Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

The Influence of Selected Factors in the Precision of Flight Operations Execution

Katarzyna BRZESKA
katarzyna.brzeska@polsl.pl (Corresponding author)
https://orcid.org/0000-0002-2580-9290
Silesian University of Technology, Faculty of Transport and Aviation Engineering Gliwice, Poland

Mateusz URZĘDOWSKI
mateusz.urzedowski@polsl.pl
https://orcid.org/0009-0008-6375-1908
Silesian University of Technology, Faculty of Transport and Aviation Engineering Gliwice, Poland

Magdalena MROZIK
magda.mrozik@polsl.pl
https://orcid.org/0000-0003-4496-8331
Silesian University of Technology, Faculty of Transport and Aviation Engineering Gliwice, Poland

Received: 8 July 2025 | Revised: 10 October 2025
Accepted: 12 December 2025 | Available online: 30 December 2025

Abstract

According to statistics, humans are the main cause of aviation accidents and incidents. The theoretical part of this paper describes selected methods of testing flight crews, including objective and behavioral methods, while the practical part, based on a quantitative method, referred to the introduction of selected emergency situations during a simulated flight on the FNPT II MCC simulator in order to check the mental load of pilots and their decision-making process (analysis using the NASA-TLX questionnaire). The final part of the article presents a summary, with particular emphasis on the research results. Conducting this type of research is crucial from the perspective of aviation safety, as it allows the precision of pilots' task performance across various situations to be determined. Although it is not possible to completely eliminate the influence of external and internal factors, knowledge of risk areas enables measures to be taken to maintain the required level of safety.

Keywords: flight operations, human factor, NASA-TLX, safety

1. Introduction

Without adequate training for flight crews and all personnel involved in conducting flight operations (including pilots, cabin crew, air traffic controllers, and ground handling staff), it would be impossible to achieve the highest level of aviation safety among all modes of transportation. Historically, the development of aviation was limited by technological unreliability. Modern aviation systems and the technological processes involved in aircraft manufacturing have reached an incomparably high level of reliability compared to the early years of aviation. Various statistics indicate that over 70% of aviation incidents, including accidents, are now caused by human error.

According to the fundamental definition, safety refers to a state that ensures survival and the continuation of existence. It is directly related to the protection of human health and life. With technological advancements, the scope of this definition has expanded, and safety is considered in different contexts across various scientific fields. In aviation, safety is a top priority

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

in every operation. (Majer P. 2012) Therefore, it is crucial to define this concept precisely. According to ICAO Document 9859, aviation safety is defined as "a state in which the risks associated with aviation activities, including the operation of aircraft or activities directly supporting flight operations, are maintained at an acceptable level and controlled". This means that every flight carries some degree of risk. However, it is essential to implement all possible measures to minimize hazards, both existing and potential, that may arise during flight operations. Maintaining and even continuously improving aviation safety remains one of the key objectives of every flight mission (ICAO 9859).

In the context of assessing a pilot's psychophysical condition, four key methods can be distinguished: subjective, objective, behavioral, and aircraft indicators (Galant M. 2017). The subjective method refers to the pilot's self-assessment. It relies on the pilot's own feelings and observations. Therefore, it is necessary for the individual being tested to be active and aware of the purpose of the assessment. Due to the subjective nature of the assessment, the participant evaluates the cognitive load they experienced during a specific task. This method utilizes the NASA-TLX questionnaire and the SWAT technique. The NASA-TLX questionnaire is a multidimensional scale designed to provide an estimated subjective evaluation of workload during a given task. It is considered the most reliable and valid method among subjective approaches to assessing psychophysical condition. Developed by a NASA research team focused on human capabilities and limitations, the questionnaire includes 10 categories that are subjectively rated by the participant on a 20-point scale. These categories are: overall workload (low to high), task difficulty (low to high), time pressure (none to rushed), performance (failure to perfect), mental/sensory effort (none to impossible), physical effort (none to impossible), frustration level (fulfilled to exasperated), stress level (relaxed to tense), fatigue (exhausted to alert), and activity type (skill-based to rule-based to knowledge-based). As a result, areas such as mental load, physical load, time pressure, performance, effort, and frustration are evaluated. The SWAT technique, developed by Reid and Nygren in 1988, can also be effectively applied to operational situations. Using this technique in studies provides a reliable measurement of load. Initially used for evaluating aircraft cockpits and other aviation crew work environments, it was later expanded to include air traffic controller load assessments as well. This tool is multidimensional, covering time load (T), mental effort load (E), and psychological stress load (S). The technique consists of two stages. In the first stage, the participant rates the levels of the three load scales from lowest to highest before the test. In the second stage, the participant rates the load levels again, following the same procedure, after completing the task. (Galnat M. 2017, Hart S. G., Staveland L. E. 1988, Luximon A., Goonetilleke R. S. 2001).

Objective methods include a group of techniques such as EEG, fNIR, GSR, eye-tracking, cardiovascular system activity measurement, and respiratory rate and depth monitoring. According to Niedermeyer and Lopes da Silva, EEG records electrical activity from the surface of the scalp, captured using metal electrodes and conductive devices. Mental workload assessment is effectively applied in studies where changes in an EEG are associated with an increase in cognitive workload and mental effort allocation. This method is classified as a non-invasive neuroimaging technique. Neuron activity reflects the brain's electrical activity, which mirrors the information transmission process. To measure the potential between electrodes, at least two electrodes are needed, although for optimal results, 19 electrodes are typically used. The brain's bioelectrical activity is recorded during an EEG and is expressed in waves with frequencies ranging from 1 Hz to 100 Hz and amplitudes from 5 μV to several hundred μV. These frequency recordings help determine a specific mental state, as each state corresponds to a particular frequency range. The main brain waves are as follows: alpha waves, representing a state of relaxation (excessive alpha waves indicate issues such as lack of motivation, discouragement, and distraction), ranging from 8 Hz to 13 Hz; beta waves, indicating readiness (subdivided into low beta, from 13 Hz to 15 Hz, related to storing and recalling information; mid beta, optimal for problem-solving with fast and effective brain activity; and high beta, indicating excessive brain arousal, manifesting in impulsivity, hyperactivity, aggression, and rebellious behavior), ranging from 12 Hz to 30 Hz; gamma waves, occurring during information association and intense thinking; theta waves, associated with processes such as meditation, hypnosis, vivid daydreaming, and emotions, also linked to pain control, where thoughts become fragmented and logic ceases to apply, reducing stress and enhancing creativity, intuition, and memory, ranging from 4 Hz to 8 Hz; and delta waves, which represent deep mental effort and occur during deep sleep, ranging from 0.5 Hz to 3 Hz. The fNIR method, also a non-invasive neuroimaging technique, records neuronal cell activity. It detects changes in blood oxygenation and flow. In this method, electromagnetic beams are passed through the skull of the subject using lasers, operating at near-infrared frequencies. Changes in light reception allow for measuring blood flow in the frontal brain area. Increased blood flow in this region is associated with heightened neuronal activity. Cognitive and perceptual tasks are used in fNIR studies to assess brain activity. These tasks include actions such as speaking or hand movement. When intense neural activity occurs in a specific area, it indicates blood flow and increased oxygen concentration, closely linked to changes in deoxyhemoglobin levels, which are visible in BOLD signal changes. The GSR method measures skin conductivity, which may be linked to changes in the human sympathetic nervous system. Using this method allows for assessing an individual's psychophysical state based on their health, arousal, anxiety, and stress levels. Increased or decreased stress causes changes in the GSR level. These factors are measurable by recording the skin's galvanic response, with the electrical resistance of the skin being the primary parameter. The skin's moisture level, influenced by sweat gland activity, determines this resistance. Eye-tracking is a method that tracks how the subject visually follows specific stimuli. In this method, an eye tracker can be used, which includes an attachment to the eyeball, or it can employ optical,

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

contactless methods for measuring eye movement, as well as utilizing electrical potentials with electrodes placed in the subject's eye. The cardiovascular system activity method primarily involves analyzing the heart rhythm and its variability, blood pressure, resting heart rate frequency, and intervals between consecutive heartbeats. The method for monitoring the frequency and depth of breathing uses acoustic signal measurements, acceleration, temperature changes, and volume. Research using this method shows that both the breathing rate and depth increase as the cognitive demands of the task rise (Shi Y., Ruiz N., Taib R., Choi E. H. C., Chen F. 2007, Zhang X. 2021).

Behavioral methods are based on objective techniques and focus on determining the relationship between a stimulus and the resulting reaction. They combine engineering methods with psychological principles. The goal is to understand the relationship between the stimulus and the response. Physiological responses that are typically observed include EEG patterns, GSR, heart rate, stress hormone levels, eye motor functions, facial muscle movements, speech parameters, and walking dynamics. Emotions, especially those associated with fear, anger, etc., are also taken into account, as they significantly influence the likelihood of creating dangerous situations, particularly in drivers on the road. This provides a basis for implementing these methods in aviation for pilots and aircrew. External physiological traits can be recorded using infrared digital cameras. Proper interpretation of the emotions experienced by the subject allows for assessing their psychophysical state. Two main categories of emotions are distinguished: neutral and negative. Neutral emotions include satisfaction, while negative emotions encompass fear, sadness, surprise, disgust, and anger. According to Mitas, one's emotional and psychophysical state is influenced by factors such as the distance between the eyes and eyebrows, the degree of eye closure, and the curvature of the mouth. Hock states that in cases of anger and disgust, the eyes of the person experiencing these emotions are mostly closed, while the eyebrows are positioned low on the face. When feeling sadness or satisfaction, the eyes are partially open, and the distance between the eyes and the mouth is notably smaller in the case of satisfaction. When someone experiences surprise or fear, their eyes are widely open, but in the case of surprise, the distance between the eyes and the eyebrows is noticeably larger than when experiencing fear (Galant M. 2017, Razin P., Kruszewski M., Niezgoda M., Kamiński T. 2016).

2. Methodology and Limitation

The completed research involved analyses conducted primarily using quantitative methods. The aim of the research was to determine the precision and accuracy of performing specific pilot tasks in various situations (external variables) under simulated conditions. An FNPT II MCC flight simulator was used for the research, and the research group consisted of licensed aviation personnel with at least a PPL(A) license. The research was conducted on a group of six people. The study involved introducing emergency situations at various stages of flight, under different meteorological conditions. The analysis of risky decisions made by pilots was based on data recorded during the study and questionnaires completed by participants after the study. The questionnaire contained questions consistent with the NASA-TLX questionnaire.

All personal data and other information enabling the identification of participants has been redacted, ensuring complete anonymity for the subjects. The data obtained from the study was used for scientific purposes only.

Despite the valuable insights obtained, the study has several limitations. The small sample size reflects the exploratory nature of the project and does not allow the findings to be generalized. In addition, the use of exclusively subjective workload assessments limits the depth of interpretation, as no objective performance or physiological measures were included. While the scenarios were standardized, they cannot fully replicate the complexity of real-world flight operations. Nonetheless, the study provides an important foundation for further research and highlights the need to investigate pilots' cognitive workload using a broader set of methodological tools.

3. Result

Knowledge of individual pilot behavior and decision-making depends on many factors. These can be divided into external factors, such as atmospheric conditions, and individual factors, resulting from personal characteristics, such as the need to impress one's flying skills in a given social group. The performance of aviation activities inevitably involves a certain degree of risk.

The ability to respond appropriately in risky situations (hereafter referred to as emergency situations) depends on many factors. In the process of training aviation personnel, training skills to deal with various emergency situations in flight operations plays a key role. The preparation of flight crews is aimed at developing correct habits that enable them to act effectively in unforeseen situations during actual flight.

Every flight operation requires prioritizing safety, so flight training and procedures largely focus on this aspect. The FNTP II MC flight simulator was used as part of the study to assess the psychophysical state of pilots. Six pilots with PPL(A) licenses and various TFH (total flight hours) flight experience participated in the study.

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

In the first phase of the study, the test subjects were informed that an emergency situation would occur during the simulator flight, but they were not given information about its specific nature. The scenarios introduced included different types of emergency situations. After the simulation, participants completed a questionnaire developed from the NASA-TLX tool.

The first part of the questionnaire concerned basic information about the participants' aviation experience. It included four questions and an assessment of individual opinions on the usefulness of simulator training in the process of training aviation personnel. The next ten questions were based on issues identified by the NASA team and served as the basis for the survey analysis.

Emergency situations of a varied nature were introduced during the study, including aircraft malfunctions, adverse weather conditions and other undesirable events that may occur during the flight. The following situations were introduced:

- Situation 1 – altimeter deactivation with deterioration of visibility after takeoff;
- Situation 2 – turning off the speedometer as visibility deteriorates after takeoff;
- Situation 3 – go-around with one non-working engine;
- Situation 4 – failure of one of the engines during an IFR approach together with incorrect setting of the rudder trimmer;
- Situation 5 – fire in one of the engines after takeoff;
- Situation 6 – failure of one of the engines.

Respondents were asked to rate, on a 20-point scale, the various parameters included in the NASA-TLX questionnaire, where extreme values indicate: 1 – none; 20 – high. As the first question, respondents were asked to rate their fatigue before the survey, as shown in Figure 1.

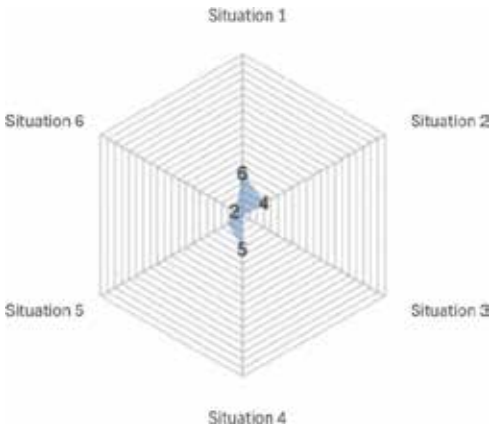

Figure 1. Fatigue level before the test. Source: the authors' own work.

Analysis of Figure 1 indicates that participants' fatigue levels were relatively low prior to the start of the study and the occurrence of emergency situations. On a 20-point scale, the values ranged from 1–6. Although fatigue may seem like a secondary factor in the context of risky situations, Dupont included it in his "scorching twelve." It negatively affects motor skills and can lead to reduced pilot performance, increasing the risk of safety violations during flight.

Then, they were asked to determine their subjective perception of workload during a given flight (Figure 2).

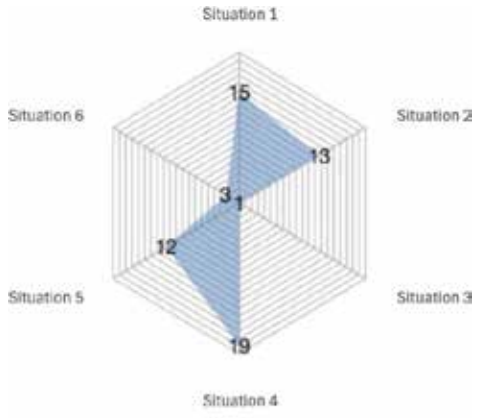

Figure 2. Workload level. Source: the authors' own work.

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

The level of workload felt by respondents during the emergency varied widely. As shown in Figure 2, the assigned values on the adopted scale span both the lowest and highest ranges. An analysis of this chart shows that the lowest values on the 20-point scale were assigned to situations 3 and 6, in which respondents rated their workload at 1 and 3, respectively. These included a flight in GA configuration with one inoperative engine and engine failure.

The highest rating (19) was given to situation 4, characterized by the highest complexity due to overlapping unfavorable factors, including difficult meteorological conditions, engine failure, and incorrect rudder trimmer settings. In the other cases studied, the ratings were similar and ranged from 12 to 15. The workload is directly related to the level of difficulty of a given task and depends on the type of emergency situation that occurs. In the studies conducted, the complexity of the failures that pilots had to face varied, as shown in Figure 3.

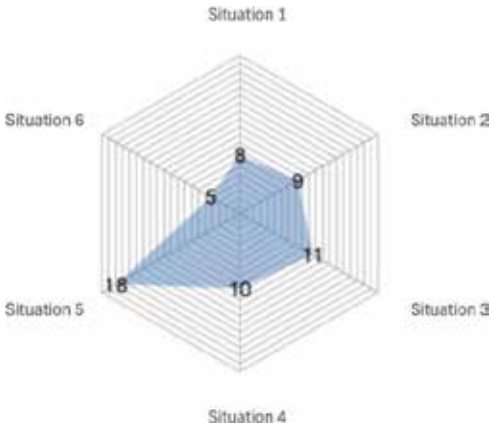

Figure 3. Degree of difficulty of the task. Source: the authors' own work.

Analysis of Figure 3 shows that most of the pilots surveyed rated the level of difficulty of the task performed during an emergency situation as low or medium. The values did not exceed 10 on the accepted scale. Situation 5, in which the pilot had to respond to an engine fire shortly after takeoff, is different. In this case, he rated the difficulty of the task as very high, assigning it a value of 18.

In emergency situations, the pilot's reaction time to conditions plays a key role. The degree of perceived time pressure, assessed in the next question, was due to breakdowns and malfunctions rather than pressure from superiors or the environment that might force immediate action. The results from this question are shown in Figure 4.

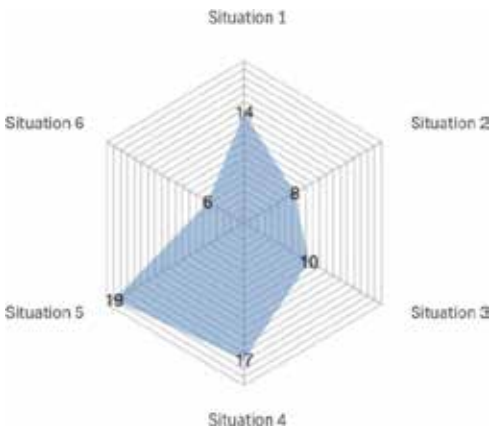

Figure 4. Perceived pressure. Source: the authors' own work.

Time pressure depends on the difficulty of the task and the risk involved in the situation. For this question, the ratings assigned on the scale varied. Pilots rated the level of time pressure as low in situations 2 and 6, while in the other four situations, the ratings were in the middle to upper range

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

In aviation, performance usually refers to the parameters that characterize an aircraft. In the context of this study, performance is understood as the end result of a given emergency situation. For the purposes of the study, it is not the performance of the simulated aircraft on the simulator that is relevant, but the results that the test pilots obtained in a given situation.

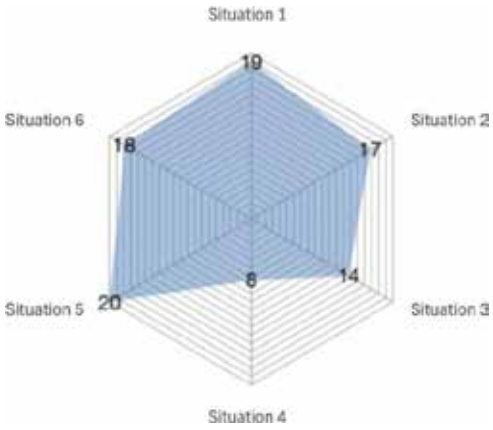

Figure 5. Performance. Source: the authors' own work.

An analysis of the chart clearly shows that five of the six pilots surveyed rated performance at the top or middle-upper end of the rating scale. In the case of situation 4, although the simulated flight ended with a successful landing at the selected airport in accordance with procedures and without major damage, the pilot rated the performance much lower. This question illustrates how subjective it is for pilots to evaluate the quality of flight performance. Although most of the ratings are close to each other, none were assigned by two pilots. Only in situation 5 was one pilot explicitly assigned the highest rating (20), considering his flight a complete success.

In the case of the mental effort put into the aerial task, the results are mostly close to each other. Only in situation 6 is the result closer to the lower range of the accepted scale. This situation, in terms of the individual evaluation of effort, is similar to the question on the evaluation of aeronautical performance in a given situation, where the pilot, in situation 4, rated performance significantly lower than the others.

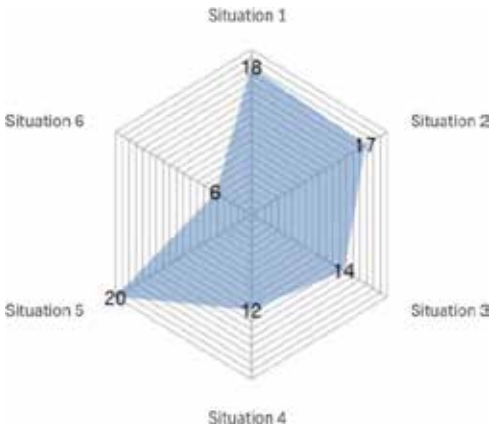

Figure 6. Situation type. Source: the authors' own work.

Based on Figure 6, most situations in the study received ratings in the middle to upper range of the rating scale, indicating that, for most pilots, the mental effort required to continue flight under difficult conditions was high. Only in the case of situation 6, concerning engine failure during the flight operation being performed, the ratings were at the lower end of the scale.

Frustration is one of the negative, difficult emotions a person can feel. The individual way of experiencing and feeling emotions is important from the perspective of the ability to cope with risky situations. From a psychological point of view, emotions and feelings are a characteristic of each person. In the context of this study, it also becomes important to examine this aspect of the frustration level.

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

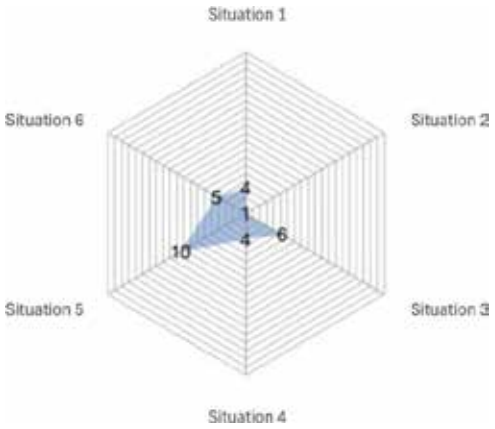

Figure 7. Frustration level. Source: the authors' own work.

An analysis of Figure 7 shows a low level of perceived frustration during an emergency. Frustration is also an individual factor that was included in the study presented in this paper. It is one of the key elements considered in emergency situations and risky situations in general. Higher levels of frustration can indicate greater difficulty with the task, but may also be due to the pilot's individual conditions.

In situation 10, in which the pilot surveyed rated the level of frustration at 10 on the accepted scale, it is not surprising that this situation directly threatened flight safety. A fire aboard an aircraft requires the pilot or crew to react immediately and land as soon as possible. Of course, this does not mean that the other situations did not require a proper approach or did not affect flight safety. However, in the other cases simulated during the study, these situations did not require immediate landing at the nearest airport. In the case of a malfunction of one of the engines in a multi-engine aircraft, it is possible to continue flying safely for an extended period of time, unlike in the event of a fire.

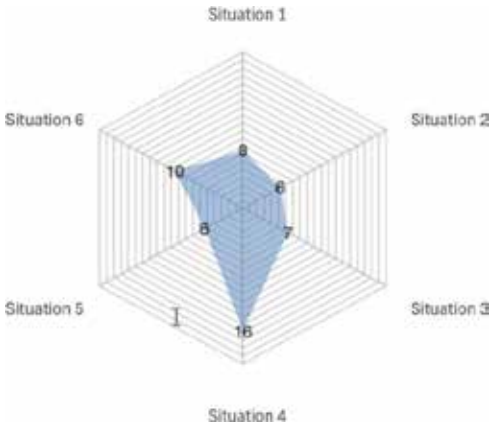

Figure 8. Stress level. Source: the authors' own work.

An analysis of Figure 8 shows that most pilots' stress levels were moderate, remaining at a medium-low level. There is, however, a significant difference in ratings for situation 4, in which pilots experienced pronounced stress related to the emergency situation. It is not surprising that pilots experienced stress in the face of such situations, but what is important is how they approached the flight going forward and how they performed their tasks. According to Figure 14, all simulated flights were successful. Nevertheless, in most cases, the pilots surveyed did not give the maximum rating on the scale. As with frustration, the perceived stress is influenced by each pilot's individual factors.

Fatigue is a result not only of the work environment and lifestyle, but also of the situation itself. In the penultimate question, pilots were asked to rate their level of fatigue after completing the survey, and thus after experiencing the emergency situation in question.

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

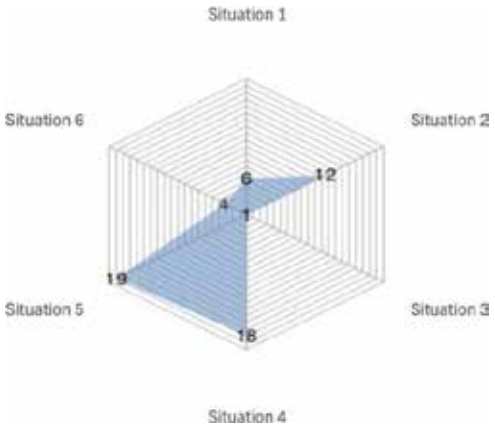

Figure 9. Fatigue level after the test. Source: the authors' own work.

Fatigue is a subjective feeling, as shown by the results in Figure 9. An analysis of the survey results shows that the ratings assigned in this question are evenly distributed. One part of the pilots reported their level of fatigue after a given emergency situation at the higher end of the scale, while the other part reported it at the lower end. It goes without saying that, in more complex or difficult tasks, fatigue levels may be higher. Such situations are especially the case when the flight involves risks that may adversely affect safety. In tasks with a high degree of difficulty, the entire process of performing the activity increases fatigue.

In the study, the least tiring situation was situation 1, in which pilots had to perform the GA procedure with one engine inoperative (rating 1). In contrast, the most fatiguing situation was situation 5, where the pilot surveyed gave a rating of 19. It is also important to examine how an emergency situation affects fatigue levels, so Figure 10 shows a comparison of this factor before and after the survey.

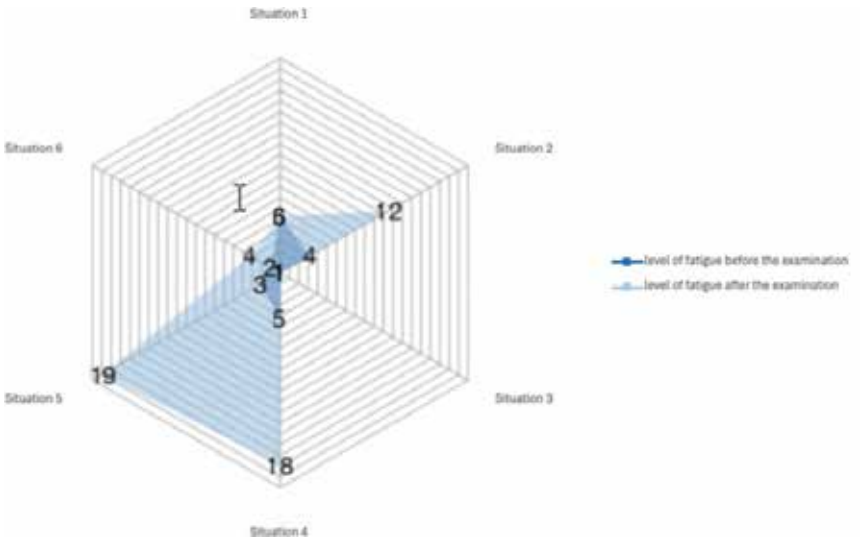

Figure 10. Comparison of fatigue levels. Source: the authors' own work.

In some situations, such as situations 1 and 3, the level of fatigue remained unchanged. In other cases, the level increased, with three situations showing a significant increase: in situation 2, fatigue increased from 4 to 12; in situation 4, from 5 to 18; and in situation 5, from 3 to 19. In situation 6, the increase was smaller, from 2 to 4. The survey results show individual perceptions of fatigue in emergency situations, as evidenced by the varying degrees of fatigue increase or lack of change across two situations.

The last question concerned the type of action to take in emergency situations. In aviation, it is crucial to follow current regulations. However, it is equally important to use one's own aviation experience. Flight training is designed to develop correct habits during normal and emergency situations. It is worth noting that as experience increases, the ability to respond effectively in various aviation situations improves significantly.

Safety & Defense 11(2) (2025)
DOI https://doi.org/10.37105/sd.264

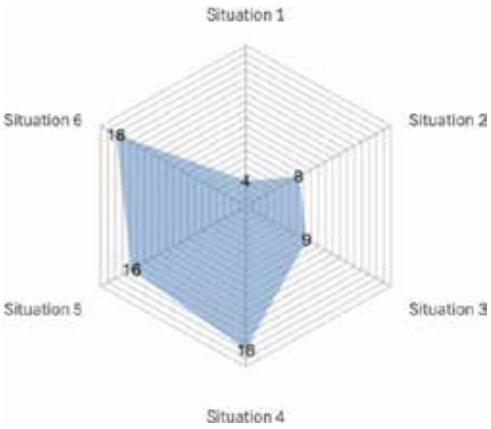

Figure 11. Type of action. Source: the authors' own work.

The actions pilots take in emergency situations largely depend on their flight experience (Figure 11). The greater the experience, the easier it is for pilots to focus on following regulations and procedures, as their movements become more automatic. Piloting activities and moving around in the cockpit become natural to him, making his actions more relaxed than those of less experienced pilots. This difference in attitude is also reflected in the pilots' individual ratings when asked to determine the type of action they would take in an emergency situation on a preset scale.

4. Conclusions

Performing flight operations always involves a certain degree of risk. Ensuring an adequate level of safety is an absolute priority for any flight. Therefore, it is extremely important to monitor the pilot's psychophysical state during flight operations. This paper addresses issues related to flight safety and the available methods for assessing the pilot's psychophysical state. Such conditions are associated with a shift from automatic, skill-based behavior to more effortful cognitive processing, which aligns with widely recognized models of pilot performance. The results also confirm the usefulness of the NASA-TLX tool for assessing perceived workload in simulated emergency scenarios. Future work should expand the sample size and incorporate objective measures, such as flight performance data and physiological indicators, to provide a more comprehensive picture of pilots' cognitive load in challenging operational environments.

A number of factors – both internal and external – are responsible for the implementation of studies to determine the accuracy, precision and overall assessment of a given flight operation.

Declaration of 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 article.

References

1. Galant M. (2017). *Ograniczanie ryzyka zagrożeń w lotnictwie ogólnym przez zastosowanie systemu monitorującego stan psychofizyczny pilota*. PhD dissertation. Department of Construction Machinery and Transport. Poznan University of Technology. Poznań.
2. Hart S. G., Staveland L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research). *Advances in Psychology*, vol. *52*.
3. Luximon A., Goonetilleke R. S. (2001). Simpli ed subjective workload assessment technique. *Ergonomics*, vol. *44*, No. 3.
4. Majer P. (2012). W poszukiwaniu uniwersalnej definicji bezpieczeństwa wewnętrznego. *Przegląd Bezpieczeństwa Wewnętrznego* 2014/4/7. Agencja Bezpieczeństwa Wewnętrznego.
5. Razin P., Kruszewski M., Niezgoda M., Kamiński T. (2016). Wybrane metody detekcji stanów zmęczenia u osób kierujących pojazdami. *Prace Naukowe Politechniki Warszawskiej. Transport*, vol. *114*.
6. *Safety Management Manual*. (2018). Doc 9859. Fourth Edition. Internation Civil Aviation Organization.
7. Shi Y., Ruiz N., Taib R., Choi E. H. C., Chen F. (2007). *Galvanic Skin Response (GSR) as an Index of Cognitive Load, Conference: Extended Abstracs Proceedings of the 2007 Conference of Human Factors in Computing Systems*. California.
8. Zhang X. (2021). *Driver Mental States Detection during Highly Automated Driving by Decoding Brain Signals*. Technischem Universitat Berlin. Berlin.