Busy and confused? High risk of missed alerts in the cockpit: An electrophysiological study

Mickael Causse [a,*], Fabrice B.R. Parmentier [b,c,d], Damien Mouratille [a,e], Dorothée Thibaut [a], Marie Kisselenko [a], Eve Fabre [a]

[a] *ISAE-SUPAERO, 10 avenue Edouard Belin, 31400 Toulouse, France*
[b] *Neuropsychology & Cognition Group, Department of Psychology & Research Institute of Health Sciences (iUNICS), University of the Balearic Islands, Palma, Spain*
[c] *Balearic Islands Health Research Institute (IdISBa), Palma, Balearic Islands, Spain*
[d] *Department of Psychology, University of Western Australia, Perth, Australia*
[e] *CLLE, Université de Toulouse, CNRS, UT2J & ENAC, France*

Keywords:
Mental workload
Auditory load
Missed alerts
Alarm profusion
ERPs
P300

ABSTRACT

The ability to react to unexpected auditory stimuli is critical in complex settings such as aircraft cockpits or air traffic control towers, characterized by high mental load and highly complex auditory environments (i.e., many different auditory alerts). Evidence shows that both factors can negatively impact auditory attention and prevent appropriate reactions. In the present study, 60 participants performed a simulated aviation task varying in terms of mental load (no, low, high) concurrently to a tone detection paradigm in which the complexity of the auditory environment (i.e., auditory load) was manipulated (1, 2 or 3 different tones). We measured both detection performance (miss, false alarm, d') and brain activity (event-related potentials) associated with the target tone. Our results showed that both mental and auditory loads affected target tone detection performance. Importantly, their combined effects had a large impact on the percentage of missed target tones. While, in the no mental load condition, miss rate was very low with 1 (0.53%) and 2 tones (1.11%), it increased drastically with 3 tones (24.44%), and this effect was accentuated as mental load increased, yielding to the higher miss rate in the 3-tone paradigm under high mental load conditions (68.64%). Increased mental and auditory loads and miss rates were associated with disrupted brain responses to the target tone, as shown by a reduced P3b amplitude. In sum, our results highlight the importance of balancing mental and auditory loads to maintain efficient reactions to alarms in complex working environment.

1. Introduction

1.1. *Reacting to auditory alarms in challenging environments*

The responsiveness to unpredictable auditory stimuli in the form of an orienting response is a fundamental aspect of our cognitive functioning (e.g., Sokolov, 1963) and is thought to have played an essential role in our survival (Garcia-Garcia et al., 2010). This mechanism enables the detection of potential dangers and the adaptation of our actions, allowing for example the sudden and unexpected sound of an approaching vehicle to break into our focus of attention as we are crossing a road. This is also of special relevance in many complex working environments (e.g., aircraft cockpits, control towers, operating rooms, nuclear power plants) where operators are expected to react promptly and accurately to auditory alarms (e.g., Guillaume, 2011). Auditory alarms are effective stimuli for conveying critical information in visually saturated environments, both because they do not require an effortful and systematic scan of a control panel (Guillaume, 2011) and because their near omnidirectional character (i.e., "gaze-free") yields a higher probability of being detected compared to visual signals (Morris and Montano, 1996; Van der Heiden et al., 2018).

In many domains, operators are required to perform tasks while remaining responsive to rare and unexpected auditory alarms (Causse, Peysakhovich, et al., 2016; Ferraro and Mouloua, 2021) that are not necessarily related to the task at hand, but that may potentially convey important information for safety (e.g., a cockpit depressurization alarm during a landing maneuver). Unfortunately, both accident reports and experimental studies have shown that operators' response to auditory

---

* Corresponding author.
E-mail address: mickael.causse@isae-supaero.fr (M. Causse).

https://doi.org/10.1016/j.brainres.2022.148035

alarms can sometimes be inappropriate (Bliss and Acton, 2003; Dehais et al., 2014; Giraudet et al., 2012; Van der Heiden et al., 2018). Sometimes, operators detect these alarms but fail to react because they misunderstand them (Dehais et al., 2014), consider them unreliable (Bliss and Acton, 2003; Bliss and Dunn, 2000), and/or underestimate their urgency (Salas and Schlesinger, 2019). However, on occasion, the absence of reaction results from a failure to detect these alarms in the first place. Growing evidence suggests that the capacity to involuntary detect and process auditory information may be modulated by the level of engagement in the task at hand (Van der Heiden et al., 2018). Selective attention allows to focus on a task and to limit distraction from extraneous stimuli. While such filtering enables to perform well on the task at hand, it can also result (when extreme) in operators missing safety-relevant signals. This phenomenon known as *inattentional deafness* can dramatically affect safety (Giraudet, St-Louis, et al., 2015; Koreimann et al., 2014; Macdonald and Lavie, 2011; Raveh and Lavie, 2015).

In the next sections, we briefly review the impact of two variables known to modulate auditory attention: mental load and auditory load, which have historically been studied independently. Yet they form an integral part of many applied settings. One important novel feature of our study is the empirical investigation of the combined effect of these factors within a simplified aviation task. A second original aspect is the combination of behavioral and electrophysiological measures, which provides a more complete view of alarm responsiveness as well as an opportunity to examine the potential relationship between these measures.

1.2. Mental load and alarm responsiveness

Operators who miss alerts are often already engaged in a mentally-demanding task (see for instance the accident of Eastern Air Lines Flight 401; NTSB, 1973). Indeed, a high mental load can affect both the detection of auditory stimuli and the allocation of attentional resources to the processing of their content (e.g., Causse, Imbert, et al., 2016; SanMiguel et al., 2008; Zhu et al., 2022). This is in line with the classical theory of an inverse relationship between mental load and attentional reserve which, when depleted, does not allow the proper processing of additional information (Kahneman, 1973). Working memory limitations certainly have a central role in this attentional bottleneck (Tombu et al., 2011): unexpected sounds that are normally automatically detected (see Parmentier, 2014, for a review) are less so when the primary task demands on working memory increases (Berti and Schröger, 2003). In numerous studies, the difficulty of the primary task is manipulated and its impact on the passive or active processing of auditory stimuli is examined via the analysis of event-related potentials (ERPs) (e.g., Ghani et al., 2020; Gibson et al., 2019; Giraudet, Imbert, et al., 2015; Kramer et al., 1995; Raabe et al., 2005; Swerdloff and Hargrove, 2020; Ullsperger et al., 2001). The tasks used to evaluate the passive processing of auditory stimuli generally consist in presenting task-irrelevant auditory stimuli (also called distractors or auditory probes), to which no active response is expected from participants. Overall, these studies have shown that an increase in task load negatively affects the processing of auditory stimuli (e.g., Berti and Schröger, 2003; Fabre et al., 2017; Lv et al., 2010; SanMiguel et al., 2008).

Studies assessing the active processing of auditory stimuli require participants to both detect a target sound and physically react to it (e.g., pressing a response button), in a way comparable to how operators are expected to behave after the occurrence of an alarm (e.g., Callan et al., 2018; Causse, Imbert, et al., 2016; Dehais, Duprès, et al., 2019; Giraudet, St-Louis, et al., 2015). Typically, participants perform a primary task (e.g., piloting task: Causse, Peysakhovich, et al., 2016; Giraudet, St-Louis, et al., 2015; Kramer et al., 1987), concurrently to an alarm detection task (referred to as a secondary task).

Several authors have suggested that brain responses to auditory stimuli are a reliable measure of the mental load elicited by the primary task (Kramer et al., 1995; Miller et al., 2011). In general, the results show that the workload impacts brain electrophysiological responses to target sounds at different stages of the information processing, as often indexed by the N100 and the P300 components (Dehais, Roy, et al., 2019; Giraudet, St-Louis, et al., 2015).

The auditory N100 is a negative component occurring 80 to 180 ms after the stimulus onset, maximal at frontal-central sites (Näätänen and Picton, 1987). It is thought to be generated in the primary and associative auditory cortices, the superior temporal gyrus, and Heschl's gyrus (Wolpaw and Penry, 1975; Zouridakis et al., 1998), and to reflect the perceptual processing of auditory change detection (Ghani et al., 2020; Näätänen et al., 1978). The P300 is a later positive component unraveling within the 250 – 450 ms time window after stimulus onset and typically increases in magnitude from the frontal to parietal electrode sites (Johnson, 1993). This component is believed to reflect both attentional and memory mechanisms (Polich, 2007). The P300 can be divided in two subcomponents, the P3a and the P3b (Snyder and Hillyard, 1976). P3a originates from stimulus-driven frontal attention mechanisms, whereas P3b originates from temporal-parietal activity (Polich, 2007). P3a is thought to reflect an automatic switch of attention from the primary task towards distractive stimuli (Polich, 2003), while P3b seems to reflect the allocation of attentional resources that promote context updating operations, memory processing (Brázdil et al., 2001; Donchin and Coles, 1988; Knight, 1996), and possibly the linking of stimulus identification and response selection (Frühholz et al., 2011; Verleger et al., 2005). According to Squires (1975), P3a (latency about 240 ms) is elicited in response to infrequent, unpredictable shifts of either intensity or frequency in a train of tones, whereas P3b component (mean latency about 350 ms) occurs only when a subject is actively attending to the tones. In this sense, P3b is considered to reflect the voluntary detection of a task-relevant stimulus (Picton, 1992; Polich and Criado, 2006), whereas P3a might be more specific to deviant auditory non-target events.

While N100 has been sometimes found to be modulated by mental load (Dehais et al., 2016; Fabre et al., 2017), the literature more frequently reports the P300 to be impacted by load or task difficulty (e.g., Kramer et al., 1995; Miller et al., 2011), including in complex applied settings such as aviation (Dehais, et al., 2019; Giraudet et al., 2015; Kramer et al., 1987) or car driving (Van der Heiden et al., 2018; Wester et al., 2008). Regarding the links between alert detection performance and brain response, Dehais, Roy, et al. (2019) found amplitude reductions of the N100, P3a and P3b for missed target sounds relative to detected ones and Giraudet et al. (2015) revealed that auditory P3b amplitude was negatively correlated with the percentage of missed auditory targets.

1.3. Multiple alarms and auditory load

Exposure to multiple alarms can also affect the identification process (Potnuru et al., 2020). In intensive health care settings, for instance, over 300 alarms can be played per patient and per day, while less than 5 % of these alarms require urgent clinical intervention (Association for the Advancement of Medical Instrumentation, 2011). This constant bombardment of alarms leads to 'alarm fatigue' (Lansdowne et al., 2016) and is also likely to create confusion, a main cause of unresponsiveness (Xiao et al., 2003). The auditory environment can vary in complexity, from quiet (i.e., no auditory stimuli other than target ones) to complex, with numerous sounds conveying independent messages competing for the operator's attention (Ho and Spence, 2005). In complex auditory environments, the processing of target stimuli becomes significantly more difficult for the operator.

An interesting way to study the detection performance of a tone target among different other auditory stimuli is to use the oddball paradigm and its variations (e.g., Ebmeier et al., 1995; Katayama and Polich, 1996). In the 2-tone paradigm, two auditory stimuli are played: a standard (non-target) tone occurring frequently, and a target tone occurring less frequently (Squires et al., 1975). Participants are required to respond to the target tone and to ignore the standard tone. The 3-tone paradigm includes an additional non-target/rare distractor tone that must be ignored (Courchesne et al., 1984; Katayama and Polich, 1996), thus three sounds must be compared. Finally, the 1-tone (or single-tone) paradigm (Cass and Polich, 1997) includes only a target tone, implying that no sound comparison must be performed. While Cass and Polich (1997) found no difference in task performance errors (target misses + false alarms) between the 1-tone and the 2-tone paradigm, other studies found a few more errors in the 2-tone than in the 1-tone paradigm (Polich et al., 1994; Polich and Heine, 1996; Polich and Margala, 1997). Katayama and Polich (1996) mentioned that the 3-tone paradigm can produce a relatively more complex stimulus construction compared to the 2-tone situation. They compared detection performance in 1-, 2-, and 3-tone paradigms versions. They found no difference in target tone detection performance, however they observed longer reaction times in the 2-tone than in the 1-tone paradigm (Polich and Heine, 1996), and in the 3-tone than in the 2-tone paradigm (Ebmeier et al., 1995; Grillon et al., 1990).

Electrophysiological studies using the oddball paradigm and its variants described above found somewhat divergent results. Some found no variation in P3b amplitude in response to target tones as the number of irrelevant tones varied (Cass and Polich, 1997; Katayama and Polich, 1996; Polich et al., 1994; Polich and Margala, 1997), while others found greater P3b responses to the target tone in the 2-tone relative to the 1-tone paradigm (experiment 1 in Cass and Polich, 1997; Polich and Heine, 1996), or in the 3-tone paradigm compared to the 2-tone paradigm (Grillon et al., 1990). In sum, the results of some behavioral and electrophysiological studies indicate a slight reduction of target processing efficiency as the number of non-target tones increases, suggesting that auditory load may affect the detection of auditory targets.

1.4. *The present study*

As described in the previous sections, both mental and auditory loads can negatively affect an operator's capacity to detect an alarm. However, these two factors have been mainly studied independently and their combined effect on detection performance is not currently known, especially when the different tones played have very close pitches. The present study combined the manipulation of mental and auditory loads to examine their impacts on target tone detection performance and brain activity (ERPs). Participants were required to perform a dual-task paradigm consisting in an aviation-inspired task and a tone detection paradigm (simulating an auditory alarm detection task). The aviation task – whose difficulty was manipulated to generate different levels of mental load (no load, low load, high load) – was taken from Giraudet et al. (2015). It reproduces a landing situation in which the pilot has to decide whether the approach should be continued or aborted based on various elements of information displayed in the cockpit (e.g., wind speed, aircraft position, etc.). Concurrently to this aviation task, participants performed a tone detection paradigm whose number of different tones varied (i.e., 1-, 2-, or 3-tone paradigm) to investigate the impact of the auditory load on the detection of the target tone. At the behavioral level, we predicted a greater number of missed target tones as both the difficulty of the aviation task (Giraudet, St-Louis, et al., 2015) and the number of different tones increased. At the electrophysiological level, we predicted a decrease in the P3b amplitude, also when both the difficulty of the aviation task and the number of different tones increased. These predictions were based on the hypothesis that (1) an increasing mental load in the aviation task might hinder the redirection of attention from the demanding aviation task towards the target tone in the tone detection paradigm; and that (2) auditory discrimination should be harder when the number of auditory distractors increases, in particular when three different tones are played (3-tone paradigm). Finally, 3) we predicted a dramatic increase in the number of missed target tones in conditions of high mental load in the 3-tone paradigm.

2. Results

2.1. *Aviation task*

A 2 (Mental Load: low load, high load) × 3 (Auditory Load: 1-tone, 2-tone, 3-tone) mixed ANOVAs revealed a significant main effect of the mental load on the number of errors committed in the aviation task, $F(1, 52) = 20.43$, $p < .001$, $\eta_p^2 = 0.27$, with more errors under high than low mental load. The main effect of the auditory load was not significant $F(1, 52) = 2.02$, $p = .141$, $\eta_p^2 = 0.07$, but it interacted with mental load, $F(2, 52) = 24.23$, $p < .001$, $\eta_p^2 = 0.46$. Whereas errors increased between low and high load in the 1- and 2-tone paradigm (LSD: $p < .001$ for both comparisons), it decreased between low and high load in the 3-tone paradigm (LSD: $p = .004$).

2.2. *Target tone detection performance*

All detection performance values are presented in supplementary material S1 (hit rate, miss rate, false alarm rate, d'). Miss rate, false alarm rate, and d' were analyzed using 3 (Mental Load: no load, low load, high load) × 3 (Auditory Load: 1-tone, 2-tone, 3-tone) mixed ANOVAs – with mental load as a within-participant factor and the auditory load as a between-participants factor.

2.2.1. *Miss rate*

We found a significant main effect of mental load, $F(2, 104) = 69.95$, $p < .001$, $\eta_p^2 = 0.57$, showing a higher percentage of misses in the high vs low mental load condition and in the low vs no mental load condition (LSD: both $p_s < 0.001$), see Fig. 1, top left. The main effect of the auditory load was significant, $F(2, 52) = 46.98$, $p < .001$, $\eta_p^2 = 0.64$, with a higher miss rate in the 3-tone paradigm than in both the 1-tone and 2-tone paradigms (LSD: both $p_s < 0.001$). Finally, the interaction term was also significant, $F(4, 104) = 4.62$, $p = .001$, $\eta_p^2 = 0.15$, showing more important effects of increased mental load on the miss rate in the 3-tone paradigm than in both the 1-tone and 2-tone paradigms.

2.2.2. *False alarm rate*

A significant main effect of mental load was found, $F(2, 104) = 13.79$, $p < .001$, $\eta_p^2 = 0.21$, with more false alarms in low and high mental load conditions than in the no mental load condition (LSD: both $p_s < 0.001$; see Fig. 1, top right). The main effect of the auditory load was significant, $F(2, 52) = 29.01$, $p < .001$, $\eta_p^2 = 0.53$, with a markedly higher number of false alarms in the 3-tone paradigm than in the 1-tone and 2-tone paradigms (LSD: both $p_s < 0.001$). The interaction term was also significant, $F(4, 104) = 8.47$, $p < .001$, $\eta_p^2 = 0.24$, while the false alarm rate increased between the no mental load and the low/high mental load conditions in the 1-tone and 2-tone paradigms (LSD: all $p_s < 0.001$), a rather opposite pattern was found in the 3-tone paradigm in which the false alarm rate decreased between the no mental load and high mental load condition (LSD: $p = .021$).

2.2.3. *Sensitivity (d')*

A significant main effect of mental load was found, $F(2, 104) = 130.44$, $p < .001$, $\eta_p^2 = 0.71$, with a lower d' in high vs low mental load condition and in low vs no mental load condition (LSD: both $p_s < 0.001$; see Fig. 1, bottom). The main effect of the auditory load was also significant, $F(2, 52) = 69.58$, $p < .001$, $\eta_p^2 = 0.73$, with a lower d' in the 3-tone than in the 2-tone paradigm and in the 2-tone than in the 1-tone paradigm (LSD: both $p_s < 0.001$). The interaction term was also significant, $F(4, 104) = 6.13$, $p < .001$, $\eta_p^2 = 0.19$, showing that the d' was low in the 3-tone paradigm even in the conditions of no mental load.

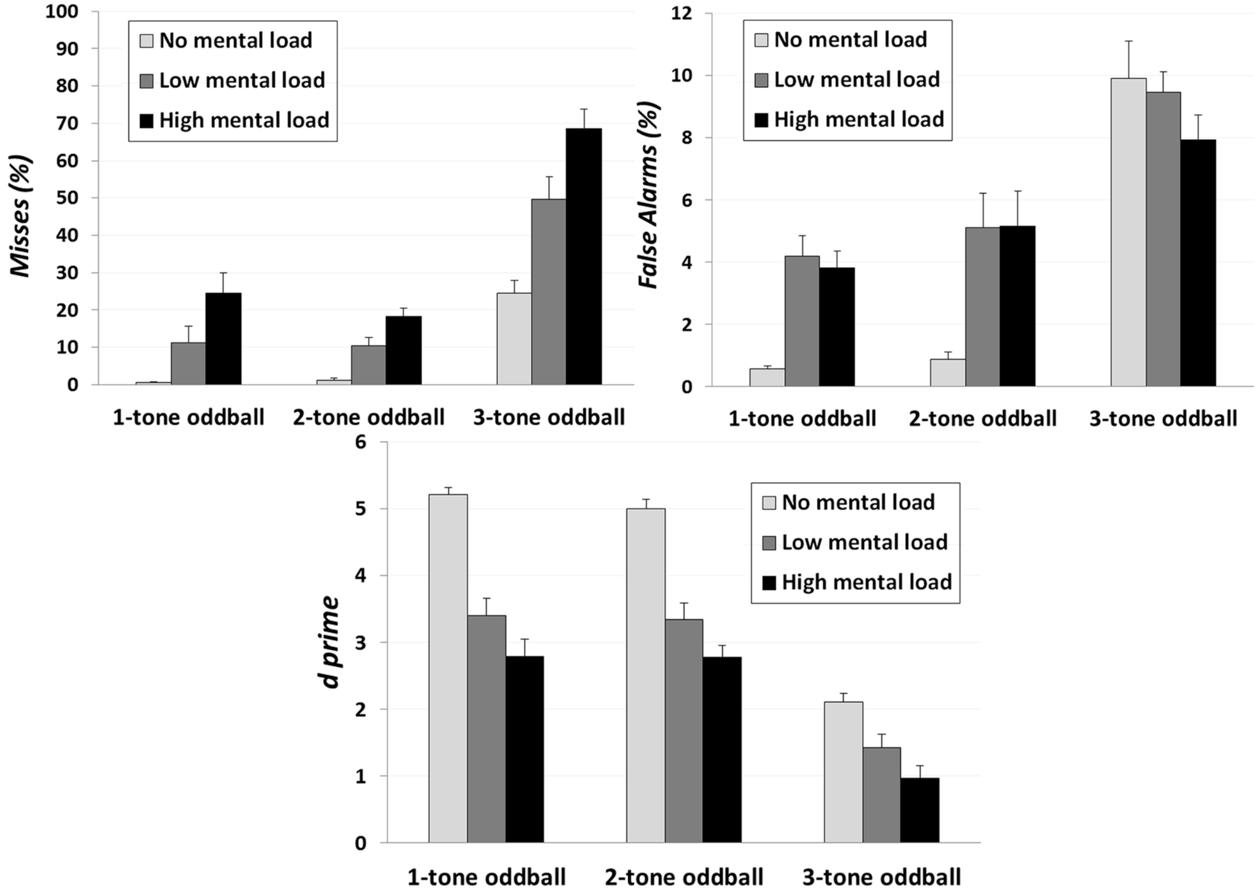


Fig. 1. Detection of the target tone in the no load (light grey), low load (medium grey) and high load (black) conditions for the 1-tone, 2-tone and 3-tone paradigms. Top left: Misses. Top right: False Alarms. Bottom: d'. Error bars represent one standard error of the mean.

2.3. Electrophysiological results

N100 and P3b values are presented in supplementary material S2. The N100 component was assessed with reference to the peak amplitude within the 80 – 150 ms time window at the Fz, Cz and Pz electrodes (e.g., Valéry et al., 2017). The P3b was assessed with respect to the mean amplitude within the 300 – 450 ms time window at Fz, Cz and Pz electrodes (Giraudet, St-Louis, et al., 2015). The amplitudes of the N100 and P3b components were analyzed at each of these electrode sites using a 3 (Electrode: Fz, Cz, Pz) × 3 (Mental Load: no load, low load, high load) × 3 (Auditory Load: 1-tone, 2-tone, 3-tone) mixed ANOVAs – with electrode and mental load as within-participant factors and the auditory load as a between-participants factor.

2.3.1. N100 component

We found a significant main effect of the electrode site, $F(2, 104) = 91.22$, $p <.001$, $\eta_p^2 = 0.64$, with greater N100 amplitudes in response to the target sound at Fz than at Cz (LSD: $p <.001$), and at Cz than at Pz (LSD: $p <.001$). The main effects of mental and auditory loads were not significant, $F(2, 104) = 0.29$, $p =.743$, $\eta_p^2 = 0.01$, and $F(2, 104) = 0.83$, $p =.438$, $\eta_p^2 = 0.03$, respectively (see Figs. 2 and 3). However, we found a significant Electrode × Auditory Load interaction, $F(4, 104) = 7.08$, $p <.001$, $\eta_p^2 = 0.21$, with reduced N100 amplitudes at Fz in the 1-tone paradigm relative to the 2-tone and 3-tone paradigms (LSD: both $p_s < 0.05$), see Fig. 3. The other interaction terms were not significant: Mental Load × Auditory Load, $F(4, 104) = 0.86$, $p =.489$, $\eta_p^2 = 0.03$; Mental Load × Electrode, $F(4, 208) = 0.19$, $p =.943$, $\eta_p^2 = 0.00$; and Mental Load × Auditory Load × Electrode, $F(8, 208) = 0.71$, $p =.676$, $\eta_p^2 = 0.03$.

2.3.2. P3b component

The analysis of the P3b amplitude revealed a significant main effect of electrode $F(2, 104) = 100.55$, $p <.001$, $\eta_p^2 = 0.66$, with a greater P3b amplitude in response to the target sound at Pz than at Cz and at Cz than at Fz (LSD: both $p_s < 0.001$). The main effect of mental load was also significant, $F(2, 104) = 20.21$, $p <.001$, $\eta_p^2 = 0.28$, with lower P3b amplitude under high than low mental load (LSD: $p =.039$) and under low than no mental load (LSD: $p <.001$), see Fig. 2. The P3b amplitude also varied with the auditory load, $F(2, 52) = 26.95$, $p <.001$, $\eta_p^2 = 0.51$, with a significantly greater P3b amplitude in the 1-tone than in the 2-tone paradigm and in the 2-tone than in the 3-tone paradigm (LSD: $p <.001$ and $p =.006$, respectively), see Fig. 3. The Electrode × Mental Load interaction was significant, $F(4, 208) = 8.99$, $p <.001$, $\eta_p^2 = 0.15$, with reduced P3b amplitude under high vs low mental load, and under low than no mental load, for all electrodes, but these effects being larger at Pz. The Electrode × Auditory Load interaction was also significant, $F(4, 104) = 3.03$, $p =.021$, $\eta_p^2 = 0.10$, with the P3b being significantly greater in the 1-tone than in the 2-tone paradigm, and in the 2-tone than in the 3-tone paradigm on all electrodes, but these effects being larger at Pz. Finally, neither the Mental Load × Auditory Load interaction nor the triple interaction were significant, $F(4, 104) = 0.64$, $p =.631$, $\eta_p^2 = 0.02$, and $F(8, 208) = 0.96$, $p =.462$, $\eta_p^2 = 0.04$, respectively.

2.3.3. Correlation between miss rate and P3b amplitude

We performed two Pearson correlations to examine the relationship between individual miss rate and P3b amplitude (Pz electrode) elicited by the target tones in the low and high mental load conditions. This analysis included the three groups of participants (i.e., the three tone detection paradigm variants). We found a significant negative

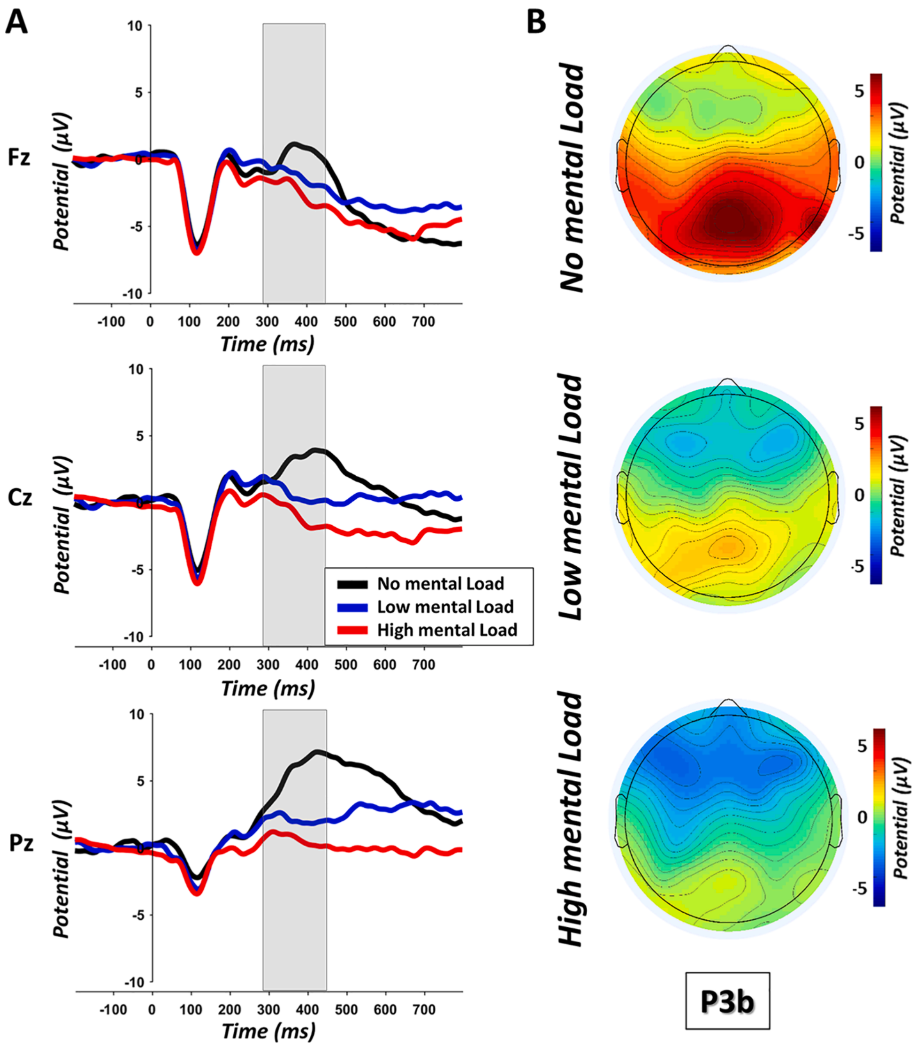


Fig. 2. (A) Grand average ERP waveforms at Fz, Cz and Pz observed in response to the target tones in the no (black line), low (blue line), and high mental load conditions (red line). (B) Scalp maps illustrating the P3b mean amplitude in the no (up), low (middle) and high mental conditions (down). A bar chart of the average P3b values at Fz, Cz and Pz during each tone detection paradigm is in supplementary material S3.

correlation between miss rate and P3b amplitude both in low ($r(52) = -0.29, p = .036$) and high load conditions ($r(52) = -0.44, p < .001$). In other words, the more the participants missed target tones, the lower was their mean P3b amplitude at the Pz electrode, see Fig. 4.

3. Discussion

3.1. *The impact of mental and auditory loads on tone detection performance*

In line with our predictions, increasing the aviation task load negatively impacted participants' performance to the parallel tone detection task. Interestingly, increased miss rate (and decreased d') were associated with an increased false alarm rate, thus this lower detection performance was not due to a lack of time to report the target tones or to a change toward a more conservative response criterion. In line with Giraudet, St-Louis, et al., (2015), we found no modulation of the N100 response due to the mental load, suggesting that early perceptual processes were not significantly impacted by this latter factor. In contrast, the P3b response decreased in amplitude as mental load increased, especially in the parietal region (Pz electrode). At least two explanations can account for the tight relationship between increased mental load and reduction of both tone detection performance and P3b amplitude. First, given the parietal contribution to the P3b generation (Polich, 2007), and the role of this cerebral structure in the control of voluntary attentional orienting (Corbetta et al., 2000), the decline of the P3b amplitude by high mental load probably reflects the increasing difficulty to voluntary orient attention towards the target tones when the aviation task demand on attentional resources increased. Attentional resources became scarce (Corbetta et al., 2000), assuming that some limited attentional resources had to be divided between the aviation task and the tone detection paradigm. Second, an increased mental load might also affect both the storage of the target tone representation in working memory and the selection of the associated response (Frühholz et al., 2011; Verleger et al., 2005).

Our results also revealed that alarm miss rate and P3b amplitude were negatively related, which is consistent with previous studies reporting lower P300 amplitudes for missed than for detected alarms (Dehais, Roy, et al., 2019), or significant correlations between the miss target tone rate and P300 amplitude (Giraudet, St-Louis, et al., 2015).

A second key finding in our study is the decrease in alarm detection as the auditory load increased, especially when three different tones were played. In the no mental load condition, while the miss rate was very low in the 1-tone (0.53 %) and 2-tone (1.11 %) paradigms (in line with Polich and Heine, 1996), it increased drastically in the 3-tone paradigm (24.44 %). This effect was accentuated as mental load increased, yielding the higher miss rate in the 3-tone paradigm under the high mental load condition (68.64 %). In this situation, participants that

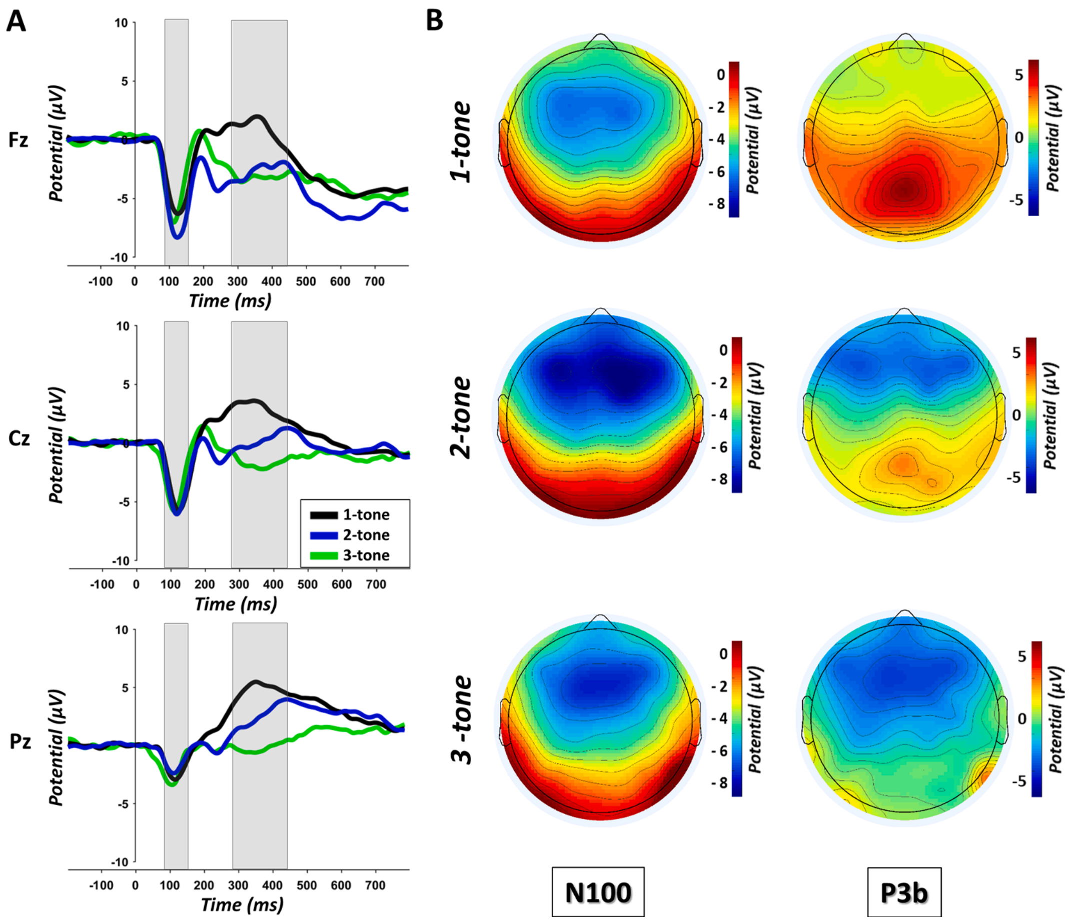


Fig. 3. (A) Grand average ERP waveforms at Fz, Cz and Pz observed in response to the target tone in the 1-tone (black line), 2-tone (blue line) and 3-tone paradigms (green line). (B) Scalp maps illustrating the N100 peak amplitude and the P3b mean amplitude in the 1-tone (up), 2-tone (middle), and 3-tone paradigms (down). A bar chart of the average P3b values at Fz, Cz and Pz during each tone detection paradigm is in supplementary material S4.

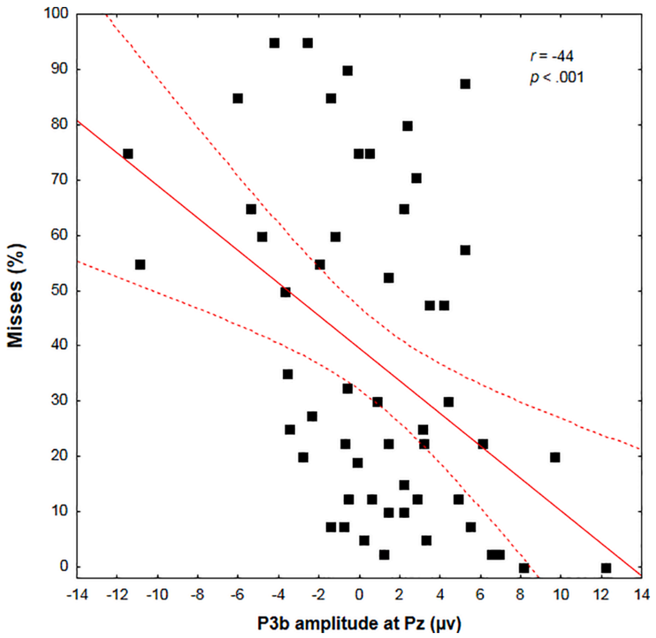


Fig. 4. Correlation between the miss rate (missed target tones) and the mean P3b amplitude for all tone detection paradigms. N = 54.

faced a difficult aviation task, also struggled to properly detect and discriminate the target tone from the two irrelevant sounds. These results contrast with the study from Katayama and Polich (1996), in which the miss rate did not decline in a 3-tone paradigm. Apart from the absence of mental load manipulation in their study, it is very likely that the main reason for this difference is that the pitches of the three tones were markedly different in their study (500, 1000, or 2000 Hz) in comparison to ours (1900, 1970, 2040 Hz). This lower pitch differences in our study certainly made the discrimination of the tones much harder than in the study of Katayama and Polich (1996). We used brief sinewave tones that were previously unknown to the participants. Without traces of these tones in long term memory, participants had to maintain these in working memory (Van Hedger et al., 2018), which was probably very difficult due to their perceptual similarity (Jackson et al., 2015). This may have contributed to a weaker memory representation of the tones in working memory and consequently a lower ability to engage an appropriate response (Frühholz et al., 2011; Verleger et al., 2005). More generally, in the 1-tone paradigm, misses can be confidently interpreted as the result of inattentional deafness. In the 2-tone and 3-tone paradigms, misses can be due to inattentional deafness or the difficulty to discriminate the target tone from the distractor tone(s). Interestingly, while errors in the aviation task increased with the load in the 1- and 2-tone conditions, it decreased in the 3-tone condition. This might reflect a disengagement from the harder tone detection task to the benefit of the aviation task when both mental and auditory load were high.

3.2. *Practical and neuroergonomics applications*

Our results provide empirical evidence of the increasing difficulty for operators to detect alarms when both mental and auditory loads increase. Importantly, aspects of our results indicate that the risk of missed alerts is amplified when an operator simultaneously deals with both types of loads. This may provide useful information to help improving the design of alarm systems in a wide range of safety–critical domains, such as transportation (e.g., semi-automated vehicles, aviation), medical care or nuclear powerplants. It is notoriously difficult to predict one's individual susceptibility to miss alerts, for example, individual working memory capacity appears to be a poor predictor (Kreitz et al., 2016). One possible measure to limit alarm unresponsiveness may consist in monitoring and balancing the operators' mental load in the working environment to spare attentional resources (Giraudet, St-Louis, et al., 2015). The real-time monitoring of an operator's mental load may possibly be implemented using a combination of performance monitoring, a physical assessment of the amount of stimulation and task difficulty faced by operators, and the measurement of key physiological indicators such as in this study. An optimum level of mental load should be sought, so as to avoid overload or under-stimulation, which can increase the risk of operators drifting "out of the loop" (Gouraud et al., 2017). Regularly exposing operators to the alarms may also help maintaining/consolidating their auditory representations in long term memory (Van Hedger et al., 2018). This may facilitate their recognition, especially when alarms are acoustically similar. Another recommendation would be to reduce the acoustic traffic jam in saturated settings, such as operating rooms or cockpits, by limiting the number and the variety of auditory alarms that can be issued during specific time periods (Edworthy, 2013; Konkani et al., 2012). Recent results show that our capacity to process auditory alarms depends in part on the relevance of the auditory stream (Scheer et al., 2018; Tellinghuisen et al., 2016), and individuals are less sensitive to auditory stimuli when instructed to disregard auditory information as opposed to monitor it. Scheer et al. (2018) suggested that requiring pilots to perform simple and frequent tasks in the auditory modality could heighten their awareness of the auditory environment and help prevent inattentional deafness. However, this must be balanced against the risk of saturating the auditory scene or interfering with the processing of other auditory signals.

Our study critically shows that brief pure tones can fail to attract operators' attention or can be difficult to discriminate, especially when operators are immerged in a mentally demanding situation. One solution may be to improve the design of alarm systems. According to Foley et al. (2020), "re-orchestring" standardized alarms with more complex sounds may be a desirable way to improve alarm efficiency and avoid confusion (Lacherez et al., 2007). Some authors recommend the use of natural auditory icons (e.g., using the sound of a heartbeat to notify to medical doctors that a patient's cardiac activity is abnormal) to maintain a correspondence between the acoustic property of the auditory natural indicator and the acoustic property of the message (Fitch and Kramer, 1992; Stevens et al., 2009). On a more general note, our study had to strike a balance between ecological validity and necessary constraints in order to collect solid and interpretable ERP data. Hence, we used brief pure tones rather than the type of longer and more complex auditory signals that are more representative of real alarms (Ho and Spence, 2005) or sounds/speech warnings encountered in the cockpit. Our results provide a first exploration and evidence of the impact of mental load and auditory load effects on responsiveness. These findings can now form the basis for future studies including more complex and ecologically-valid sounds (Winkler, 2003). Finally, evaluating and validating alarm systems using brain electrophysiological measurements may be a promising way to enhance alarms and reduce missed alerts/inattentional deafness. It may be useful for industry stakeholders and practitioners to set up alarm evaluation platforms and perform more systematically empirical experiments similar to the present one to determine the efficiency of specific alarms using behavioral and cerebral data (in particular P3b component).

4. Conclusion

In this paper, we presented novel results on the combined effects of both mental and auditory loads, two factors that have previously been assessed separately. We highlighted an interaction of these two factors on detection performance and ERPs. Past studies generally highlighted effects of mental load on ERP amplitudes, but rarely related them directly to the tone detection performance. Overall, our results showed that the detection of auditory target stimuli is hindered by both high mental and auditory loads, and that the combination of these factors yields the greater reduction in detection performance. Our electrophysiological results showed that the N100 response appears to be relatively immune to the impact of these two factors. In contrast, the amplitude of the P3b reduced as both type of loads increased, and was correlated with individual alarm miss rates. Most likely, mental and auditory loads appear to affect the attention orienting mechanism, the identification process, the storage of the target sound representation in working memory, and the subsequent selection of the associated response. Taken together, our behavioral and ERP results support the idea that the inability of operators to react to some auditory alarms can be due to attentional and cognitive limitations, contrary to the well-known cry-wolf effect whereby operators or pilots deliberately ignore alarm considered as unreliable due to a high rate of false positives (Bliss et al., 1995, 1999).

Measuring the P3b response may constitute an efficient tool to evaluate the ability of an individual to process auditory alarms in high load contexts. It may also help evaluate alarm designs and objectively identify those that can be efficient despite high load contexts. Our data also confirmed that studying alarm responsiveness can be performed with a 1-tone detection paradigm, which is simple to implement and integrate as part of applied studies of inattentional deafness. In contrast, the classical 2-tone detection paradigm (i.e. oddball task), which requites to compare two sounds may be more suited to study the "change deafness" phenomenon (Vitevitch, 2003). Numerous domains could benefit from a better consideration of the limitations of the human brain in terms of responsiveness to auditory alarms, including environments where human and artificial systems share control (Janssen et al., 2019; Van der Heiden et al., 2018).

5. Material & methods

5.1. *Participants*

Sixty participants (15 females; $M_{age} = 23.1$ years old, $SD = 3.6$) from the University of Toulouse participated in the present study. They were all knowledgeable in the field of aeronautic and had previous experience with flight simulators. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. They received no financial compensation for their participation in the study. They were equally distributed among three different groups, performing respectively the 1-tone, 2-tone, or 3-tone paradigm. The sample size per group was determined from previous studies (Dehais, Roy, et al., 2019; Giraudet, St-Louis, et al., 2015; Katayama and Polich, 1996).

5.2. *Ethics statement*

The study was conducted in accordance with the Declaration of Helsinki (1973, revised in 1983) and was approved by a national ethics committee (CEEI/IRB00003888). After being informed of their rights, all participants gave their written consent.

5.3. *Tasks*

Two different tasks were used in this study: an *aviation task* and a *tone* *detection paradigm*. Our main measure of interest was the performance in the tone detection paradigm, which was compared across three levels of mental load. In the *no-load* condition, participants performed the tone detection paradigm only, thus mental load was minimal. The aviation task was displayed, but participants had to entirely ignore it and to fixate a green cross that was displayed at the center of the screen. In the *low-* and *high-load* conditions, participants performed the tone detection paradigm while concurrently performing the aviation task in which the mental load was low or high, respectively. Participants performed a total of 600 trials, with 200 trials for each level of mental load (no–, low- and high- load). Low and high load conditions were performed first in random order, and the no load condition was always performed after. Within each condition, trials order was counterbalanced. The aviation task was displayed on a 24′ computer screen and the tones were played via speakers placed on both sides of the screen. Both tasks were programmed in Matlab, using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli and Vision, 1997). Mental load was a within-participant factor and the tone detection paradigm condition (i.e., number of tones) a between-participants factor (i.e., three different groups).

5.3.1. *Aviation task*

This task was designed to reproduce a decision-making task performed by pilots during the landing phase, but it can be performed by non-pilot subjects. The participants were presented with short video clips displaying a Primary Flight Display (PFD), an instrument located in front of pilots in aircraft cockpits and where the main flight parameters are displayed. Compared to an actual cockpit display, the PFD was simplified (various parameters were omitted). The information available on this simplified PFD included the aircraft's vertical and horizontal position relative to the landing field (represented by two moving cursors; one on a vertical axis, on the right of the artificial horizon, and the other on a horizontal axis, below the artificial horizon, see Fig. 5). In addition, three indicators were displayed on the top left of the screen: the heading ("Cap", optimal value = "180"), wind speed ("Vent", optimal value = "0.00"), and magnetic deviation (displayed below the heading, optimal value = "0"). The magnetic deviation is particular, its value must be subtracted (when negative) or added (when positive) to the heading value. The three indicators were static during the video, only the two cursors were moving.

Each trial consisted of a 2- to 4.5-second video clip, followed by a 2-second response time window during which participants had to respond by pressing a corresponding button (see Fig. 5: right (accept landing) or left (abort landing). The landing had to be accepted in 50 % of the trials. At the onset of the 2-second response window, the moving cursors were frozen, and the artificial horizon within the PFD disappeared, thereby prompting participants to respond. Participants were asked to respond based on the information displayed during this response window (see below).

In low mental load trials, participants were asked to make their decision based solely on the final position of the two cursors marking the aircraft vertical and horizontal position. The three indicators located on the upper left corner, which participants were instructed to ignore, appeared in green and their values deviated by less than five units from their optimal ones. The landing was considered possible when the deviation of each cursor was in a [−2, +2] interval on the arbitrary scale (see Fig. 5). In the high load trials, the cursors moved two times faster than in the low load trials, and the three indicator values appeared in red, which indicated a degradation of the aircraft situation. In this case, participants had to make their decisions based on both the cursors and the indicators, checking whether heading and wind values deviated by more than 5 units from their optimal value (acceptable range for wind: 0–5; acceptable range for heading: 175–185). An illustration of the low- and high-load trials as well as the rules to make a decision are presented in Fig. 6.

To increase motivation and engagement in the task, the participants' score (number of correct decisions) was displayed on the upper right corner of the PFD (see Fig. 6). For each correct landing response, the score increased by one point. When participants missed a response, or made an incorrect response, this score remained unchanged.

5.3.2. *1-, 2-, and 3-tone detection paradigms*

Depending on the group they were assigned to, participants performed the 1-, 2, or 3-tone version of the tone detection paradigm. An illustrative video of the 3-tone paradigm is available in supplementary material S5. During each 2 to 4.5-seconds video, one 50 dB (SPL) tone lasting 100 ms was randomly played somewhere between 500 ms after the onset of the video and 500 ms before the end of the video (see Fig. 5). Note that only one tone type was presented during each video, and the term auditory load is used in the present study to describe the difficulty to identify a target tone, while other very similar tones can be also played. In each trial, the auditory stimulus could be a standard tone (sound frequency = 1900 Hz), a target tone (sound frequency = 1970 Hz), or a deviant tone (sound frequency = 2040 Hz), depending on the

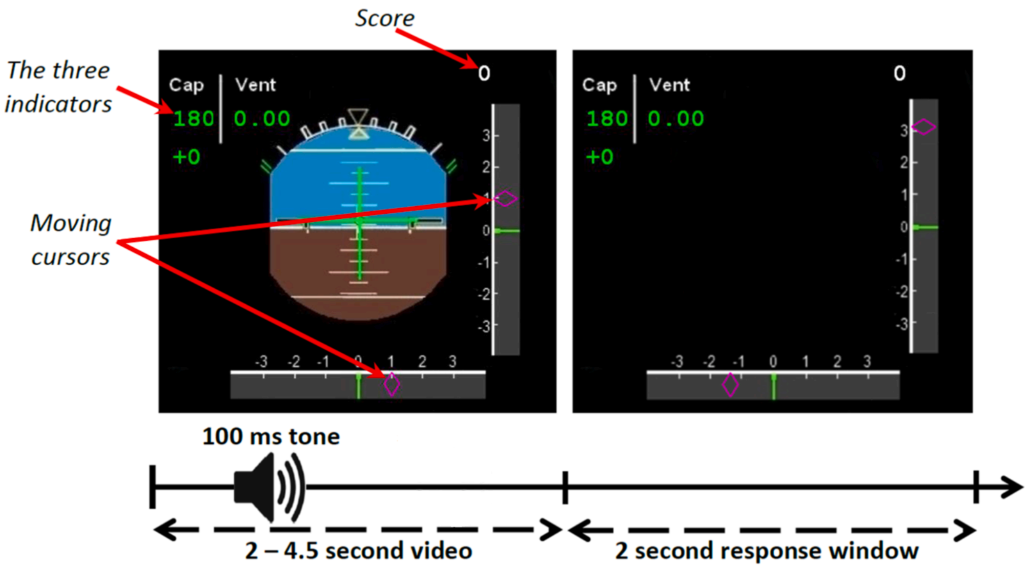
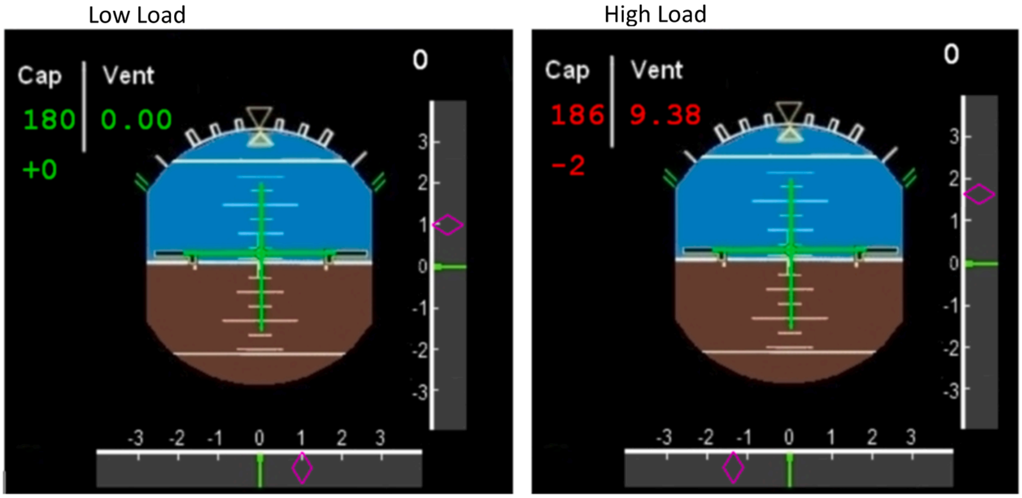


Fig. 5. Illustration of a trial in the low load condition. A 2 to 4.5-second video was displayed during which a tone was played, followed by a 2-second response window during which the participant responded to the landing task, and then pressed again a button if he/she detected a target tone.



Difficulty	Wind and heading indicators difference from optimal values	Cursors positions to accept the landing
Low load	Both < 5	Both cursors in a [-2, 2] interval
High load	One < 5 and one > 5	One cursor in a [-2, 2] interval and one cursor in a [-1, 1] interval
	Both > 5	Both cursors in a [-1, 1] interval

Fig. 6. Top: Still capture of examples of videos used in the low- and high-load conditions (left and right pictures, respectively). Bottom: table indicating the criteria to base landing decision on (as a reminder, the landing task had to be ignored in the no load condition). In the low-load example (top left), indicator values are presented in green, indicating that they are optimal. Landing is acceptable, for the cursors marking the aircraft's horizontal and vertical positions are both within the [−2, +2] range. In the high-load example (top right), the wind value is 9.38, thereby deviating by more than 5 units from the optimal 0 value; the heading value is 184 (i.e., 186 – 2 magnetic deviation units = 184) and thus deviates by<5 units from the optimal 180 value. In this situation, with one indicator outside the acceptable range and the other within it, landing should be carried out only if one of the two cursors falls within [−2, +2] range while the other falls within a more conservative [−1, +1] range. The vertical position is deviating by less than + 2 units but the horizontal position is deviating by more than −1 unit, consequently, the landing conditions are not met and landing should be aborted. In the event that both heading and wind values are outside the acceptable range, tolerance on cursor deviation is even lower and landing should only accepted if both cursors fall within the [−1,+1] range.

tone detection paradigm performed by the participant. The pitch of the tones was inspired by a study of Kolev et al. (1997). In the 1-tone paradigm, only targets tones were played (n = 60). In the 2-tone paradigm, the tones were targets (n = 60; probability = 0.30) or standards (n = 140; probability = 0.70). In the 3-tone paradigm, the tones were targets (n = 40; probability = 0.20), deviants (n = 20; probability = 0.10), or standards (n = 140; probability = 0.70). The number of target tones was the same across the three levels of mental load (e.g., a participant that performed the 1-tone paradigm was administered with 60 target tones in the no mental load condition, 60 target tones in the low mental load condition, and 60 target tones in the high mental load condition). The pitch of the target tone was maintained (i.e., 1970 Hz) across the three tone detection paradigms to avoid introducing a bias that could have influenced ERP responses. We also maintained the same probability (i.e., 0.70) for the standard tones across the 2- and the 3-tone paradigms. During the response window, participants were instructed to respond to the aviation task first (except in the no load condition during which the aviation task was presented but had to be ignored), and then to report target tones, when detected, by pressing the central button on the response box. They were asked to ignore the standard tones (in the 2- and 3-tone paradigms) and the deviant tones (in the 3- tone paradigm). To minimize motion artifacts on the EEG signal, we separated the video clip period, during which tones were played and ERP analysis performed, from the 2-s response time window. The performance to the alarm detection task was measured with the percentage of miss and false alarms (all performance data are available in supplementary material S1).

5.4. *Procedure*

Participants were seated in a comfortable armchair in a sound-dampened room. They completed the consent form and were instructed to keep their forearms stable on the chair's arms, with their two hands resting on the response box (Cedrus RB-740). Participants were introduced with the tone(s) corresponding to their tone detection paradigm, and we ensured that they were able to recognize it/them. Then, participants underwent a training consisting of 10 trials of the no mental load, low mental load, and high mental load conditions. They performed the tone-detection task as during the actual experiment. Following this training, participants were fitted with the EEG electrode cap as well as the Electro-OculoGraphic (EOG) electrodes (see description below). They then performed the low mental load and high mental load conditions (random order, 40 min), during which they had to perform the aviation task and the tone detection paradigm concurrently. Participants were instructed to give equal importance to the aviation task and tone detection paradigm. Next, they performed the no mental load condition (20 min) in which they performed the tone detection paradigm only (while ignoring the aviation task). The whole procedure lasted approximately 100 min, including 20 min to install the EEG (see Fig. 7).

5.5. *EEG data acquisition and processing*

Electroencephalographical (EEG) data were recorded with a BioSemi ActiveTwo system (https://www.biosemi.com) from 64 Ag/AgCl active electrodes (Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, C1, C3, C5, T7, TP7, CP5, CP3, CP1, *P*1, *P*3, P5, P7, P9, PO7, PO3, O1, Iz, Oz, POz, Pz, CPz, Fpz, Fp2, AF8, AF4, AFz, Fz, F2, F4, F6, F8, FT8, FC6, FC4, FC2, FCz, Cz, *C*2, C4, C6, T8, TP8, CP6, CP4, CP2, *P*2, *P*4, *P*6, P8, P10, PO8, PO4, and O2) mounted on a cap and placed on the scalp according to the international 10–20 system, plus two sites below each eye to monitor eye movements. Two additional electrodes placed close to Cz —

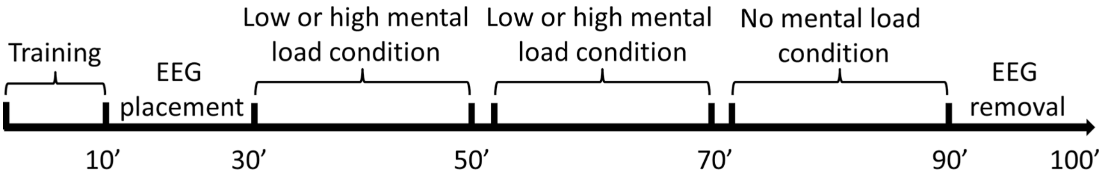


Fig. 7. The procedure timeline.

the common mode sense (CMS) active electrode and the driven right leg passive electrode – were used to drive the participants’ average potential as close as possible to the AD-box reference potential (Metting Van Rijn et al., 1990). Electrode impedance was kept below 5 kΩ for scalp electrodes and below 10 kΩ for the four eye channels. Skin–electrode contact, obtained using conductive gel, was monitored, keeping voltage offset from the CMS below 25 mV for each measurement site.

All the signals were DC-amplified and digitized continuously at a sampling rate of 512 Hz, using an anti-aliasing filter (fifth-order sinc filter) with a 3-dB point at 104 Hz. No high-pass filtering was applied online. Data were analyzed with the EEGLAB toolbox (Delorme and Makeig, 2004). EEG data were re-referenced offline to the average activity of the two mastoids and bandpass filtered (0.1–40 Hz, 12 dB/octave), given that the low-pass filter was not effective in completely removing the 50 Hz artifact for some participants. An independent component analysis was performed to isolate eye-blinks and movements related artifacts that were removed from the signal. A visual inspection of the data was performed to reject residual artifacts intervals. Data were segmented into epochs from −200 to 800 ms that were time-locked to the onset of the target tone. A 200 ms pre-stimulus baseline was used in all analyses. Data with excessive blinks were adaptively corrected using independent component analysis. Segments including artifacts (e.g., excessive muscle activity) were eliminated offline before data averaging. A total of 15 % of data were lost due to artifacts. Five participants were removed from the analysis due to the poor quality of the EEG signal. In the end, 19 participants, 18 participants and 18 participants were respectively tested in the 1-tone, 2-tone and 3-tone variants of the tone detection paradigm. In order to investigate different processing stages of the target sounds, we focused on the N100 and the P3b responses.

CRediT authorship contribution statement

Mickael Causse: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Supervision, Resources. **Fabrice B.R. Parmentier:** Writing – review & editing. **Damien Mouratille:** Software, Data curation, Formal analysis. **Dorothée Thibaut:** Formal analysis. **Marie Kisselenko:** Formal analysis. **Eve Fabre:** Methodology, Data curation, Formal analysis, Writing – original draft, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement.

The datasets generated for this study are available on request to the corresponding authors.

Acknowledgments

Fabrice Parmentier was supported by research grant PID2020-114117 GB-I00 funded by the Spanish Ministry of Science and Innovation (MICINN), the Spanish State Agency for Research (AEI) and the European Regional Development Fund (ERDF) “A Way of Making Europe”. The authors would like to thank Johanna Clerc for her help with data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brainres.2022.148035.

References

Association for the Advancement of Medical Instrumentation. (2011). A siren call to action: Priority issues from the medical device alarms summit. *Clinical Alarms AAMI Summit Proceedings*, 5–6.

Berti, S., Schröger, E., 2003. Working memory controls involuntary attention switching: Evidence from an auditory distraction paradigm. Eur. J. Neurosci. 17 (5), 1119–1122.

Bliss, J.P., Acton, S.A., 2003. Alarm mistrust in automobiles: How collision alarm reliability affects driving. Appl. Ergon. 34 (6), 499–509.

Bliss, J.P., Dunn, M.C., 2000. Behavioural implications of alarm mistrust as a function of task workload. Ergonomics 43 (9), 1283–1300. https://doi.org/10.1080/001401300421743.

Bliss, J.P., Gilson, R.D., Deaton, J.E., 1995. Human probability matching behaviour in response to alarms of varying reliability. Ergonomics 38 (11), 2300–2312. https://doi.org/10.1080/00140139508925269.

Bliss, J.P., Freeland, M.J., Millard, J.C., 1999. Alarm related incidents in aviation: A survey of the aviation safety reporting system database. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 43 (1), 6–10.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10 (4), 433–436.

Brázdil, M., Rektor, I., Daniel, P., Dufek, M., Jurák, P., 2001. Intracerebral event-related potentials to subthreshold target stimuli. Clin. Neurophysiol. 112 (4), 650–661.

Callan, D.E., Gateau, T., Durantin, G., Gonthier, N., Dehais, F., 2018. Disruption in neural phase synchrony is related to identification of inattentional deafness in real-world setting. Hum. Brain Mapp. 39 (6), 2596–2608. https://doi.org/10.1002/hbm.24026.

Cass, M., Polich, J., 1997. P300 from a single-stimulus paradigm: Auditory intensity and tone frequency effects. Biol. Psychol. 46 (1), 51–65. https://doi.org/10.1016/S0301-0511(96)05233-7.

Causse, M., Imbert, J.-P., Giraudet, L., Jouffrais, C., Tremblay, S., 2016a. The Role of Cognitive and Perceptual Loads in Inattentional Deafness. Front. Hum. Neurosci. 10, 344. https://doi.org/10.3389/fnhum.2016.00344.

Causse, M., Peysakhovich, V., Fabre, E.F., 2016b. High working memory load impairs language processing during a simulated piloting task: An ERP and pupillometry study. Front. Hum. Neurosci. 10, 240. https://doi.org/10.3389/fnhum.2016.00240.

Corbetta, M., Kincade, J.M., Ollinger, J.M., McAvoy, M.P., Shulman, G.L., 2000. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. Nat. Neurosci. 3 (3), 292–297. https://doi.org/10.1038/73009.

Courchesne, E., Kilman, B.A., Galambos, R., Lincoln, A.J., 1984. Autism: Processing of novel auditory information assessed by event-related brain potentials. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 59 (3), 238–248. https://doi.org/10.1016/0168-5597(84)90063-7.

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., Tremblay, S., 2014. Failure to Detect Critical Auditory Alerts in the Cockpit Evidence for Inattentional Deafness. *Human Factors: The Journal of the Human Factors and Ergonomics*. Society. https://doi.org/10.1177/0018720813510735.

Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R.N., 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. https://doi.org/10.3390/s19061324.

Dehais, F., Roy, R. N., Gateau, T., & Scannella, S. (2016). *Auditory alarm misperception in the cockpit: An EEG study of inattentional deafness*. 177–187.

Dehais, F., Roy, R.N., Scannella, S., 2019b. Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. Behav. Brain Res. 360, 51–59. https://doi.org/10.1016/j.bbr.2018.11.045.

Donchin, E., Coles, M.G., 1988. Is the P300 component a manifestation of context updating. Behav. Brain Sci. 11 (3), 357–427. https://doi.org/10.1017/S0140525X00058027.

Ebmeier, K.P., Steele, J.D., MacKenzie, D., O’Carroll, R., Kydd, R.R., Glabus, M.F., Blackwood, D., Rugg, M.D., Goodwin, G.M., 1995. Cognitive brain potentials and regional cerebral blood flow equivalents during two-and three-sound auditory "oddball tasks". Electroencephalogr. Clin. Neurophysiol. 95 (6), 434–443. https://doi.org/10.1016/0013-4694(95)00173-5.

Fabre, E. F., Peysakhovich, V., & Causse, M. (2017). Measuring the Amplitude of the N100 Component to Predict the Occurrence of the Inattentional Deafness Phenomenon. In *Advances in Neuroergonomics and Cognitive Engineering* (pp. 77–84). Springer.

Ferraro, J.C., Mouloua, M., 2021. Effects of automation reliability on error detection and attention to auditory stimuli in a multi-tasking environment. Appl. Ergon. 91, 103303 https://doi.org/10.1016/j.apergo.2020.103303.

Foley, L., Anderson, C. J., & Schutz, M. (2020). *Re-Sounding Alarms: Designing Ergonomic Auditory Interfaces by Embracing Musical Insights*. *8*(4), 389. https://doi.org/10.3390/healthcare8040389.

Frühholz, S., Godde, B., Finke, M., Herrmann, M., 2011. Spatio-temporal brain dynamics in a combined stimulus–stimulus and stimulus–response conflict task. NeuroImage 54 (1), 622–634. https://doi.org/10.1016/j.neuroimage.2010.07.071.

Garcia-Garcia, M., Yordanova, J., Kolev, V., Domínguez-Borràs, J., Escera, C., 2010. Tuning the brain for novelty detection under emotional threat: The role of increasing gamma phase-synchronization. NeuroImage 49 (1), 1038–1044. https://doi.org/10.1016/j.neuroimage.2009.07.059.

Ghani, U., Signal, N., Niazi, I., Taylor, D., 2020. ERP based measures of cognitive workload: A review. Neurosci. Biobehav. Rev.

Gibson, Z., Butterfield, J., Rodger, M., Murphy, B., Marzano, A., 2019. Use of Dry Electrode Electroencephalography (EEG) to Monitor Pilot Workload and Distraction Based on P300 Responses to an Auditory Oddball Task. In: Ayaz, H., Mazur, L. (Eds.), Advances in Neuroergonomics and Cognitive Engineering. Springer International Publishing, pp. 14–26.

Giraudet, L., Saint-Louis, M., & Causse, M. (2012). Electrophysiological correlates of inattentional deafness: No hearing without listening. *Human Factors: A View from an Integrative Perspective*. Human Factors and Ergonomics Society Europe Chapter, Toulouse, France.

Giraudet, L., Imbert, J.-P., Bérenger, M., Tremblay, S., Causse, M., 2015a. The neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and EEG/ERP measures. Behav. Brain Res. 294, 246–253. https://doi.org/10.1016/j.bbr.2015.07.041.

Giraudet, L., St-Louis, M.-E., Scannella, S., Causse, M., 2015b. P300 event-related potential as an indicator of inattentional deafness? PLoS ONE 10, e0118556. https://doi.org/10.1371/journal.pone.0118556.

Gouraud, J., Delorme, A., Berberian, B., 2017. Autopilot, mind wandering, and the out of the loop performance problem. Front. Neurosci. 11, 541. https://doi.org/10.3389/fnins.2017.00541.

Grillon, C., Courchesne, E., Ameli, R., Elmasian, R., Braff, D., 1990. Effects of rare non-target stimuli on brain electrophysiological activity and performance. Int. J. Psychophysiol. 9 (3), 257–267. https://doi.org/10.1016/0167-8760(90)90058-L.

Guillaume, A. (2011). Intelligent auditory alarms. In *The sonification handbook* (T. Hermann, A. Hunt, J.G. Neuhoff, pp. 493–508).

Ho, C., Spence, C., 2005. Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. J. Exp. Psychol. Appl. 11 (3), 157.

Jackson, M.C., Linden, D.E., Roberts, M.V., Kriegeskorte, N., Haenschel, C., 2015. Similarity, not complexity, determines visual working memory performance. J. Exp. Psychol. Learn. Mem. Cogn. 41 (6), 1884. https://doi.org/10.1037/xlm0000125.

Janssen, C.P., van der Heiden, R.M.A., Donker, S.F., Kenemans, J.L., 2019. Measuring Susceptibility to Alerts While Encountering Mental Workload. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings, pp. 415–420. https://doi.org/10.1145/3349263.3351524.

Johnson, R., 1993. On the neural generators of the P300 component of the event-related potential. Psychophysiology 30 (1), 90–97. https://doi.org/10.1111/j.1469-8986.1993.tb03208.x.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Katayama, J., Polich, J., 1996. P300 from one-, two-, and three-stimulus auditory paradigms. Int. J. Psychophysiol. 23 (1–2), 33–40. https://doi.org/10.1016/0167-8760(96)00030-X.

Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3? Perception 36 ECVP Abstract Supplement. PLoS ONE.

Knight, R.T., 1996. Contribution of human hippocampal region to novelty detection. Nature 383 (6597), 256–259.

Kolev, V., Demiralp, T., Yordanova, J., Ademoglu, A., Isoglu-Alkaç, Ü., 1997. Time–frequency analysis reveals multiple functional components during oddball P300. NeuroReport 8 (8), 2061–2065.

Koreimann, S., Gula, B., Vitouch, O., 2014. Inattentional deafness in music. Psychol. Res. 78 (3), 304–312.

Kramer, A.F., Sirevaag, E.J., Braune, R., 1987. A psychophysiological assessment of operator workload during simulated flight missions. Human Factors: The Journal of the Human Factors and Ergonomics Society 29 (2), 145–160. https://doi.org/10.1177/001872088702900203.

Kramer, A.F., Trejo, L.J., Humphrey, D., 1995. Assessment of mental workload with task-irrelevant auditory probes. Biol. Psychol. 40 (1), 83–100. https://doi.org/10.1016/0301-0511(95)05108-2.

Kreitz, C., Furley, P., Simons, D.J., Memmert, D., 2016. Does working memory capacity predict cross-modally induced failures of awareness? Conscious. Cogn. 39, 18–27. https://doi.org/10.1016/j.concog.2015.11.010.

Lacherez, P., Limin Seah, E., Sanderson, P., 2007. Overlapping melodic alarms are almost indiscriminable. Hum. Factors 49 (4), 637–645. https://doi.org/10.1518/001872007X215719.

Lansdowne, K., Strauss, D.G., Scully, C.G., 2016. Retrospective analysis of pulse oximeter alarm settings in an intensive care unit patient population. BMC Nursing 15 (1), 36. https://doi.org/10.1186/s12912-016-0149-3.

Lv, J.-Y., Wang, T., Qiu, J., Feng, S.-H., Tu, S., Wei, D.-T., 2010. The electrophysiological effect of working memory load on involuntary attention in an auditory–visual distraction paradigm: An ERP study. Exp. Brain Res. 205 (1), 81–86. https://doi.org/10.1007/s00221-010-2360-x.

Macdonald, J.S., Lavie, N., 2011. Visual perceptual load induces inattentional deafness. Atten. Percept. Psychophys. 73 (6), 1780–1789.

Miller, M.W., Rietschel, J.C., McDonald, C.G., Hatfield, B.D., 2011. A novel approach to the physiological measurement of mental workload. Int. J. Psychophysiol. 80 (1), 75–78. https://doi.org/10.1016/j.ijpsycho.2011.02.003.

Morris, R., Montano, S., 1996. Response times to visual and auditory alarms during anaesthesia. Anaesth. Intensive Care 24 (6), 682–684. https://doi.org/10.1177/0310057X9602400609.

Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. 42 (4), 313–329. https://doi.org/10.1016/0001-6918(78)90006-9.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology 24 (4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x.

NTSB. (1973). *Aircraft accident report: Eastern Air Lines, Inc* (NTSB/AAR-73/14); Issue NTSB/AAR-73/14)). http://libraryonline.erau.edu/online-full-text/ntsb/aircraft-accidentreports/AAR73-14.pdf.

Parmentier, F., 2014. The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. Psychol. Res. 78 (3), 321–338. https://doi.org/10.1007/s00426-013-0534-4.

Pelli, D.G., Vision, S., 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat. Vis. 10, 437–442.

Picton, T.W., 1992. The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9 (4), 456–479. https://doi.org/10.1097/00004691-199210000-00002.

Polich, J., 2003. Theoretical overview of P3a and P3b. Detection of Change 83–98.

Polich, J., 2007. Updating P300: An integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148.

Polich, J., Criado, J.R., 2006. Neuropsychology and neuropharmacology of P3a and P3b. Int. J. Psychophysiol. 60 (2), 172–185. https://doi.org/10.1016/j.ijpsycho.2005.12.012.

Polich, J., Eischen, S.E., Collins, G.E., 1994. P300 from a single auditory stimulus. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 92 (3), 253–261. https://doi.org/10.1016/0168-5597(94)90068-X.

Polich, J., Heine, M.R., 1996. P300 topography and modality effects from a single-stimulus paradigm. Psychophysiology 33 (6), 747–752.

Polich, J., Margala, C., 1997. P300 and probability: Comparison of oddball and single-stimulus paradigms. Int. J. Psychophysiol. 25 (2), 169–176.

Potnuru, P., Epstein, R.H., McNeer, R., Bennett, C., 2020. Development and Validation of an Algorithm for the Identification of Audible Medical Alarms. Cureus 12 (11). https://doi.org/10.7759/cureus.11549.

Raabe, M., Rutschmann, R.M., Schrauf, M., & Greenlee, M. W. (2005). Neural correlates of simulated driving: Auditory oddball responses dependent on workload. In *Foundations of Augmented Cognition* (D. D. Schmorrow (Mahwah, NJ: Lawrence Erlbaum Associates, Inc.), pp. 1067–1076).

Raveh, D., Lavie, N., 2015. Load-induced inattentional deafness. Atten. Percept. Psychophys. 77 (2), 483–492.

Salas, B., & Schlesinger, J. J. (2019). *Enhancing the Sounds of Urgency: Lessons From Music and Aviation on Human Response to Critical Auditory Stimuli*.

SanMiguel, I., Corral, M.-J., Escera, C., 2008. When loading working memory reduces distraction: Behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. J. Cognit. Neurosci. 20 (7), 1131–1145. https://doi.org/10.1162/jocn.2008.20078.

Scheer, M., Bülthoff, H.H., Chuang, L.L., 2018. Auditory Task Irrelevance: A Basis for Inattentional Deafness. Hum. Factors 60 (3), 428–440. https://doi.org/10.1177/0018720818760919.

Snyder, E., Hillyard, S.A., 1976. Long-latency evoked potentials to irrelevant, deviant stimuli. Behavioral Biology 16 (3), 319–331.

Sokolov, E.N., 1963. Higher Nervous Functions: The Orienting Reflex. Annu. Rev. Physiol. 25 (1), 545–580. https://doi.org/10.1146/annurev.ph.25.030163.002553.

Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalogr. Clin. Neurophysiol. 38 (4), 387–401.

Swerdloff, M.M., Hargrove, L.J., 2020. Quantifying Cognitive Load using EEG during Ambulation and Postural Tasks. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 2849–2852. https://doi.org/10.1109/EMBC44109.2020.9176264.

Tellinghuisen, D.J., Cohen, A.J., Cooper, N.J., 2016. Now hear this: Inattentional deafness depends on task relatedness. Atten. Percept. Psychophys. 78 (8), 2527–2546.

Tombu, M.N., Asplund, C.L., Dux, P.E., Godwin, D., Martin, J.W., Marois, R., 2011. A unified attentional bottleneck in the human brain. Proc. Natl. Acad. Sci. 108 (33), 13426–13431.

Ullsperger, P., Freude, G., Erdmann, U., 2001. Auditory probe sensitivity to mental workload changes–an event-related potential study. Int. J. Psychophysiol. 40 (3), 201–209.

Valéry, B., Scannella, S., Peysakhovich, V., Barone, P., Causse, M., 2017. Can an aircraft be piloted via sonification with an acceptable attentional cost? A comparison of blind and sighted pilots. Appl. Ergon. 62, 227–236. https://doi.org/10.1016/j.apergo.2017.03.001.

Van der Heiden, R.M., Janssen, C.P., Donker, S.F., Hardeman, L.E., Mans, K., Kenemans, J.L., 2018. Susceptibility to audio signals during autonomous driving. PLoS ONE 13 (8), e0201963.

Van Hedger, S.C., Heald, S.L., Nusbaum, H.C., 2018. Long-term pitch memory for music recordings is related to auditory working memory precision. Q. J. Exp. Psychol. 71 (4), 879–891.

Verleger, R., Jaśkowski, P., Wascher, E., 2005. Evidence for an integrative role of P3b in linking reaction to perception. J. Psychophysiol. 19 (3), 165–181.

Vitevitch, M.S., 2003. Change deafness: The inability to detect changes between two voices. J. Exp. Psychol. Hum. Percept. Perform. 29 (2), 333.

Wester, A.E., Böcker, K.B.E., Volkerts, E.R., Verster, J.C., Kenemans, J.L., 2008. Event-related potentials and secondary task performance during simulated driving. Accid. Anal. Prev. 40 (1), 1–7. https://doi.org/10.1016/j.aap.2007.02.014.

Winkler, I., 2003. Change Detection in Complex Auditory Environment: Beyond the Oddball Paradigm. In: Polich, J. (Ed.), Detection of Change: Event-Related Potential and fMRI Findings. Springer, US, pp. 61–81. https://doi.org/10.1007/978-1-4615-0294-4_4.

Wolpaw, J.R., Penry, J.K., 1975. A temporal component of the auditory evoked response. Electroencephalogr. Clin. Neurophysiol. 39 (6), 609–620.

Xiao, Y., Seagull, F.J., Nieves-Khouw, F., Barczak, N., Perkins, S., 2003. Why are They not Responding to our Alarms? Another Analysis of Problems with Alarms. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 47 (3), 386–390. https://doi.org/10.1177/154193120304700329.

Zhu, R., Wang, Z., Ma, X., You, X., 2022. High expectancy influences the role of cognitive load in inattentional deafness during landing decision-making. Appl. Ergon. 99, 103629 https://doi.org/10.1016/j.apergo.2021.103629.

Zouridakis, G., Simos, P.G., Papanicolaou, A.C., 1998. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. Brain Topogr. 10 (3), 183–189.

Edworthy, J., 2013. Medical audible alarms: a review. J. Am. Med. Inf. Assoc. 20 (3), 584–589.

Konkani, A., Oakley, B., Bauld, T.J., 2012. Reducing hospital noise: a review of medical device alarm management. Biomedical Instrumentation & Technology 46 (6), 478–487.

Fitch, W.T., Kramer, G., 1992. Sonifying the body electric: Superiority of anauditory over a visual display in a complex, multivariate system. In: Kramer, G. (Ed.), Auditory Display, sonification, audification and auditory interfaces. TheProceedings of the First International Conference on Auditory Display. Addison-Wesley, Santa FéInstitute, Santa Fé, pp. 307–326.

Stevens, C.J., Brennan, D., Petocz, A., Howell, C., 2009. Designing informative warning signals: Effects of indicator type, modality, and task demand on recognition speed and accuracy. Advances in Cognitive Psychology 5, 84.

Metting Van Rijn, A.C., Peper, A., Grimbergen, C.A., 1990. High-quality recording of biolectric events: Part 1 Interference reduction, theory and practice. Med. Biol. Eng. Comput. 28, 389–397. https://doi.org/10.1007/BF02441961.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21.