



HAL
open science

EEG characterization of dynamic complex processes and rare events to understand operator's activity in aeronautical context

Bertille Somon, Aurélie Campagne, Mick Salomone, Bruno Berberian

► To cite this version:

Bertille Somon, Aurélie Campagne, Mick Salomone, Bruno Berberian. EEG characterization of dynamic complex processes and rare events to understand operator's activity in aeronautical context. *Measuring Behavior* 2024, May 2024, Aberdeen, United Kingdom. hal-04724517

HAL Id: hal-04724517

<https://hal.science/hal-04724517v1>

Submitted on 7 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EEG characterization of dynamic complex processes and rare events to understand operator's activity in aeronautical context

B. Somon¹, A. Campagne², M. Salomone² and B. Berberian¹

1. ONERA/DTIS, 13661 Salon Cedex Air, France. {bertille.somon ; bruno.berberian}@onera.fr
2. Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France. {salomomi; aurelie.campagne}@univ-grenoble-alpes.fr

Introduction

In recent years, more and more researchers have questioned the relevance of lab-based controlled experiments to characterize operational everyday life activity. Several criticisms arise not only at the conceptual, but also at the methodological level. At the conceptual level, researchers have pointed out that the complexity, stakes and engagement level required in lab-based experiments is usually very reduced compared to what can be observed in real-life environments. Moreover, laboratory experimental paradigms mainly assessed in a controlled manner the effect of a specific factor on a specific cognitive function, but few of them evaluated the applicability of these results in more operational contexts, and considered the interactions and covariations between cognitive mechanisms which can arise in such contexts [1]. Researchers have tried to prevent this by developing novel controlled experimental paradigms such as the Multi-Attribute Task Battery (MATB-II; [2]) or adaptation of the Overcooked! game [3], which tend to mimic everyday-life activities and introduce increased levels of complexity. Notably, this complexity is based on the concomitant use of several tasks which challenge various cognitive functions. At the methodological level, experimental designs, especially in cognitive neuroscience, tend to assess cognitive functions and their variations at a discrete event-based level. Most studies require the repetition of tens of stimuli to: i) reduce the impact of the intra- and inter-individual variability; ii) ensure the statistical power and reproducibility of the results [4]; and iii) ensure the effect of a specific variation on a specific function [5]. Even though some research has focused on continuous [6] or block instead of event-based analyses of activity, these measures are currently often used to monitor very global operator states, such as fatigue, vigilance, or cognitive workload. Additionally, event-related potentials (ERP) estimation remains the most employed electroencephalographic (EEG) analysis method and requires the experimenters to introduce task-irrelevant events or specific stimuli in various types and modality, likely to bias the task in hand. This has been applied in a few brain-computer interface (BCI) or operational research, where cues [7] have been used.

This question of continuous or dynamic monitoring is particularly relevant in complex and/or risky operational situations. Recent advanced technologies and the advent of Artificial Intelligence (AI) have led to increase the automation level of systems. We now interact daily with automated systems in various activity sectors, particularly in the field of aeronautics and rail transports and, more recently in the field of the automotive with the growing accepted use of automated cars [8]. Although the benefits of such automated systems are undoubtable [9], several studies have also demonstrated that there are drawbacks to the change in operator activity driven by this automation [10]. Operators who previously performed specific actions now shifted to a supervisory role and therefore had to adapt their routines and skills, leading to new unforeseen difficulties in human-system interactions. These difficulties, such as a loss of motor skills, attention decrements, increased complacency and lack of system understanding, are coined the out-of-the-loop (OOL) performance problem [11]. Although this phenomenon has been widely studied, especially in the human factors for aeronautics community, it remains very difficult to characterize, and even more to compensate. Two difficulties arise: a theoretical and a technical one. At the theoretical level, the OOL has been described in terms of situation awareness or high-level constructs and concepts, providing low explainability regarding the whys and wherefores of this phenomenon. The folk-model-like description of the OOL prevents us from identifying properly the cognitive mechanisms implicated in its emergence and anticipate their variations [12]. Still, the OOL has been identified to emerge in highly complex environments in which the operator remains a supervisor of the automated system. In this context, when removed from the control loop, the operator is often unable to detect errors from the automated system and compensate them whenever necessary. Thus, we have proposed that the performance (or error) monitoring mechanism, widely studied by the cognitive neuroscience community, might be a good concept to characterize the OOL [12].

Additionally, we believe that using the methods of the cognitive neuroscience community, such as cerebral activity recording methods (e.g., EEG), can help us have a better understanding and characterization of these applied difficulties. Nevertheless, we have been faced with the technical difficulty related to the OOL characterization. Studies on the neurophysiological correlates of performance monitoring activities have focused mainly on event-related potentials time-locked to participants' responses [13]. However, the long time-scale of emergence of the OOL phenomenon requires more continuous measures to be performed. Additionally, the activity usually studied in performance monitoring is the Error-Related Negativity (ERN; [14]), which is very difficult to reconcile with the very reliable systems used in the aeronautics context. Finally, these activities are time-locked specifically to a participants' responses. In the context of the OOL, the system is highly automated and requires no response from the participant which limits the use of these measures. In this short paper, we describe the framework and the experimental protocols we used as well as the brain activity analysis approach we adopted to characterize the supervisory activity of an automated system and the out-of-the loop phenomenon, taking into account the difficulties associated with the study context and previously described. The limits of this approach and adaptations to improve characterization of the brain activity underlying the cognitive mechanisms involved are also explained.

Rationale

Concerning the performance monitoring framework, we have seen through a review of the literature that, even though performance monitoring associated to system supervision was not a very studied topic, there were more and more studies related to understanding the brain activity associated with another human agent's error detection [15]. These studies partly emerged from the Joint Action community and social neuroscience, in order to understand when, how and why we detect the errors committed by another human being [16]. One interesting aspect of these studies is the fact that researchers evidenced performance monitoring brain components based on EEG measures during supervision and error observation [13]. This knowledge led us to develop an experimental paradigm in order to compare the brain activities associated with human and artificial agent error monitoring. These activities not being fully characterized, we had to proceed in a step-by-step approach, from a lab-based to a more applied environment, which was our target context. In a first experiment, we used a well-known experimental paradigm in the cognitive neuroscience community – the Eriksen flanker task – that we adapted as an aeronautics-based supervision task [17]. In a second step, we tried to transfer the observed results and to adapt the EEG analysis to a more dynamic and ecological task, consisting of the supervision of an aeronautics-based conflict avoidance simulator [18]. Additionally, we also assessed through these studies, the effect of time-on-task on performance monitoring brain correlates and their evolution during the OOL phenomenon.

Experimental studies

In this section, we will present briefly the paradigms we have used to characterize brain activity related to supervision activity and its evolution across time during the OOL phenomenon.

Material and methods

In a first experiment (see [17] for full details), we recorded the EEG activity of 17 healthy participants (12 men; 27.5+/-4.78 y.o.) during another human agent or system supervision in a modified supervision arrowhead version of the Flanker task [19] [20]. During this task, participants had to supervise the decisions of another agent (human or automated system) on the orientation of a central target arrow on a screen among or without distracters. Participants performed a total of 40 blocks of 72 trials over 2 experimental sessions. Each trial (see Figure 1a) presented a stimulus for 10ms followed by a short reaction time window after which the other agent's response was displayed. Finally, the question "Error?" was displayed, and the participant had one second to tell whether the agent's response was correct or not. The participant supervised a human agent (introduced to him at the beginning of the task) for half of the blocks and an artificial agent for the other half. The type of agent performing the task was provided at the beginning of each block. In reality, all blocks were performed by a computer, which accuracy was 66.6% and reaction times were based on the participant's own reaction times to the same flanker task performed previously. This ploy was used in order to ensure the same number of error and correct trials, as

well as prevent effects of trust or differential learning across participants. Additionally, two levels of difficulty were introduced for the flanker task, as a follow up of a previous experiment [20], but will not be discussed here.

For the two experimental sessions, participants' subjective (difficulty levels), behavioral (error detection rate and d') as well as neurophysiological (75 Ag-AgCl active electrode EEG) activity were recorded. The EEG signal was amplified using an actiCHamp system (Brain Vision, LLC), digitized at a 24-bit rate and sampled at 1kHz, with a 0.05 μ V resolution. Classical signal pre-processing tools (i.e., band-pass filtering, artifact removal with ICA) were used in order to clean the signal (see [17] section 2.3.2. for a full description). In this first study, consistently with the literature on error supervision, event-related potentials were time-locked to the other agent's response display and analyzed at fronto-central sites [21]. In accordance with the literature, mean peak amplitude of the N2 and P3 ERPs (likened to the observational versions of the error-related negativity and positivity [13]) were computed and analyzed with a repeated-measures analysis of variance with accuracy – error vs. correct – difficulty – easy vs. difficult – and agent type – human vs. system – as within-subject factors. Going further, given a potential spatio-temporal dynamic in supervisory activity and inter-individual variances, a more robust cluster-based permutation test [22] (i.e., requiring no a priori on ERP location or latency) was performed.

In the second experiment (see [18] for full details), we recorded the EEG activity of 20 healthy participants (7 women; 27.75 \pm 1.42y.o.) performing a dynamic conflict avoidance simulator supervision task. Participants had to assess whether simulator choices to avoid obstacles by turning right or left was correct or erroneous depending on the context surrounding the aircraft. Participants completed a total of 12 experimental blocks of 25 trials over 2 sessions. Each trial (see Figure 1b) started with a first initialization phase lasting 2 to 5 seconds during which the aircraft flew straight, then conflicting obstacles arrived and were detected by the simulator, then the simulator's avoidance decision was displayed. At this moment the simulation was frozen, providing no new evidence for the participant to identify the correct or erroneous response. After 1 second the participant was asked to provide his response stating whether the simulator's response was correct or erroneous. Once the participant had given his response, the simulation was started again and the simulator performed the avoidance. Finally, feedback was displayed to the participant, indicating the correctness of the simulator's response. As described more precisely in [18], and as a follow up of the previous experiment, there were two difficulty levels corresponding to obstacles being aligned or dispersed throughout the entire screen (respectively easy and difficult condition). Across the entire sessions, the accuracy of the simulator was set at 60%, in order to ensure a sufficient number of trials to be analyzed for both accuracy conditions.

For the two sessions, participants' subjective (difficulty levels), behavioral (hit rate) as well as neurophysiological (64 Ag-AgCl active electrode EEG) activity were recorded. The signal was acquired and preprocessed with similar parameters as in the first experiment. (see [18] section 2.3.3. for a full description). In this second study, a first classical analysis on ERPs was performed in order to see the transferability of previous results to more ecological and dynamic tasks. Given the low transferability of previously observed results to this dynamic task, we completed this classical analysis with time-frequency analyses allowing to assess more dynamically brain activity variations. Both analyses were time-locked to the display of the system's response. To avoid a priori assumptions about the supervisory brain activity in the studied dynamic environment, we performed a cluster-based permutation test to identify the specific spatio-temporal clusters differentiating the various experimental conditions. EEG analyses were averaged across participants depending to the system accuracy – error vs. correct – task difficulty – easy vs. difficult – but also to the moment of the experiment – beginning vs. end – as within-subject factors. For all analyses, only the participant correct evaluations of the simulator's responses were analyzed. The impact of time-on-task (moment of the experiment) on the performance monitoring activity was assessed by averaging activity of the first two blocks (beginning) and the last two blocks (end) of each session.

Results

In the first experiment (see [17] for a full description of results), despite an effect of the difficulty level on both the subjective feedback ($diff_{easy} = 3.35 \pm 0.46$, $diff_{difficult} = 4.60 \pm 0.58$; $t[14] = 3.073$, $p < .01$) and the stimuli detectability ($d'_{easy} = 2.42$, $d'_{difficult} = 1.57$; $F(1,16) = 17.24$, $p < .001$, $\eta^2_p = .52$), this effect was not observed on the error detection rate ($F(1,16) = 0.48$, $p = .5$). Interestingly there was no difference either of the type of agent (human vs. System) on neither the error detection rate, nor the d' measures.

At the physiological level (see Figure 2a), we identified a significant N2-P3 component at the FCz electrode associated with error detection. The P3 was present only for error detection (main effect of accuracy on peak P3 amplitude; $F(1,16) = 69.42, p < .005, \eta^2_p = .81$). Additionally, we observed a significant effect of difficulty on both the N2 ($F(1,16) = 4.60, p < .05, \eta^2_p = .22$) and P3 ($F(1,16) = 15.77, p < .005, \eta^2_p = .50$) components. This classical ERP analysis revealed no significant effect of the type of agent supervised on these performance monitoring components. Yet, another more robust analysis (cluster-based permutation test) revealed a significant difference of activity with a broad positive component, differentiating significantly between human agent and system error detection. Interestingly, this component had the same characteristics as the P3 in terms of shape and latency.

In the second experiment (see [18] for a full description of results), similarly to the first experiment, participant subjectively reported a significant difference between the easy and difficult conditions ($diff_{easy} = 1.07 \pm 0.13, diff_{difficult} = 2.12 \pm 0.15; t[17] = -5.0675, p < .005$) but there was no effect on the objective performances ($t[17] = -0.35, p = .73$) with the average hit rate equal to $97.61 \pm 0.57\%$. Additionally, the moment of the experiment (beginning vs. end) showed no effect on behavioral performances.

At the physiological level (see Figure 2b), the classical ERP analysis revealed no cleared component associated with the observation of system's correct or erroneous responses. Cluster-based permutation tests identified a small negative posterior cluster differentiating between error-related and correct-related brain activity. However, the shape, location and latency of this cluster were difficult to associate to usual performance monitoring ERP. In contrast, the use of a time-frequency decomposition (see Figure 2c), better suited to the temporal variances of brain activity, allowed us to identify a brain cluster in the theta frequency band whose activity was significantly increased in response to system errors compared to correct system responses. Additionally, theta activity has been modulated by the moment of the experiment within a cluster, with a higher left-lateralized fronto-temporo-parietal theta activity at the beginning compared to the end of the experiment, regardless of system response and also following system errors. Finally, several clusters associated with low alpha activity also showed increased amplitude at the beginning of the experiment, and decreased amplitude for correct responses compared to system errors at the beginning.

Discussion and way forwards

The goal of these works was to characterize the brain activity variations related to system performance monitoring activity and their evolution during the OOL phenomenon. The use of brain measurements appears of major interest to better understand and characterize supervision activity with regard to subjective and behavioral data, likely to be less sensitive to modulation factors such as task difficulty, as illustrated by our two studies and the literature, even in more applied contexts [23] [24]. In both experiments, task difficulty modulated subjective feedback as well as physiological activity related to performance monitoring (for both ERPs and time-frequency measures) but did not show any effect on supervision performances. Likewise, in the second experiment, the moment of the experiment (beginning vs. end) showed a significant effect on brain markers related to performance monitoring activity and vigilance state (respectively in the theta and alpha frequency bands) but did not show any impact on performances.

However, although current EEG measures – especially ERPs – constitute good biomarkers to characterize performance monitoring activity in lab-based environments, as also illustrated in other prior studies [24], their transferability to more operational and dynamic contexts, remains difficult. In this context (second study), we were unable to reproduce the ERP results related to the supervision activity of system's erroneous and correct responses highlighted in the first study. Greater temporal variability in the timing of system error detection and greater uncertainty in identifying correct and erroneous responses in this dynamic and more complex environment to process may help to justify this result, given the sensitivity of ERP markers related to supervisory activity with task difficulty. Nevertheless, in the second study, we observed specific error-related variations in the frequency domain, as well as their evolution across time. Based on this result, part of the difficulty we encountered in translating the ERP results in the second study could come from the fact that operational environments would involve, compared to more controlled laboratory environments (as in the first study), several cognitive functions with variable temporality depending on the dynamic evolution of the stimulation environment, which would make

the analysis of the neurophysiological correlates associated with each function more difficult [5]. If we wish to better understand the mechanisms at play during the supervision activity and the precursors of its degradation during the OOL phenomenon in more operational and dynamic environments, it appears necessary i) to apply and develop new methods for processing and analyzing EEG signals and ii) to associate other metrics (e.g. data fusion) to better identify and individualize the cognitive functions involved and the associated neural correlates.

Concerning the EEG signal analysis methods, we have seen through our two studies that it is possible to extract new information with new signal processing or analysis methods. During the first experiment the use of a statistical analysis method without *a priori* spatio-temporal of the EEG activity, namely the cluster-based permutation test, allowed us to observe a difference in amplitude between human and system error detection, on what we can link to the P3 component. Given the functional role of the P3 regarding information processing, this result would suggest the establishment of some form of complacency during system supervision. This complacency is a phenomenon that usually occurs during system supervision and would be considered a precursor to the OOL phenomenon [12]. Additionally, the time-frequency processing method also provided more information during the more dynamic experiment. Indeed, the complacency result observed in the first experiment echoes with the time frequency results observed during the second experiment, where we observed an increased theta activity at the beginning compared to the end of the experiment, especially for error detection. Theta activity has generally been associated to cognitive control and performance monitoring [25]. However, this analysis method does have its limitations. As mentioned earlier, in an ecological context, there is more variability in the cognitive processes involved and their temporality. This can make their detection difficult to the benefit of more massive activities representing global operator states (such as vigilance, fatigue, cognitive workload). In this study, the clusters we observed could be composed of several dozen electrodes. These clusters, which spread over several brain regions, could therefore reflect more general states, making it more difficult to characterize complex cognitive processes. Complementary solutions could improve our understanding of the data EEG and performed analyses.

The use of methods such as blind source separation methods or feature identification algorithms (such as Independent/Principal/Canonical Component Analysis – ICA, PCA, CCA – for instance) could also be an interesting avenue to disentangle various sources of activity and therefore might allow to better decipher the different cognitive mechanisms related to the supervision activity and the precursor of the OOL phenomenon. The use of methods allowing to improve the spatial resolution of the surface EEG data such as a surface Laplacian transformation [24], estimating current source density, could also contribute to this objective. The application of this method on the data from a previous study allowed to better distinguish the influence of task difficulty on the different response-locked and feedback-locked performance monitoring ERPs (see [20] for full details). Nevertheless, all these methods do not completely eliminate the difficulty of properly labelling the cognitive process associated with a given brain activity. Thus, these methods could be combined with source reconstruction algorithms (e.g., beamforming), which have the advantage of associating cerebral activity with a specific brain region or structure. In addition to reducing the size of clusters, these analyses can be used to refine the identification of brain activity and thus the characterization of the cognitive processes involved. Finally, we have seen in the second experiment that the data was much noisier in a dynamics experimental context compared to a controlled one. One possibility could be to use more recent data cleaning algorithms, such as the artificial subspace reconstruction (ASR) or neural networks that have proven very useful recently [26].

Concerning the data fusion, one of the objectives is to use additional metrics allowing in particular to better characterize the cognitive processes involved during the supervision activity to better dissociate them on the EEG patterns. Monitoring ocular activity constitutes a candidate of interest in this perspective. In the literature, ocular activity is identified as a window into vigilance, attentional processes as well as complacency [27]. Throughout visual scene exploration, ocular activity characterizes attentional spreading as well as preferential treatment or neglect of specific elements [28]. Ocular activity might thus be a good candidate to identify and characterize the dynamics of exploration and information processing of a visual scene, leading to decision making. This tool would therefore provide additional information for understanding the cognitive operations of operators during activity. This contribution would in turn nourish the understanding and characterization of cerebral markers obtained with EEG during dynamic tasks. The EEG-eye tracking coupling also presents a major methodological interest in the

processing of EEG signals, in particular in resolving the problem of spatio-temporal overlap of neuronal activities during the close succession of eye fixations during the visual exploration of information during a dynamic and complex task. Thus, we are currently setting up another experiment aiming at using joint eye-tracking-EEG activity to time-lock performance monitoring and have a better characterization of it related brain components in a dynamic conflict avoidance task (see the [ASTRID ANR Project EMOOL](#)). Additionally, recent studies have proposed that the fixation rate could be a good marker for detecting the OOL phenomenon [29]. Notably, variations of trust and complacency tend to modify the ocular activity of operators in highly automated environments [27].

Conclusion

This set of studies aimed to question the neurocognitive mechanisms underlying performance supervision activity and the nature of the precursors at the origin of the out-of-the-loop phenomenon, a widely used concepts which lacks neurocognitive foundations to be characterized more thoroughly, and thus anticipated and countered. We have seen that with this type of concept rising from interactions with highly automated and very reliable systems, it is difficult to define experimental protocols allowing the use of neurophysiological methods classically used in the literature. After showing difficulties in reproducing lab-based measures in more operational contexts, we have faced the difficulty of the number of trials, the dynamic of the task and the time-locking events. We propose promising avenues to address, step by step, these difficulties (signal processing tools, innovative, analyses, physiological data fusion, etc.) and aim for our final goal.

Ethical Statement

Studies presented in this paper were approved by a local ethics committee (CERNI Grenoble, n°IRB00010290-2016-09-13-12 and n°IRB00010290-2017-07-04-20-CERNI_AvisConsultatif-2017-06-13-04) and conducted according to the principles expressed in the revised Declaration of Helsinki.

References

1. P.A. Hancock, "Whither Workload? Mapping a Path for Its Future Development," in *Human Mental Workload: Models and Applications*, Dublin, Ireland, 2017, pp. 3-17.
2. Yamira Santiago-Espada, Robert R. Myer, Kara A. Latorella, and James R., Jr. Comstock, "The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide.," NASA Langley Research Center, Technical Memorandum 2011.
3. M. Carroll et al., "On the Utility of Learning about Humans for Human-AI Coordination," in *Advances in Neural Information Processing Systems 32*, 2019, pp. 5174-5185.
4. Guiomar Niso et al., "Good scientific practice in EEG and MEG research: Progress and perspectives.," *NeuroImage*, vol. 257, no. 119056, August 2022.
5. Steven J. Luck, *An introduction to the event-related potential technique.*. Cambridge, Massachusetts; London, England: MIT press, 2014.
6. Cameron D. Hassall, Yan Yan, and Laurence T. Hunt, "The neural correlates of continuous feedback processing," *Psychophysiology*, vol. 60, no. 12, p. e14399, 2023.
7. Raphaëlle Nina Roy, Stéphane Bonnet, Sylvie Charbonnier, and Aurélie Campagne, "Efficient workload classification based on ignored auditory probes: a proof of concept.," *Frontiers in Human Neuroscience*, vol. 10, no. 00519, October 2016.
8. Carl Benedikt Frey and Michael A. Osborne, "The future of employment: How susceptible are jobs to computerisation?," *Technological Forecasting and Social Change*, vol. 114, pp. 254-280, January 2017.
9. Earl L. Wiener and Renwick E. Curry, "Flight-deck automation: Promises and problems.," *Ergonomics*, vol. 23, no. 10, pp. 995-1011, 1980.

10. S. W. A. Dekker and D. D. Woods, "MABA-MABA or abracadabra? Progress on human–automation co-ordination.," *Cognition, Technology & Work*, vol. 4, pp. 240-244, 2002.
11. Mica R. Endsley and Esin O. Kiris, "The Out-of-the-Loop Performance Problem and Level of Control in Automation," *Human Factors*, vol. 37, no. 2, pp. 381-394, June 1995.
12. B. Berberian, B. Somon, A. Sahaï, and J. Gouraud, "The out-of-the-loop Brain: A neuroergonomic approach of the human automation interaction.," *Annual Reviews in Control*, vol. 44, pp. 303-315, 2017.
13. Markus Ullsperger, Adrian G. Fischer, Roland Nigbur, and Tanja Endrass, "Neural mechanisms and temporal dynamics of performance monitoring.," *Trends in Cognitive Sciences*, vol. 18, no. 5, pp. 259-267, March 2014.
14. Anja Riesel, Anna Weinberg, Tanja Endrass, Alexandria Meyer, and Greg Hajcak, "The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks.," *Biological Psychology*, vol. 93, no. 3, pp. 377-385, July 2013.
15. Bertille Somon, Aurélie Campagne, Arnaud Delorme, and Bruno Berberian, "Performance Monitoring Applied to System Supervision," *Frontiers in Human Neuroscience*, vol. 11, no. 360, July 2017.
16. Margherita Adelaide Musco, Elisa Zazzera, Eraldo Paulesu, and Lucia Maria Sacheli, "Error observation as a window on performance monitoring in social contexts? A systematic review," *Neuroscience and Biobehavioral Reviews*, vol. 105077, February 2023.
17. Bertille Somon, Aurélie Campagne, Arnaud Delorme, and Bruno Berberian, "Human or not human? Performance monitoring ERPs during human agent and machine supervision.," *NeuroImage*, vol. 186, pp. 266-277, 2019.
18. Bertille Somon, Aurélie Campagne, Arnaud Delorme, and Bruno Berberian, "Brain mechanisms of automated conflict avoidance simulator supervision.," *Psychophysiology*, vol. 60, no. 2, p. e14171, 2023.
19. Barbara A. Eriksen and Charles W. Eriksen, "Effects of noise letters upon the identification of a target letter in a nonsearch task," *Perception & Psychophysics*, vol. 16, no. 1, pp. 143-149, 1974.
20. Bertille Somon, Aurélie Campagne, Arnaud Delorme, and Bruno Berberian, "Evaluation of performance monitoring ERPs through difficulty manipulation in a response-feedback paradigm.," *Brain Research*, vol. 1704, pp. 196-206, 2019.
21. Stefanie Enriquez-Geppert, Carsten Konrad, Christo Pantev, and René J. Huster, "Conflict and inhibition differentially affect the N200/P300 complex in a combined go/nogo and stop-signal task," *NeuroImage*, vol. 51, no. 2, pp. 877-87, June 2010.
22. Eric Maris and Robert Oostenveld, "Nonparametric statistical testing of EEG- and MEG-data.," *Journal of neuroscience methods*, vol. 164, no. 1, pp. 177-190, 2007.
23. Luca Longo and M. Chiara Leva, Eds., *Human Mental Workload: Models and Applications*. Dublin, Ireland: Springer, 2017.
24. Liesbet Van der Borght, Femke Houtman, Boris Burle, and Wim Notebaert, "Distinguishing the influence of task difficulty on error-related ERPs using surface Laplacian transformation.," *Biological Psychology*, vol. 115, pp. 78-85, March 2016.
25. J. F. Cavanagh and M. J. Frank, "Frontal theta as a mechanism for cognitive control.," *Trends in Cognitive Sciences*, vol. 18, no. 8, pp. 414-421, August 2014.
26. S. Blum, N. S. J. Jacobsen, M. G. Bleichner, and S. Debener, "A Riemannian Modification of Artifact Subspace Reconstruction for EEG Artifact Handling.," *Frontiers in Human Neuroscience*, vol. 13, no. 00141, April 2019.
27. Gal Ziv, "Gaze Behavior and Visual Attention: A Review of Eye Tracking Studies in Aviation.," *The International Journal of Aviation Psychology*, vol. 26, no. 3-4, pp. 75-104, July 2016.

28. Pieter J.A. Unema, Sebastian Pannasch, Markus Joos, and Boris M. Velichkovsky, "Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration.," *Visual Cognition*, vol. 12, no. 3, pp. 473-494, 2005.
29. G. Di Flumeri et al., "Brain-Computer Interface-Based Adaptive Automation to Prevent Out-Of-The-Loop Phenomenon in Air Traffic Controllers Dealing With Highly Automated Systems.," *Frontiers in Human Neuroscience*, vol. 13, no. 00296, September 2019.