From Laboratory to Cockpit: Evaluating the Predictive Value of Cognitive Tasks on Flight Simulator Performance

Sébastien Scannella[a] and Quentin Chenot[b]
Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, France
{*sebastien.scannella, quentin.chenot*}*@isae-supaero.fr*

Keywords: Neuroergonomics, Flight Simulator, Executive Functions, Complex Tasks.

Abstract: Comprehending the cognitive mechanisms underpinning success in demanding daily tasks is imperative for the human factors field. It not only necessitates a foundational grasp of the human brain but also furnishes invaluable insights for managing and averting human errors. This is particularly true in critical systems such as airplane, surgery, and more recently with autonomous vehicles. In order to provide additional information of the link between flying activity and high level cognitive functions, we investigated the relations between common executive function task performance, more complex and ecological task performance and flight simulator performance. Our results suggest that the unitary nature and lack of real-life legitimacy of common laboratory executive function tasks limit their ability to explain flight performance in the simulator —except for set shifting. Conversely, more ecological and dynamic tasks that engage executive functions tend to explain a larger variance in flying activity. Further research is planned to refine these predictive models and understand the underlying cognitive mechanisms.

1 INTRODUCTION

Understanding the cognitive processes that contribute to performance in complex, real-world tasks is important in the human factors' domain. Beyond the need for fundamental understanding of the human brain, it provides additional insight into handling and preventing human errors (Leiden et al., 2001; Koechlin, 2014). This extends to professionals operating in high-pressure environments, such as airplane pilots or surgeons—to name a few—where human lives are at risk. Of particular interest, executive functions (EFs) have been shown to be the most important functions to achieve efficient and adaptable behavior that is required in such tasks (Causse et al., 2011; Smit et al., 2021; Panganiban and Matthews, 2014). However, our understanding of executive functioning in real operating settings is constrained by the experimental settings.

On the one hand, complex cognitive processes have been investigated in numerous studies under controlled laboratory conditions for planning, inhibition, working memory, or switching (Cristofori et al., 2019; Logue and Gould, 2014; Sorel and Pennequin, 2008; St Clair-Thompson and Gathercole, 2006; Friedman and Miyake, 2017; Miyake et al., 2000). Although they have brought precious knowledge about how the brain can efficiently provide such high-level behaviors, they present at least two limits. First, they often lack ecological validity because they are designed to study an isolated process on a simplified computer task. Most of the time stimuli are presented in blocs of conditions, and\or one at a time, which is never encountered in such a way in real life. In addition, due to the isolation of one process over the others, there is a limitation in the dynamics and interactions among them, which consequently limits the complexity. On the other hand, real life complex task performance is also a significant research field (Smit et al., 2021). Aircraft pilots, for instance, have been involved in studies aiming at understanding auditory attentional processes (Dehais et al., 2019a) or more widely mental state (Gateau et al., 2018) in real aircrafts. These studies, however, are less numerous than laboratory ones, due to several difficulties. The operational setting is often expensive and difficult to access to (*e.g.,* airplanes, airfield, control tower, etc.). The measurement tools to assess cognitive processes are exposed to internal (participants' movements) and external (light, electromagnetic fields) interferences. Finally, what determines the quality of measurement

---

[a] https://orcid.org/0000-0001-9547-8303
[b] https://orcid.org/0000-0001-6439-9952

Scannella, S. and Chenot, Q.
From Laboratory to Cockpit: Evaluating the Predictive Value of Cognitive Tasks on Flight Simulator Performance.
DOI: 10.5220/0012918500004562

In *Proceedings of the 2nd International Conference on Cognitive Aircraft Systems (ICCAS 2024)*, pages 13-20
ISBN: 978-989-758-724-5

in real-life situations is also its biggest flaw, namely, the lack of reproducibility and the presence of numerous confounding factors.

To get the best of the two worlds, researchers have developed complex tasks in simulators or with pseudo-ecological environments (Causse et al., 2011; Dehais et al., 2019b; Kennedy et al., 2010; Scannella et al., 2018; Yesavage et al., 2011). The Multi-Attribute Task Battery (MATB-II) for instance, is a computer-based task designed to evaluate operator performance and workload (Santiago-Espada et al., 2011). The main interest of this battery lays in its multitasking aspects. Unlike unitary process evaluation, MATB requires the simultaneous performance of monitoring, dynamic resource management, and tracking tasks involving working memory, inhibition, and switching. However, it can be argued that the subtasks of the MATB are interdependent solely in relation to time-sharing. Allocating cognitive resources to one subtask diminishes the resources available for the others; a bad performance on one of them, however, do not directly affect the state of the others. In an attempt to design a pseudo-ecological task to study complex skill acquisition, Mané and Donchin (Mané and Donchin, 1989) have created the Space Fortress (SF) video game. In this 2D game, the participant must control a ship and shoot at a central fortress to destroy it. Simultaneously, the participant must earn points and missiles by accomplishing several subtasks, like identifying a predefined sequence of special characters or mines types. According to the authors, SF has been designed to rely on skills such as memory, attention, dual-tasking ability, and psychomotor control and speed, although these relationships have not been rigorously tested yet. Latter, this video game has shown successful transfer learning to aircraft pilot trainees' performance (Gopher et al., 1994).

Regarding flight performance *per se*, in the review of Smit et al., (Smit et al., 2021), the authors have shown that multiple EFs together with other cognitive abilities are associated with most of the measures of flying, navigating, and communicating, and furthermore can predict flight performance. This suggests a general involvement of cognition (including EFs) in flight management. However, these results do not allow for the identification of critical EFs for handling a plane, except for the working memory (Durantin et al., 2016; Causse et al., 2011).

The present study sought to bridge the gap between laboratory cognitive assessments and real-world task performance by examining the predictive value of a range of cognitive tasks on flight simulator performance. Laboratory executive function tests and pseudo-ecological tasks were performed by pilots before assessing their performance on different flight scenarios in a flight simulator, including standard and complex airfield patterns (system failures, bad weather, etc.). Because of the complexity of the flight scenario, we hypothesized that pseudo-ecological tasks (SF and MATB) would be higher correlates of simulator performance than the EF laboratory tasks.

2 METHODS

2.1 Participants

Thirty Private Pilots (4 women, mean age = 22y, mean flight hours = 59.4) holding a Private Pilot License (PPL) or in the process of obtaining it, were included. The study has been approved by the local ethic committee of EUROMOV, Montpellier University, IRB-EM: 2203C, in 2022.

2.2 Experimental Protocol

Participants were involved in a three-session protocol across three days. In the first session they practiced the SF and MATB tasks. In the second session they underwent the executive function battery. In the last session they flew the four flight scenarios in the simulator.

2.3 Executive Function Tasks

Participants first underwent a battery of nine executive function tasks that has been created according to the literature (Friedman and Miyake, 2017). It included three inhibition tasks, three updating tasks, and three switching tasks. Note that the task codes come from the millisecond test library (https://www.millisecond.com/download/library/) and were administrated through the *Inquisit* software (V.6). The tasks have been translated in French for the purposes of this experiment and the modified code files are available on https://osf.io/fm58p/.

2.3.1 Inhibition

Antisaccade. During this task, the participant must focus on a fixation cross in the center of the screen. A yellow square flashes (*i.e.*, a visual cue) on either the right or the left side of the cross. After the flash, an arrow appears on the opposite side of the flash, pointing either left, right or up. The participant must respond to which direction the arrow is pointing through the

arrow keys on the keyboard. Note that we did not measure saccades with an eye-tracker. The dependent variable on this task is the proportion of correct responses.
Stop Signal. In this task, the participant must focus on a fixation cross in the center of the screen. Then, an arrow appears, pointing either left or right. The participant has to press the corresponding arrow key. However, in some trials, the participant hears an auditory signal which indicates that she\he has to inhibit her\his response. This task is extracted from Verbruggen et al. (Verbruggen et al., 2019) and default parameters were used. The dependent variable on this task was the stop signal RT (response time).
Stroop. In this task, the participant saw colored stimuli (figures and words). Their goal was to indicate the color of the stimuli. This test is adapted for computers, and keyboard keys are associated with colors (Scarpina and Tagini, 2017). There were congruent (*e.g.*, the word RED written in red) and incongruent (*e.g.*, the word RED written in blue) trials. The dependent variable on this task was the RT difference between congruent and incongruent trials.

2.3.2 Updating

Keep Track. During this task, the participant must memorize and update words that are specific to certain categories (amongst 6 in total). The words are presented one by one in the center of the screen. The trials on this task include to keep-track of 3 to 4 words simultaneously and are randomized. The dependent variable was the proportion of correctly recalled words.
Letter Memory. In this task, the participant views a series of letters that appear one at a time at the center of the screen. Their goal is to memorize the four last letters. This task is modified from Friedman et al. (Friedman et al., 2008) in which participants have only to memorize the three last letters. This modification choice was done according to a pilot study that revealed a ceiling effect with only 3 letters to memorize. The dependent variable was the proportion of correctly recalled letters.
Dual N-back. In this task, the participant needs to follow a sequence of stimuli in two modalities at the same time (visual and auditory). N-value was set to two (*i.e.*, 2-back task). The participant must determine whether the position of the square (visual) was the same as the one observed two trials before in a 3 x 3 grid; and simultaneously determine whether the heard letter is the same as the one presented two trials before. Note that this is the only task that differs from Friedman et al. (Friedman et al., 2008) study in which they choose a spatial n-back. The reason is that the spatial n-back is not available on millisecond test library, and the dual n-back is the closest that we found in this library. The dual n-back task from the present study comes from Jaeggi et al. (Jaeggi et al., 2010) and we used identical parameters with only the 2-back. The dependent variable was the proportion of correct responses (yes and no).

2.3.3 Switching

Number Letter. During this task, the participant sees a 2 x 2 matrix on the computer screen. A pair of characters (ex: '7C') is presented and the participants have to respond either on the letter (consonant *vs.* vowel) or the digit (odd *vs.* even) depending on the position of the characters on the matrix that will randomly change. The dependent variable was the difference of RT between switching trials and non-switching trials.
Color Shape. In this task, red or green circles or triangles are presented to the participant. The goal is to respond to the type of stimuli depending on the cue (S for Shape *vs.* C for Color). The trials were randomized and the dependent variable on this task was the difference of RT between switching trials and non-switching trials.
Category Switch. During this task, the participants are asked to categorize a word in terms of (a) living criterion (living *vs.* non-living) or (b) size criterion (smaller *vs.* larger than a basketball). A cue determined which categorization needs to be performed, with a heart associated to the living criterion, and a cross associated to the size criterion. The trials were randomized and the dependent variable on this task was the difference of RT between switching trials and non-switching trials.

2.4 Pseudo-Ecological Tasks

2.4.1 MATB II

In the Multi-Attribute Task Battery II (MATB-II, see https://github.com/VrdrKv/MATB/blob/master), the participant must manage a computer-based system featuring tasks intended to simulate those performed during aircraft piloting. The main objective is to maximize performance on several tasks on the same screen. We used a four-task version (see Figure 1.b) that included: (1) managing visual alarms (*Monitoring*); (2) managing target tracking using a joystick (*Tracking*); (3) managing fuel pumps (*Fuel Management*); and (4) managing radio communications (*Communication*). In the *Monitoring* task, sliders are moving randomly between two extreme values along with two boolean generic alarms (red or green). The participant must press the corresponding button (from

F1 to F6) whenever one slider is stuck or when an alarm switches from green to red. In the *tracking* task, the participant must use the joystick to keep a moving reticle as close as possible to the center of the target. In the *Fuel Management* task, participants have to open or close 8 different pumps to optimize the total amount of fuel in two main tanks. Finally, in the *Communication* task, participant have to set virtual radio frequencies according to simulated air-traffic control tower auditory messages.

The dependent variable was the cognitive-motor performance of the best session as measured by a global mean z-score including the reaction time in the *Monitoring* task, the mean distance from the center in the *Tracking* task and the mean distance from the optimal level in the two main tanks of the *Fuel Management* task. Due to the low number of relevant *Communication* events, this task has not been taken into account in the scoring.

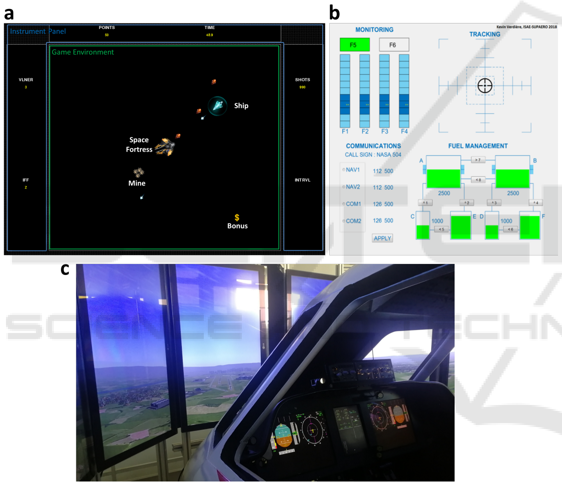

Figure 1: Complex tasks a. Space Fortress, b. MATB-II and c. ISAE-SUPAERO flight simulator.

2.4.2 Space Fortress

For the purpose of this study, a Python-based (ver. 2.7) version of SF was chosen from the github.com/CogWorks website (see Figure 1.a). The main goal of the game is to control a spaceship in space with no gravity (first task). The second goal is to destroy the fortress (second task). In addition, the player has to memorize three specific letters that will help him to identify and destroy different types of mines (type-1 and type-2) that regularly appear in the game (third task). Simultaneously, the player must keep focusing on sequences of symbols (*e.g.* # and $) that appear continuously throughout the game in order to capture bonuses (fourth task). To obtain the higher possible score, the participant has to perform the four sub-tasks in parallel. To this goal, participants were given the following instruction: ”You will maximize your score by performing all sub-tasks to the best of your abilities”. A detailed description of rules is available in our previous study (Chenot et al., 2022).

During the first session, participants practiced the game. This practice started with reading a document that described the task (environment and rules). After reading, participants were asked to play the SF game for approximately 12 minutes (4 × 3 min games). They practiced one task at a time following the variable-training methods used in previous studies (Boot et al., 2010). More precisely, participants were asked: to control the ship and only focus on destroying the fortress (game 1); to capture bonuses only (game 2); to destroy mines only (game 3); and finally to perform all tasks together (game 4). Note that participants were informed about the points' distribution prior to play. The experimenter made sure that participants understood and had experienced all the rules of the SF game before starting the evaluation. After practice, the participant performed the experimental session which constituted of ten SF games (3 minutes each). They performed five ”monotask” games (only flying the ship and destroy the fortress) and five ”multitask” games (all tasks together) alternating between one another.

The dependent variable was the z-score of the best game total score, corresponding to the sum of points including all sub-tasks.

2.5 Flight Simulator

In the flight simulator session, pilots were installed on the left seat of the ISAE-SUPAERO flight simulator *Pegase* which simulates an Airbus A320 (see Figure 1.c). Four flight scenarios were created according to the flight rules (Visual Flight Rules; VFR *vs.* Instrument Flight Rules; IFR) and the difficulty (Low *vs.* High). The four scenarios consisted in flight traffic patterns in the vicinity of Toulouse-Blagnac airport and the difficulty was manipulated according to visibility, landing type and failure (see Table 1). Flight Simulator performance was evaluated with a composite score taking into account compliance with flight parameters, as well as approach and landing quality. Each z-scored metric was assigned a specific weight in the final composite score to reflect its relative importance in flight operations. See Table 2 for a detailed description of the global flight performance calculation.

Table 1: Flight Simulator Scenarios.

Type	Diff	Visibility	Land	Failure
VFR	Low	Clear	Visual	None
	High	Cloudy	Visual	Alt+Speed
IFR	Low	Clear	ILS	None
	High	Fog+Night	ILS	Engine

VFR: Visual Flight Rules. IFR: Instruments Flight Rules. ILS: Instrument Landing System. Alt: Altitude.

Table 2: Flight Simulator Performance Scoring.

Parameters weights	Metric	Flight Phase
Altitude: 15%	Area under curve	Crosswind
		Downwing
		Base Leg
Speed: 15%	Area under curve	Crosswind
		Downwing
		Base Leg
Time: 15%	Time Difference	Crosswind
		Downwing
		Base Leg
Landing: 40%	Max G's	Landing
	Euclidian Distance	Landing
Approach: 15%	Flight Path Deviation	Approach

2.5.1 Statistical Analyses

In order to test the hypothesis that complex task performance (SF and MATB-II) would be a better correlate of the flight performance compared to executive function scores, correlation analyses have been carried-out using the global executive function score (mean z-score of the nine task performance), SF score (z-score in the best game), MATB-II score (mean z-score of three subtasks in the best game) and flight simulator performance score (mean weighted z-score of altitude, speed, heading, approach and landing metrics) of the four scenarios. Two additional correlation coefficient analyses have been done in order to statistically validate the best flight simulator performance predictive variable. All scores have been z-scored and analyses have been carried out using R studio (V4.2.2). The main hypotheses were tested by the following correlations:

- Executive Function tasks Flight simulator performance (EF and flight simulator composite score).
- Pseudo-ecological tasks *versus* Flight simulator (a. SF and flight simulator composite score; b. MATB and flight simulator composite score).
- Correlation coefficient comparisons between these three correlations.

Additional correlations analyses were also performed:

- Executive functions *versus* pseudo ecological tasks (a. EF and MATB; b. EF and SF);
- Between the two pseudo ecological tasks (EF *versus* SF).

Finally, a complementary analysis between individual EF tasks and simulator performance (Inhibition, Updating and shifting) was done.

3 RESULTS

Main hypothesis results are summarized in Figure 2.a, b and c. Concerning correlations with the flight simulator performance, SF was the higher predictive variable, albeit not reaching statistical significance ($r = 0.35$; $p = 0.06$), while executive functions and MATB scores did not significantly correlate with simulator score ($r = 0.2$; $p = 0.29$ and $r = -0.035$; $p = 0.86$, respectively). A first correlation coefficient comparison (R *cocor* package) revealed that SF was a better predictor of the flight simulator performance than the MATB (one-tailed $z = 2.01$, $p < 0.05$). A second correlation coefficient however did not reveal a better predictive value of SF compared to EF (one-tailed $z = 0.75$, $p = 0.23$).

As shown in Figure 2.f, despite the difference between MATB and SF in correlating with the flight simulator performance, performance in these two tasks was highly correlated with each other ($r = 0.59$; $p < 0.001$). Similarly, our analysis revealed moderate correlation ($r = 0.3$; $p = 0.09$; in the best case) between EFs and complex tasks (SF and MATB). Looking at the relation between each three executive functions and the flight simulator performance (see Figure 2.g, h and i), we found a significant predictive value of the shifting performance ($r = 0.39$; $p < 0.05$) but not for the Inhibition or updating scores ($r = 0.11$; $p = 0.57$ and $r = 0.00$; $p = 0.99$ respectively).

4 DISCUSSION

Although based on a limited sample, this study underscores the complex relationship between laboratory cognitive tasks and "close-to-real-world" performance. It suggests that Space Fortress—the more ecological task—is the closest to the flying one, as attested by its relative predictive power for flight simulator performance. This task has been first developed

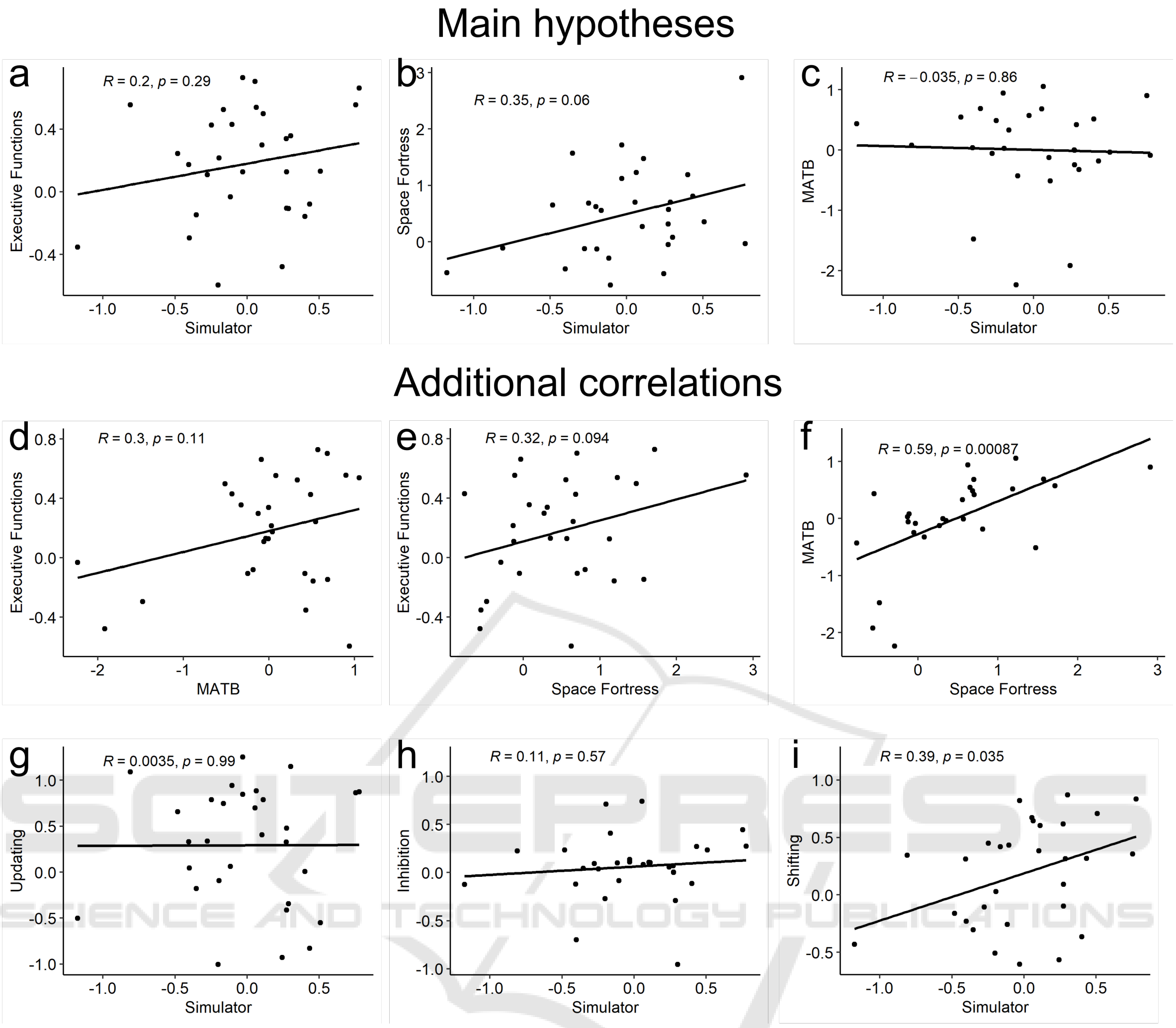

Figure 2: Correlation results for the main hypotheses (a, b and c) and additional comparisons (from d to i). All scores are z-scores. Solid lines stand for the linear fits.

by psychologists to study complex skill acquisition (Mané and Donchin, 1989) and latter showed transfer learning to real-life flying activity (Gopher et al., 1994). As a consequence, our hypothesis was that this task would be a better correlate to the flying activity in the flight simulator compared to more abstract EF tasks. Indeed, it involves visuo-motor coordination, working memory, inhibition and alternating between four different, dynamic, and interconnected subtasks. All these cognitive processes may be needed to handle a flight in complex situations.

A second interesting results is the unexpected non-predictable value of the MATB for the simulator performance (close to 0% of explained variance), despite its high correlation level with Space Fortress (around 35% of explained variance). These results suggest that these two complex tasks exhibit both overlaps and notable differences. Like SF, the MATB involves visual tracking, working memory and shifting from one subtask to another. A major difference, however, is that MATB subtasks are not interconnected; meaning that poor performance on one subtask does not affect other subtasks but only the global score. In SF, missing bonuses or ignoring mines type prevent from getting points and ammo to shoot at the Fortress. For instance, the fortress and mine subtasks cannot be achieved without a proper performance on the bonus subtask.

Similarly, the global score obtained in executive functioning in our sample of pilots poorly predicted the flight performance. This may appear contradictory with previous studies showing that reasoning and working memory, as attested by simplified computer-based tasks, can significantly predict flight simulator performance (Causse et al., 2011; Smit et al., 2021). Looking more closely at the subscores of the three executive functions (*i.e,* inhibition, updating, shifting) we found that only shifting could explain part of the simulator performance. One explanation could stand in the fact that working memory in flight management has been mostly shown to be used for radio communication (Morrow et al., 2003; Smit et al., 2021), which was absent in our simulator task.

4.1 Limits

It should be noted that this protocol is part of a larger study on cognitive training of airplane pilots. The present article focuses on the first flight simulator session only, and subsequent results from this larger study may be addressed in a future publication. The sample size of pilots is not large enough to prevent from false positive\negative correlation effects or to test more advanced prediction models (*e.g.,* multiple regression models). Finally, in this first approach study, neither the flight experience, nor the video game experience have been taking into account, although these could be covariates that explain a part of variance in complex task or flight simulator performance.

4.2 Conclusion

As a conclusion, our results suggest that the ability to handle the different subtasks in our most difficult flight simulator scenario seems to rely on being able to handle interconnected tasks by switching efficiently between them and taking into account their interdependence rather than just relying on working memory, inhibition or unrelated multitasking components.

Further research is planned to refine these predictive models with a higher statistical power and more exhaustive flight simulator performance assessment. It will allow for more sophisticated models taking into account pilot's demographics and performance in subtasks of the flying activity.

ACKNOWLEDGEMENTS

The authors wish to thank Stephane Perrey for his help in obtaining the local committee approval. This study is part of a research project funded by the French procurement agency (DGA).

REFERENCES

Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., Gratton, G., Voss, M. W., Prakash, R., Lee, H., et al. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta psychologica*, 135(3):349–357.

Causse, M., Dehais, F., and Pastor, J. (2011). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *The International Journal of Aviation Psychology*, 21(3):217–234.

Chenot, Q., Hamery, C., Lepron, E., Besson, P., De Boissezon, X., Perrey, S., and Scannella, S. (2022). Performance after training in a complex cognitive task is enhanced by high-definition transcranial random noise stimulation. *Scientific Reports*, 12(1):4618.

Cristofori, I., Cohen-Zimerman, S., and Grafman, J. (2019). Executive functions. *Handbook of clinical neurology*, 163:197–219.

Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R. N., and Lotte, F. (2019a). Monitoring pilot's mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions. *Sensors*, 19(6):1324.

Dehais, F., Roy, R. N., and Scannella, S. (2019b). Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural brain research*, 360:51–59.

Durantin, G., Scannella, S., Gateau, T., Delorme, A., and Dehais, F. (2016). Processing functional near infrared spectroscopy signal with a kalman filter to assess working memory during simulated flight. *Frontiers in human neuroscience*, 9:707.

Friedman, N. P. and Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86:186–204.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of experimental psychology: General*, 137(2):201.

Gateau, T., Ayaz, H., and Dehais, F. (2018). In silico vs. over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-bci. *Frontiers in human neuroscience*, 12:187.

Gopher, D., Well, M., and Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human factors*, 36(3):387–405.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., and Meier, B. (2010). The concurrent validity of the n-back task as a working memory measure. *Memory*, 18(4):394–412.

Kennedy, Q., Taylor, J. L., Reade, G., and Yesavage, J. A. (2010). Age and expertise effects in aviation decision making and flight control in a flight simulator. *Aviation, space, and environmental medicine*, 81(5):489–497.

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1655):20130474.

Leiden, K., Laughery, K. R., Keller, J., French, J., Warwick, W., and Wood, S. D. (2001). A review of human performance models for the prediction of human error. *Ann Arbor*, 1001:48105.

Logue, S. F. and Gould, T. J. (2014). The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition. *Pharmacology Biochemistry and Behavior*, 123:45–54.

Mané, A. and Donchin, E. (1989). The space fortress game. *Acta psychologica*, 71(1-3):17–22.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and 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):49–100.

Morrow, D. G., Menard, W. E., Ridolfo, H. E., Stine-Morrow, E. A., Teller, T., and Bryant, D. (2003). Expertise, cognitive ability, and age effects on pilot communication. *The International Journal of Aviation Psychology*, 13(4):345–371.

Panganiban, A. R. and Matthews, G. (2014). Executive functioning protects against stress in uav simulation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 994–998. SAGE Publications Sage CA: Los Angeles, CA.

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., and Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide. Technical report, NASA.

Scannella, S., Roy, R. N., Laouar, A., and Dehais, F. (2018). Auditory neglect in the cockpit: using erps to disentangle early from late processes in the inattentional deafness phenomenon. In *Neuroergonomics*, pages 207–208. Elsevier.

Scarpina, F. and Tagini, S. (2017). The stroop color and word test. *Frontiers in psychology*, 8:557.

Smit, D., Geppert, S. E., Daneshnia, N., Lafont, A., and Dehais, F. (2021). Executive functions and flying performance in pilots. In *Neuroergonomics Conference 2021*.

Sorel, O. and Pennequin, V. (2008). Aging of the planning process: The role of executive functioning. *Brain and cognition*, 66(2):196–201.

St Clair-Thompson, H. L. and Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly journal of experimental psychology*, 59(4):745–759.

Verbruggen, F., Aron, A. R., Band, G. P., Beste, C., Bissett, P. G., Brockett, A. T., Brown, J. W., Chamberlain, S. R., Chambers, C. D., Colonius, H., et al. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *elife*, 8:e46323.

Yesavage, J. A., Jo, B., Adamson, M. M., Kennedy, Q., Noda, A., Hernandez, B., Zeitzer, J. M., Friedman, L. F., Fairchild, K., Scanlon, B. K., et al. (2011). Initial cognitive performance predicts longitudinal aviator performance. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(4):444–453.