Auditory Alarm Misperception in the Cockpit: An EEG Study of Inattentional Deafness

Frédéric Dehais[(⊠)], Raphaëlle N. Roy, Thibault Gateau,
and Sébastien Scannella

Human Factors and Neuroergonomics Department, ISAE, 10 av E. Belin,
31055 Toulouse Cedex 4, France
{frederic.dehais,raphaelle.roy,thibault.gateau,
sebastien.scannella}@isae.fr

Abstract. Missing auditory alarms is a critical safety issue in many domains such as aviation. To investigate this phenomenon, we designed a scenario involving three flying scenarios corresponding to three different level of difficulty along with an oddball paradigm in a motion flight simulator. This preliminary study was conducted with one pilot equipped with a 32-channel EEG. The results shown that manipulating the three levels of task difficulty led respectively to rates of 0, 37, and 54 % missed alarms. The EEG analyses revealed that this decrease in performance was associated with lower spectral power within the alpha band and reduced N100 component amplitude. This latter finding suggested the involvement of inattentional deafness mechanisms at an early stage of the auditory processing. Eventually, we implemented a processing chain to enhance the discriminability of ERPs for mental state monitoring purposes. The results indicated that this chain could be used in a quite ecological setting (i.e. three-axis motion flight simulator) as attested by the good results obtained for the oddball task, but also for more subtle mental states such as mental demand and stress level and the detection of target, that is to say the inattentional deafness phenomenon.

Keywords: Inattentional deafness · Stress · ERP · EEG · Alarm perception · Aviation · Passive brain computer interface

1 Introduction

Auditory alarms are used to alert the human operators of impeding off-nominal situations. These auditory warnings present several advantages as long as their detection doesn't produce head movement. For instance, they lead to faster responses than their visual counterpart. However, many studies and safety analyses reported the absence of response to these auditory alarms in several critical domains [3,9,20]. This inability to detect alarms may find several explanations. For instance, poor warning systems design are known to trigger spurious alarms thus leading to the so-called "Cry Wolf effect": the human operator consciously neglects such warnings especially under high workload settings [27].

Moreover, the design of the auditory signal itself may fail to capture attention due to its lack of salience or if it is too loud and distracting [21]. Another explanation is to consider the inattentional deafness hypothesis that states that unattended auditory stimuli may fail to reach awareness under high visual load conditions [19]. Since flying is a demanding task [4] that largely involves the processing of visual cues [5], the occurrence of unexpected auditory alarms could remain unnoticed [6,7].

This hypothesis is supported by a recent experiment that involves a simplified landing decision task based on the analysis of visual indicators under different load conditions while continuous electroencephalography (EEG) measurements were performed. During the task, a tone was presented, either a standard one, which participants were told to ignore, or a deviant one ("the alarm") which participants were told to overtly report. The analysis of the event related potentials (ERPs) showed that a drastic diminution of the late auditory component (P300) amplitude was concomitant with the occurrence of inattentional deafness [12]. However, the underlying concepts of scarce attentional resources may not be sufficient to fully account for the inattentional deafness phenomenon [16]. A complementary approach is to consider the existence of the visual dominance over hearing as demonstrated in several paradigms [26]. Reinforcement learning since childhood has probably lead the human beings to rely on visual information rather than on auditory ones. It is admitted that 80 % of the data to handle the flight safely are visual ones, thus leading the pilot to trust visual cues. Indeed, when visual information conflict with auditory ones, or when task demand is high, automatic gating mechanisms may take place and inhibit the processing of auditory information [17,18]. In order to verify this hypothesis, we recorded in a previous study electrophysiological measurements while participants supervised a simplified automated landing sequence by considering both visual and auditory signals. Our results revealed evidence of an early visual-to-auditory gating mechanism that occurs when visual parameters ("land") were contradictory to the auditory alarm ("go around"). This mechanism attenuates early auditory processing (N100) and could provide an alternative explanation to the inattentional deafness phenomenon in aeronautics [23].

In the present study, we intended to investigate auditory misperception with EEG in a motion flight simulator. The main objective of this study was to implement realistic experimental scenarios that could cause inattentional deafness under more ecological conditions than in previous work [12,23]. We therefore manipulated three different level of difficulty and test their effectiveness over the auditory signal detection. These three scenarios were tested on one pilot equipped with a 32-channel EEG. Eventually, in order to assess the feasibility of performing mental state estimation in the cockpit, we tested the use of machine learning algorithms on the collected data set.

2   Material and Method

2.1   Flight Simulator

We used the ISAE three-axis motion (roll, pitch, and height) flight simulator designed by the French Flight Test Center to conduct the experiment. It simulates a twin-engine aircraft flight model. The user interface is composed of a Primary Flight Display, a Navigation Display, and an upper Electronic Central Aircraft Monitoring Display page. The flight simulator is equipped with classical actuators such as the side-stick, the rudder, the throttle, the flaps levers and a complete autopilot to control the flight. Two stereophonic speakers, located under the displays on each side of the cabin, were used to broadcast continuous radio communications and the engine sound (70 dB), as well as to trigger the oddball sounds.

2.2   Scenarios

The participant performed three scenarios that differed from each other in the level of difficulty. In each of these scenarios, we used an oddball paradigm (stimulation with matlab psychtoolbox-3) with a total of 280 auditory stimuli: 70 targets (normalized pure tone, 1100 Hz at 88 dB) and 210 non-targets (normalized pure tone, 1000 Hz at 88 dB). The pilot used a Cedrus response pad, located below the throttle lever, to respond to auditory targets only, in each scenario (Figs. 1 and 2). The three scenarios lasted for about eight minutes each and were defined as follows:

– "Level 0" scenario: This was the reference scenario. In this experimental condition, the autopilot was engaged to level off the plane at a constant speed. The only task for the pilot was to respond to auditory target stimuli.
– "Level 1" scenario: This scenario consisted of performing a night manual approach and landing at Blagnac Airport (Toulouse, France) following the official procedure. The plane was initiated 30 nautical miles (NM) from Toulouse VOR DME (a radio beacon) at an altitude of 5000 ft. The pilot had to steer the aircraft to a heading of 210 degrees for 17 NM descending to an altitude of 3000 ft. He then had to steer 270 degrees until his radio-navigation beacon indicated a radial of 310 degrees. He eventually had to align to a heading of 144 degrees to capture the lateral axis of Toulouse Blagnac airport. In a distance of 5NM from Toulouse VOR DME, the pilot could initiate his descent to land safely on the landing ground.
– "Level 2" scenario: This scenario was identical to the previous one to the exception that we triggered a burning left-engine event at the beginning. A red light was flashing in the cockpit and we used a Bundle FOG 400 device to smoke out slightly the cabin every one minute. Moreover, this night landing was performed with no visibility. Eventually, the pilot was told that he had 8 min to perform this emergency landing.

Fig. 1. Participant in the motion flight simulator.

Fig. 2. Level 2 scenario, with smoke and red flashing warning. (Color figure online)

3   Participant and Protocol

The participant was a pilot (24 year old, 50 flight hours) from Institut Superieur de l'Aeronautique et de l'Espace (ISAE) campus, recruited by local advertisement and did not receive any payment for his participation. As a PPL pilot, he reported normal or corrected-to-normal vision and normal audition. Once the participant was told about the purpose of the experiment, he was sat in the flight simulator, and the EEG was set. He then completed a 30 min training

session in which he performed manual approaches and landings. Eventually, he was then trained to perform the oddball task for five minutes. After the training, the experiment was started: the simulator motion was engaged to reproduce realistic flight sensations, and a continuous radio communication was also broadcasted to reproduce more ecological flight conditions.

3.1   Data Acquisition and Processing

A self-report of stress and mental demand level was collected using a visual analog scale (1 for very low, 7 for very high) after the experiment. Reaction times and accuracy for the detection of auditory alarms were collected using a Cedrus RB response-pad and stored in a matlab file.

EEG data were recorded continuously with a BioSemi EEG system (BioSemi©, Amsterdam) from 32 "active" (preamplified) Ag-AgCl scalp electrodes located according to the International 10/20 system, at a 512 Hz sampling rate and with a 0–104 Hz band-pass filter. The data were then re-referenced offline to the algebraic average of the left and right mastoids, down-sampled to 500 Hz, and filtered with a band-pass of 0.1–40 Hz. An independent component analysis using EEGlab[1] (13.4.4b version) was performed to isolate and reject eye blinks and movements. Data were later segmented into 1200 ms epochs starting 200 ms before the onset of each sound. Spectral power analyses were carried out on the continuous data in the 8–12Hz band, and ERPs were computed using EEGlab.

Moreover, in order to assess whether these EEG data could be used for mental state monitoring purposes, a preliminary test of machine learning techniques was performed on the data of this subject. The processing chain included a 100 Hz resampling and a baseline correction using the first 200 ms of each epoch. Next, a spatial filtering step that has been proven to work efficiently to enhance the discriminability of ERPs for mental state monitoring purposes [22]. Contrary to Roy and collaborators, only one filter was kept here. The filter that was selected had the higher associated eigenvalue. Lastly, the filtered signal (60 samples) was classified using a robust Linear Discriminant Analysis (LDA) (i.e. with a shrinkage estimation of the covariance matrices; [24]) and a random 10-fold cross-validation procedure to check the LDA performance. We evaluated the feasibility of discriminating between two signals:

– at the single-trial level;
– averaging two trials, when applicable.

Three analyses were performed:

– The first one was designed to allow the assessment of the performance of the processing chain that was originally designed for laboratory settings by Roy and collaborators [22]. This first test was performed on the ERP discrimination of targets (hits) and distractors in the level 0 scenario (test 1).

---

[1] www.sccn.ucsd.edu/eeglab.

Therefore, in this case the algorithm attempts to classify the **targets** vs. the **distractors**.

– The second test was performed on the binary estimation of the stress and mental demand using only the correctly detected targets for the level 0 and the level 1 scenarios (test 2). In this case, the algorithm attempts to classify the task demand, that is to say the **level 0** vs. the **level 2** conditions.
– The third test was performed on the binary estimation of the detection of targets in the level 2 scenario (test 3). Here, the algorithm attempts to classify the target **hits** vs. **misses**.

4   Results

4.1   Subjective and Behavioral Measurements

The Table 1 reports the participant's subjective ratings (stress, mental demand) for each scenario.

Table 1. Subjective rating

	Level 0	Level 1	Level 2
Stress	1	3	5.5
Mental demand	1	4	6.5

The analyses of the behavioral results revealed that the level of difficulty of the scenarios impacted the pilot's alarm detection rate. In the first scenario (level 0: level off flight) the pilot missed no alarm whereas he missed 38 % of them in the "level 1" condition (night landing). This rate increased in the third scenario ("level 2" condition: night landing with smoke in the cockpit) as the pilot misssed 58 % of alarms (Fig. 3).

4.2   Physiological Measurements

Scenario Load. The spectral analysis of the continuous EEG data showed a decrease of alpha power at parietal sites as a function of stress and mental demand (Fig. 4). More precisely, a decrease of the 10 Hz band was observed between "level 0" ($11.65\mu V^2/Hz$) and "level 1" ($9.5\mu V^2/Hz$). In addition, a larger decrease of the 10 Hz was observed between the "level 1" ($9.5\mu V^2/Hz$) and the "level 2" scenario ($3.5\mu V^2/Hz$).

Load × Sound Interaction. When looking at the interaction between the scenraio type and the sound type, we found that the sound-related N100 amplitude was affected differentially by the scenario depending on the sound type. Scalp topographies at 114 ms revealed a decrease of the fronto-central N100 amplitude for the more demanding scenario compared to the reference one for both the
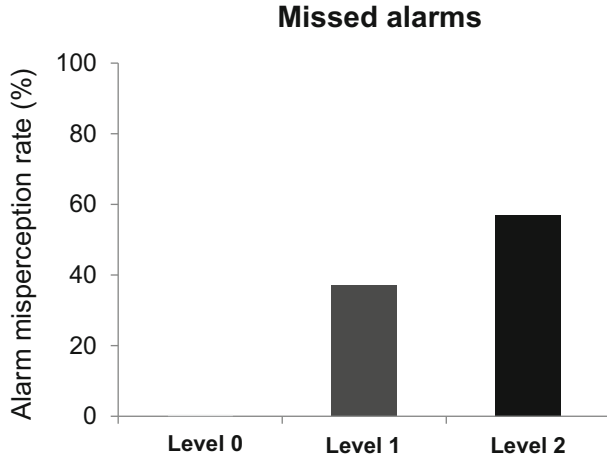
Missed alarms

Fig. 3. Behavioral results of the participant.

standard sound and the alarm (Fig. 5). However, only the alarm-related N100 amplitude decreased in the "level 1" condition compared to the reference one. In addition, the N100 amplitude in the "level 0" scenario was larger for the alarm than for the standard sound while it was the opposite in the "level 1" scenario.

4.3 Deafness Classifier

The results obtained thanks to the passive BCI chain detailed in the Data acquisition and processing section are given in Table 2 for each mental state estimation procedure that was performed. For the first test, the attempt to classify the targets vs. the distractors in the level 0 condition, the best performance was obtained using the average of two trials, with 70 % of accuracy. This demonstrates that the chosen processing chain can work well in an ecological setting for a simple oddball task. As regards the more subtle mental state estimation tests, for both the workload level estimation and the detection estimation, the best performance is reached with a single trial, with respectively 61 % and 60 % of accuracy. Although far from perfect, these results show that it is possible to estimate detection and engagement in an ecological setting with performances above chance level.

5   Discussion

The main objective of this paper was to investigate inattentional deafness under realistic settings. We manipulated three levels of difficulty across the scenarios. In the reference scenario, the flying task was easy and not stressful as the pilot only had to respond to auditory alarms. In the second one, the task was more demanding and stressing as it involved a night approach and landing while responding to alarms. The last one was designed to be the most stressing and
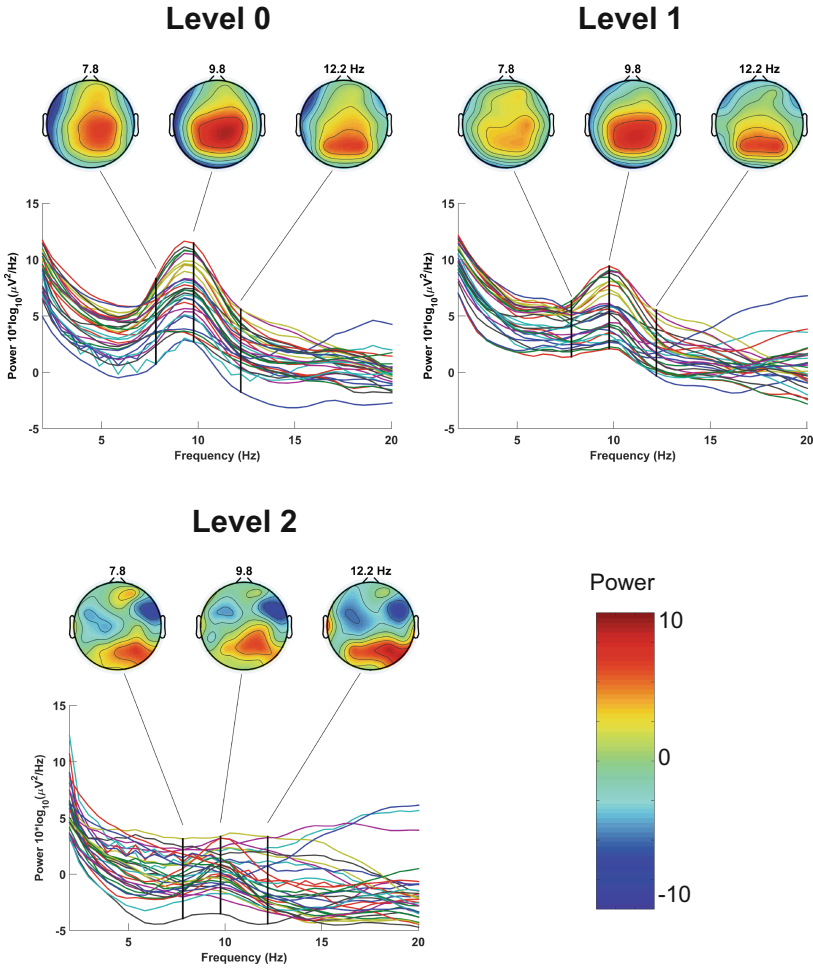
Fig. 4. Power spectral results on continuous EEG data for the three scenarios.

demanding as it involved a night approach and landing with a simulated engine failure (i.e. flashing warning and smoke in the cockpit). The participant's subjective results confirmed our assumptions as his rating indicated an increased level of stress and mental demand across the scenarios. As expected, the behavioral results revealed that the rate of missed alarms increased across these three scenarios (respectively of 0, 37, and 54 %). Consistently with these findings, the EEG frequency analysis revealed a decrease in spectral power within the alpha band across the three scenarios. Indeed, decrease in the this frequency band is associated with higher task demand [28] and stress situations [1]. Taken together our subjective, behavioral and electrophysiological results suggested that, when combined, increased task demand and stress were efficient to induce a high rate
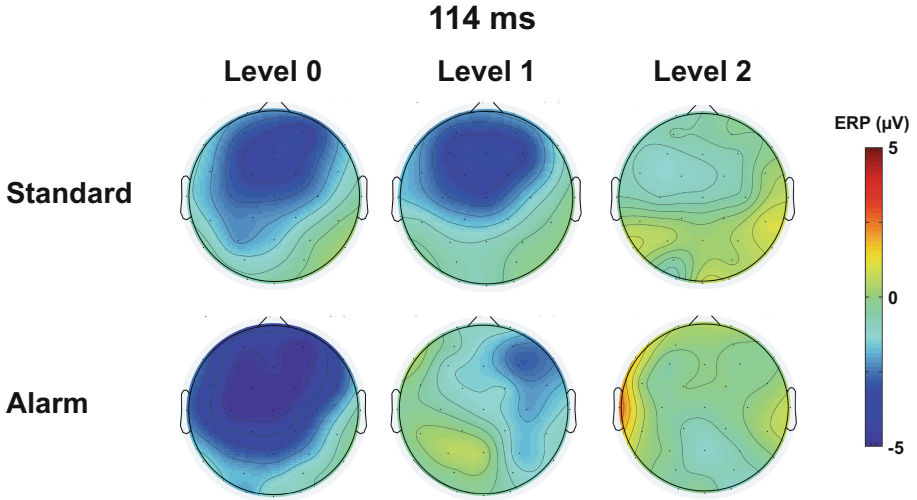
114 ms

Fig. 5. Scalp topographies plotted at 114 ms post-stimulus onset.

Table 2. Mental State Estimation: average *(standard deviation)* performance for the 3 tests.

	Description		Algorithm performance	
Test	Classes	Condition	1 trial	2 trials
test 1	Targets vs Distractors	Level 0	60.60 *(2.68)*	**69.5 (3.56)**
test 2	Level 0 vs. Level 2	Hits	**61.42 (5.16)**	59.13 *(5.38)*
test 3	Hits vs. Misses	Level 2	**59.74 (5.96)**	58.14 *(2.48)*

of auditory misperception. They confirmed previous findings that increased task demand [11,12,16,23] and stress [7] impair attentional abilities. However, our experimental design did not allow us to disentangle the contribution of each of these factors to provoke this phenomenon.

The ERP analyses highlighted findings at the perceptual level with the N100 component. This ERP is the early electrophysiological signal of the stimulus processing in auditory primary cortices around 100 ms post-stimulus. In accordance with previous findings [15,25], the N100 amplitude was higher during the processing of the auditory alarm (low probability stimulus), than during the processing of the frequent sound in each conditions. More interestingly, the N100 amplitude was reduced as the level of difficulty increased, concomitantly with lower behavioral responses to the auditory alarm. It is well admitted that the N100 amplitude is an index of auditory-attention allocation [14]. This result may therefore suggest that an inattentional deafness phenomenon has occurred at the perceptual level (i.e. 100 ms). If this phenomenon had occurred at 300 ms (which was not the case) when the alarm reached awareness, one could have argued that it was a conscious and voluntary inhibition of the alarm

(e.g. cry wolf effect). However, our measurements did not allow us to identify the neural correlates of inattentional deafness as we did not find differences between misses and hit.

Moreover, this study investigated the possibility of using a processing chain designed to enhance the discriminability of ERPs for mental state monitoring purposes [22]. The results indicated that this chain could be used in a quite ecological setting (i.e. the flight simulator) as attested by the good results obtained for the oddball task, but also for more subtle mental states such as stress level and the detection of target, that is to say the inattentional deafness phenomenon. Therefore, one could imagine the design of an adaptive cockpit that would take this information of stress level and inattention to alarms into account to implicitly adapt itself with a set of counter-measures. However, here only one subject was recorded and therefore it remains to be evaluated whether the performance of this chain would maintain itself or even increase for several participants. Moreover, it would be interesting to assess the possibility to use mental state markers from other recording modalities such as fNIRS [2,10,13] and eye movements [8].

References

1. Alonso, J., Romero, S., Ballester, M., Antonijoan, R., Mañanas, M.: Stress assessment based on eeg univariate features and functional connectivity measures. Physiol. Meas. **36**(7), 1351 (2015)
2. Ayaz, H., Shewokis, P.A., Bunce, S., Izzetoglu, K., Willems, B., Onaral, B.: Optical brain monitoring for operator training and mental workload assessment. Neuroimage **59**(1), 36–47 (2012)
3. Bliss, J.P.: Investigation of alarm-related accidents and incidents in aviation. Int. J. Aviat. Psychol. **13**(3), 249–268 (2003)
4. Causse, M., Péran, P., Dehais, F., Caravasso, C.F., Zeffiro, T., Sabatini, U., Pastor, J.: Affective decision making under uncertainty during a plausible aviation task: an fmri study. NeuroImage **71**, 19–29 (2013)
5. Dehais, F., Causse, M., Pastor, J.: Embedded eye tracker in a real aircraft: new perspectives on pilot/aircraft interaction monitoring. In: Proceedings from The 3rd International Conference on Research in Air Transportation. Federal Aviation Administration, Fairfax (2008)
6. Dehais, F., Causse, M., Régis, N., Menant, E., Labedan, P., Vachon, F., Tremblay, S.: Missing critical auditory alarms in aeronautics: evidence for inattentional deafness?. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56, pp. 1639–1643. Sage Publications (2012)
7. Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., Tremblay, S.: Failure to detect critical auditory alerts in the cockpit evidence for inattentional deafness. Hum. Factors J. Hum. Factors Ergon. Soc. **56**(4), 631–644 (2014)
8. Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., Gateau, T.: Automation surprise in aviation: real-time solutions. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 2525–2534. ACM (2015)

 9. Dehais, F., Tessier, C., Christophe, L., Reuzeau, F.: The perseveration syndrome in the pilot's activity: guidelines and cognitive countermeasures. In: Palanque, P., Vanderdonckt, J., Winckler, M. (eds.) HESSD 2009. LNCS, vol. 5962, pp. 68–80. Springer, Heidelberg (2010)
10. Gateau, T., Durantin, G., Lancelot, F., Scannella, S., Dehais, F.: Real-time state estimation in a flight simulator using fnirs. PloS one **10**(3), e0121279 (2015)
11. Giraudet, L., Imbert, J.P., Tremblay, S., Causse, M.: High rate of inattentional deafness in simulated air traffic control tasks. Procedia Manuf. **3**, 5169–5175 (2015)
12. Giraudet, L., St-Louis, M.E., Scannella, S., Causse, M.: P300 event-related potential as an indicator of inattentional deafness? PLoS one **10**(2), e0118556 (2015)
13. Hernandez-Meza, G., Slason, L., Ayaz, H., Craven, P., Oden, K., Izzetoglu, K.: Investigation of functional near infrared spectroscopy in evaluation of pilot expertise acquisition. In: Schmorrow, D.D., Fidopiastis, C.M. (eds.) AC 2015. LNCS, vol. 9183, pp. 232–243. Springer, Heidelberg (2015)
14. Hink, R.F., Van Voorhis, S., Hillyard, S., Smith, T.: The division of attention and the human auditory evoked potential. Neuropsychologia **15**(4), 597–605 (1977)
15. Kramer, A.F., Trejo, L.J., Humphrey, D.: Assessment of mental workload with task-irrelevant auditory probes. Biol. Psychol. **40**(1), 83–100 (1995)
16. Kreitz, C., Furley, P., Simons, D.J., Memmert, D.: Does working memory capacity predict cross-modally induced failures of awareness? Conscious. Cogn. **39**, 18–27 (2016)
17. Lebib, R., Papo, D., de Bode, S., Baudonnière, P.M.: Evidence of a visual-to-auditory cross-modal sensory gating phenomenon as reflected by the human p50 event-related brain potential modulation. Neurosci. Lett. **341**(3), 185–188 (2003)
18. Macaluso, E., Driver, J.: Multisensory spatial interactions: a window onto functional integration in the human brain. Trends Neurosci. **28**(5), 264–271 (2005)
19. Macdonald, J.S., Lavie, N.: Visual perceptual load induces inattentional deafness. Attention Percep. Psychophysics **73**(6), 1780–1789 (2011)
20. Meredith, C., Edworthy, J.: Are there too many alarms in the intensive care unit? an overview of the problems. J. Adv. Nurs. **21**(1), 15–20 (1995)
21. Peryer, G., Noyes, J., Pleydell-Pearce, K., Lieven, N.: Auditory alert characteristics: a survey of pilot views. Int. J. Aviat. Psychol. **15**(3), 233–250 (2005)
22. Roy, R.N., Bonnet, S., Charbonnier, S., Jallon, P., Campagne, A.: A comparison of erp spatial filtering methods for optimal mental workload estimation. In: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, pp. 7254–7257. IEEE (2015)
23. Scannella, S., Causse, M., Chauveau, N., Pastor, J., Dehais, F.: Effects of the audiovisual conflict on auditory early processes. Int. J. Psychophysiol. **89**(1), 115–122 (2013)
24. Schäfer, J., Strimmer, K.: A shrinkage approach to large-scale covariance matrix estimationand implications for functional genomics. Stat. Appl. Genet. Mol. Biol. **4**(1), 1175–1189 (2005)
25. Singhal, A., Doerfling, P., Fowler, B.: Effects of a dual task on the n100–p200 complex and the early and late nd attention waveforms. Psychophysiology **39**(02), 236–245 (2002)
26. Sinnett, S., Spence, C., Soto-Faraco, S.: Visual dominance and attention: the colavita effect revisited. Percept. Psychophysics **69**(5), 673–686 (2007)
27. Wickens, C.D., Rice, S., Hutchins, S., Keller, M.D., Hughes, J., Clayton, K.: False alerts in the air traffic control traffic conflict alerting system: is there a "cry wolf" effect? Hum. Factors J. Hum. Factors Ergon. Soc. **51**, 446–462 (2009)
28. Wilson, G.F.: An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int. J. Aviat. Psychol. **12**(1), 3–18 (2002)