Human Factors in Design, Engineering, and Computing, Vol. 159, 2024, 1288–1296
https://doi.org/10.54941/ahfe1005696

Cognitive Cost Assessment in Aeronautical Tasks: New Objective and Sensitive Method

Marianne Jarry, Jean-Christophe Hurault, Grégory Froger, Anne-lise Marchand, and Colin Blättler

Centre de Recherche de l'Ecole de l'air et de l'espace, Salon-de-Provence, France

ABSTRACT

The future fighter aircraft, such as the Next Generation Fighter central to the Future Combat Air System, are being developed in response to evolving Multi-Domain conflicts. AI-based onboard systems will be limited in their tasks for technical and ethical reasons, leaving some tasks to be managed by the aircrew. This increases the cognitive load on pilots. Assessing this cognitive cost can help identify particularly costly tasks and develop better training and human-system interfaces. This study aims to validate a method for evaluating cognitive cost using the complex span task protocol on basic aeronautical tasks performed by novices. The protocol involves alternating between memorization and processing during flight simulations. Results indicate that the high difficulty phase is more cognitively costly than the low difficulty phase. Evaluating performance using this method during simulated flights provides an objective indicator of cognitive cost and can be extended to more complex tasks, offering a time-efficient and non-invasive alternative to physiological measurements.

Keywords: Cognitive load, Cognitive cost evaluation, Complex span, Aeronautic

INTRODUCTION

The context of conflicts is evolving towards what is called the Multi-Domain, involving increased management of data such as communications and asset positions. In response, it is necessary to develop new types of aircraft, such as the Next Generation Fighter (NGF) in the European project Future Combat Air System (FCAS). With the integration of Artificial Intelligence (AI) systems on board, interactions between humans and systems are becoming more complex. AI-based onboard systems will be limited in their tasks for technical and/or ethical reasons, thus leaving some tasks to be managed by the aircrew. Alongside these tasks is the overall management of the onboard system, further increasing an already high cognitive cost. Assessing this cognitive cost would help identify particularly costly tasks and develop training and human-system interfaces accordingly. This study aims to validate an objective methodology for evaluating the cognitive cost of aeronautical tasks.

The present study is part of a broader project dedicated to the comprehensive assessment of the most cognitively demanding aeronautical tasks for ab initio pilots in real flight, with the overarching objective of developing ground training to remediate these difficulties. Aviation, as a

© 2024. Published by AHFE Open Access. All rights reserved.

dynamic and intricate system, exposes pilots to discrete and simultaneous activities right from the outset of their training. These activities cover tasks such as retaining navigation information in working memory while concurrently monitoring the sky or engaging in mental calculations (e.g., drifts, fuel). For novices, these multitasking situations impose a cognitive cost (Wickens, 2002), potentially compromising both aviation safety and skill acquisition. While a unanimous definition of cognitive cost remains elusive (Longo et al., 2022), this study considers this cost to represent the portion of an individual's cognitive resources mobilized when performing a task. In ecological settings, working memory plays an essential role in task execution (Baddeley, 1996, 2000; Cowan, 1999), making working memory models apt for assessing the cognitive cost of tasks. Notably, an overload of working memory has been associated with less effective learning (Sweller, 1994) and, in extreme cases, the inability to perceive potential dangers (e.g., inattentional blindness risk; Kreitz et al., 2016).

Various approaches already exist to assess cognitive cost, each with its merits and drawbacks (Hancock & Matthews, 2019; Longo et al., 2022). In addition to subjective measures (e.g., NASA TLX; Hart & Staveland, 1988), numerous technological tools have been developed for the objective assessment of cognitive cost (e.g., nIRS, EEG; Dehais et al., 2020). However, constraints associated with these tools, such as navigability and cabin space, hinder their application in real-world activities. Thus, finding an objective, non-invasive and operational-friendly method to evaluate cognitive cost is essential.

The method employed is adapted from the approach used by real-flight instructors with novice pilots. Traditionally, instructors ask students questions during specific flight phases, and an inability to answer signifies cognitive overload. However, this method lacks standardization and does not permit a quantifiable measurement of cognitive cost. To address this, a method adapted from the Time-Based Resource Sharing (TBRS) model by Barrouillet and Camos (2007) is proposed. This model enables an objective measurement of the cognitive cost of a task by assessing a complex span of working memory. In our study, the complex span involved a processing task inserted between each item presentation (see Figure 1), resembling the method employed by instructors. According to the TBRS model, the complex span is influenced by two factors: the time interval between item presentations (longer intervals result in lower span) and the cognitive cost of the processing task (higher cost leads to lower span). By controlling the duration between item presentations, only the cognitive cost of the processing task varies the span.

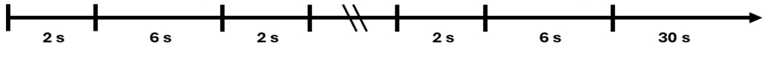

Figure 1: Complex span task evaluation procedure. The individual is asked to memorize a list of items presented sequentially. A processing task is inserted between the items. After the last item, the individual must recall the list of items in the order they were presented.

The experiment of this study aimed to evaluate an objective measurement method for cognitive cost that overcomes constraints of the existing methods of cognitive cost evaluation. In this study cognitive cost of aeronautical tasks on basic aeronautical tasks performed by novices early in their training was evaluated. As mentioned, the complex span task protocol involves alternating between memorization and processing. Here, processing tasks correspond to navigation task including two flight phases: a low difficulty phase (i.e., following a navigation leg) and a high difficulty phase (i.e., turning points with many actions to do). The cost is measured by asking participants to perform a letter memorization task in a flight simulator during these flight phases. The number of letters recalled in the correct order determines the participants' complex span in each experimental condition (i.e., the higher the number of letters, the better the complex span, the lower the cognitive cost), interpreted as an indicator of the current task's cognitive cost.

To validate the method an effect of cognitive load should be observed on the complex span (in this experiment, the number of letters recalled). A large complex span is then expected in the control group (group performing only the memorization task). A smaller complex span than the control group on average is expected for the experimental group: more precisely, the number of letters recalled in low cognitive load is expected to be higher than in high cognitive load.

METHOD

Participants

29 individuals participated split in two groups: 15 undergraduates constituted the control group, and 14 pilot trainees constituted the experimental group. The average age of participants was 23 years (SD = 2.5). All participants had normal or corrected vision.

Stimuli and Apparatus

Navigation Task

The experimental group performed the memorization task while carrying out a navigation task in a flight simulator used for pilot training (see Figure 2). Participants are seated in a reproduction cockpit of a Cirrus SR20 (side-by-side seats, screens with all aircraft parameters, stick and throttle to control aircraft attitude). Images of the flight environment were projected onto 3 large 2*2-meter white screens.

Memorization Task

In the memory task, participants verbally recalled items captured by the computer's microphone. Each item was listed by the experimenter, each 10 seconds.

Figure 2: Simulator display used for navigation task.

PROTOCOL

Navigation Task

The experimental group performed the memorization task while carrying out a navigation task in a flight simulator. Participants had to do a navigation task, which required them to maintain aircraft attitude (altitude, speed), follow flight routes (heading) and manage radio communications to ensure safety in the air environment. Two moments were identified as requiring low and high cognitive load, respectively: “‘transit phases’“, during which the individual follows a straight trajectory, and “‘turning point phases’“, in which several actions have to be performed (e.g. radio communications, taking information from the cockpit) (see Figure 3).

Navigation task consisted of 5 legs composed of transit phases intersected by turning point phases. As such, each participant of the experimental group did 5 low cognitive load phases and 5 high cognitive load phases. Five of the ten lists were presented during the low cognitive load phases and the other five were presented during the high cognitive load phases.

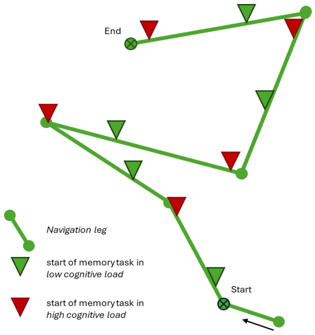

Figure 3: Diagram of the navigation followed by participants in the experimental group. After a take-off and stabilization phase, the experiment began at the black cross. All participants performed the same navigation. A letter sequence was presented at each green (low cognitive load) and red (high cognitive load) plot. The experiment ended after the landing phase.

Memorization Task

Participants heard a list of 10 consonants. Each consonant was listed once every 10 seconds. After hearing the 10 consonants, participants had 20 seconds to recall the list in the same order (Figure 4). Each group were exposed to 10 lists of 10 consonants in the memory task.

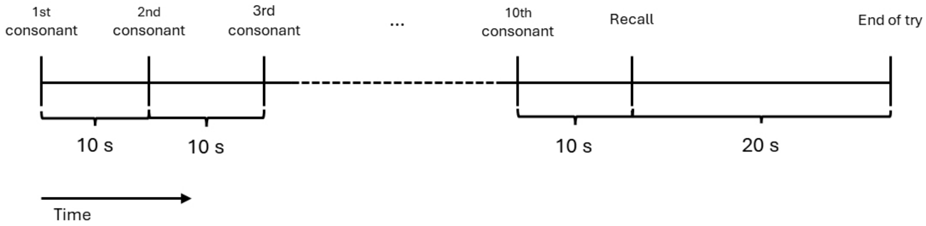

Figure 4: Sequence of memory task: 10 consonants and recall phase.

Participants in the control group performed only the memorization task. Each participant performed the task in an isolated room to avoid all sources of disturbance and carried out the task as described in Figure 4.

To compare the performance of the experimental group during the high and low phases with that of the control group, the lists projected during these phases are identified in the same way for the control group (i.e. 5 lists for the high phases and 5 lists for the low phases). This distinction was not made explicit to the participants and is only used in data processing (low vs. high).

RESULTS

Participants' responses were recorded using a microphone and corrected by the experimenters. The number of consonants recalled was counted. A large number of consonants recalled indicates a high span, conversely, a small number of consonants recalled indicates a low span. Thus, the lower the span, the greater the cognitive load. A large complex span is then expected in the control group (group performing only the memorization task). A smaller complex span than the control group on average is expected for the experimental group: more precisely, the number of letters recalled in low cognitive load is expected to be higher than in high cognitive load.

A repeated-measures ANOVA was performed on the participants' complex spans, with the inter-group factor Group with two modalities, Control and Experimental, and the intra-group factor Cognitive Load with two modalities, High and Low (see Figure 5). A significant effect of the Group variable is observed, $F(1, 27) = 42.088$, $p<.001$. A significant effect of the Cognitive Load variable was observed, $F(1, 27) = 29.9$, $p<.001$. A significant interaction effect is observed between the variables Group and Cognitive Load, $F(1, 27) = 16.299$, $p<.001$. Tukey's post-hoc tests were performed, with the following significant results: Experimental High ($M = 6$; $SD = 1.5$) vs Control High ($M = 9$; $SD = 0.7$), $p<.001$; Experimental Low ($M = 8$;

SD = 0.9) vs Control Low (M = 9; SD = 0.7), p<.01; Experimental High vs Experimental Low, p<.001; Experimental High vs Control Low, p<.001; Control High vs Experimental Low, p<.05.

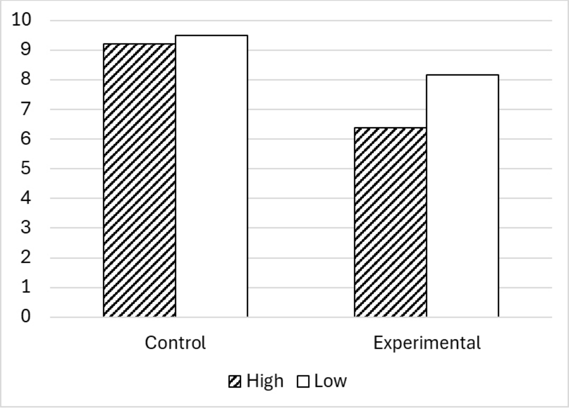

Figure 5: Graph showing complex spans for each group as a function of high (hatched) vs. low (white) cognitive cost.

No effect of the Cognitive Load variable is observed in the Control group, only in the Experimental group. A significant difference is therefore observed between the Control and Experimental groups (for both modalities of the Cognitive Load variable). In addition, the span of the Experimental group in the high cognitive load condition is smaller than the span of the Experimental group in the low cognitive load condition. These results are discussed in the following section.

DISCUSSION

The aim of this study was to evaluate an objective measurement method for cognitive cost, specifically of aeronautical tasks on basic aeronautical tasks performed by novices early in their training. Participants had to fly a plane in a flight simulator while retaining a list of letters. A control group only had to memorize the list of letters. Cognitive cost in each experimental condition corresponded to the complex span (i.e., number of letters correctly recalled).

First it was expected that the complex span would be higher for the control group than the experimental group. It was also expected that the complex span would be higher for the experimental group in low cognitive cost than for high cognitive cost. The results support both hypotheses. Therefore, the cognitive cost of the processing task varied the span. These results are consistent with other findings (Froger, 2021; Froger et al., 2018) suggesting the feasibility of diagnosing the cognitive cost of tasks by adapting a protocol commonly used in the laboratory to an environment mimicking an aeronautical context, and the sensitivity of the method to distinguish cognitive cost between relatively similar tasks. Indeed, even in the transit phase (low-cost) the cognitive cost is measured as higher than with the least-costing task : items to be memorized are better recalled if no other task is performed than if at least one other task is performed, even with a low cognitive cost task. The number of items recalled seems to be a good indicator of the cognitive cost of the task performed simultaneously. Taken together, these results thus indicate that the method can be tested in an ecological environment such as real flight.

This study was limited to assessing cognitive cost by measuring the complex span of a memory task in a single-task situation (control group) or in a double-task situation with varying cognitive costs (experimental group). The possible evolution of cognitive cost as a function of the individual's expertise on specific tasks was not measured in detail. Indeed, regular practice of an activity (set of tasks) calls on different processes to lead to expertise. For example, as novices, pilots in training are confronted with several discrete tasks that will be integrated into a single supertask when they become experts. According to Sweller (1994), this supertask cognitive structure is described by several names, including schema, chunk or template. Task performance begins with preparation, which can be described as the adoption of a task-specific cognitive configuration that manages the rules for solving the task (Rogers & Monsell, 1995). The process of adopting and maintaining a cognitive configuration relies on executive functions (Miyake et al., 2000; Monsell, 2015), and is modelled within the central executive processor (Baddeley, 2000). In a multitasking situation, the cognitive cost is mainly related to the impossibility of simultaneously activating the rules for solving two discrete tasks (Pashler, 1994; Liepelt et al., 2011). This cost is present until sufficiently repeated exposure to a multitasking situation benefits from the integration mechanism consisting in the creation of a supertask encompassing the solving rules of all tasks within a single rule. As long as this supertask has not been acquired, multitasking situations generate a higher cognitive cost. Acquiring expertise therefore involves training for specific tasks, which calls for the development of specific training programs that include assessment of the cognitive cost of specific tasks, enabling this cost to be reduced through sufficient practice, and finally transferring expertise to the operational environment. This study contributes to the effort to develop and validate a non-invasive, operational method for assessing cognitive cost. Such a method could be made available to flight instructors to monitor a pilot's progress more objectively in ecology (in-flight) training.

Furthermore, this study is part of a process of anticipating the development of increasingly complex on-board systems such as FCAS developed in response to Multi-Domain context. Indeed, being able to measure the cognitive cost of a task and its evolution would make it possible to assess the cognitive impact of each task on the individual (in this case the pilot) and propose remedial solutions. These solutions would make it possible to reduce the cognitive cost of a task and thus limit its cognitive impact on the individual in an operational environment. The aim is not to replace current training courses, but to supplement them. Another aim is to anticipate the cognitive impact of future Human-System Interactions within the Multi-Domain context.

DECLARATION OF INTEREST STATEMENT

The authors report there are no competing interests to declare.

ACKNOWLEDGMENT

Many thanks to the students who helped us carry out this study. Many tanks to all the participants, for their time and cooperation.

FUNDING

This work was supported by the Agence de l'Innovation de Défense under Grant [number 2023_65_0085].

REFERENCES

Baddeley, A. (1996). Exploring the central executive. The Quarterly Journal of Experimental Psychology Section A, 49(1), 528. https://doi.org/10.1080/713755608

Baddeley, A. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4(11), 417423. https://doi.org/10.1016/S1364-6613(00)01538-2

Barrouillet, P., & Camos, V. (2007). The time-based resource-sharing model of working memory. The Cognitive Neuroscience of Working Memory, 455, 5980. https://doi.org/10.1093/acprof: oso/9780198570394.003.0004

Camos, V. (2017). Domain-specific versus domain-general maintenance in working memory: Reconciliation within the time-based resource sharing model. Psychology of Learning and Motivation, 67, 135171. https://doi.org/10.1016/bs.plm.2017.03.005

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. Frontiers in neuroscience, 14, 268. https://doi.org/10.3389/fnins.2020.00268

Froger, G. (2021). Vers une gestion experte des tâches simultanées en aéronautique militaire (Doctoral dissertation, Aix-Marseille).

Froger, G., Blättler, C., Dubois, E., Camachon, C., & Bonnardel, N. (2018). Time-interval emphasis in an aeronautical dual-task context: A countermeasure to task absorption. Human Factors, 60(7), 936–946. https://doi.org/10.1177/0018720818783946

Hancock, P. A., & Matthews, G. (2019). Workload and Performance: Associations, Insensitivities, and Dissociations. Human Factors, 61(3), 374–392. https://doi.org/10.1177/0018720818809590

Kreitz, C., Furley, P., Memmert, D., & Simons, D. J. (2016). The influence of attention set, working memory capacity, and expectations on inattentional blindness. Perception, 45(4), 386–399. https://doi.org/10.1177/0301006615614465

Liepelt, R., Strobach, T., Frensch, P., & Schubert, T. (2011). Improved intertask coordination after extensive dual-task practice. The Quarterly Journal of Experimental Psychology, 64(7), 12511272. https://doi.org/10.1080/17470218.2010.543284

Longo L, Wickens CD, Hancock G and Hancock PA (2022). Human mental workload: A survey and a novel inclusive definition. Front. Psychol. 13:969140. https://doi.org/10.3389/fpsyg.2022.969140

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "'frontal lobe'" tasks: A latent variable analysis. Cognitive Psychology, 41(1), 49100. https://doi.org/10.1006/cogp.1999.0734

Monsell, S. (2015). Task-set control and task switching. In J. M. Fawcett, E. F. Risko, & A. Kingstone (Eds.), The handbook of attention (pp. 139172). Boston Review. https://doi.org/10.1093/med: psych/9780198528883.003.0002

Pashler, H. (1994). Dual-task interference in single tasks: Data and theory. Psychological Bulletin, 116(2), 220. https://doi.org/10.1037/0033-2909.116.2.220

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124(2), 207. https://doi.org/10.1037/0096-3445.124.2.207

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction, 4(4), 295312. https://doi.org/10.1016/0959-4752(94)90003-5

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 3(2), 159177. https://doi.org/10.1080/14639220210123806