Air Traffic Controllers' Acceptability of Physiological and Behavioural Data Processing for Air Traffic Management

An exploratory online survey

Anaïs Servais, Léa Cressonnier, Clément Lheureux, Tim Decrock, Mathis Breton, Christophe Hurter
École Nationale de l'Aviation Civile, ENAC
Toulouse, France
anais.servais@hotmail.com, lea.cressonnier@gmail.com, clement.lheureux@alumni.enac.fr,
tim.decrock@gmail.com, mathis.breton@laposte.net, christophe.hurter@enac.fr

***Abstract*—A significant increase in air traffic is expected in the coming years. Air traffic controllers (ATCOs) will play a pivotal role in ensuring the efficient management of this growing traffic by optimizing flight paths and addressing various constraints, including environmental concerns. Consequently, they bear a high level of responsibility, as human errors can lead to incidents or even accidents. Reducing the risk of human error is paramount in the field of human factors in aeronautics. One possible approach involves detecting and monitoring the mental states of operators, including factors such as fatigue, mental distraction, cognitive load, and stress, by measuring their physiological and behavioral parameters, such as eye movements, heart rate, and brain oscillations. However, the acceptability of these measurements by ATCOs has received limited investigation. To address this gap and provide insights for future research in human factors, we conducted a survey to gather the opinions of 174 French professional and student ATCOs. Our findings represent an initial step toward gaining a deeper understanding of the factors influencing their acceptance of such measures, while also identifying prevailing trends within the field.**

***Keywords*—Air traffic controllers, Air traffic management, Acceptability, Physiological data, Physiological measurements**

I. Introduction

EUROCONTROL (2021) announced a significant growth in air traffic in the upcoming years. Air traffic controllers (ATCOs) will play a crucial role in ensuring the smooth flow of this traffic by optimizing trajectories while considering various temporal and environmental constraints. This high responsibility is accompanied by the potential for human error, which can lead to incidents or even accidents. For example, ATCOs frequently report degraded attentional states [1], and alterations in attention constitute the most common cause of incidents among ATCOs [2]. In the field of human factors, particularly in aeronautics, a critical issue is to reduce the risk associated with human error.

In recent years, thanks to the development of new devices for recording physiological parameters, an increasingly utilized strategy involves measuring operators' physiological or behavioral metrics (e.g., eye movements [3], [4], pupil diameter, heart rate, temperature, brain oscillations, emotions [5]). These measurements offer real-time information about the operators' state with high temporal resolution. They assist in inferring both the physical and cognitive state of the operators before, during, or after specific events within their control sector. Studies in this direction have demonstrated that physiological or behavioral markers enable the identification of various distinct mental states (see examples below). Numerous studies are available in the field of aeronautics, particularly in air traffic management (ATM) [6], [7].

For example, high mental workload has been associated with specific brain oscillations measured using electroencephalography (EEG) [3], [6], as well as hemodynamic changes in prefrontal brain regions detected through functional near-infrared spectroscopy (fNIRS) [8]. Vigilance [9] and stress [3] have also been linked to changes in EEG rhythms, and stress is further associated with increased skin conductance [10]. Eye behavior, including saccades and blinks, often recorded with eye-tracking devices or, more rarely, with electro-oculography [11], [12], has been investigated as a marker of mind-wandering [4] or cognitive workload [3]. Heart rate and blood pressure have shown positive correlations with stress [13] and the density of traffic in the sector [14]. While this list of works is not exhaustive, it illustrates the strong interest in these methods and their potential for enhancing safety. Such measures could be valuable for monitoring an operator's mental state (e.g., triggering an alarm in the event of physiological parameters indicating a degraded attentional state), analyzing incidents to better understand their causes, or during the training of student ATCs (e.g., visually comparing their scanning technique with that of experts [15]).

However, before implementing such devices, the approval of air traffic controllers (ATCs) and their unions is essential. While this strategy appears beneficial for enhancing safety, the recording of physiological parameters in the workplace is a complex matter. Valuable insights can be gained from other workplaces where wearable body or brain sensors are already in use, such as in the utilities sectors, construction, healthcare, education, and government [16]. The implementation of these

SESAR Innovation Days
27-30 November 2023
Seville

tools, if not executed thoughtfully, can lead to feelings of dehumanization and a loss of meaning for the operators. This, in turn, may result in depressive symptoms such as negative thoughts (i.e., rumination) and a decline in motivation [17]. Part of this reaction can be attributed to the negative aspects reported by employees, including concerns about the confidentiality of their personal data (”feeling spied on,” ”lack of trust,” ”fear of inaccurate data,” as indicated by previous studies) [16]. However, it’s worth emphasizing that workers also often experience benefits, such as reduced stress, a subjective improvement in their health, and a heightened sense of safety [16]. Beyond these pros and cons, the level of acceptability for recording physiological signals varies significantly depending on the specific use case and environmental factors [18]. For instance, acceptability tends to be higher when the system is implemented to ensure the well-being and safety of employees themselves (e.g., maintaining an acceptable stress level) rather than solely focusing on productivity or efficiency [18].

Until now, the perspective of air traffic controllers (ATCOs) regarding their willingness to embrace physiological measures has been underexplored. To the best of our knowledge, only one study has investigated the acceptability of devices for assessing hazardous states of awareness (HSAs) among ATCOs [19]. The authors initially gathered critical incidents from workshops involving 11 professional ATCOs and airplane pilots. Subsequently, they used these critical incident data to develop a questionnaire about the acceptability of awareness measures. The questionnaire was completed by a sample of 100 ATCOs and airplane pilots. The results indicated that operators were willing to adopt technology for monitoring HSAs, even if it was considered intrusive, as long as it was perceived as useful. ATCOs reported that such devices could assist them in enhancing their vigilance by signaling alarms when they encountered difficulties. However, a significant concern was raised regarding the legal complexities associated with these recordings, particularly among the older respondents.

This unique study is now over twenty years old. It is likely that, given the evolution of technologies and their changing societal considerations, the opinions of ATCOs have also evolved. Furthermore, the fact that the survey was constructed using critical incidents might have resulted in a higher level of acceptability than what would be encountered in their day-to-day work shifts. As it is essential to comprehend the factors that influence the acceptability of physiological measurement devices among ATCOs, with the goal of preserving their motivation and well-being at work, we believe there is an urgent need for obtaining new data about their perceptions. This will enable future studies in the field of human factors to consider the determinants of acceptability for the implementation of operator-centered systems.

This study aims to explore the willingness of Air Traffic Controllers (ATCOs) to accept the recording and processing of their physiological signals and the influencing factors. An online survey was designed, and responses were collected from a substantial sample of both expert and student ATCOs in France (n = 174).

II. METHODS

A. Participants

Two hundred and eighty-six participants (286) clicked on the survey link to start responding. Among them, 65 participants did not meet the inclusion criteria, and 47 provided incorrect answers during the validation trial. Thus, the analyses were conducted on the responses of the remaining 174 participants. Among these, 74% were professionals, 13.29% were qualifying (in the process of obtaining their approval for operating), 9.83% were students at ENAC, and 2.89% were instructors at ENAC. For analysis purposes, participants were divided into two groups: professional ATCOs (n = 128) and others who were not currently in positions (n = 46, including students, qualifying individuals, or instructors). Demographic data for the two groups can be found in Table 1.

TABLE I. DEMOGRAPHIC DATA OF THE WHOLE SAMPLE, AND OF THE TWO SUBGROUPS.

	Total sample	ATCOs	Others
N	174	128	46
Age (yr)	38.93 ± 10.88	43.20 ± 7.64	27.24 ± 9.84
Gender ($F\%$)	38.15	38.28	37.78
Experience (yr)	NA	18.88 ± 7.65	NA

B. Materials and procedure

The survey was conducted using a LimeSurvey server hosted at ENAC-Lab in Toulouse, France, and data collection took place from May to July 2023. Participants had the option to complete the survey on their laptops, digital tablets, or smartphones. Recruitment of participants was carried out through mailing lists of ENAC students and French air traffic control centers and towers. Upon clicking on the questionnaire link, participants were directed to a page providing information about the study’s objectives. They were then asked to confirm whether they met the inclusion criteria, which required them to be at least 18 years old, fluent in French, and either students or professionals in air traffic control. If any of these criteria were not met, participants were redirected to a ”thank you” page. In case all criteria were satisfied, participants were presented with an electronic consent form for agreement. Following consent, participants received instructions for completing the survey.

To ensure that participants read the questions and were genuinely answering the survey, we inserted a validation question among the real ones: participants were explicitly asked to choose a given irrelevant proposition (e.g., “For this multiple-choice question, choose the proposition saying that the sun is shining today.”). If participants gave a wrong response to the validation question, all of their responses were discarded. To avoid participants answering randomly when they did not have a clear opinion, they had the option to choose the “no answer” option for the given question.

On average, participants took 22.5 minutes to answer the survey. In conformity with ethics rules, participants were allowed to give up and delete their responses at any time. The collected data were strictly anonymous. This study was approved by the Ethics Committee for Research of the University of Toulouse (agreement 2023–640).

C. Structure of the survey

The full questionnaire is freely available at OSF[1]. Here, we explain the main structure of the survey.

1) Demographic data and familiarity with devices to measure physiological parameters: First, we collected personal information, including age and gender. We also inquired about participants' experience as ATCOs, including their status (student, qualifying, professional, or other) and, for professionals, the number of years of experience. Following this, participants were asked whether they had ever used physiological measurement devices at work and, if so, to provide a description of their experience.

2) Open questions: Participants were requested to list three potential advantages that physiological measurements could offer for their job, assuming that these measurements would be reliable in assessing mental states (e.g., attention, fatigue, stress, cognitive workload, etc.). Likewise, participants were asked to identify three potential drawbacks that physiological measurements might have for ATCOs.

3) Scales assessing acceptability: Participants provided their acceptability ratings on a scale ranging from 0 (totally disagree) to 100 (totally agree) using a slider. They rated their acceptability across various categories, including different types of physiological parameters (heart rate, eye movements, pupil dilation, electrodermal activity, and brain activity), various purposes (health, safety, efficiency, shift management, ergonomics), potential concerns (confidentiality, surveillance, administrative consequences, trust issues, competition, health risks, automation, or artificial intelligence development), and the potential detection of various degraded mental states (mind-wandering, mental overload, attentional tunneling, situation awareness, hypovigilance, stress, fatigue, and visual scan path). It's important to note that definitions and explanations were provided for the different concepts to help participants' understanding. Additionally, images of various recording devices were included in the survey for reference.

As previous studies have shown [16], workers often have significant concerns about the protection of their data. Consequently, we also asked participants to indicate their level of agreement with the transmission of their data to various levels within the hierarchical structure, following the French system (algorithm on anonymous data, themselves for informative feedback, the supervisor, subdivisions, QS, BEA safety investigators).

To assist researchers in human factors to better understand the types of devices, along with their associated constraints and benefits, that ATCOs would be more likely to accept (i.e. devices they consider suitable for their workplace), we presented various tools and asked participants to rate their acceptability on a scale from 0 to 100 for each of them. For recording eye movements, we offered options such as remote infrared cameras and wearable eye-tracking glasses. The devices provided for measuring heart rate included the connected watch, the chest belt, and a remote webcam. For electrodermal activity monitoring, we proposed either the connected watch or a wrist patch. Regarding brain activity, electroencephalography (EEG) was always assumed to be the method, but we suggested options such as a headband, a comfortable helmet, or a medical EEG cap. To identify the critical factors influencing participants' extremely low acceptability ratings, we invited participants to explain their choices when they rated an item below 10.

Given that the acceptability of physiological measurements in a daily work context can differ significantly from other situations, we inquired whether participants would be willing to accept physiological measurements during their training or as part of a temporary research study. Lastly, we asked participants about the overall impact of physiological measurements on their well-being at work, whether it would be positive, negative, or have no effect.

[1] https://osf.io/w5vqg/?view_only=6cfc8b3df84f47159d0ad47e88a095c5

D. Analyses

Raw data were exported as an Excel file (publicly available on the OSF repository[1]). All analyses were performed with JASP 0.18 and Python 3.

The responses to open-ended questions concerning the three advantages and three drawbacks of physiological measurements underwent a semantic analysis with cross-validation conducted by two independent judges. The core idea expressed in each response, sometimes articulated as a complete sentence by participants, was summarized in one or two words by these two independent evaluators. Subsequently, congruence between the two evaluators was verified, and any discrepancies were discussed. Once the concept words were extracted, they were translated into English and visualized as word clouds. This visualization helped depict the primary concerns and benefits spontaneously mentioned by ATCOs regarding physiological measurements at their workplace. It's worth noting that comments freely provided by the participants were translated into English by the author when cited.

For each primary factor (purposes, concerns, physiological parameters, recording devices, mental states), we conducted statistical tests to determine the most and least accepted options. Due to the substantial violation of the normality assumption in our dataset (as evidenced by the violin plots depicting kernel density estimation—the distribution shape of the data), we employed Friedman non-parametric tests (utilizing JASP 0.18) for repeated measures, followed by Conover's post-hoc tests with Holm's correction for p-values to address multiple comparisons. Kendall's W was reported as an effect size estimate. We maintained the Type I error risk at 5% with α-threshold 0.5.

To examine the influence of professional experience, we conducted acceptability comparisons between the two groups previously defined in Section II-A (Table 1) using Mann-Whitney tests, which can accommodate different sample sizes. Given our hypothesis of future ATCOs showing greater acceptability compared to experienced ones, we conducted a one-tailed test. Effect sizes were reported as rank-biserial correlations. To account for multiple tests, we applied the Bonferroni-Holm procedure for correction [20].

Similarly, to investigate the impact of previous experience with physiological measurements at work on acceptability, we compared two subgroups within our sample: ATCOs who had previous experience with sensors in their professional environment (n = 18) and those who did not (n = 109; one ATCO did not respond to this question and was therefore excluded from these analyses). Due to the small sample size in one of the groups, we did not perform inferential statistics, as the Mann-Whitney test has limited power with small samples. Instead, we compared the median acceptability for different factors in a table.

III. Results

A. Spontaneous Opinions on Pros and Cons

Figure 1 presents word clouds illustrating the frequencies at which advantages and drawbacks were spontaneously reported by all the participants. Interestingly, we found that our participants expressed similar concerns as workers in other professions in previous studies [16]. In particular, our participants appeared concerned about the privacy of their data, mentioning terms such as ”confidentiality”, ”intimacy”, and ”intrusive”. The feeling of surveillance, akin to the ’Big Brother’ effect, was also observed, with participants using words like ”stressful,” ”surveillance,” ”policing,” and ”prejudicial” quite frequently. However, it’s worth exploring the drawbacks specific to the air traffic control domain. Here, we primarily noted concerns related to potential discomfort and distraction associated with wearing such devices, as evidenced by terms like ”annoying,” ”distractive,” and ”discomfort.” Another significant factor was the perceived unsuitability of these devices for the work environment, as indicated by the term ”unsuitable.” Additionally, it became evident that participants had reservations about the reliability of physiological measurements. For example, they considered them ”useless” and ”individual,” suggesting a belief that measurements may not be consistent across individuals. To a somewhat lesser extent but still notable, some responses raised concerns about the potential adverse effects of physiological measurements. Participants feared the possible loss of their ”medical certification,” which is crucial for their job, as well as an increase in ”automation” based on the collected data.

In addition to these potential drawbacks, participants identified various benefits, including enhanced ”safety,” improved management of their ”fatigue” and ”stress,” and better monitoring of their ”hypovigilance” and ”cognitive workload.” Specific to the field of air traffic control, participants frequently mentioned the potential for enhancing ”shift management” and ”team coordination.” Participants also believed that physiological measurements could contribute to their ”self-awareness,” and they recognized ”research” and ”ergonomics” as potential use cases.

These results offer valuable insights into their preconceptions. We will now examine the results concerning their level of acceptability, which will be discussed in relation to their preconceptions later in the discussion.

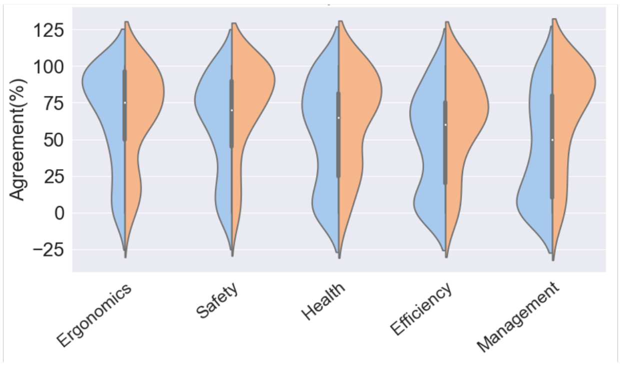
Figure 1. Words clouds showing the frequencies of spontaneous drawbacks and advantages spontaneously reported among all participants.

B. Acceptability

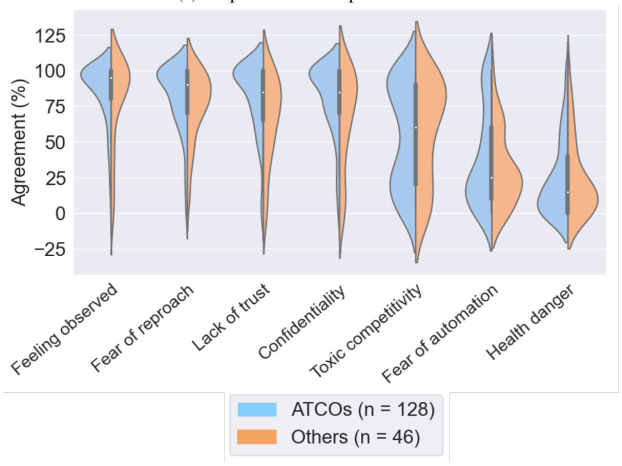
Note that all the statistics mentioned in this section III-B were computed on the entire sample ($n = 174$). However, we decided to present visualizations on violin plots with a separation between the two subgroups defined in section II-A as it allows us to visualize the different shapes of distributions that will be discussed later (Section IV).

1) Agreement with purposes and concerns: While acknowledging the advantages of physiological measurements, their acceptability is not guaranteed, as it depends on striking the right balance with the perceived drawbacks. Acceptability varies significantly based on the intended purposes of implementing physiological measurements, which encompass operator’s health, safety, team management, ergonomics, and operator’s efficiency; see Figure 2a) according to the Friedman test, $\chi^2(4) = 58.58, p < .001, W = .11$. Interestingly, post hoc comparisons revealed that the improvement of ergonomics was significantly better accepted than any other purpose ($t(548) = 3.30, p = .006$ vs. health; $t(548) = 5.78, p < .001$ vs. management; $t(548) = 6.25, p < .001$ vs. efficiency), except safety ($t(548) = 1.41, p = .32$) which was also relatively well accepted. ATCOs had higher acceptability rate for safety purposes than both efficiency ($t(548) = 4.83, p < .001$) and team management purposes ($t(548) = 4.36, p < .001$), which were both globally the least accepted reasons to implement physiological measurements at work.

The significance of certain concerns was notably higher than others (see Figure 2b). According to the Friedman test, $\chi^2(6) = 348.81, p < .001, W = .45$, the four most prominent concerns were data confidentiality, the sense of being under observation, a lack of trust, and the apprehension of potential repercussions from higher-ups. These four concerns were statistically tied and significantly differed from concerns related to competition, health risks, or automation in pairwise comparisons (all p-values were $p < .001$). Furthermore, the concern about competition was significantly greater than concerns about health risks ($t(774) = 6.41, p < .001$) or automation ($t(774) = 3.64, p = .002$). Lastly, the concern about automation surpassed concerns about health risks ($t(774) = 2.77, p = .04$), which, in fact, did not appear to be a significant concern.



(a) Purposes of the implementation



(b) Concerns about the recordings

Figure 2. Violin plots of the agreement scores of the participants with different purposes of physiological measurements, and concerns about them. The violin plots are split into two subgroups: the professionals ATCOs currently in a position in blue, and the other participants who were students, qualifying, or instructors in orange. As sample sizes differed between the two subgroups, to facilitate visual comparisons, violin distributions were plotted for the proportions in each subgroup rather than the counts. The boxplots represent the medians (white dot) and quartiles for the whole sample ($n = 174$)–i.e., data used for statistics in Section III-B.

2) Acceptability of physiological parameters and recording devices: Participants were inclined to accept the recording of certain physiological parameters more than others, $\chi^2(4) = 15.42, p = .004, W = .03$ (Figure 3). Specifically, thanks to pairwise post hoc comparisons, we observed that the recording of eye movement was significantly better accepted than recording the heart rate $t(604) = 3.19, p = .01$ and brain activity $t(604) = 3.50, p = .005$.

However, the type of device that is used for recording also seemed to have an influence on the acceptability. For eye movement, remote infrared cameras reached median acceptability of 60% compared to wearable eye-tracking glasses that were massively not accepted (median acceptability 5%). For heart rate recording, participants tend to accept wearing a smartwatch (median acceptability 65%), but way less wearing a belt (median acceptability 20%) or being remotely recording via webcam image processing (median acceptability 0%). None of the proposed devices for brain activity recording, all wearable, was well accepted (average median acceptability across devices 6%).

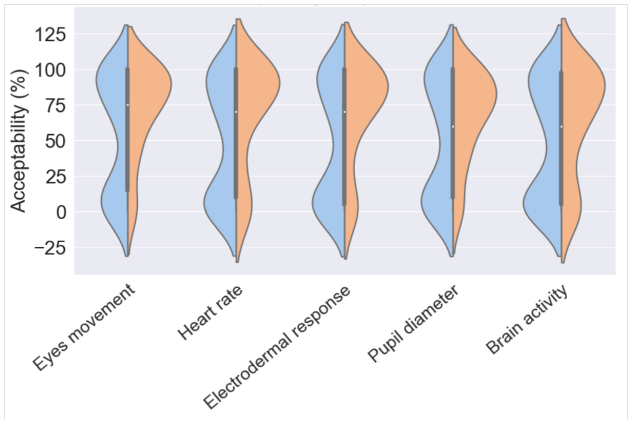


Figure 3. Physiological parameters: Violin and box plots of the acceptability (on a scale from 0 to 100) for the recording of different physiological parameters that were proposed in our survey. Legend and explanations of the graphs are the same as Figure 2.

3) Acceptability of mental states inferences: The Friedman test revealed a significant impact on the acceptability depending on the mental state that is inferred from the physiological data, $\chi^2(4) = 119.70, p < .001, W = .12$ (Figure 4). It is important to note that this main effect was driven by the fact that mind-wandering was significantly less accepted than every other mental state according to post hoc analyses ($t(1015) = 7.76, p < .001$ vs. mental overload; $t(1015) = 7.44, p < .001$ vs. attentional tunneling; $t(1015) = 5.20, p < .001$ vs. situation awareness; $t(1015) = 7.84, p < .001$ vs. hypovigilance; $t(1015) = 6.46, p < .001$ vs. stress; $t(1015) = 9.20, p < .001$ vs. fatigue; $t(1015) = 8.57, p < .001$ vs. visual scan path). All the other pairwise comparisons were not significant, except the situation awareness which had a lower acceptability than fatigue ($t(1015) = 4.00, p = .001$) and visual scan path ($t(1015) = 3.37, p = .02$).

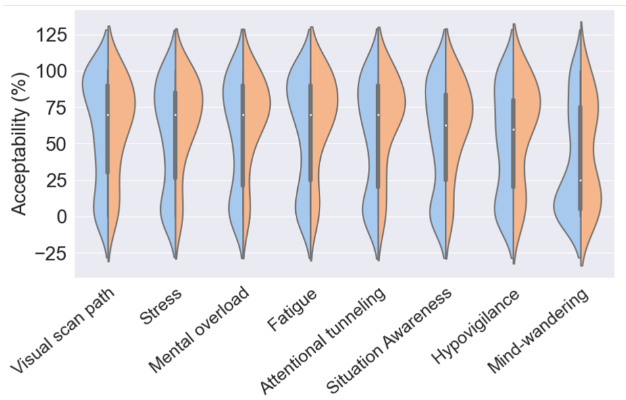


Figure 4. Monitoring of mental states: Violin and box plots of the acceptability (on a scale from 0 to 100) for the recording with the aim of monitoring certain degraded mental states. Legend and explanations of the graphs are the same as Figure 2.

4) Confidentiality and data access: As previous studies already suggested in other workplaces, confidentiality of data is a key point. Our results showed that ATCOs have a very high acceptability for physiological measurements when they are the only ones to access their own personal data: the median acceptability for informative feedback about their physiological state was 100%. Participants also agreed with recordings to constitute a pool of anonymous data that could be used to improve the tools and interfaces (median acceptability 65%). As soon as another person, superior, colleague, or safety center, might access the data, the median acceptability fell to 0%, reflecting strong opposition.

5) Well-being at work: When participants were asked how physiological measurements would impact their well-being at work, 126 participants (72% of the participants) indicated that they would feel worse, while only one participant responded that they would feel better. Other participants mentioned that it would not have any impact ($n = 17$), or chose not to respond ($n = 30$). The proportion of individuals who anticipated feeling worse was significantly higher than the expected chance level of 33.33% based on a binomial test of frequencies ($p < .001$).

C. Comparison between Qualified and Student ATCOs

Here, we investigate if experienced ATCOs have a lower acceptability than the other participants. Indeed, according to Mann-Whitney, their agreement with purposes of the implementation was significantly lower for team management ($W = 1405, p < .001, r_{rb} = -.39$), for safety ($W = 1844, p = .006, r_{rb} = -.26$), and for efficiency ($W = 1826, p = .008, r_{rb} = -.25$) purposes, but not for health or ergonomics improvements. Interestingly, their agreement with disadvantages was also significantly higher for the lack of trust from the hierarchy ($W = 3253, p < .001, r_{rb} = .36$), the concerns about the confidentiality of the data ($W = 3464, p = .003, r_{rb} = .27$). No significant difference was found between groups for the agreement of the other proposed potential disadvantages.

Given the strong bimodal distributions for the group of professional ATCOs (see violin plots in Figures 2, 3, and 4), we wanted to verify if the two peaks were not related to age–that is the high acceptability for younger participants (matching the distributions of the other subgroup containing students and qualifying participants, who are younger, see Table I). To eliminate the possible confounding factor of age, we, therefore, have conducted the same analyses by dividing our sample into two age groups (below or above the median age). We did not find any significant difference either for purposes or for disadvantages.

D. Impact of previous experience

It's important to note that only 14% of ATCOs currently in positions have prior experience with physiological measurements at work (18 participants compared to 109 who did not), which constitutes a minority. Table II shows that the subgroup of participants who had previous experience with devices for physiological measurement at work had systematically higher median acceptability for all types of physiological parameters and all types of mental states (except visual scan path) than the subgroup of participants who had never experienced it. On average, the median acceptability of ATCOS who have prior experience with physiological measurement was higher than the other group by 12 points (/100) for physiological parameters and 8.75 points (/100) higher for the monitoring of mental states. Interestingly, the group of ATCOs who already used physiological recording at work perceived the different purposes of physiological measurements as more positive than the other subgroup (their median agreement was on average 19 points higher). Nevertheless, even though having experience with physiological measurement seems to be associated with higher acceptability and higher perception of the positive aspects, it does not erase the perception of all the negative aspects.

TABLE II. MEDIAN ACCEPTABILITY OF ATCOS (IN POSITION) WHO HAVE PREVIOUS EXPERIENCE WITH PHYSIOLOGICAL RECORDING AT WORK ('YES') AND THOSE WHO DON'T ('NO'). THE LAST COLUMN REPRESENTS THE DIFFERENCE BETWEEN THE TWO GROUPS–IT IS THE SUBTRACTION OF THE MEDIAN OF THE UNEXPERIENCED GROUP FROM THE MEDIAN OF THE EXPERIENCED GROUP.

Prior experience	Yes	No	Difference
Mind-wandering	30	22.5	+7.5
Mental overload	80	70	+10
Attentional tunneling	75	67.5	+7.5
Situation awareness	70	50	+20
Hypovigilance	75	65	+10
Stress	70	60	+10
Fatigue	80	70	+10
Visual scan path	65	70	−5
Heart rate	70	60	+10
Eye movement	80	70	+10
Pupil dilatation	60	50	+10
Electrodermal activity	70	50	+20
Brain activity	60	50	+10
Health	77.5	60	+17.5
Safety	80	70	+10
Efficiency	60	50	+10
Management	67.5	35	+32.5
Ergonomics	100	75	+25
Confidentiality	92.5	90	+2.5
Feeling observed	85	100	−15
Fear of reproach	90	85	+5
Lack of trust	92.5	85	+7.5
Competitivity	35	62.5	−27.5
Health danger	10	15	−5
Fear of automation	25	25	0

E. Acceptability for a temporary research study

In the context of a temporary study, 78% of the participants agreed with the use of physiological measurements. This was a significant proportion of positive responses compared to the chance level (of 50% for yes-no choice) according to the binomial test on frequencies ($p < .001$).

IV. DISCUSSION

Our motivation is to gain a more comprehensive understanding of the operators' opinions to advocate for a human factors approach that places them at the center of the system. As one of the ATCOs who participated in our study expressed, 'Currently, there is a complete lack of understanding of our work by the administration. They assume they know, and as a result, push forward forcefully. [...] Ergonomics is always an afterthought.' This quote perfectly illustrates the goal of our exploratory study, which is to counteract such misconceptions by gathering insights directly from the operators themselves.

As indicated by previous studies [16], [18], [19], the use of physiological recording for safety purposes has generally been well-received. In the context of air traffic control, it appears even more critical, as accident prevention is directly aligned with the core objective of their job. In line with the participant's comment mentioned earlier, it is noteworthy that ergonomics was one of the reasons for implementing physiological measurements that was both spontaneously mentioned by the participants and received the highest level of acceptance. Among the potential reasons for utilizing physiological recording in ATM, enhancing efficiency or managing staff (i.e., sector ungrouping, team shifts, etc.) were significantly less accepted. While this effect has been observed in other workplaces for efficiency purposes [18], it was unexpected that ATCOs would reject the use of physiological measurements for team management, especially considering they spontaneously mentioned it as a potential advantage. The reality is that, even if they recognize it as a possible advantage for improving ATM, we hypothesized that they fear more the negative consequences that sharing these data with others, in particular the hierarchy, could have.

Indeed, our findings support this hypothesis, as we have clearly demonstrated that data privacy is a major concern. Beyond being one of the primary concerns among our participants, a complete aversion was indicated by a median acceptability rating of zero when it came to granting access to data by others (superiors, colleagues, or safety centers). While concerns about data confidentiality may appear to be a common thread across professions [16], the ATM field presents specific nuances. ATCOs must adhere to medical requirements to maintain their work licenses, and they appear to have particular concerns about the potential misuse of collected data (see Figure 2b), which could have adverse implications for their employment. For instance, some participants expressed fear of being diagnosed with a cardiovascular condition that could render them unfit for duty. This may explain why certain physiological parameters, such as heart rate or brain activity, were less accepted than others, like eye movement.

Based on our survey, there are other significant concerns, including the subjective feeling of a lack of trust from the administration, the fear of reproach, and the sensation of being under observation. Participants used phrases such as 'Big Brother is watching,' 'intrusive,' 'policing,' 'surveillance,' and even 'prejudicial' to express these concerns. Furthermore, our results revealed that experienced ATCOs, who are familiar with the work environment, appear to be even more apprehensive about issues related to confidentiality and the lack of trust from the administration than future ATCOs who have not yet experienced real working conditions. Our data do not provide insights into the causes of this effect, but it's important to acknowledge that the current atmosphere in the French ATM system does not appear conducive to the acceptance of physiological recordings. For future studies, it would be valuable to identify the specific workplace factors that are perceived as constraints or obstacles by ATCOs.

A significant majority of participants (72%) believe that physiological recordings will negatively impact their well-being at work. To justify their responses, participants commonly cited concerns about feeling under surveillance, as well as physical discomfort (e.g., 'uncomfortable,' 'restriction of freedom of movement,' 'it looks heavy'), and the unsuitability of current devices for their workstations. Therefore, when selecting devices, we strongly recommend that human factors studies consider both comfort and the subjective sense of intrusion. Our results did not reveal strong preferences for either remote or wearable devices. Wearable devices were frequently regarded as uncomfortable or unsuitable (e.g., due to concerns about additional helmets or glasses not fitting over existing headphones). In contrast, remote devices were associated with a stronger sense of intrusion, resembling cameras. As it stands, existing sensors and devices do not appear to be fully compatible in their current versions [21]. However, we remain hopeful that technological advancements will soon address this issue.

While reviewing the comments freely provided by some participants, we observed that there are still misconceptions about physiological measurements. These misconceptions pertain to both their functioning and the interpretations that can be derived from the data. Here are a few examples: 'These devices are designed for medical care; I am not unwell,' 'We are not machines or robots,' and 'This seems like science fiction'. When we consider this observation in light of our results, which indicated that participants who had prior experience with physiological measurements at work exhibited higher acceptance, we can suggest guidelines for the future design of training and educational programs for ATCOs. We believe that the topic of physiological measurements should be more comprehensively integrated into the curriculum to provide accurate information and dispel misconceptions. Additionally, we recommend familiarizing both current and future ATCOs with these devices, potentially through internships or simulations. Exposure is crucial, especially because it not only influences acceptability (i.e., the prospective judgment) but also plays a pivotal role in acceptance (i.e., the attitude toward the device after its introduction) [22]. Furthermore, for workers, short-term experiments conducted as part of research projects could serve as a useful exposure strategy, considering that 78% of the participants expressed their willingness to participate in such studies.

A. *Limitations*

First, it is important to interpret our results with caution as our interviews were exclusively conducted with French air traffic controllers (ATCOs), limiting the generalizability of our findings. Cultural factors can significantly influence individuals' relationships with technology, and these findings may not be universally applicable. Second, our assessment of acceptability was based on a predetermined, non-exhaustive list of factors. While there was substantial congruence between the factors spontaneously mentioned by participants and those we used for evaluating acceptability, it is noteworthy that the inclusion of additional factors could enhance our understanding more comprehensively. Third, upon examining the violin plots, we observed a bimodal distribution in acceptability for most evaluated factors. One subgroup exhibited very low acceptability, while another displayed very high acceptability. This suggests that our sample of professional ATCOs can be divided into two subgroups with distinctly opposing opinions. However, our study did not enable us to uncover the underlying factors driving these two extreme viewpoints.

V. Conclusion

As observed in other workplaces, the inevitable future integration of physiological data in air traffic management (ATM) warrants careful consideration. To ensure a successful implementation, we emphasize the crucial role of placing humans at the center to prevent feelings of dehumanization. Our results primarily recommend the incorporation of physiological measures with the aim of enhancing safety and/or ergonomics, all while prioritizing data confidentiality and user comfort. However, before implementation, we advocate for educational initiatives on physiological devices and recordings. We hope that our preliminary exploratory findings will contribute to the development of suitable methodologies and serve as guidance for researchers in the field of human factors.

Acknowledgment

We would like to thank Julie Saint-Lot, Jean-Paul Imbert, Yannick Migliorini, and Nadine Matton for their feedback on a preliminary version of the survey. A special thank you to Julie Saint-Lot and Yannick Migliorini for their help in sharing the survey with diverse centers to recruit participants. This work was financed by the European Union's Horizon 2020 within the framework SESAR 2020 research and innovation program under grant agreement N. 894238, project Transparent artificial intelligence and automation to air traffic management systems, ARTIMATION, N. 101114838, Trustworthy intelligent system for remote digital tower, TRUSTY, and N. 101114765 COntroller adaptive Digital Assistant, CODA.

References

[1] Yannick Migliorini, Jean-Paul Imbert, Raphaëlle N Roy, Alex Lafont, and Frédéric Dehais. Degraded states of engagement in air traffic control. *Safety*, 8(1):19, 2022.

[2] Anthony M Pape, Douglas A Wiegmann, and Scott A Shappell. Air traffic control (atc) related accidents and incidents: A human factors analysis. 2001.

[3] Chama Belkhiria and Vsevolod Peysakhovich. Eog metrics for cognitive workload detection. *Procedia Computer Science*, 192:1875–1884, 2021.

[4] Anaïs Servais, Florine Riedinger, Emmanuel Barbeau, and Christophe Hurter. A new protocol to study mind-wandering for air traffic controllers: A pilot study. *Aviation Psychology and Applied Human Factors*, 2, 2023.

[5] Anne-Marie Brouwer, Loïs van de Water, Maarten Hogervorst, Wessel Kraaij, Jan Maarten Schraagen, and Koen Hogenelst. Monitoring mental state during real life office work. In *Symbiotic Interaction: 6th International Workshop, Symbiotic 2017, Eindhoven, The Netherlands, December 18–19, 2017, Revised Selected Papers 6*, pages 18–29. Springer, 2018.

[6] Pietro Arico, Gianluca Borghini, Gianluca Di Flumeri, Alfredo Colosimo, Ilenia Graziani, Jean-Paul Imbert, Géraud Granger, Railene Benhacene, Michela Terenzi, Simone Pozzi, et al. Reliability over time of eeg-based mental workload evaluation during air traffic management (atm) tasks. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7242–7245. IEEE, 2015.

[7] Gianluca Di Flumeri, Francesca De Crescenzio, Bruno Berberian, Oliver Ohneiser, Jan Kramer, Pietro Aricò, Gianluca Borghini, Fabio Babiloni, Sara Bagassi, and Sergio Piastra. Brain–computer interface-based adaptive automation to prevent out-of-the-loop phenomenon in air traffic controllers dealing with highly automated systems. *Frontiers in human neuroscience*, 13:296, 2019.

[8] Hasan Ayaz, Patricia A Shewokis, Scott Bunce, Kurtulus Izzetoglu, Ben Willems, and Banu Onaral. Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59(1):36–47, 2012.

[9] Marika Sebastiani, Gianluca Di Flumeri, Pietro Aricò, Nicolina Sciaraffa, Fabio Babiloni, and Gianluca Borghini. Neurophysiological vigilance characterisation and assessment: Laboratory and realistic validations involving professional air traffic controllers. *Brain sciences*, 10(1):48, 2020.

[10] Gianluca Borghini, Andrea Bandini, Silvia Orlandi, Gianluca Di Flumeri, Pietro Aricò, Nicolina Sciaraffa, Vincenzo Ronca, Stefano Bonelli, Martina Ragosta, Paola Tomasello, et al. Stress assessment by combining neurophysiological signals and radio communications of air traffic controllers. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 851–854. IEEE, 2020.

[11] Chama Belkhiria, Atlal Boudir, Christophe Hurter, and Vsevolod Peysakhovich. Eog-based human–computer interface: 2000–2020 review. *Sensors*, 22(13):4914, 2022.

[12] Anaïs Servais, Raphaël Poveda, Hanna Gerony, Emmanuel Barbeau, and Christophe Hurter. Can we use eog to identify when attention switches away from the outside world to focus on our mental thoughts? In *Proceedings of the 1st International Conference on Cognitive Aircraft Systems - ICCAS*, pages 24–29. INSTICC, SciTePress, 2023.

[13] Susana Rodrigues, Joana S Paiva, Duarte Dias, Marta Aleixo, Rui Filipe, and João Paulo Silva Cunha. A wearable system for the stress monitoring of air traffic controllers during an air traffic control refresher training and the trier social stress test: a comparative study. *The Open Bioinformatics Journal*, 11(1), 2018.

[14] Joachim Vogt, Tim Hagemann, and Michael Kastner. The impact of workload on heart rate and blood pressure in en-route and tower air traffic control. *Journal of psychophysiology*, 20(4):297–314, 2006.

[15] Ricardo Palma Fraga, Ziho Kang, Jerry M Crutchfield, and Saptarshi Mandal. Visual search and conflict mitigation strategies used by expert en route air traffic controllers. *Aerospace*, 8(7):170, 2021.

[16] Lauren C Tindale, Derek Chiu, Nicole Minielly, Viorica Hrincu, Aline Talhouk, and Judy Illes. Wearable biosensors in the workplace: Perceptions and perspectives. *Frontiers in Digital Health*, 4:800367, 2022.

[17] Kateryna Maltseva. Wearables in the workplace: The brave new world of employee engagement. *Business Horizons*, 63(4):493–505, 2020.

[18] Jesse V. Jacobs, Lawrence J. Hettinger, Yueng hsiang Huang, Susan Jeffries, Mary F. Lesch, Lucinda A. Simmons, Santosh K. Verma, and Joanna L. Willetts. Employee acceptance of wearable technology in the workplace. *Applied ergonomics*, 78:148–156, 2019.

[19] Terry L Dickinson, Peter J Milkulka, Doris Kwan, Amy A Fitzgibbons, Florence R Jinadu, Frederick G Freeman, Mark W Scerbo, and AT Pope. User acceptability of physiological and other measures of hazardous states of awareness. Technical report, 2001.

[20] Massimiliano Giacalone, Zirilli Agata, Paolo Carmelo Cozzucoli, and Angela Alibrandi. Bonferroni-holm and permutation tests to compare health data: methodological and applicative issues. *BMC medical research methodology*, 18:1–9, 2018.

[21] Carmelo P Morabito, Jason Boggs, and Kathryn Feltman. End-user panel evaluating the feasibility and acceptability of physiological sensors for operator state monitoring in the aircraft.

[22] Jens Schade and Bernhard Schlag. Acceptability of urban transport pricing strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 6(1):45–61, 2003.