Current Research in Behavioral Sciences 11 (2026) 100223

Contents lists available at ScienceDirect

Current Research in Behavioral Sciences

journal homepage: www.elsevier.com/locate/crbeha

Windows into the mind: From neurophysiological to multimodal measurements in the cockpit

Anneke Hamann *, Nils Carstengerdes

Institut für Flugführung, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Lilienthalplatz 7, Braunschweig, 38108, Germany

ARTICLE INFO

Keywords:
Aviation
Mental workload
Operator state monitoring
Neurophysiology
EEG
fNIRS

ABSTRACT

Increasing automation in modern aircraft is intended to alleviate the physical and cognitive demands put on the pilots and thereby to reduce human error. With new technologies available, it might be possible to design adaptive assistance systems that offer support tailored to the situation and the pilots' needs. In order to do so, however, the systems need information about the pilots' state and situation. Neurophysiological measurements like electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) seem a powerful tool for enabling adaptive automation by monitoring pilots' cognitive states. Such measurements need to be reliable and valid in order to gain interpretable and robust results. Our research, however, suggests that this may not be as easy as it seems. In this paper, we give a short overview of the methodological approach of our research and results, identify problems and research gaps concerning the validity and reliability of neurophysiological assessments and discuss possible mitigations to make these measurements the robust and dependable sources of information we wish for.

1. Introduction: aviation, automation, adaptation

The history of aviation is one of rapid increases in technological advances and ever higher levels of automation (Chartered Institute of Ergonomics and Human Factors [CIEHF], 2020). These developments go hand in hand with decreasing numbers of crew in the cockpit, following the assumption that higher levels of automation result in tasks that can be achieved by fewer human pilots (Billings, 1991; Kyle et al., 2025). Historically, the task allocation between human and machine has followed simple guidelines based on the capabilities of the actors (Fitts, 1951), resulting in a static allocation. Later on, dynamic allocations arose in the form of adaptable automation which the humans can changes based on their current needs, or adaptive automation in which the task allocation is changed by the machine, triggered by pre-defined situational, environmental or task-related factors (Parasuraman et al., 1992; Scerbo, 1996). With the emergence of more sophisticated technologies such as machine learning algorithms, the aviation industry and research alike now hope to develop concepts of adaptive automation that are tailored to the pilots and their current needs, and are thus more flexible and more supportive than traditional systems (CIEHF, 2020; Kyle et al., 2025).

Since flying an aircraft is a cognitively demanding task that requires extensive and repeated specialised training (Feary, 2018; Kearns et al., 2016; Zarlengo, 2025), a sudden or unforeseen decline in the pilots' cognitive capabilities or visual and motor skills can lead to dangerous situations, often with fatal consequences. In traditional two-crew cockpits, the second pilot provides a redundancy and safety net, albeit an imperfect one (Comisión de Investigación de Accidentes e Incidentes de Aviación Civil [CIAIAC], 2024). In new concepts of Reduced Crew Operations, i.e. one pilot remaining in control during low-load segments of the flight while the other(s) rest, or Single Pilot Operations, i.e. one pilot in charge of the whole flight, this safety net will thin out or disappear. Thus, the need arises to provide the pilots with assistance provided by automation. To date, there is no assistance system in aircraft that adapts to pilots' states and needs. If adaptation is done, it is based on flight parameters, such as the Airbus autopilot philosophy in which the autopilot switches modes depending on which flight phase (climb, cruise, descend) is detected. In modern cars, there are drowsy driving warning systems that alert the driver if deviations from normative driving is detected, but those systems are usually based on steering behaviour (Albadawi et al., 2022). In highly automated contexts like aircraft, where only few manual inputs to the aircraft are expected during cruise flight, an adaptation solely based on behaviour would not work well. Moreover, there are more factors that may induce the need

* Corresponding author.
E-mail address: anneke.hamann@dlr.de (A. Hamann).

https://doi.org/10.1016/j.crbeha.2026.100223
Received 17 March 2026; Received in revised form 10 July 2026; Accepted 29 July 2026
Available online 31 July 2026

for further assistance, such as medical emergencies in the cockpit (Evans and Radcliffe, 2012), startle and surprise (Landman et al., 2017), spatial disorientation (Dixon et al., 2025), boredom and attention lapses (Trampitsch-Vink et al., 2025) or high task demands resulting in high mental workload and stress (Masi et al., 2023). Depending on the actual situation, assistance may need to look very differently: For full pilot incapacitation events, automation should be able to take over completely and land the aircraft safely at the next suitable airport. In less severe cases, the pilots may need assistance in assessing the situation and aircraft status, building up situation awareness again, need guidance in difficult decision-making processes like which alternate airport to choose for landing in sub-optimal conditions. The functionalities of the onboard assistance system can thus vary from fully autonomous landings (Fuchs et al., 2026) to decision aides (Djartov and Würfel, 2025) to simple auditory and visual alerting and highlighting of critical displays and information that have been missed (Lounis et al., 2020; Riedinger et al., 2026). Ideally, an assistance system will offer multiple functionalities that are triggered adaptively, tailored to the specific circumstances of the situation.

In order to do so, however, the system needs information about the pilots, their current state and needs. In their quest for detailed information about the human operator, researchers have turned to physiological measurements (Borghini et al., 2014; Boumann et al., 2023; Causse, Deniel, et al., 2025). Peripheral measurements like electrocardiography or electrodermal activity are often used to indicate stress or high arousal (Borghini et al., 2014). Neurophysiological measurements like electroencephalography (EEG) are used to gain more detailed insights into the humans' brain activity and cognitive concepts like mental workload or mental fatigue (Borghini et al., 2014). With the emergence of wearable functional near-infrared spectroscopy (fNIRS) devices, this interest in neurophysiological measurements has peaked again, indicated by a growing body of research in the last years (van Weelden et al., 2022). EEG and fNIRS are promising measurements for the use in the cockpit because of their low intrusiveness, possibility to gain data continuously, and of course their objectivity. Thus, they have huge advantages compared to self-report or performance data. Yet, there is one question that needs to be answered before sophisticated adaptive assistance systems should be fed with (neuro-) physiological data: How precise are the information we can gain from said measurements?

2. Mental workload and the validity of EEG measures

To give an example and highlight the problems with valid neurophysiological measurements, in this paper[1] we will focus on one concept: Mental workload. Mental workload is one of the most frequently researched concepts in aviation, despite controversies about its definition and conceptualisation (Dekker and Hollnagel, 2004; Parasuraman et al., 2008). In general, mental workload is defined as the part of one's cognitive resources needed to accomplish a task (O'Donnell and Eggemeier, 1986). The more difficult the task, the more cognitive resources one needs to allocate to its accomplishment. If one has no spare resources left, task performance will decline and errors will occur. In aviation, it is generally believed that a medium level of mental workload is ideal (Martins, 2016): Too low mental workload can border on boredom and can prompt disengagement from the task, causing loss of attention and out-of-the-loop phenomena when the pilot is suddenly asked to actively engage again. On the other hand, too high mental workload can lead to declined performance and errors that could lead to catastrophic outcomes. In sum, it would be ideal to keep the pilots' mental workload at a comfortable medium level to ensure optimal performance (Ghaderi et al., 2023; Martins, 2016). Future adaptive assistance systems could help with this, if the pilots' current mental workload level could be assessed.

There is a large and growing body of research on mental workload assessment using EEG. Researchers usually focus on frequency band analyses for this, i.e. the decomposition of the EEG signal into its frequency bands. Changes in the composition of the signal are used to assess changes in cognitive states, e.g. between an easy and a difficult task. This way, data can be collected continuously and without the need to insert additional stimuli into the cockpit environment to elicit responses (as is done in analyses of event-related potentials). Indeed, the literature suggests a classical, "tell-tale" pattern of changing signal compositions for mental workload: Increasing mental workload is usually indicated by an increase in frontal theta activity, accompanied by a decrease in parietal alpha activity (Dehais et al., 2019; Dussault et al., 2004; Grissmann et al., 2017; Owen et al., 2005). This pattern likely arises from the involvement of the frontoparietal network (FPN) in the human brain. This network is formed by the prefrontal cortex (PFC) and the posterior parietal cortex (PPC). The PFC controls top-down processing for higher cognitive functions, such as decision-making, selective attention, integration of working memory functions and response inhibition (Friedman and Robbins, 2022; Lara and Wallis, 2015; Miller and Cohen, 2001; Owen et al., 2005), while the PPC is responsible for sustained attention within this network (Duncan, 2010; Posner and Dehaene, 1994; Sridharan et al., 2008). Increasing task demands, increasing effort to meet them and thus increasing mental workload seems to result in an activation of the FPN to control attention and executive functions. Yet, upon closer inspection, this "tell-tale" pattern of frontal and parietal activation with increasing mental workload cannot be found in every publication (Hamann, 2023). While this is to be expected in research, the reasons for failing to detect the expected patterns should concern us.

For example, there are other concepts that elicit very similar cognitive activity, such as mental fatigue. Mental fatigue is described as a sense of weariness usually induced by long monotonous, yet demanding tasks (Charbonnier et al., 2016; Grandjean, 1979; Hamann, van Klaren, et al., 2025). Not unlike mental workload, it is also based on the idea of resource depletion. The longer one needs to focus on a task, the more cognitive resources one spends until these resources are used up. If the individual does not or cannot take a break to replenish their cognitive resources, they will experience a subjective feeling of fatigue and a general unwillingness to spend further effort (Grandjean, 1979) as well as more negative emotions and worsened emotion regulation (Grillon et al., 2015; Rosa et al., 2021). In EEG measurements, mental fatigue usually manifests in increasing frontal theta activity, accompanied by increasing parietal alpha activity (Tran et al., 2020; Trejo et al., 2007; Wascher et al., 2014), the same regions involved in mental workload. This is not just inconvenient when one wants to differentiate the concepts. There is research indicating that confounding mental workload and mental fatigue (when a strenuous task is performed over a long period of time) decreases the accuracy of mental workload classification (Roy et al., 2013). Why? Because decreasing alpha activity with mental workload and increasing alpha activity with mental fatigue may cancel each other out and result in no detectable changes (Roy et al., 2013). And mental fatigue is not the only concept that can interfere with a mental workload assessment. Increasing parietal alpha activity is, for example, also found during episodes of mind wandering and out-of-the-loop phenomena (Benedek et al., 2014; Gouraud et al., 2017). Similarly, sleep deprivation is characterized by decreasing alpha and increasing theta activity (Ahn et al., 2016; Wilson et al., 2007).

These examples should suffice to illustrate that different cognitive states impact the same EEG frequency bands and can therefore influence an assessment. In addition, task characteristics can also interfere with the assessment. Frequent task switching, for example between

[1] The present paper constitutes an extension of an earlier version presented at the conference "Measuring Behavior 2024": Hamann and Carstengerdes (2024). A window into the mind? On usefulness and challenges of neurophysiological measurements in the cockpit. In A. Spink, G. Riedel, K. Truong, & L. Robinson (Eds.), *Proceedings of Measuring Behavior 2024, the 13th International Conference on Methods and Techniques in Behavioral Research* (pp. 27–33).

navigation and communication tasks as is common in the cockpit, also affects the alpha frequency band (Puma et al., 2018) and could be responsible for "missing" changes in parietal alpha activity (Hamann and Carstengerdes, 2022). On top of that, emotional responses, especially negative ones, impact frontal theta activity (Grissmann et al., 2017) and have the potential to interfere with the assessment as well. These examples highlight why valid mental workload assessment may theoretically be achievable in the laboratory, but becomes rather difficult rather fast when turning towards realistic tasks. An algorithm trained to spot the typical, "tell-tale" mental workload pattern from laboratory studies can be confused easily when other factors come into play in real-world settings. And yet, very few studies consider these problems and explicitly control for confounding factors or discuss limitations of their findings in this regard (Brouwer et al., 2015; Hamann, 2023).

3. A systematic approach to achieving validity

In order to bridge this gap and find valid neurophysiological measures of mental workload – or at least approach the limitations of the measures – we did a series of studies, systematically focusing on mental workload (Hamann and Carstengerdes, 2022) and mental fatigue (Hamann and Carstengerdes, 2023a) while controlling for the influence of the respective other and comparing the results (Hamann, 2023; Hamann and Carstengerdes, 2023b). In the following, we detail our approach. The described experiments were approved by the ethics commission of the German Psychological Society (DGPs) and conducted in accordance with the declaration of Helsinki.

3.1. Step one: convergent validity – combining and comparing measurement methods

Because of the aforementioned limitations of EEG measurements concerning mental workload, in a first step, we added a new data source for neurophysiological measures, fNIRS. Combining fNIRS and EEG measurements has certain advantages. The higher temporal resolution of EEG and the higher spatial resolution of fNIRS measurements complement each other and can lead to more accurate results of when and where exactly changes happen, if analysed accordingly. And there is also the possibility to analyse convergent validity: If both measurements point towards the same result, it is more likely that the finding is true and not influenced by a confounding factor. fNIRS measures the cortical activation based on changing oxygen concentrations in the cortical blood flow. Overly simplified, increasing brain activity increases the consumption of oxygen and thus the flow of oxygenated blood in the activated region, while the levels of deoxygenated blood in said region decrease simultaneously. Because of the different optical properties of oxygenated and deoxygenated blood, fNIRS can make this process visible and give an indication of the changes in cortical activation (for a more thorough explanation see Hamann, 2023). The majority of studies indicate increasing (pre-) frontal cortical oxygenation (more oxygenated, less deoxygenated blood) with increasing mental workload (Causse et al., 2017; Gateau et al., 2018; Geissler et al., 2020).

Of course, we were not the first with the idea of combining EEG and fNIRS for assessing mental workload. As early as 2009, researchers tried to combine EEG and fNIRS data to classify mental workload (Hirshfield et al., 2009). In the following years, the number of publications combining the two measurements for mental workload assessment grew and authors usually found that the combination improved the assessment (Aghajani et al., 2017; Banville et al., 2019; Liu et al., 2015, 2017; Matthews et al., 2015; Roy et al., 2019; Saadati et al., 2020). The researchers usually utilized controlled laboratory settings and working memory tasks such as the n-back task with distinct difficulty levels, and still report a very diverse range of success on identifying changes in mental workload, from 49 % (Liu et al., 2017) to > 80 % accuracy (Aghajani et al., 2017; Saadati et al., 2020). It seems that even under laboratory conditions, mental workload assessment is still far from perfect.

Moreover, all of these studies failed to control for (or failed to report) influences of mental fatigue, sleepiness or other possible influencing variables. This is problematic because fNIRS seems just as susceptible to confounding influences as EEG. The same pattern of increasing (pre-) frontal cortical activation with increasing mental workload is also found with increasing mental fatigue (Chuang et al., 2018; Nguyen et al., 2017; Nogueira et al., 2022). This could indicate that frontal cortical activation, as captured by fNIRS, is an indication of general demand, but rather unspecific. Moreover, unfortunately, most studies are performed on frontal and prefrontal regions, and there are only few published findings on parietal activation to this date (Hamann and Carstengerdes, 2023a). Thus, it is still unclear how fNIRS signals change in parietal areas, if the same pattern of FPN activation as for EEG can be expected, and if parietal activation could help disentangle mental workload from other cognitive states. In sum, while fNIRS is an interesting addition to EEG, the results gained from the measurement cannot necessarily differentiate mental workload from other types and sources of demand. This is highly problematic for the development of assistance systems based on neurophysiology. Assistance for a fatigued pilot could look very different from assistance for an overloaded (or underloaded) pilot, and a system unable to tell the difference may adapt in the wrong direction. Moreover, if both effects indeed cancel each other out (Roy et al., 2013), an overloaded and fatigued pilot may not be offered any assistance because the system could no longer detect "unwanted" patterns of activation.

3.2. Step two: internal validity – inducing a concept while controlling for confounds

It seemed that adding a data source alone would not solve the problem. If one wants to disentangle the neurophysiological correlates of mental workload from those of other concepts, one needs to make sure to induce only mental workload and control for other influences.

For example, when inducing different levels of mental workload by means of task difficulty, collecting self-reported mental workload (e.g. via NASA-TLX, Hart and Staveland, 1988; ISA, Tattersall and Foord, 1996; DLR-WAT, Brandenburger et al., 2023) and task performance can help to verify the success of the manipulation. While most studies that are undertaken use this kind of quality control (see e.g. Da Tao et al., 2025; Matthews et al., 2015, Geissler et al., 2020; Grissmann et al., 2017), there are studies that do not actually report these metrics. Unfortunately, this tendency to neglect manipulation checks seems to emerge especially in research focused on machine learning algorithms for mental workload classification (Hassan et al., 2025; Penumaka et al., 2026; Zheng et al., 2025), thus make it difficult to interpret the neurophysiological results and quality of the algorithms.

In addition to that, controlling for confounds is just as important. The literature suggests that effects of mental fatigue become visible in (and thus problematic for) EEG recordings around 45 min into performing a task (Trejo et al., 2007) (and possibly even earlier in fNIRS [Ahn et al., 2016]). To our knowledge, with the exception of Roy et al. (2013), no studies have explicitly manipulated and analysed this relationship. If one does not want to rule out confounding effects, keeping the task duration below this seems a good idea for obtaining only mental workload-associated data free from contaminations of mental fatigue. Adding self-reported mental fatigue questionnaires (e.g. via F-ISA, Hamann and Carstengerdes, 2020) to the experimental design and taking the data into account during analysis also seems wise. Moreover, controlling for participants' sleepiness by collecting information about their sleep routines prior to the experiment, by recording their self-reported sleepiness (e.g. via KSS, Akerstedt and Gillberg, 1990) and by excluding too sleepy participants from the analyses also help to achieve a clean mental workload dataset. Yet, there is a body of research that explicitly utilizes sleep deprivation on top of mental workload manipulations (see e.g. Holm et al., 2009; Pontiggia et al., 2026; Wilson et al., 2007). While this kind of research in and of itself is of value in showing the combined detrimental effects of unwanted cognitive states, the results should not be used to inform mental workload assessments without further validation of the used metrics.

3.3. Step three: sensitivity and specificity – testing measures on different concepts

While controlling for confounds is important, these other "unwanted" cognitive states can help to improve the mental workload assessment we aim for. Ideally, a good measurement is not just sensitive to changes in the construct of interest, but also specific to it (Matthews et al., 2015). It should thus not vary with changes in other concepts. In other words, once we are reasonably sure we have captured the neural correlates of "pure" mental workload, we should test them against another cognitive state like mental fatigue in order to rule out that our measures vary with both. For this, a second data collection is necessary that follows the ideas of Step Two in this paper, but with an experiment tailored to the previously unwanted state.

3.4. Our results: as valid as can be?

In our series of studies, we heeded our own advice. We performed simultaneous fNIRS and EEG measurements to investigate the convergent validity between the measurements (Step One). In order to do so, we used compatible devices and chose a montage in which EEG electrodes and fNIRS optodes covered the same areas. There is software available for such purposes, like the fNIRS Optodes' Location Decider fOLD (Zimeo Morais et al., 2018), and we highly recommend making use of such tools as well as documenting the exact measurement locations in order to foster replicability.

Moreover (Step Two), in our first study (Hamann and Carstengerdes, 2022), we took great care to induce mental workload in four levels by means of increasing the difficulty of an adapted n-back task that was tailored to the flight context. We recruited 35 participants for this experiment. We analysed self-reported mental workload ratings and task performance in order to ensure that we had actually induced four distinct levels of mental workload. We succeeded in the self-report, but due to a ceiling effect in the low difficulty levels, we only had three distinct categories of task performance. Nevertheless, we considered this a successful manipulation. We also controlled for influences of mental fatigue by keeping the duration of the task to approx. 45 min and by randomizing the difficulty levels to rule out confounding between mental workload and mental fatigue. We collected KSS ratings prior to starting the experiment in order to exclude participants with high levels of initial sleepiness (none were excluded) and a second rating in the end. We also continuously assessed subjective mental fatigue via F-ISA ratings and thus analysed the progression of mental fatigue. Our experimental design prevented the accumulation of excessive mental fatigue (only a slight increase of 0.69 from 2.09 ["low"] to 2.77 ["medium/relaxed wakeful"]) and no interaction between mental fatigue and mental workload was visible in the data. Thus, we considered the control for confounds successful.

Finally, in order to see which neurophysiological changes were unique to mental workload and could not be mistaken for mental fatigue, we also needed to research mental fatigue (Step Three). Thus, we conducted a subsequent experiment in which we induced mental fatigue and controlled for confounding with mental workload (Hamann and Carstengerdes, 2023a). By comparing the results of both studies, we would be able to see which neurophysiological measures were sensitive to mental workload, i.e. would vary with increasing mental workload, and specific to it, i.e. would not vary with increasing mental fatigue. In order to make the results comparable we designed the second experiment to be as similar as possible to the first. We used the same measurement equipment and montage, the same laboratory and flight simulator and the same cognitive task. We induced mental fatigue by increasing time on task (90 min) and kept mental workload constant by applying only one moderate difficulty level derived from our first study (details on our methods can be found in Hamann and Carstengerdes, 2022, 2023a). We also chose the same measures, and data processing and analysis steps we had used in our first study. This time, we ensured that self-reported mental workload and performance remained constant across the duration of the experiment (they did, apart from a practice effect in the beginning of the experiment) and that self-reported mental fatigue increased. It did, by more than twice the size of the previous study: 1.42 points, from 1.61 ("low") to 3.03 ("medium/relaxed wakeful").[2] Again, we considered our manipulation and experimental control successful.

As our aim in this paper is to highlight the methods rather than the results, we will only briefly discuss our results here and refer the interested reader to the individual publications (Hamann and Carstengerdes, 2022, 2023a, 2023b) as well as a comparative in-depth discussion in (Hamann, 2023). In short, we analysed and compared the fNIRS data (oxygenated blood HbO, deoxygenated blood HbR) and EEG frequency bands (frontal theta 4–8 Hz, and parietal alpha 8–13 Hz and parietal beta 13–30 Hz), as well as two commonly used mental workload indices (Task Load Index TLI; Smith et al., 2001; Engagement Index EI; Pope et al., 1995) between both experiments (Hamann, 2023; Hamann and Carstengerdes, 2023b). An overview of our findings is shown in Fig. 1. We found that frontal measures (i.e. cortical oxygenation, theta band power and TLI) were sensitive to changes in mental workload, and that by combining EEG and fNIRS data, we could differentiate all four induced levels of mental workload, more than with the separate measurements. However, the frontal EEG measures proved sensitive to increasing mental fatigue as well, thus lacking specificity to mental workload. fNIRS may prove a viable option to differentiate the concepts, but our results were somewhat inconclusive (for a discussion see Hamann, 2023). Parietal activity (i.e. alpha and beta band power and EI) lacked sensitivity to changing mental workload altogether. In sum, even though we took great care to control for confounding, and could thus compare "pure" effects of mental workload and mental fatigue, it was quite difficult to tell the concepts apart. What is evident is that increasing frontal cortical activation indicates increasing demand, but cannot be used to explain its cause exactly, even under laboratory conditions.

4. Implications for pilot monitoring

4.1. The need for more precise measurements

Unfortunately, in the cockpit it is unlikely to find pure mental workload. During long-haul flights, mental fatigue may well arise from time on task, even if mental workload is on a constant and low level. Rostering and long shifts can lead to sleepiness that can further decrease pilots' cognitive capabilities (Alaminos-Torres et al., 2023; Caldwell, 2012). And frequent task switching is inherent to piloting an aircraft. In short, in the real world outside the laboratory, there is a multitude of factors that influence pilots' cognitive state, attention and underlying brain activity. Neurophysiological measurements may not be ideal for disentangling these complicated relationships, differentiating cognitive concepts and measuring mental workload "in the wild". We may be able

[2] Note the different initial mental fatigue ratings. We kept the time of day and the lighting conditions constant between both experiments. Due to the differences in task difficulty already during the practice sessions, participants may have felt more fatigued early on in the mental workload study than the mental fatigue study, or just anchored their self-report differently. We chose to interpret the difference between start and end of each study, not the absolute ratings towards the end, but we are aware of this limitation. This will be addressed in future research.

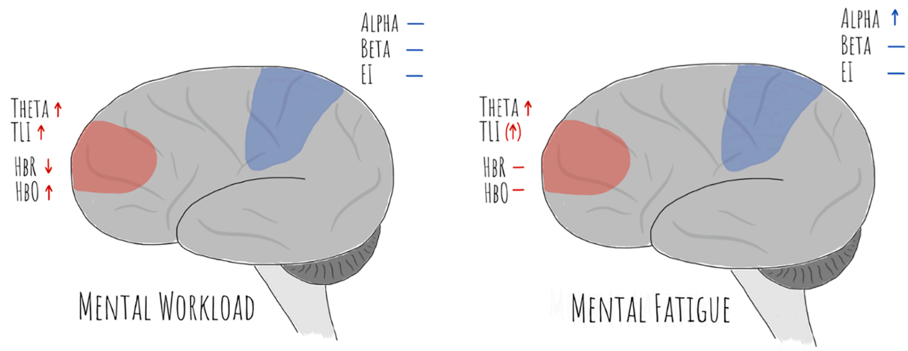

Fig. 1. Schematic comparison of the results of our previous studies for mental workload and mental fatigue (Hamann and Carstengerdes, 2022, 2023a, 2023b). Red indicates frontal cortical areas, blue indicates parietal areas. Significant increases are marked ↑; significant decreases are marked ↓; no significant changes are marked –; brackets indicate significant overall effect but no significant post-hoc tests. Figure taken from Hamann, 2023, p. 55 (Hamann, 2023).

to capture increasing or decreasing demand in a pilot, such as activation or deactivation in the FPN and its underlying structures, but not the exact reason for this change in demand. This may sound a little disappointing, but does not mean we should abandon neurophysiological measurements altogether. The fact alone that we can gain insights into a pilot's current cognitive demand and monitor changes should be impressive enough. And maybe this ability is already sufficient for our purposes. Which level of detail on the pilot's state we eventually need depends on the specific functionalities of the assistance system. If the assistance system is meant to only inform the pilot (or a colleague) about drifting into an unwanted state, similar to drowsiness detection systems in modern cars (Albadawi et al., 2022), it may be sufficient to detect changes in cognitive demand. The decision if and how to act upon this information will be up to the pilot. However, if the assistance system is meant to act independently, change the level of automation, take over tasks or perform automated emergency landings (Fuchs et al., 2026; Vizioli et al., 2023), the data leading to this need to be highly accurate. They need to correctly reflect the reason for the change in the pilot's state (overloaded, fatigued, tired, physically impaired etc.) as well as the pilot's ability to still control their aircraft. A too conservative assessment that results in many false negatives would not provide much benefit over today's operations and likely not warrant the effort of collecting all the data. A too liberal assessment that results in false positives, on the other hand, would have an immense impact on the flight and on traffic in the vicinity. It would likely also result in a rapid loss of trust and disuse from the affected pilots (Dixon et al., 2007; Hinss et al., 2022; Lee and Seppelt, 2012).

4.2. *Towards multimodal measurements*

If more precision is necessary in assessing the pilots' cognitive states and the reasons for changes, it may not be enough to focus only on neurophysiological data. Even a combination of EEG and fNIRS seems insuffcient to draw conclusions about the pilot beyond resonable doubt. However, there are more data sources available that can theoretically be utilized.

There is a large and ever-growing body of research on peripheral physiological parameters for assessing mental workload, such as heart rate and heart rate variability, respiration, electrodermal activity or body temperature (Borghini et al., 2014; Boumann et al., 2023; Wang et al., 2024). Peripheral physiology has repeatedly shown the potential to measure changes in mental workload, but the parameters show varying ability to differentiate multiple levels of mental workload (Boumann et al., 2023; Luzzani et al., 2025). This may not come as a complete surprise. If neurophysiological measures cannot identify mental workload beyond doubt, why should peripheral physiological parameters that are even further away from the origin of mental workload, the brain? Nevertheless, the research clearly shows that mental workload influences multiple physiological signals. Adding these to our assessments may prove beneficial (Da Tao et al., 2025; Hristov et al., 2026; Vanneste et al., 2021). Another alternative is the usage of peripheral physiology to further enhance the quality of the neurophysiological measures. Especially the fNIRS signal is contaminated by heart rate and respiration (Brigadoi et al., 2014). By collecting simultaneous fNIRS and peripheral physiological data and using the latter to remove systemic noise from the former, the validity and reliability of the measurement could potentially be improved (Bauernfeind et al., 2014; Scholkmann et al., 2022).

Of course, we could also utilize the pilots' behaviour. In manual flight, there are many different parameters one could use to assess pilots' flight performance, such as deviations from ideal altitudes, speeds, or measures of flight stability (Chenot et al., 2025). However, within highly automated commercial aircraft, the use of autopilot functions complicates these kinds of assessment. Cognitive tasks like scanning instruments, gathering information and making decisions have largely replaced manual flight control (Billings, 1991). Cognitive processes are difficult to observe, but one type of behaviour remains that is still visible: Eye movements. Eye-tracking can be used to assess changes in eye movements and gaze behaviour in relation to mental workload (Peißl et al., 2018), and even to infer flight phases (Peysakhovich et al., 2022). But gaze behaviour also changes with other mental and physiological phenomena like anxiety, spatial disorientation or hypoxia (Peißl et al., 2018; Ziv, 2016), with expertise (Lounis et al., 2021; Ziv, 2016) and are dependent on the pilots' tasks and role in the cockpit (Causse, Mercier, et al., 2025; Faulhaber et al., 2022). Finally, looking at an instrument does not equal processing the information displayed on it. This phenomenon of inattentional blindness, often also called "looking without seeing", can further complicate the interpretation of eye-tracking parameters (Beanland and Pammer, 2010; Mack, 2003; White and O'Hare, 2022). Thus, similarly to peripheral and neurophysiological measures, eye-tracking alone may not be able to validly and reliably assess the pilots' state.

Fortunately, we are not limited to assessing the human in the cockpit. In an aircraft, there is an abundance of other sources of information readily available. Covariates like the time of day, duration of the flight, weather, aircraft system health indications, flight phase or frequency of radio communications could bridge the gap between the pilots' cognitive demand and its origin (Hinss et al., 2022). This information alone may not tell us if and what kind of assistance the pilots need – highly skilled and trained personnel may be able to cope with adverse weather conditions, difficult approaches or certain system failures without needing further assistance – but the data could add context about the situation to ambiguous changes in (neuro-) physiology or behaviour. This way, it might become possible to interpret changes in the pilots' state with more confidence.

Indeed, there are first public multimodal datasets that include simultaneous recordings of pilots' physiological and behavioural data, aircraft data, self-report and observer ratings (see for example Hamann, Rankova, and Carstengerdes, 2025; Xu et al., 2025). For now, the question remains how these data can be combined for a comprehensive assessment of the pilot's state, which data sources are necessary, which weight should be given to which data under which circumstances, and how valid and reliable such assessments can be when applied to different situations and pilots. But with the collection and publication of such datasets, we believe it is only a matter of time until researchers will find ways to utilize the information for enhanced pilot state monitoring.

5. Conclusion

In the present work, we wanted to highlight the possibilities but also the limitations of neurophysiological measurements in assessing mental workload in the cockpit by giving insights from our own attempts of a very structured and controlled investigation that still did not provide a definite answer. While there is definite potential in assessing changes in mental workload by analyzing changing brain activity, there are many confounding factors like other states or task characteristica that challenge the validity of the assement. To overcome this challenge, it may be beneficial to add more data sources in order to add context to the measurement. To illustrate this with an example, increasing frontal cortical activation in combination with an increasing heart rate and aircraft system failure messages may be a good indication of mental workload, while the same pattern of frontal activation after multiple uneventful hours of flight is likely due to mental fatigue. Such a multimodal approach and combination of the overall state of both pilot (via the pilots' physiology or behaviour) and aircraft (via covariates) could help to achieve the vision of an adaptive assistance system that is tailored to the pilots' current needs without relying too much on the power of one measurement. Ultimately, we need to be are aware of the capabilities and limitations of our methods, single or multimodal measurements alike, instead of just assuming validity, and to apply them carefully where they are suitable and useful.

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.

References

Aghajani, H., Garbey, M., Omurtag, A., 2017. Measuring mental workload with EEG+fNIRS. Front. Hum. Neurosci. 11, 359. https://doi.org/10.3389/fnhum.2017.00359.
Ahn, S., Nguyen, T., Jang, H., Kim, J.G., Jun, S.C., 2016. Exploring neuro-physiological correlates of drivers' Mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. Front. Hum. Neurosci. 10, 219. https://doi.org/10.3389/fnhum.2016.00219.
Akerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37. https://doi.org/10.3109/00207459008994241.
Alaminos-Torres, A., Martínez-Álvarez, J.R., Martínez-Lorca, M., López-Ejeda, N., Marrodán Serrano, M.D., 2023. Fatigue, work overload, and sleepiness in a sample of Spanish commercial airline pilots. Behav. Sci. (Basel) 13 (4), 300. https://doi.org/10.3390/bs13040300.
Albadawi, Y., Takruri, M., Awad, M., 2022. A review of recent developments in driver drowsiness detection systems. Sensors 22 (5), 2069. https://doi.org/10.3390/s22052069.
Banville, H., Parent, M., Tremblay, S., Falk, T.H., 2019. Toward mental workload measurement using multimodal EEG–fNIRS monitoring. Neuroergonomics. Elsevier, pp. 245–246. https://doi.org/10.1016/B978-0-12-811926-6.00057-9.
Bauernfeind, G., Wriessnegger, S.C., Daly, I., Müller-Putz, G.R., 2014. Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals. J. Neural Eng. 11 (5), 56010. https://doi.org/10.1088/1741-2560/11/5/056010.
Beanland, V., Pammer, K., 2010. Looking without seeing or seeing without looking? Eye movements in sustained inattentional blindness. Vision. Res. 50 (10), 977–988. https://doi.org/10.1016/j.visres.2010.02.024.
Benedek, M., Schickel, R.J., Jauk, E., Fink, A., Neubauer, A.C., 2014. Alpha power increases in right parietal cortex reflects focused internal attention. Neuropsychologia 56 (100), 393–400. https://doi.org/10.1016/j.neuropsychologia.2014.02.010.
Billings, C.E., 1991. Toward a human-centered aircraft automation philosophy. Int. J. Aviat. Psychol. 1 (4), 261–270. https://doi.org/10.1207/s15327108ijap0104_1.
Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F., 2014. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. 44, 58–75. https://doi.org/10.1016/j.neubiorev.2012.10.003.
Boumann, H., Hamann, A., Biella, M., Carstengerdes, N., Sammito, S., 2023. Suitability of physiological, self-report and behavioral measures for assessing mental workload in pilots. In: Harris, D., Li, W.-C. (Eds.), Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics, Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics, 14017. Springer, Nature Switzerland, pp. 3–20. https://doi.org/10.1007/978-3-031-35392-5_1.
Brandenburger, N., Dreßler, A., Grippenkoven, J.D., 2023. *DLR Workload Assessment Tool (DLR-WAT): Official English Version*. https://elib.dlr.de/201450/.
Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., Gagnon, L., Boas, D.A., Cooper, R.J., 2014. Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. Neuroimage 85 (0 1), 181–191. https://doi.org/10.1016/j.neuroimage.2013.04.082. *Pt 1*.
Brouwer, A.-M., Zander, T.O., van Erp, J.B.F., Korteling, J.E., Bronkhorst, A.W, 2015. Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls. Front. Neurosci. 9, 136. https://doi.org/10.3389/fnins.2015.00136.
Caldwell, J.A., 2012. Crew schedules, sleep deprivation, and aviation performance. Curr. Dir. Psychol. Sci. 21 (2), 85–89. https://doi.org/10.1177/0963721411435842.
Causse, M., Chua, Z.K., Peysakhovich, V., Del Campo, N., Matton, N., 2017. Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. Sci. Rep. 7 (1), 5222. https://doi.org/10.1038/s41598-017-05378-x.
Causse, M., Deniel, J., Schwartz, F., Duchevet, A., Matton, N., Imbert, J.-P., Cegarra, J., 2025a. Cognitive incapacitation in aviation: a narrative review. Theor. Issues. Ergon. Sci. 26 (4), 525–543. https://doi.org/10.1080/1463922X.2025.2475431.
Causse, M., Mercier, M., Lefrançois, O., Matton, N., 2025b. Impact of automation level on airline pilots' flying performance and visual scanning strategies: a full flight simulator study. Appl. Ergon. 125, 104456. https://doi.org/10.1016/j.apergo.2024.104456.
Charbonnier, S., Roy, R.N., Bonnet, S., Campagne, A., 2016. EEG index for control operators' mental fatigue monitoring using interactions between brain regions. Expert. Syst. Appl. 52, 91–98. https://doi.org/10.1016/j.eswa.2016.01.013.
Chartered Institute of Ergonomics & Human Factors. (2020). *The human dimension in tomorrow's aviation system: White Paper.* Chartered Institute of Ergonomics & Human Factors (CIEHF).
Chenot, Q., Riedinger, F., Dehais, F., Scannella, S., 2025. Assessing and visualizing pilot performance in traffic patterns: a composite score approach. Safety 11 (2), 37. https://doi.org/10.3390/safety11020037.
Chuang, C.-H., Cao, Z., King, J.-T., Wu, B.-S., Wang, Y.-K., Lin, C.-T, 2018. Brain electrodynamic and hemodynamic signatures against fatigue during driving. Front. Neurosci. 12, 181. https://doi.org/10.3389/fnins.2018.00181.
Comisión de Investigación de Accidentes e Incidentes de Aviación Civil, 2024. https://www.transportes.gob.es/recursos_mfom/comodin/recursos/in-001-2024_final-report_nm.pdf.
Da Tao, Huang, J., Zhang, Q., Wang, T., Qu, X., 2025. Assessment of drivers' mental workload: exploring the roles of multimodal physiological measures, driving measures and their combinations. Adv. Eng. Informatics 68, 103796. https://doi.org/10.1016/j.aei.2025.103796.
Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R.N., Lotte, F., 2019. 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.
Dekker, S., Hollnagel, E., 2004. Human factors and folk models. Cognit. Technol. Work 6 (2), 79–86. https://doi.org/10.1007/s10111-003-0136-9.
Dixon, J.B., Clark, T.K., Endsley, T.C., 2025. Findings from interviews with pilots on spatial disorientation: training, temporal dynamics and countermeasures. Front. Physiol. 16, 1620737. https://doi.org/10.3389/fphys.2025.1620737.
Dixon, S.R., Wickens, C.D., McCarley, J.S., 2007. On the independence of compliance and reliance: are automation false alarms worse than misses? Hum. Factors 49 (4), 564–572. https://doi.org/10.1518/001872007X215656.
Djartov, B., Würfel, J., 2025. Navigating decisions in the cockpit: the Intelligent pilot Advisory System. In: Coursaris, C.K., Beringer, J., Léger, P.-M., Öz, B. (Eds.), Progress in IS. The Design of Human-Centered Artificial Intelligence for the Workplace. Springer, Nature Switzerland, pp. 305–325. https://doi.org/10.1007/978-3-031-83512-4_18.
Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. (Regul. Ed.) 14 (4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004.
Dussault, C., Jouanin, J.-C., Guezennec, C.-Y, 2004. Eeg and ECG changes during selected flight sequences. Aviat. Space Environ. Med. 75 (10), 889–897.
Evans, S., Radcliffe, S.-A, 2012. The annual incapacitation rate of commercial pilots. Aviat. Space Environ. Med. 83 (1), 42–49. https://doi.org/10.3357/ASEM.3134.2012.
Faulhaber, A.K., Friedrich, M [Maik], Kapol, T., 2022. Absence of pilot monitoring affects scanning behavior of pilot flying: implications for the design of single-pilot cockpits. Hum. Factors 64 (2), 278–290. https://doi.org/10.1177/0018720820939691.

Feary, M., 2018. A first look at the evolution of flight crew requirements for emerging market aircraft. In: Aiaa Aviation Technology, Integration, and Operations Conference 2018: Held at the AIAA Aviation Forum 2018. Curran Associates Inc.

Fitts, P.M., 1951. Human Engineering For an Effective Air-Navigation and Traffic-Control System. National Research Council.

Friedman, N.P., Robbins, T.W., 2022. The role of prefrontal cortex in cognitive control and executive function. Neuropsychopharmacology 47 (1), 72–89. https://doi.org/10.1038/s41386-021-01132-0.

Fuchs, A., Bejarano, C., Metge, A., Ruano, S., Cordero, J.M., Perillo, A., Vaiopoulos, P., Fedrizzi, G., Vicario, A.G., 2026. Safe return-to-land operations in future cockpits: an analysis of cases and mitigation technologies. In: Harris, D., Li, W.-C., Streitz, N.A., Konomi, S. (Eds.), Lecture Notes in Computer Science. HCI International 2025 – Late Breaking Papers, Lecture Notes in Computer Science. HCI International 2025 – Late Breaking Papers, 16334. Springer, Nature Switzerland, pp. 54–73. https://doi.org/10.1007/978-3-032-12392-3_4.

Gateau, T., Ayaz, H., Dehais, F., 2018. In silico vs. Over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-BCI. Front. Hum. Neurosci. 12, 187. https://doi.org/10.3389/fnhum.2018.00187.

Geissler, C.F., Domes, G., Frings, C., 2020. Shedding light on the frontal hemodynamics of spatial working memory using functional near-infrared spectroscopy. Neuropsychologia 146, 107570. https://doi.org/10.1016/j.neuropsychologia.2020.107570.

Ghaderi, M., Ben Abdessalem, H., Frasson, C., 2023. An analysis of mental workload involved in piloting tasks. In: Krouska, A., Troussas, C., Caro, J. (Eds.), Lecture Notes in Networks and Systems. Novel & Intelligent Digital Systems: Proceedings of the 2nd International Conference (NiDS 2022), 556. Springer International Publishing, pp. 211–220. https://doi.org/10.1007/978-3-031-17601-2_21.

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.

Grandjean, E., 1979. Fatigue in industry. Br. J. Ind. Med. 36 (3), 175–186. https://doi.org/10.1136/oem.36.3.175.

Grillon, C., Quispe-Escudero, D., Mathur, A., Ernst, M., 2015. Mental fatigue impairs emotion regulation. Emotion 15 (3), 383–389. https://doi.org/10.1037/emo0000058.

Grissmann, S., Faller, J., Scharinger, C., Spüler, M., Gerjets, P., 2017. Electroencephalography based analysis of working memory load and affective valence in an N-back task with emotional stimuli. Front. Hum. Neurosci. 11, 616. https://doi.org/10.3389/fnhum.2017.00616.

Hamann, A., 2023. *A Systematic Investigation of EEG and fNIRS Measures For the Assessment of Mental Workload in the Cockpit* [Dissertation]. Technische Universität Dresden, Dresden, Germany. https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-869293.

Hamann, A., & Carstengerdes, N. (2020). *Fatigue instantaneous self-assessment (F-ISA): development of a short mental fatigue rating* (DLR-IB-FL-BS-2020-64). Deutsches Zentrum für Luft- und Raumfahrt, Institut für Flugführung.

Hamann, A., Carstengerdes, N., 2022. Investigating mental workload-induced changes in cortical oxygenation and frontal theta activity during simulated flights. Sci. Rep. 12 (1), 6449. https://doi.org/10.1038/s41598-022-10044-y.

Hamann, A., Carstengerdes, N., 2023a. Assessing the development of mental fatigue during simulated flights with concurrent EEG-fNIRS measurement. Sci. Rep. 13 (1), 4738. https://doi.org/10.1038/s41598-023-31264-w.

Hamann, A., Carstengerdes, N., 2023b. Don't think twice, it's all right? – an examination of commonly used EEG indices and their sensitivity to mental workload. In: Harris, D., Li, W.-C. (Eds.), Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics, Lecture Notes in Computer Science. Engineering Psychology and Cognitive Ergonomics, 14017. Springer, Nature Switzerland, pp. 65–78. https://doi.org/10.1007/978-3-031-35392-5_5.

Hamann, A., Carstengerdes, N., 2024. A window into the mind? On usefulness and challenges of neurophysiological measurements in the cockpit. In: Spink, A., Riedel, G., Truong, K., Robinson, L. (Eds.), Proceedings of Measuring Behavior 2024, the 13th International Conference on Methods and Techniques in Behavioral Research, pp. 27–33.

Hamann, A., Rankova, E., & Carstengerdes, N. (2025a). *Multimethod data on mental workload in pilots*. Open Science Framework (OSF).

Hamann, A., van Klaren, C., Zon, R., Dehais, F., Carstengerdes, N., van Miltenburg, M., Cabrera Castillos, K., 2025b. The state of the art in assessing mental fatigue in the cockpit using head-worn sensing technology. Front. Neuroergonomics 6, 1673268. https://doi.org/10.3389/fnrgo.2025.1673268.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Advances in Psychology. Human Mental Workload, Advances in Psychology. Human Mental Workload, 1st ed., 52. Elsevier Textbooks, pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hassan, J., Reza, S., Ahmed, S.U., Anik, N.H., Khan, M.O., 2025. Eeg workload estimation and classification: a systematic review. J. Neural Eng. 22 (5), 051003. https://doi.org/10.1088/1741-2552/ad705e.

Hinss, M.F., Brock, A.M., Roy, R.N., 2022. Cognitive effects of prolonged continuous human-machine interaction: the case for mental state-based adaptive interfaces. Front. Neuroergonomics 3. https://doi.org/10.3389/fnrgo.2022.935092. Article 935092.

Hirshfield, L.M., Chauncey, K., Gulotta, R., Girouard, A., Solovey, E.T., Jacob, R.J.K., Sassaroli, A., Fantini, S., 2009. Combining electroencephalograph and functional near infrared spectroscopy to explore users' Mental workload. In: Schmorrow, D.D., Estabrooke, I.V., Grootjen, M. (Eds.), *Lecture notes in computer science Lecture notes in artificial intelligence: Vol. 5638. Foundations of augmented cognition: Neuroergonomics and operational neuroscience; 5th international conference, FAC 2009, held as part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009; proceedings*, 5638. Springer, pp. 239–247. https://doi.org/10.1007/978-3-642-02812-0_28.

Holm, A., Lukander, K., Korpela, J., Sallinen, M., Müller, K.M.I., 2009. Estimating brain load from the EEG. Scientif. World J. 9, 639–651. https://doi.org/10.1100/tsw.2009.83.

Hristov, H., Minchev, Z., Shoshev, M., Kancheva, I., Koleva, V., Vakarelsky, T., Dimitrov, K., Prodanov, D., 2026. Sensing cognitive responses through a non-invasive brain-computer interface. Sensors (Basel) 26 (6), 1892. https://doi.org/10.3390/s26061892.

Kearns, S.K., Mavin, T.J., Hodge, S., 2016. Competency-based Education in aviation: Exploring alternate Training Pathways. Routledge. https://doi.org/10.4324/9781315563220. https://www.taylorfrancis.com/books/9781134801800.

Kyle, A.R., Rouser, B., Paul, R.C., Jurewicz, K.A., 2025. Quantifying pilot performance and mental workload in Modern Aviation systems: a scoping literature review. Aerospace 12 (7), 626. https://doi.org/10.3390/aerospace12070626.

Landman, A., Groen, E.L., van Paassen, M.M., Bronkhorst, A.W., Mulder, M., 2017. The influence of surprise on upset recovery performance in airline pilots. Int. J. Aerosp. Psychol. 27 (1–2), 2–14. https://doi.org/10.1080/10508414.2017.1365610.

Lara, A.H., Wallis, J.D., 2015. The role of prefrontal cortex in working memory: a Mini review. Front. Syst. Neurosci. 9, 173. https://doi.org/10.3389/fnsys.2015.00173.

Lee, J.D., Seppelt, B.D., 2012. Human factors and ergonomics in automation design. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics. Wiley, pp. 1615–1642. https://doi.org/10.1002/9781118131350.ch59.

Liu, Y., Ayaz, H., Onaral, B., Shewokis, P.A., 2015. Neural adaptation to a working memory task: a concurrent EEG-fNIRS study. In: Schmorrow, D.D., Fidopiastis, C.M. (Eds.), Lecture Notes in Computer Science: Vol. 9183, Foundations of Augmented Cognition. Springer, pp. 268–280.

Liu, Y., Ayaz, H., Shewokis, P.A., 2017. Mental workload classification with concurrent electroencephalography and functional near-infrared spectroscopy. Brain-Comput. Interfaces 4 (3), 175–185. https://doi.org/10.1080/2326263X.2017.1304020.

Lounis, C., Peysakhovich, V., Causse, M., 2020. Flight eye tracking assistant (FETA): proof of concept. In: Stanton, N. (Ed.), Advances in Intelligent Systems and Computing. Advances in Human Factors of Transportation, Advances in Intelligent Systems and Computing. Advances in Human Factors of Transportation, 964. Springer International Publishing, pp. 739–751. https://doi.org/10.1007/978-3-030-20503-4_66.

Lounis, C., Peysakhovich, V., Causse, M., 2021. Visual scanning strategies in the cockpit are modulated by pilots' expertise: a flight simulator study. PLoS One 16 (2), e0247061. https://doi.org/10.1371/journal.pone.0247061.

Luzzani, G., Pogliano, M., Buraioli, I., Colavincenzo, M., Martorana, S., Guglieri, G., Demarchi, D., 2025. ECG, respiration, fNIRS, and eye tracking for stress and mental workload monitoring in Human-machine interaction. IEEe Access. 13, 122726–122741. https://doi.org/10.1109/ACCESS.2025.3588384.

Mack, A., 2003. Inattentional blindness. Curr. Dir. Psychol. Sci. 12 (5), 180–184. https://doi.org/10.1111/1467-8721.01256.

Martins, A.P.G., 2016. A review of important cognitive concepts in aviation. Aviation 20 (2), 65–84. https://doi.org/10.3846/16487788.2016.1196559.

Masi, G., Amprimo, G., Ferraris, C., Priano, L., 2023. Stress and workload assessment in aviation-a narrative review. Sensors 23 (7), 3556. https://doi.org/10.3390/s23073556.

Matthews, G., Reinerman-Jones, L.E., Barber, D.J., Abich, J., 2015. The psychometrics of mental workload: multiple measures are sensitive but divergent. Hum. Factors 57 (1), 125–143. https://doi.org/10.1177/0018720814539505.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu Rev. Neurosci. 24, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167.

Nguyen, T., Ahn, S., Jang, H., Jun, S.C., Kim, J.G., 2017. Utilization of a combined EEG/NIRS system to predict driver drowsiness. Sci. Rep. 7, 43933. https://doi.org/10.1038/srep43933.

Nogueira, M.G., Silvestrin, M., Barreto, C.S.F., Sato, J.R., Mesquita, R.C., Biazoli, C., Baptista, A.F., 2022. Differences in brain activity between fast and slow responses on psychomotor vigilance task: an fNIRS study. Brain ImAging Behav. 16, 1563–1574. https://doi.org/10.1007/s11682-021-00611-8.

O'Donnell, R.D., Eggemeier, F.T., 1986. Workload assessment methodology. In: Boff, K. R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of Perception and Human Performance. John Wiley & Sons.

Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E., 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 25 (1), 46–59. https://doi.org/10.1002/hbm.20131.

Parasuraman, R., Bahri, T., Deaton, J.E., Morrison, J.G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems: Progress report* (NAWCADWAR-92033-60). Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2008. Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. J. Cogn. Eng. Decis. Mak. 2 (2), 140–160. https://doi.org/10.1518/155534308X284417.

Peißl, S., Wickens, C.D., Baruah, R., 2018. Eye-Tracking measures in aviation: a selective literature review. Int. J. Aerosp. Psychol. 28 (3–4), 98–112. https://doi.org/10.1080/24721840.2018.1514978.

Penumaka, S.M., Kumar Goru, H., Reddy Gadi, N.K., Sirigudi, G., 2026. Multi-level EEG signal analysis for mental workload classification using hybrid CNN-transformer-BiLSTM model. In: 2026 Third International Conference on Networking and Communications (ICNWC). IEEE, pp. 1–5. https://doi.org/10.1109/ICNWC68145.2026.11518156.

Peysakhovich, V., Ledegang, W., Houben, M., Groen, E., Eivazi …, S., 2022. Classification of flight phases based on pilots' visual scanning strategies. In: Shic, F., Kasneci, E., Khamis, M., Gellersen, H., Krejtz, K., Weiskopf, D., Blascheck, T., Bradshaw, J., Vrzakova, H., Binaee, K., Burch, M., Kiefer, P., Bednarik, R., Mardanbegi, D., Clarke, C., Kothari, R., Rajanna, V., Jayarathna, S., Villanueva, A. (Eds.), 2022 Symposium on Eye Tracking Research and Applications. ACM, pp. 1–7. https://doi.org/10.1145/3517031.3529641.

Pontiggia, A., Quiquempoix, M., Fabries, P., Beauchamps, V., Jacques, C., Guillard, M., van Beers, P., Malle, C., Gomez-Merino, D., Koulmann, N., Chennaoui, M., Sauvet, F., 2026. Robust multimodal mental workload classification: a cross-physiological condition machine learning approach. Comput. Methods Programs Biomed. 277, 109251. https://doi.org/10.1016/j.cmpb.2026.109251.

Pope, A.T., Bogart, E.H., Bartolome, D.S., 1995. Biocybernetic system evaluates indices of operator engagement in automated task. Biol. Psychol. 40 (1–2), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3.

Posner, M.I., Dehaene, S., 1994. Attentional networks. Trends Neurosci. 17 (2), 75–79. https://doi.org/10.1016/0166-2236(94)90078-7.

Puma, S., Matton, N., Paubel, P.-V., Raufaste, É., El-Yagoubi, R., 2018. Using theta and alpha band power to assess cognitive workload in multitasking environments. Int. J. Psychophysiol. 123, 111–120. https://doi.org/10.1016/j.ijpsycho.2017.10.004.

Riedinger, F., Suied, C., Keane, R., Sciabica, J.-F., Dehais, F., 2026. Rough auditory alarms mitigate inattentional deafness during piloting-like task. Hum. Factors 68 (8), 1078–1097. https://doi.org/10.1177/00187208261431654.

Rosa, E., Gronkvist, M., Kolegard, R., Dahlstrom, N., Knez, I., Ljung, R., Willander, J., 2021. Fatigue, emotion, and cognitive performance in simulated long-duration, single-piloted flight missions. Aerosp. Med. Hum. Perform. 92 (9), 710–719. https://doi.org/10.3357/AMHP.5798.2021.

Roy, R.N., Bonnet, S., Charbonnier, S., Campagne, A., 2013. Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2013, pp. 6607–6610. https://doi.org/10.1109/EMBC.2013.6611070.

Roy, R.N., Moly, A., Dehais, F., Scannella, S., 2019. EEG and FNIRS connectivity features for mental workload assessment. Neuroergonomics. Elsevier, pp. 327–328. https://doi.org/10.1016/B978-0-12-811926-6.00098-1.

Saadati, M., Nelson, J., Ayaz, H., 2020. Convolutional neural network for Hybrid fNIRS-EEG mental workload classification. In: Ayaz, H. (Ed.), Advances in Intelligent Systems and Computing. Advances in Neuroergonomics and Cognitive Engineering, Advances in Intelligent Systems and Computing. Advances in Neuroergonomics and Cognitive Engineering, 953. Springer International Publishing, pp. 221–232. https://doi.org/10.1007/978-3-030-20473-0_22.

Scerbo, M.W., 1996. Theoretical perspectives on adaptive automation. In: Parasuraman, R., Mouloua, M. (Eds.), Human Factors in Transportation. Automation and Human Performance: Theory and Applications, 1st ed. CRC Press, pp. 37–63.

Scholkmann, F., Tachtsidis, I., Wolf, M., Wolf, U., 2022. Systemic physiology augmented functional near-infrared spectroscopy: a powerful approach to study the embodied human brain. Neurophotonics 9 (3), 30801. https://doi.org/10.1117/1.NPh.9.3.030801.

Smith, M.E., Gevins, A., Brown, H., Karnik, A., Du, R., 2001. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. Hum. Factors 43 (3), 366–380. https://doi.org/10.1518/001872001775898287.

Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc. Natl. Acad. Sci. U.S.A. 105 (34), 12569–12574. https://doi.org/10.1073/pnas.0800005105.

Tattersall, A.J., Foord, P.S., 1996. An experimental evaluation of instantaneous self-assessment as a measure of workload. Ergonomics 39 (5), 740–748. https://doi.org/10.1080/00140139608964495.

Trampitsch-Vink, L.S., Uitdewilligen, S., Peukert, M., & Ott, F. (2025). *Navigating monotony in high reliability organisations: Developing a comprehensive taxonomy of boredom coping strategies in aviation professionals.* https://doi.org/10.2139/ssrn.5550643.

Tran, Y., Craig, A., Craig, R., Chai, R., Nguyen, H., 2020. The influence of mental fatigue on brain activity: evidence from a systematic review with meta-analyses. Psychophysiology 57 (5), e13554. https://doi.org/10.1111/psyp.13554.

Trejo, L.J., Knuth, K., Prado, R., Rosipal, R., Kubitz, K., Kochavi, R.L., Matthews, B., Zhang, Y., 2007. Eeg-based estimation of mental fatigue: convergent evidence for a three-state model. In: Schmorrow, D.D., Reeves, L.M. (Eds.), Lecture Notes in Computer Science: Vol. 4565, Foundations of Augmented Cognition. Springer, pp. 201–211. https://doi.org/10.1007/978-3-540-73216-7_23.

van Weelden, E., Alimardani, M., Wiltshire, T.J., Louwerse, M.M., 2022. Aviation and neurophysiology: a systematic review. Appl. Ergon. 105, 103838. https://doi.org/10.1016/j.apergo.2022.103838.

Vanneste, P., Raes, A., Morton, J., Bombeke, K., van Acker, B.B., Larmuseau, C., Depaepe, F., van den Noortgate, W., 2021. Towards measuring cognitive load through multimodal physiological data. Cognit. Technol. Work 23 (3), 567–585. https://doi.org/10.1007/s10111-020-00641-0.

Vizioli, A.D.B., Bonelli, S., Mallozzi, G., Lieb, T.J., Martins, A.P.G., Friedrich, M [Max], Capasso, P.J., Senatore, C., Godano, F., Contissa, G., Sartor, G., dos Reis, R.J., 2023. Single-pilot incapacitation in commercial aviation - evaluation of an operational concept. J. Phys.: Conf. Ser. 2526 (1), 12077. https://doi.org/10.1088/1742-6596/2526/1/012077.

Wang, P., Houghton, R., Majumdar, A., 2024. Detecting and predicting pilot mental workload using heart rate variability: a systematic review. Sensors 24 (12), 3723. https://doi.org/10.3390/s24123723.

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., Heuer, H., Gutberlet, I., 2014. Frontal theta activity reflects distinct aspects of mental fatigue. Biol. Psychol. 96, 57–65. https://doi.org/10.1016/j.biopsycho.2013.11.010.

White, A., O'Hare, D., 2022. In plane sight: inattentional blindness affects visual detection of external targets in simulated flight. Appl. Ergon. 98, 103578. https://doi.org/10.1016/j.apergo.2021.103578.

Wilson, G.F., Caldwell, J.A., Russell, C.A., 2007. Performance and psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty. Int. J. Aviat. Psychol. 17 (2), 219–247. https://doi.org/10.1080/10508410701328839.

Xu, R., Cao, S., Barnett-Cowan, M., Bulbul, G., Irving, E., Niechwiej-Szwedo, E., Kearns, S., 2025. An in-flight multimodal data collection method for assessing pilot cognitive states and performance in general aviation. MethodsX 15, 103589. https://doi.org/10.1016/j.mex.2025.103589.

Zarlengo, L., 2025. Flight simulators and the effectiveness of transfer of training in aviation. Scientia Moralitas Int. J. Multidiscip. Res. 10 (1), 105–125. https://doi.org/10.5281/zenodo.18041049.

Zheng, L., Jiang, Y.M., Liu, Z., Yang, W., 2025. EEG-based monitoring of pilot training: transferring full-cap representations to headphone-style electrode positions. Proc. ACM Hum. Comput. Interact. 9 (8), 277–291. https://doi.org/10.1145/3773070.

Zimeo Morais, G.A., Balardin, J.B., Sato, J.R, 2018. Fnirs Optodes' Location decider (fOLD): a toolbox for probe arrangement guided by brain regions-of-interest. Sci. Rep. 8 (1), 3341. https://doi.org/10.1038/s41598-018-21716-z.

Ziv, G., 2016. Gaze behavior and visual attention: a review of eye tracking studies in aviation. Int. J. Aviat. Psychol. 26 (3–4), 75–104. https://doi.org/10.1080/10508414.2017.1313096.