in the same measure.

Keywords: workload, nasa tlx, safety, aeronautical.

1. Introduction

After the pandemic, there was an increase in vacancies in the job market, leading to difficulties in retaining engineers and production operators in medium-sized companies in the aeronautical sector, because many professionals make a career in smaller companies and then seek opportunities in large companies or a job that offers home office and better quality of life.

In this sense, the top management of the companies seeks to verify the working conditions and find out what is generating greater fatigue for workers.

It is necessary to balance the workload of operators in order to improve performance and increase engagement. However, in the aeronautical industry, there is high pressure with the deadline, and lack of study of the workload level of the tasks by operator and sector. In addition, everyday life is hectic, requiring practical and functional methods.

In daily tasks, professionals perform different levels of workload (with implications such as fatigue, stress, illness and accidents). In order to maintain acceptable performance, task owners need to scale the system workload for design and operation (Hart 2006).

Identifying areas where users experience significant levels of mental workload and attempting to regulate this through system redesign can minimize human error (Davenport and Beck 2001; Longo et al. 2022).

Among the workload evaluation methods, the NASA Task Load Index (NASA_TLX) stands out for its practicality, today there is even a cell phone application for the calculation. This consists of six subscales: mental, physical, temporal, frustration, effort, and performance. These dimensions were selected after an extensive analysis of primary factors that help define the subjective experience of workload for different people working with a variety of activities ranging from simple laboratory tasks to flying an aircraft.

The weights of each subscale are decided by each participant at the beginning of a study. The first part is followed by the pairwise comparison of the subscales and the second part by the score (from one to one hundred) of each subscale. In the end, each subscale rating provided by the participant during the study is multiplied by the appropriate weight, developing a composite tailored to individual workload definitions (Hart 2006).

This study is important as selecting an appropriate and practical workload measure is still difficult due to the multidimensional nature of workload and the fact that different measures are selectively appropriate for different issues, tasks and environments (Hart et al. 1988).

In this field, the analysis oftraining, performance strategies development, and subjective workload evaluations can be may lead to the development of human performance models suitable for multitasking supervisory control systems (Hart, Battiste, and Lester 1984).

Thus, while human adaptability and creativity are essential for the effective functioning of complex systems, human capabilities and limitations are also a limiting factor in overall system performance. For this reason, operator workload is an important factor that must be considered when assessing the suitability and feasibility of operating requirements, system designs, and training procedures.

The objective of this study is to evaluate the mental and operational task-load of operators in the aeronautical sector.

2. Materials and Method

2.1. Definitions workload

Workload is a term used to describe the cost for the human being to fulfil task requirements in the human-machine system. This "cost" can be reflected in the depletion of attentional, cognitive or response resources, the inability to carry out additional activities, emotional stress, fatigue or decreased performance. A workload can be physical or mental, so the more demanding and complex the task, the harder you have to work to accomplish it (Lean and Shan 2012; Vidulich and Tsang 2015).

2.2. Method for workload analysis

Thus, the mental workload is based on the principle that the level of mental exhaustion will be associated with the worker's ability to perform his or her job. There may be underload (referring to the underutilization of the worker's abilities) or overload (associated with the use of capabilities beyond the worker's physical and psychological limits).

Cooper and Harper (1969) were authors who initially worked to create subjective scales. This study was basically divided into two stages:

The Sheridan-Simpson scale is a modification of the original Cooper-Harper scale, to which three subjective load assessment dimensions (stress, effort and load) were added. This scale was later modified by Wright-Patterson, originating the SWAT method (Subjective Workload Assessment Technique).

The second stage gives rise to the NASA-Ames method, developed at the NASA research laboratory in 1981. This lab used bipolar scales to find the minimum number of dimensions needed to indicate individual differences in mental workload. In 1987, this method originated the Nasa-TLX (Task Load Index), used in this study (Hart and Staveland 1988).

2.3. NASA TLX Method

Workload assessment techniques abound, however the most commonly used methods use subjective ratings. The NASA TLX method was proposed as a classification technique to reduce variability between subjects. That's because workload sources are numerous and vary by task. The proposed technique is multidimensional, it considers specific sources of workload relevant to a given task in order to calculate the global workload classification (Hart and Staveland 1988; Vargas et al. 2022).

Rating scales can be one-dimensional or multidimensional, requiring judgments about various task-related or psychological variables (Hart and Staveland 1988).

Rating scales are the most practical and applicable measure of workload. This is because they are easy to implement and score, appropriate in most environments, acceptable to most operators, and a way to validate assumptions (Hart and Wickens 1990). Therefore, the dimensions of the workload vary according to the tasks and the subject's perception of carrying them out (Guimarães and Carvalho 2013).

Table 1 shows the subscales and definitions used during the interviews.

Table 1. Rating Scale and Definition (1st Phase) (Hart and Staveland 1988).

Rating Scale Definitions		
Title	**Endpoints**	**Descriptions**
Mental Demand	Low / High	How much mental and perceptual activity was required (e. g., deciding, calculating, thinking, searching, etc.)?
Physical Demand	Low / High	How much physical activity was requirired (e.g., pulling, pushing, controlling, etc.)?
Temporal Demand	Low / High	How much time pressure did you feel due to the rate or pace at which the task elements occurred?
Effort	Low / High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Performance	Good / Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)?
Frustation Level	Low / High	How insecure, discouraged, stressed versus secure, relaxed and content did you feel during the task?

In applying the method, subscale ratings ranging from 1 to 100 in 5-point increments (Fig. 1) are given either verbally or by selecting a position along a scale presented on a form or on a computer screen. In addition, raters quantify at the outset the relative importance of each factor, based on their experience. These values, which range from 0 to 5, are used to weight the magnitude ratings when calculating the overall workload score. Diagnostic information is identified by analysing variations in subscale ratings, as well as by the weight assigned to each factor (Hart and Wickens 1990).

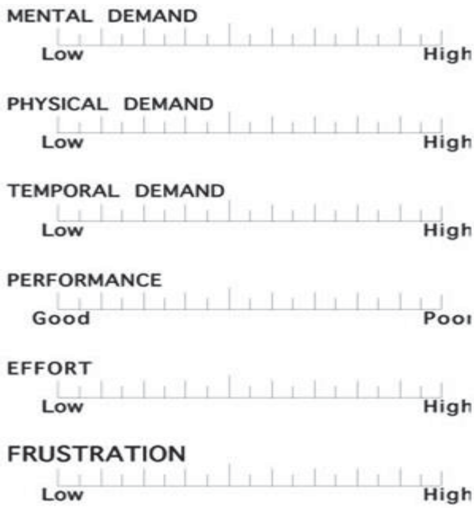

Fig. 1. Rating Scale (2nd Phase) (Hart and Staveland 1988).

3. Results

This experiment involved the collaboration of four engineers (Structure, two of Stress and Manufacturing Designer; with 9 to 17 years of experience in the area) and four operators (between 18 and 35 years of experience in the area) who work in the production of large aircraft.

3.1. Interview and data collection

In order to use the NASA-TLX, the experiment is divided into two phases, the first for comparison and weighting of subcriteria, and the second for apportionment between 1 and 100 of each subscale. In this way, it is possible at the end to multiply the relative weight of the subscale with the absolute value assigned to it.

Prior to the interviews, a meeting was held with experienced professionals in the areas (operation and engineering) to define the common tasks that could be used for workload analysis were verified.

For the group of operators, the following tasks were selected: 1) Conceptualization of the part and 2) Detailing of the part. For the group of engineers, the tasks were: 3) drilling and 4) driving.

The first task, Conceptualization of the part, the actions were: Initial design of the engineering part, evaluation of the use in the project, analyse the function of the part and its integration with the system, definition of material.

On the second task, Detailing of the part, the actions were: Dimensioning of all measurements of the piece, definition of tolerance based on the use of context, definition of drawing notes (rules and exceptions), analysing whether it will have heat treatment and/or painting.

Third task: Drilling is a machining process that aims to generate holes, most often cylindrical, in a part, through movement. Drilling needs to be precise in location and hole diameter, as it is practically irreversible.

Forth task, driving: Solid Rivets driving. The hammer applies continuous blows to the rivet's factory head, where on the other side of the rivet, with the use of a bucking bar, the rivet expands and forms a uniform counterhead, fixing the component properly.A brief introduction describing the purpose of the simulation and the survey was given to participants prior to the experiment.

A brief introduction describing the purpose of the simulation and the survey was given to participants prior to the experiment.

During the interviews, the questions and definitions were explained, allowing time for responses. The answers were marked on papers distributed to the collaborators.

Subsequently, the NASA-TLX application provided by Google Play app was used for data input and method application. Below, an image of one of the outputs (Fig. 2).

In task 1, the collaborators took one minute and twenty seconds in the first phase and fifty-eight seconds in the second. In task 2, the collaborators took one minute and twenty-eight seconds in the first phase and fifty seconds in the second.

In task 3, the collaborators took one minute and ten seconds in the first phase and one minute in the second. In task 4, the collaborators took one minute and thirty seconds in the first phase and fifty-five in the second.

Fig. 2.NASA-TLX app

3.2. Analysis of results

Tables 2 and Table 3 show the results of tasks 1 and 2 with the group of production operators, highlighting mental task-load for task 1 and temporal demand for task 2.

This result makes sense, because for task 1 (part design) it requires greater mental task-loads, more concentration on the part of the engineer, thinking about technical solutions and redesigning the model to meet the customer's requirements.

Table 2. Task 1 (part design) - Group of Engineers

	Men tal	Phys ical	Temp oral	Perform ance	Eff ort	Frustra tion
1	300	0	195	120	150	120
2	320	15	165	120	0	140
3	160	0	475	120	210	80
4	300	0	200	140	50	30
M	270	3	258	125	102	92
%	16%	0%	15%	7%	6%	5%

Table 3. Task 2 (part nailing) - Group of Engineers

	Ment al	Phys ical	Tem poral	Perfor mance	Effort	Frustra tion
1	320	30	150	300	0	100
2	240	0	280	125	150	65
3	70	0	475	60	150	300
4	165	0	260	75	90	40
M	198	7	291	140	97	126
%	12%	0%	17%	8%	6%	7%

For task 2 (part nailing), the time demand may actually be greater due to the shorter time dedicated to this phase and because there is already a smaller margin for changing the delivery date of the part. Being a job that requires greater research and detailing of the piece, and if the previous phases do not finish on the scheduled date, this task will also suffer over time.

In the end, as shown in Table 4, the temporal demand subscale represented a greater workload. This is due to the need to meet customer deadlines and manage the schedule with little margin for error.

Table 4. Total and percentage by subscale

Men tal	Physi cal	Temp oral	Perfor mance	Effort	Frustra tion
468	11	550	265	200	218
25%	1%	29%	14%	10%	11%

The sum of the results was performed in excel. Tables 5 and Table 6 reveal the results for tasks 1 and 2 with the group of engineers. It is noteworthy that in task 1, the mental and temporal task-loads were higher and tied, and for task 2, the physical demand was greater.

Table 5. Task 1 (drilling) - Production Operators Group

	Ment al	Phys ical	Tem poral	Perform ance	Effor t	Frustr ation
1	360	130	400	120	60	0
2	150	0	280	75	65	350
3	190	60	195	80	50	280
4	270	225	85	140	210	60
M	242	103	240	103	96	172
%	13%	5%	13%	5%	5%	9%

For task 3 (drilling), the mental and temporal demand was greater.According to the operators it requires a lot of precision for the hole, a human error could be irreversible. As for the time demand, the operators' project is in the final phase of assembling the plane, that is, with short deadlines and close to the end.

Table 6. Task 2 (driving) - Production Operators Group

	Mental	Physical	Temporal	Performance	Effort	Frustration
1	60	320	270	25	320	130
2	45	240	220	60	80	110
3	100	300	120	50	200	0
4	0	285	240	200	360	55
M	51	286	212	83	240	73
%	3%	15%	11%	4%	13%	4%

For task 4 (driving), the physical demand is greater, as the drilling has already been done, requiring manual labor to drive rivets.

In the end, as shown in Table 7, the temporal demand subscale represented a greater workload. This occurs due to the phase of the project that the operators work on, being the final assembly line of the aircraft, with deadlines close to the delivery of the final product to the customer.

Table 7. Total and percentage by subscale

Mental	Physical	Temporal	Performance	Effort	Frustration
293	390	452	187	336	246
15%	20%	24%	10%	18 %	13%

In addition, the python language was used to plot the Boxplot of the data collected with the engineers, in tasks 1 (Fig. 4) and task 2 (Fig. 5).

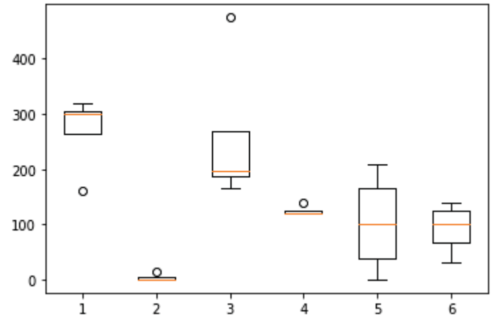

Fig. 3. Boxplot - Group of Engineers - Task 1 (part design)

Surprisingly, four outliers were obtained for the subscales: Mental, Physical, Temporal and Performance.

This is due to the high variation of responses, one reason being that the experience in the area by respondents ranges from 9 to 17 years.

For example, for the third engineer, the mental demand was less required, the temporal demand was more required and the performance was more required than for the others (he has ten years of experience in the area). The second engineer pointed out a physical demand, and for the others it was zero, since the task is the design of a part in a computer software.

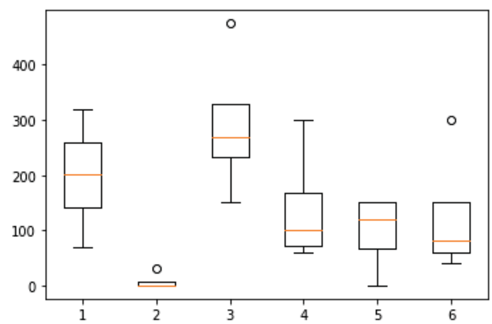

Fig. 4.Boxplot - Group of Engineers - Task 2 (part nailing)

For the first engineer, the task of detailing the part requires physical effort, unlike the others who indicated zero. And for the third engineer, the time demand and frustration were much greater than the others.

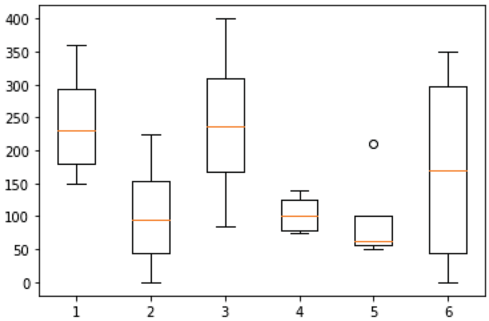

Fig. 5.Boxplot - Production Operators Group - Task 1 (drilling)

It is possible to notice a high variability of the data regarding the level of Frustration and an outlier regarding the level of Effort.

In this case, one of the production operators selected the effort subscale three times, showing that he demanded mental and physical task-load together to accomplish the task, he has thirty-three years of experience.

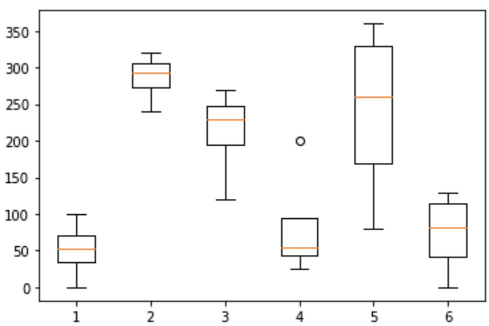

Fig. 6.Boxplot - Production Operators Group - Task 2 (driving)

Boxplot graphs were also plotted for the data collected with the group of production operators (Fig. 6 and Fig. 7).

It is possible to notice a high variability of the data regarding the level of Effort and an outlier regarding the level of Performance.

In this case, a production operator with thirty-three years of experience selected the Performance subscale four times, showing that he demanded high performance to perform the task..

This result was not expected either, three outliers were obtained for the subscales: Physical, Temporal and Frustration.

4. Discussion

As a result, for the group of engineers, the task of designing the part requires greater mental task-load, as calculations and analysis are required. For the same group, for the task of detailing the part, the time demand was greater, as the delivery time is shorter than in the previous phases of the project.

As for the group of production operators, in the drilling task, the mental and temporal task-load was greater, as they need a lot of precision to hit the holes in the aircraft. And, for the driving task, the most demanding subscale is the time to finish the job.

5. Conclusions

The analyses were presented to the managers who follow the tasks of the two groups, responding to the validation of the result.

The assessment of the workload is important, as even the demands of tasks considered simple may be exceeding the capabilities of the operators, this occurs when apparently human limitations reflect poorly designed controls or views, inadequate automation or insufficient training, for example. The individual operator's performance, depending on whether the workload is high or low, will reflect on the overall effectiveness of the system.

Based on subjective measurements and their association with other ways of analysing aspects of the worker-work relationship, a series of investigations can be carried out that go beyond measuring the workload.

Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

Davenport, Thomas H., and John C. Beck. (2001). "The Attention Economy." *Ubiquity* 2001 (May): 1. https://doi.org/10.1145/376625.376626.

Guimarães, Bruno, and Maria Carvalho. (2013). "Avaliação Da Carga Mental de Trabalhadores de Uma Empresa de Instalação Civil Em Recife, PE." *Lecturas Educación Física y Deportes (Buenos Aires)* 18: 1–6.

Hart, Sandra G. (2006). "Nasa-Task Load Index

(NASA-TLX); 20 Years Later." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50 (9): 904–8. https://doi.org/10.1177/154193120605000909.

Hart, Sandra G., Vernol Battiste, and Patrick T. Lester. (1984). "POPCORth A Supervis,Ory Control Simulation, for. Workload and Performance Research." Washington, D.C. https://ntrs.nasa.gov/citations/19850006206.

Hart, Sandra G., E. James Hartzell, James W. Voorhees, Nancy M. Bucher, and R. Jay Shively. (1988). "An Integrated Approach to Rotorcraft Human Factors Research." Washington, D.C. https://ntrs.nasa.gov/citations/19880007270.

Hart, Sandra G., and Lowell E. Staveland. (1988). "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In *Adv. Psychol.*, 52nd ed., 139–83. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hart, Sandra G., and Christopher D. Wickens. (1990). "Workload Assessment and Prediction." In *Manprint*, 257–96. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-0437-8_9.

Lean, Ying, and Fu Shan. (2012). "Brief Review on Physiological and Biochemical Evaluations of Human Mental Workload." *Human Factors and Ergonomics in Manufacturing & Service Industries* 22 (3): 177–87. https://doi.org/10.1002/hfm.20269.

Longo, Luca, Christopher D. Wickens, Gabriella Hancock, and P. A. Hancock. (2022). "Human Mental Workload: A Survey and a Novel Inclusive Definition." *Frontiers in Psychology* 13 (June). https://doi.org/10.3389/fpsyg.2022.883321.

Vargas, Vanessa Bertholdo, Mario T. Crema, Mayara Gomes Bovo, Moacyr Machado Cardoso Junior, and Jefferson de Oliveira Gomese. (2022). "Workload Analysis of Health Workers During COVID-19 Vaccination and Organizations of Queues at UBS in City of Franca (SP-BR)." In *Safety and Reliability Conference (ESREL 2022)*, 7. Dublin, Ireland. https://doi.org/10.3850/978-981-18-5183-4_J01-01-069-cd.

Vidulich, Michael A., and Pamela S. Tsang. (2015). "The Confluence of Situation Awareness and Mental Workload for Adaptable Human–Machine Systems." *Journal of Cognitive Engineering and Decision Making* 9 (1): 95–97. https://doi.org/10.1177/1555343414554805.