BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING
FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF AEROSPACE ENGINEERING
LETECKÝ ÚSTAV

HAPTIC AND VISUAL JOYSTICK GUIDANCE

HMATOVÉ A VIZUÁLNÍ NAVÁDĚNÍ JOYSTICKU

MASTER’S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR **BSc. Leonardo Manuel Vargas Lambarry**
AUTOR PRÁCE

SUPERVISOR **doc. Ing. Pavel Zikmund, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2025

Assignment Master's Thesis

Institut:	Institute of Aerospace Engineering
Student:	**BSc. Leonardo Manuel Vargas Lambarry**
Degree programm:	Aerospace Technology
Branch:	no specialisation
Supervisor:	**doc. Ing. Pavel Zikmund, Ph.D.**
Academic year:	2024/25

As provided for by the Act No. 111/98 Coll. on higher education institutions and the BUT Study and Examination Regulations, the director of the Institute hereby assigns the following topic of Master's Thesis:

Haptic and visual joystick guidance

Brief Description:

This work is focused on research on the interaction between the pilot and the aircraft. With the development of systems and electronics, there is an opportunity to implement "Human–centered design" elements into the cockpits of small aircraft. This work is related to this, as it involves comparing visual and tactile guidance methods and verifying their impact on the pilot's workload.

Master's Thesis goals:

The objectives of the thesis follow the assignment within the industrial project. The following tasks will be completed as part of the thesis:
– conducting the experiment in combined visual and haptic joystick guidance.
– evaluation of additional task effect on guidance methods.
– statistical evaluation of the experiment.

Recommended bibliography:

STANLEY, Andrew A.; KUCHENBECKER, Katherine J. Evaluation of tactile feedback methods for wrist rotation guidance. IEEE Transactions on Haptics, 2012, 5.3: 240-251.

WICKENS, Christopher D. Processing resources and attention. In: Multiple task performance.CRC Press, 2020. p. 3-34.

OLIVARI, Mario, et al. An experimental comparison of haptic and automated pilot support systems. In: AIAA modeling and simulation technologies conference. 2014. p. 0809.

Deadline for submission Master's Thesis is given by the Schedule of the Academic year 2024/25

In Brno,

L. S.

doc. Ing. Jaroslav Juračka, Ph.D.
Director of the Institute

doc. Ing. Jiří Hlinka, Ph.D.
FME dean

Declaration of authenticity

I declare that I have prepared this thesis independently and that I have listed all sources used in accordance with the applicable copyright law in the list of used literature and sources of information.

Place, date

Signature

Acknowledgement

To my beloved family, my wonderful girlfriend and my devoted friends, thank you for your support, patience, and love during this stage of my life.

I would also like to thank Doc. Ing. Pavel Zikmund, Ph.D. for the professional supervision of my graduate thesis.

Abstract

LAMBARRY, Leonardo. Haptic and visual joystick guidance. Brno, 2025. Master's thesis diploma project. FSI VUT v Brně. Thesis supervisor: doc. Ing. Pavel Zikmund, Ph.D.

With the growing interest in multisensory navigation systems in aviation, the need to explore its benefits and applications has become more relevant. Mainly because it can contribute to the mitigation of risks associated with pilot workload in flight environments. This thesis explores the performance of visual, haptic, and combined guidance modalities in simulated aviation tasks under varying levels of cognitive load. The key metrics for the study are average error (AE), time to reach target position (TRTP) and the perceived workload. The device used for this experiment is a joystick haptic device, with directional feedback through a motorized sliding element beneath the user's finger. For the experimental methodology, twelve participants performed navigation tasks using visual, tactile, and combined cues within a flight simulator developed in MATLAB and LabVIEW. Statistical analysis using general linear model repeated measures ANOVA showed that visual guidance consistently outperformed the other modalities, producing faster and more accurate responses with lower workload. Combined guidance offered occasional performance enhancement but increased cognitive demands. Haptic feedback was associated with a higher error and greater perceived workload, especially under cognitive stress. These findings highlight the importance of multisensory interface design and workload management in aviation. This is essential for maintaining safe flight operations. Limitations due to consumer grade hardware and the absence of physiological workload monitoring were acknowledged.

Key Words: haptic guidance, workload, flight simulator, joystick, multitasking, aviation safety.

Contents

1. Introduction

Tactile navigation in the aviation industry has transformed the way pilots interact with flight controls and instrumentation by increasing the safety and precision of airborne operations. By integrating tactile feedback mechanisms into cockpit interfaces, such as control yokes, throttles, and instrument panels, pilots can receive sensory cues that enhance situational awareness and facilitate intuitive decision making while airborne. This technology enables pilots to effectively interpret flight data without compromising visual attention from the external environment, allowing them to maintain spatial orientation and execute precise maneuvers, particularly in adverse weather conditions or low visibility scenarios. As tactile navigation continues to develop, its integration into modern aircraft design represents significant progress in aviation safety and efficiency, improving the performance of pilots and aircraft systems at the same time.

This research will explore the efficiency of visual guidance when compared with tactile. The interactions between the two sensorial systems will also be considered and analyzed. The hypothesis is that for basic tasks visual guidance will be more effective, but as the workload of visual modality starts to increase, its performance will get worse. Tactile guidance will take over in terms of efficiency once the workload of the pilot is increased, as the pilot will start using multisensorial information processing. As Wickens suggests, vibrotactile direction signals may support subtasks like navigation, and reduce the overall workload [1].

The intention of this research is to evaluate the performance of haptic guidance devices for navigation. Several papers have proven already how tactile systems often show a shorter reaction time when compared to other sensorial cues. This has been tested under multiple conditions and for different purposes. A. Chan and Matthew S. Prewett for instance conducted similar investigations into the matter, their publications will be analyzed later since their results are relevant for this study. Nevertheless, reaction time is not the only aspect that must be considered to evaluate the performance of a system. When it comes to navigation, parameters such as the time to reach the desired position and the average error of the responses must be measured as well. It is not possible to ignore any parameter, since all of them may determine the fulfillment of the pilot's tasks.

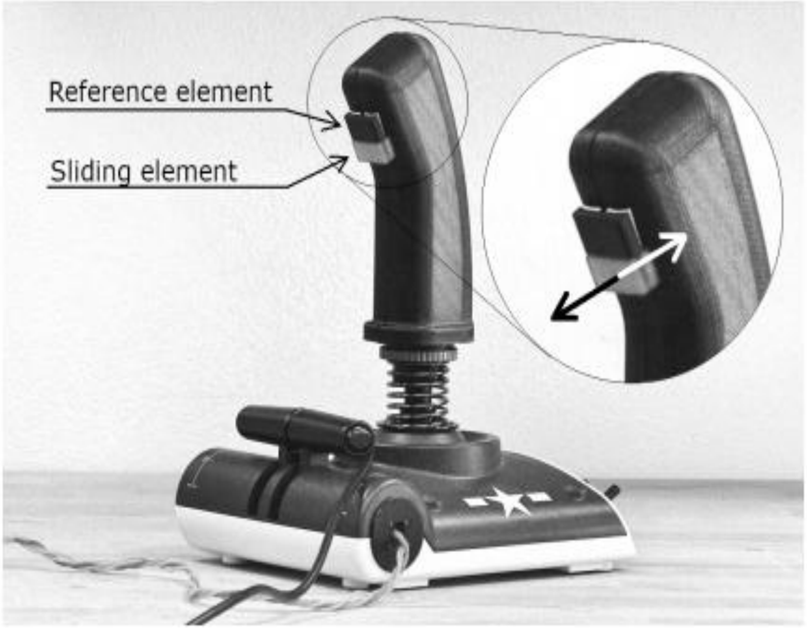


Fig. 1 Joystick with tactile guidance device attached [2]

The device used for this experiment consists of a joystick that has attached a sliding element located under the operator’s index finger. This sliding element performs a translational movement front or back to indicate a change in trajectory and gives guidance instructions. The joystick has a reference element on top of the sliding element with the same shape and size. On the neutral position, there is a null offset between them. Once the sliding element moves in or out of the handle surface of the joystick, under the user’s fingers, the subject will need to move the handle in the corresponding direction to match the sliding element with the reference. The movement description and the elements of the joystick can be seen in *figure 1*. The size of the position difference between the sliding element and the reference front surfaces is proportional to the joystick deflection required to reach the target position. The sliding element is powered by two SG90 digital servo motors, each providing a maximum force of 20 N, with a movement range of 8 mm. [3]

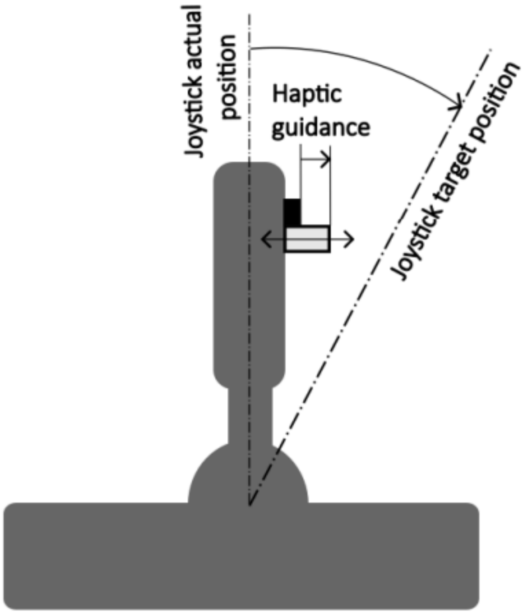


Fig. 2 Illustrative demonstration of haptic guidance function [2]

A previous research conducted by P. Zikmund about the learning effect in joystick tactile guidance already proposed which variables must be measured to properly evaluate the performance of the device. [2] These variables will be the time to reach the target position TRTP, measured from the target position generation to the first reach of the target position tolerance interval. The second criteria is the average error AE between the target and the joystick position, measured from the first target interval until the new desired position is generated. The joystick range is from -1 to +1. AE gives the average difference between the target and actual position. Theoretically it can reach a max. value of 2. Hence, it will be considered as a nondimensional unit, since the joystick range is normalized. An illustrative demonstration of the joystick range is shown in *figure 2*. In Addition to TRTP and AE, a new variable will be included, which is the workload level. This can be explained as the mental and physical effort perceived by the user during the realization of a certain task.

This study is part of an ongoing research led by professor P. Zikmund. In previous stages of the project, similar works measured the performance of the proposed device. However, since this investigation also considers visual guidance, the experiment methodology will have additional steps. This prototype is intended to assist pilots when it comes to navigation, hence it is important to simulate the flight conditions that the pilot faces. For that purpose, a flight simulator was developed using LabVIEW software. The interface was designed so that the participants get immersed in the tasks realization to resemble the flight environment.

An important clarification is that this thesis is part of a collaborative project. Assistance during the execution phase was provided by Jan Bredemeyer, a fellow classmate, and professor P. Zikmund. J. Bredemeyer supported the experimental methodology by helping carry out participant sessions, monitoring the proper saving of data, and assisting in exporting and organizing the raw data into Excel format. These collaborative efforts ensured the integrity and reliability of the collected data. P. Zikmund assisted as well by carrying out some of the participants' sessions. He provided the haptic device as well and the MATLAB code responsible for saving raw data and generating the random target positions for the tracking tasks. The MATLAB code was used in previous stages of the project, and it was modified to fulfill the needs of this study. Overall, the execution phase required the collective effort of everyone involved, mainly due to the coordination of the participants. Everything else presented in this work was done by the author.

2. Background research

2.1 Tactile stimuli & reaction time

Understanding the basis of tactile perception is fundamental when exploring guidance systems and human machine interactions. This subchapter explores how tactile perception affects reaction time in interactive tasks. It covers the basics of haptic sensing, the variability of response times to different stimuli, and how consumer grade technology may affect measurement accuracy. It also discusses the benefits of multisensory feedback, and the learning effects associated with repeated exposure to tactile guidance.

2.1.1 Overview of tactile perception

The sense of touch encompasses a complex system involving various receptors in the joints, muscles, and skin, each responding to different stimuli and contributing to perceptual qualities like mass, pressure, and texture. Despite common beliefs that vision and hearing dominate our understanding of the world, cases have demonstrated the significance of touch in communication and perception. Mechanoreceptors in the skin specialize on different sensations such as light touch, pressure, and vibration. Additionally, muscle and skeletal mechanoreceptors provide information on limb position, movement, and external forces, crucial for balance and interaction with the environment. Overall, touch perception is a process involving a combination of sensory inputs from various receptors, shaping our understanding of the world around us. [4]

Haptic devices serve as interfaces for processing both input and output information, often complementing visual or auditory displays for feedback. However, when used independently, they require careful design to ensure effective interaction. Key components of haptic interaction include indicating control points, executing commands, and providing feedback to users, utilizing physical attributes like force, shape, texture, and temperature. Various interaction techniques facilitate direct correspondence between real and virtual environments, such as moving pointers, objects, or applying forces. While virtual environments offer limitless possibilities, it is necessary to take precautions when it comes to the design of guidance cues, as they can enhance experienced users' understanding while potentially confusing naive users. Clear guidelines are essential to maximize the clarity of information presented to users. [5]

A different perspective that must be considered with respect to the tactile feedback on the accuracy of guidance is presented by Voudouris [6]. This author states that tactile stimuli on a moving hand are consistently suppressed, possibly due to the brain's limited capacity to process irrelevant sensory input. The author set out to explore whether individuals can simultaneously enhance tactile signals relevant to movement while executing purposeful reaching actions. Through conducted experiments, it was observed that participants could dynamically adjust tactile sensitivity by concurrently suppressing irrelevant signals and amplifying those pertinent to movement during goal reaching tasks.

Juravle and Spence focused on sensory suppression within intricate motor activities like juggling. Their study tasked participants with detecting interruptions in continuous signals across various modalities (haptic, auditory). The findings revealed a notable reduction in participants' ability to discern gaps in tactile stimulation while engaged in juggling, illustrating movement induced suppression of tactile sensation. This suggests that humans may initiate tactile suppression in the brain prior to executing motor commands. Such tactile suppression could pose challenges for employing tactile guidance in aviation, as individuals may inadvertently suppress tactile cues while focusing on other demanding tasks. Temporary change in sensitivity to tactile stimuli should be registered in the experiment, since participants will be in contact with the sliding element for long periods. If this factor has adverse effects on the performance of haptic guidance over time, it will be seen within the sessions. [7]

2.1.2 Sensorial Stimuli Reaction Time

Sensorial stimuli are nowadays commonly used for enhancing the usability and effectiveness of human machine interfaces. By incorporating visual, auditory, and tactile stimuli it is possible to provide users with more intuitive and engaging interactions. Visual stimuli, such as graphical user interfaces and augmented reality displays, enable users to process information quickly and make informed decisions. Auditory feedback, including alerts and notifications, enhances situational awareness and aids multitasking scenarios. Tactile stimuli, such as haptic feedback, provides users with physical feedback, improves precision and enhances the sense of immersion. Together, these sensorial stimuli contribute to the creation of more intuitive human machine interfaces that enhance the user experience and efficiency across various domains, from consumer electronics to industrial control systems. Just as an example, lately there has been an increasing use of tactile stimuli modality. It improves the reaction time of aerial vehicle operators. When used in tactile navigation displays it has shown to result in better driving performance and to reduce workload for the driver.

There are several studies that compare response times to visual, auditory and tactile stimulus in human machine interactions. The most relevant result showed that the time response to tactile stimuli is faster compared to the other two cases. On average, the reaction time of tactile stimuli was 28% shorter than the auditory, which correspondingly is 5% shorter compared to visual stimuli. These studies not only reveal relevant data regarding the response to different kinds of stimuli, but also the influence that diverse factors may have on the reaction times, altering them considerably. Just as an example, for the case of tactile stimuli, the mentioned study used a vibrator to send the signal to the participants. This device was worn by the participants in different parts of the body, to study if this had any influence. After several successful evaluations where the device was placed randomly either on the wrist or the ankle of different test subjects, it was concluded that the location of the tactile stimuli has no influence over the reaction time. [8]

Regarding the test subjects of the study, the influence that certain characteristics of the participants may have on the results was also analyzed. Age, education level and time spent on the computer are some of the parameters considered when conducting the tests on diverse participants. The procedure of the study consisted in asking the participants to press a button on a keypad with their finger after receiving the stimuli. For the visual modality the signal was displayed on a computer screen. As for the auditory signal, it was emitted from a built-in speaker. Lastly, the tactile signal was produced by a vibrator worn directly on the skin by the participants.

Age has a significant influence on response times for all kinds of stimulus types. This was proven by classifying the participants with respect to their age into four different groups. Each group had a range of 10 years, in the following manner: 11-20, 21-30, 31-40, 41-50. General results showed that reactions improve with an increase of age up to 21-30 years. From this point on, the reaction time increases in relation to age, showing the slowest reactions with the 41-50 years old group.

For the education classification, three different groups were made, one for each level of education that participants had concluded. The first group was for primary education, which refers to the initial stage of formal education typically provided to children for the foundation of basic skills and knowledge acquisition to develop essential cognitive, social, and emotional abilities. Secondary education corresponds to the second group and consists of academic knowledge, critical thinking skills, and practical competencies. Tertiary education group is the last classification, and emphasizes critical thinking, advanced skills development, and specialized knowledge acquisition relevant to specific disciplines or industries. This classification revealed that with higher education level; participants had shorter reaction time for all kinds of stimuli. Tertiary education group had considerably faster reactions compared to the secondary group. Correspondingly, the secondary group was faster in comparison with the first education group.

Another aspect that was evaluated as well is the time each subject spends on the computer. For this analysis, each participant had to answer how much time they spend daily on the computer by choosing between the next options: less than 2 hours, 2 to 4 hours, 4 to 6 hours, more than 6 hours. This last evaluation may be seen as the time any test subject spends doing any activity. Being more familiar with the system due to extensive use of it will have an influence on how fast a user may perceive and respond to the signals. This test showed that in general, the longer time spent on the computer in daily life, the shorter the response time. [8] The results after measuring all the response times with its standard deviation can be observed in *table 1*.

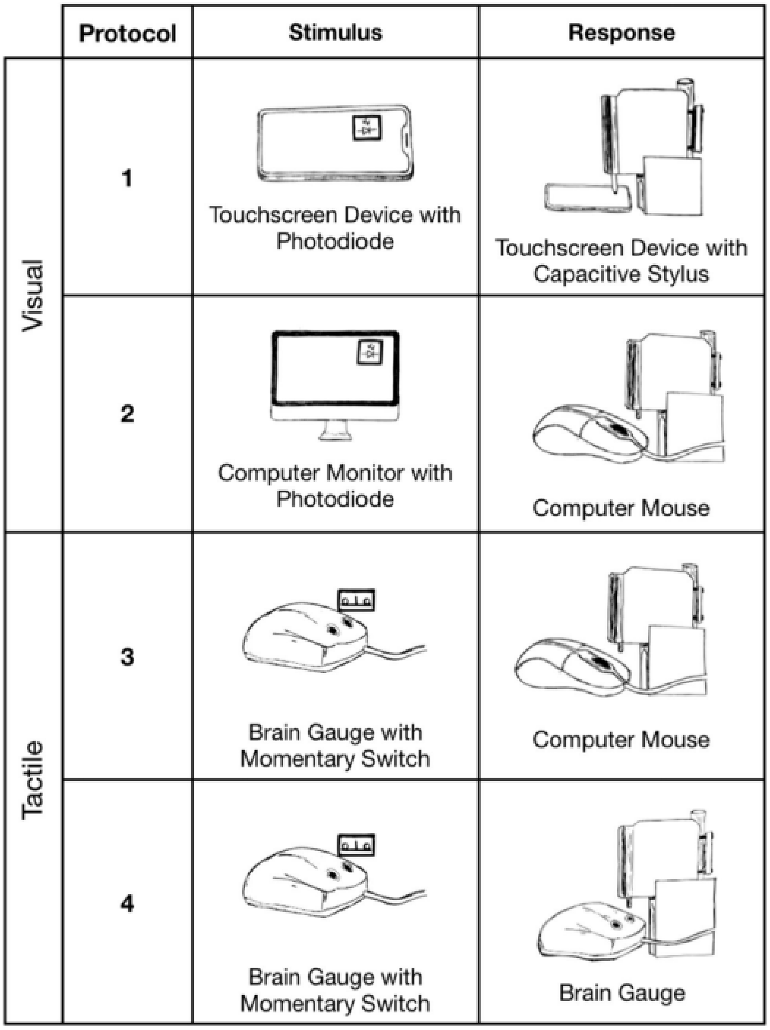
Table 1: Response time and standard deviation for each modality [8]

Stimulus Modality	**Response time**	**Standard deviation**
Tactile	0.385	0.071
Auditory	0.493	0.178
Visual	0.517	0.181

2.1.3 Reaction Time Delay due to consumer grade technology

Reaction time tests carried out in consumer grade technology commonly follow all the same procedures. They consist of presenting a visual stimulus on a computer monitor and asking the test subject to react to the signal by pressing a keypad or computer mouse as fast as possible. Yet there are inherent delays and variables with reaction time tests which mainly rely on consumer grade technology. These delays and variables are associated with the computer specs and often result in significant errors in reaction time measurements. To demonstrate the influence that software and hardware may have over reaction time tests, a study was carried out to compare reaction times using different device configurations with the same non-human interface to completely discard the human bias delay.

Two different types of stimuli were tested, tactile and visual, with two configurations for each modality, making a total of four different conditions. The first visual configuration used a computer monitor to send the signal, a photodiode for the detection of it and lastly a computer mouse to deliver the mechanical response. The second visual condition used a touchscreen device for the signal, a photodiode once again as a receptor and a touchscreen device with a capacitive stylus to send back the response. For the tactile modality, the signal was sent using a brain gauge device, which is a commercial tactile simulator that delivers vibrotactile stimuli and can send back responses as well. Both the third and fourth remaining configurations use this device to send the signal, and a momentary switch to detect it. The only difference is that for the third case the mechanical response is sent back using a computer mouse, while the fourth case uses the same brain gauge as a response device. [9] All the protocol descriptions for the visual and the tactile stimulus and responses can be seen in *figure 3*.



	Protocol	Stimulus	Response
Visual	1	Touchscreen Device with Photodiode	Touchscreen Device with Capacitive Stylus
	2	Computer Monitor with Photodiode	Computer Mouse
Tactile	3	Brain Gauge with Momentary Switch	Computer Mouse
	4	Brain Gauge with Momentary Switch	Brain Gauge

Fig. 3 Reaction time protocols [9]

The Brain gauge is a device designed specifically to measure the reaction time and/or reaction time variability of the tested subject. Even if it is commercially available, it is often used as a laboratory research tool due to its precision and accuracy. It is a laboratory grade hardware, and the reaction times measured from this device will be used as reference.

The reaction time test conducted by consumer grade technology is dependent on the CPU processing speed, which can vary depending not only on the number of cores, but also on the RAM memory of the device, and the programs running in the background. These variables often introduce delays of 15 msec without considering possible unnoticed malware. Even with the same hardware, the device firmware may also have a huge influence on the reaction delays. The use of wireless or USB peripherals connected to the computer also introduces a latency due to the communication protocols. Touch sensing peripheral hardware usually employs a variety of methods related to the sensing mechanism, having different response lag values each and varying from 50 to 200 msec. The last consideration that must be recalled is the refresh rate of the monitor screen used. This may affect mainly the stimulus delivery time. [9]

To entirely remove the human factor delays, automated hardware was implemented to perform the reaction time tests. For each one of the four conditions, a sensor was placed to perceive the stimuli. A mechanical device was then configured to respond automatically after the signal was detected with a fixed delay of 100 msec. The mechanical device is triggered at the middle point of the pulse, with an expected delay of 5 msec. For the four conditions a total delay of 105 msec is then expected.

Results showed different average latencies for each condition, even when the reaction time response was entirely automated with an additional fixed delay. This means that the latency for each configuration depends entirely on the software and hardware limitations for all the devices used. The average latencies measured in milliseconds for all types of responses are graphically summarized in *figure 4*. Visual stimulus with touchscreen response devices had the highest latency with 399 msec. This same stimulus with a USB mouse response instead reduced the latency to 80.1 msec. Tactile stimulus modality seems to have a smaller latency, with the USB mouse response showing a reaction time of 30.7 msec. As expected, the configuration with the smaller reaction time error of 5.6 msec was the fourth and last, with the use of a laboratory grade hardware to generate and respond to the signal. [8] Results also showed a variability for each condition. There resides the importance of setting up several trials to reduce the variability effect over the experiment results.

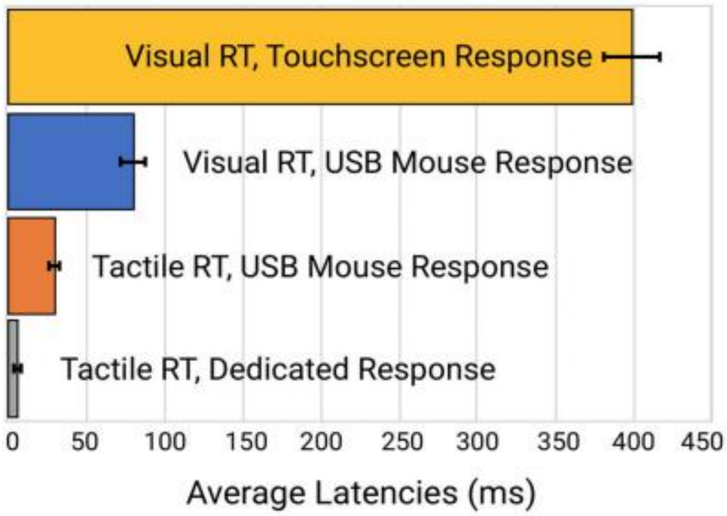


Fig. 4 Latencies for each hardware configuration [9]

2.1.4 Task enhancement with multisensorial signals

A variety of studies have investigated the impact that vibrotactile signals have on task performance. Modern systems tend to increase their complexity due to the amount of information they are capable of processing. As technology improves, systems become more interconnected and interdependent. This connectivity allows for greater functionality but also introduces more challenges for the operator, causing a task overload in most cases.

A key solution to reduce the overall workload is to efficiently display the information. The use of multisensory devices, for example, will allow the user to keep the same amount of information received by the user while reducing the cognitive workload. Good performance is achievable then by reducing the reaction time to signals or alerts, thus improving decision making. Multisensory devices can provide information through visual, tactile and auditory channels. Each of them offers its own advantages over the others, and their implementation will highly depend on their purpose. For the case of this study, it is important to recall if vibrotactile feedback will somehow improve the interpretation of information when accompanied by visual signals or by substituting them.

A meta-analysis conducted by M. Prewett proposed three different manners in which it is possible to evaluate if tactile feedback enhances tasks performance and information processing. The purpose of this meta-analysis was to have a quantitative review to summarize multiple experimental studies which examine the use of vibrotactile signals in different scenarios for task performance. Certain criteria were followed so that all chosen papers will be relevant for the quantitative review fulfilling contents and statistics. [10]

The first comparison is made between a standard condition in which only visual signals are being employed to carry out the tasks, and the same condition but with tactile signals added to the previous ones. Since tactile signals are processed faster and easier, the visual overload is reduced while performing the task, improving its performance. This comparison showed that in the absence of other types of stimuli, vibrotactile signals are effective to support tasks like visual search, navigation, target acquisition and piloting.

The second analysis is done between a standard condition with only visual signals employed, compared with the same condition but with visual cues being replaced entirely by tactile signals. The main aspect to evaluate in this configuration is the response time between the type of stimulus. Since the information given is the same for the task, it's only possible to analyze which one will have a faster reaction time. For this configuration, the results were inconclusive, since the effectiveness of substituting visual stimulus with vibrotactile will depend entirely on the type of information transmitted and the context in which it is being used. In cases where visual information is easy to read and to comprehend, the vibrotactile replacement shows no benefit. However, when the tasks have a higher workload and require more attention, the vibrotactile signals are more effective. The same applies to situations in which the visibility is low.

The third and last configuration uses both tactile and visual signals simultaneously to display the same information and then compares it with a standard condition employing just visual signals. Redundant signals will create a multisensory environment, increasing efficiency with simultaneous non conflictive tasks. This comparison showed a slight increase in performance for multisensory conditions when it comes to directional signals. [10]

There are two main theories that discuss the benefits of multisensory display of information. The Multiple Resource Theory from Wickens focuses more on workload, and how multisensorial channels deal with this information. The Prenav model by van Erop on the other hand, emphasizes how to efficiently reduce the cognitive overload, by introducing automatic or intuitive responses.

The Multiple Resource Theory (MRT) proposed by Wickens suggests that humans have multiple channels of information processing resources, which can be categorized into visual, auditory, and tactile. These channels operate simultaneously but have limited capacity, leading to potential interference when multiple tasks use the same resource. Yet, multisensory displays are effective with multiple non conflicting tasks. Wickens' theory emphasizes the importance of considering the demands placed on different channels of information when designing interfaces or tasks to optimize human performance and minimize errors. This theory has been studied in relation to the improvement of pilot performance. It has been shown how information from a sensory channel may be efficiently offloaded to another one. With pilot context, while realizing visual scanning, vibrotactile direction signals may support subtasks like navigation, reducing the overall workload. [1]

The Prenav model, developed by van Erp, studies human performance in multitasking scenarios. It suggests that humans have a limited capacity for processing information, and this capacity should be divided among different tasks based on their priority levels. The model has three key components: perception, cognition, and action. Perception refers to the gathering of information, followed by cognition which consists of processing and interpreting the gathered information. Lastly, action involves the execution of the responses. The Prenav model focuses on the importance of having task priorities and to properly use the cognitive resources to optimize performance in multitasking situations. It also proposes that vibrotactile senses are highly intuitive with fast reaction times, which may reduce cognitive overload in activities that could be automated, like navigation. [11]

2.1.5 Tactile guidance learning effect

Before testing the overall impact in performance during pilot aircraft interactions due to tactile guidance methods, it is also important to recall the influence that learning effect may have on it. Pilots rely on muscle memory, which means that the performance of haptic feedback may rely on the learning phases. Sometimes the lack of participants training increases results variance during comparative experiments. [12]

A laboratory experiment introduced a method to measure the effect on the learning effect. This helps to understand the learning curve, and to determine the necessary training to properly conduct comparisons between haptic and visual guidance for aircraft control. Only this way is possible to confirm if there is a benefit for introducing haptic feedback.

One of the parameters which influence performance during training is the time intervals between each training session. Wang et al. [13] did an investigation about the training time interval duration influence on a tactile orientation task. The experiment consisted of two groups training with different time intervals. The first group trained once a day while the second one did it once a week. Results showed that both groups have the same improvements after five sessions. From this experiment it was concluded that training intervals may be flexible if sessions are planned and consistently followed.

Other experiments of similar nature show how multiple sessions within the same day may be counterproductive due to fatigue and overtraining, leading to negative effects on perceptual learning. The consolidation of learned information during sleep must take place to prevent such deficits in learning. The experiment focuses on tactile guidance accuracy and reaction time, and how these two parameters are affected by training sessions. The test procedure consisted of asking twelve undergraduate and graduate volunteer students aged 19 to 26 years to guide the joystick to 20 randomly generated positions without visual feedback. For the first task each position was generated as a constant random position with a uniform distribution over the joystick front back travel range with a uniform duration between 3 and 6 seconds. The second task had instead continuously changing target positions. All participants were asked to use their dominant hand. [2]

The subjects repeated the experiment in 12 sessions with a break between sessions of at least 8 hours. The mean time between sessions was 45 hours, since there were weekend breaks between sessions. Each session included both tasks described previously. For both types of tasks is possible to observe how the average error between the target and the actual joystick position varies over time. This is graphically shown in *figure 5* for task 1 and *figure 6* for task 2. After every session the participants were requested to fill in a questionnaire where they had to indicate their workload for each task using the Bedford scale. This consists of a range from 1 to 10, 1 meaning insignificant workload and 10 meaning that it is not possible to complete the task.

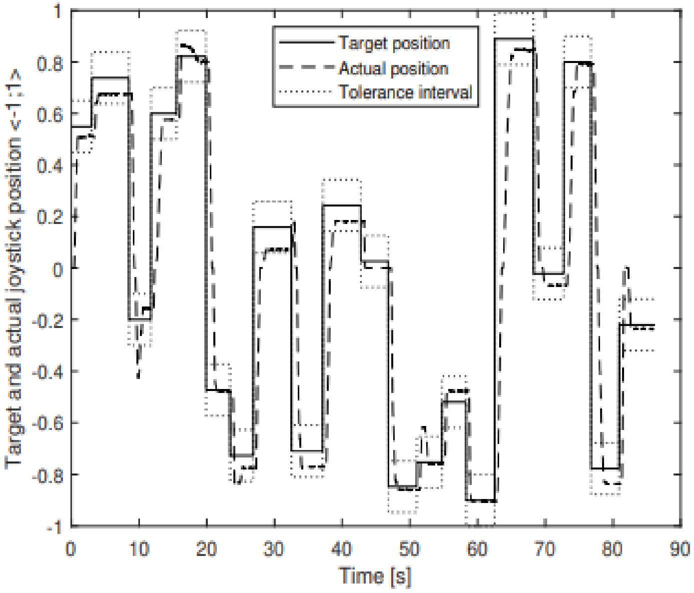


Fig. 5 Set of measured data from task 1 [2]

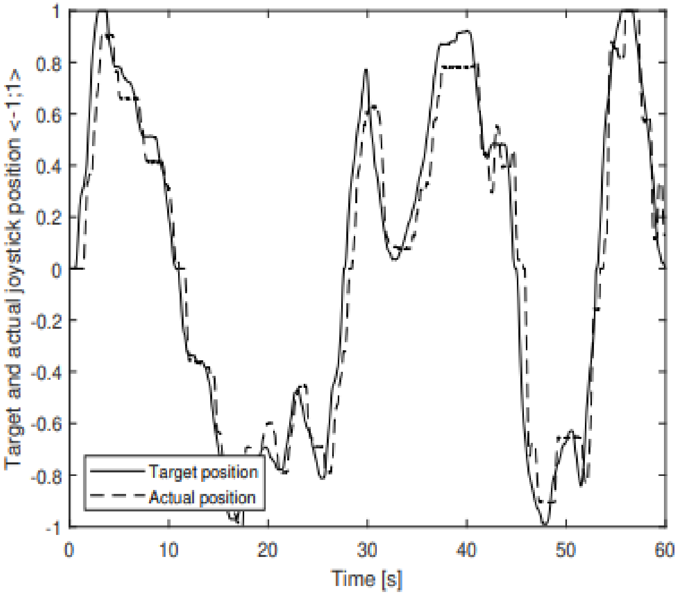


Fig. 6 Set of measured data from task 2 [2]

The learning effect was evaluated with the time to reach the target position (TRTP) and the average error (AE) with respect to the desired position. The TRTP was measured from the time of the target position generation and the time it took the participant to reach it. A tolerance valid interval of +-5% was set of the joystick range around the target position. As an additional parameter, the reaction delay (RD) was also measured, and represents the time interval between the generation of the new target position and the beginning of the response. The effect of training interval duration between sessions was not evaluated for this experiment, but every interval included at least one night, to provide sufficient time for knowledge consolidation and to prevent overtraining.

The mean values of all parameters were analyzed by one-way repeated measures ANOVA to investigate the influence of the learning effect. The result showed that the reaction delay (RD) was not significantly affected by training. The perception of the tactile input was estimated at 0.241 s. The time to reach the target position (TRTP) did not show a significant learning effect either. The average error (AE) was observed to be the parameter with the most significant learning effect up to session 6th for task 1 and to session 7th for task 2. After these sessions, the participant's improvement seems to be insignificant. The experiment also revealed an improvement in the workload evaluation after the training sessions. [2]

The parameters used to measure performance improvements on the pilots during tactile guidance training include accuracy, reaction time, workload and situational awareness. When comparing visual and tactile navigation tasks with respect to situational awareness, the research concluded that visual display supports global awareness, while tactile display supports local guidance instead and has a lower mental workload rating. [14] In general humans have shorter reaction times to tactile stimuli than to visual one, which motivates the application of tactile methods in the aviation domain. Since tactile interactions are suited for local guidance, by reducing the mental workload in combined tasks it is possible to achieve better situational awareness. [15]

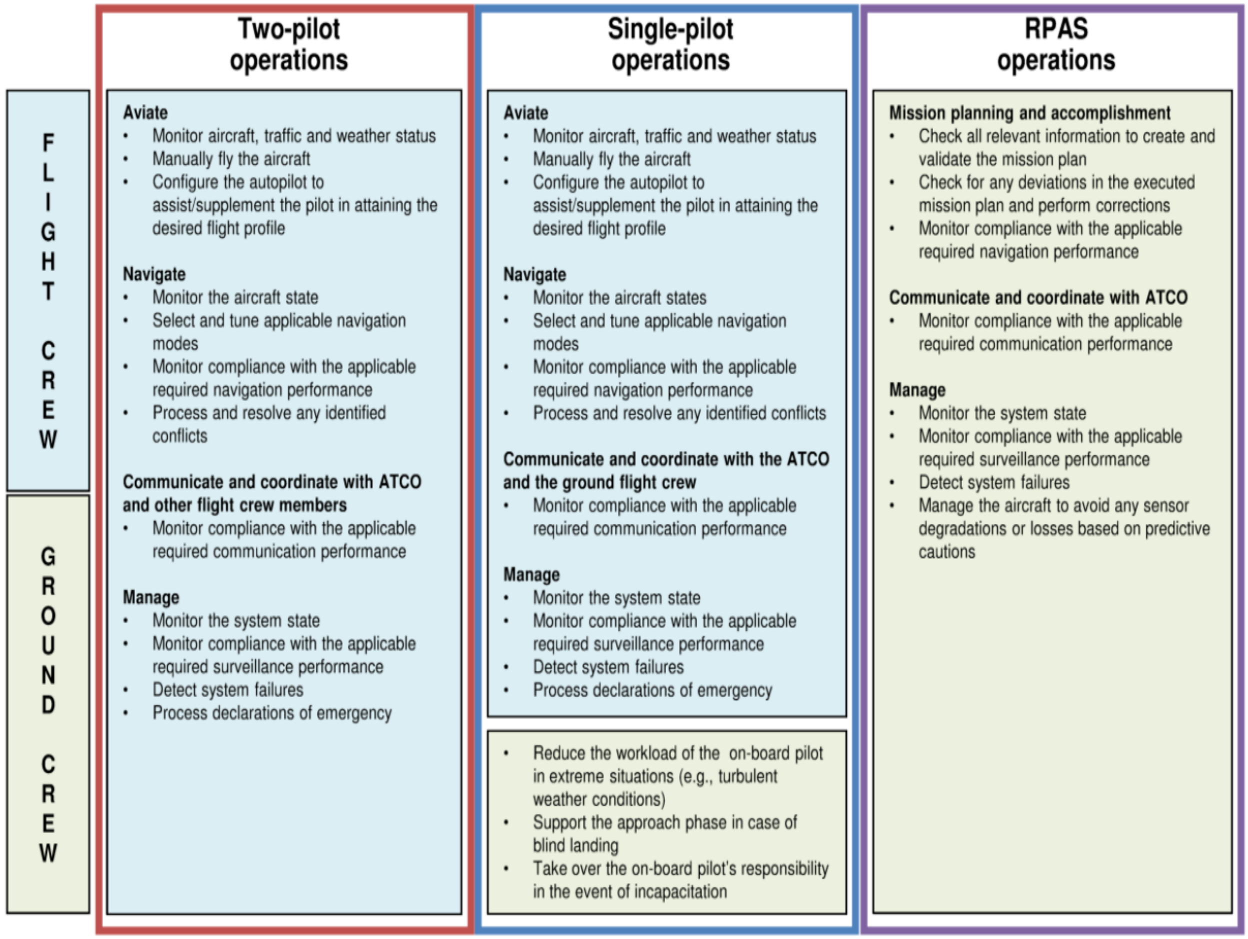
2.2 Multitasking in aviation

This section examines the challenges pilots face when managing multiple tasks simultaneously during flight. It focuses on cognitive load, the main responsibilities pilots must overcome, and how these can impact performance. Additionally, it explores common factors contributing to general aviation accidents, highlighting the risks associated with cognitive overload in flight environments.

2.2.1 Pilot’s cognitive load and main tasks

This type of haptic signal is designed to facilitate motor control over a vehicle or simulation. When the tactile signal provides direction information, it is expected to enable navigation and allow more visual attention to read the avionics or pay attention to the surroundings. Possible outcomes may be reaction times, information processed and navigation pattern completion.

An investigation was conducted to identify the main tasks pilots have on a standard flight, to properly replicate the flight conditions and to evaluate the influence that tactile guidance may have on the participant’s workload. For the current two pilot operations, the cabin crew consists of the captain and the co-pilot, who are responsible for the tasks of aviation, navigation, communication and aircraft managing from takeoff to landing. The captain is in charge of following the flight plan by monitoring the flight path for any possible deviation. The co-pilot on the other hand has the responsibility to navigate, control radio communications and to assist the captain if it is necessary. With a proper strategy the workload is efficiently shared between the cabin crew to maintain proper situational awareness. *Figure 7* displays a full compilation of operations performed by both the flight crew and the ground crew for commercial flights under standard conditions. These responsibilities follow the civil aviation safety regulations. [16]



Two-pilot operations

Aviate
- Monitor aircraft, traffic and weather status
- Manually fly the aircraft
- Configure the autopilot to assist/supplement the pilot in attaining the desired flight profile

Navigate
- Monitor the aircraft state
- Select and tune applicable navigation modes
- Monitor compliance with the applicable required navigation performance
- Process and resolve any identified conflicts

Communicate and coordinate with ATCO and other flight crew members
- Monitor compliance with the applicable required communication performance

Manage
- Monitor the system state
- Monitor compliance with the applicable required surveillance performance
- Detect system failures
- Process declarations of emergency

Single-pilot operations

Aviate
- Monitor aircraft, traffic and weather status
- Manually fly the aircraft
- Configure the autopilot to assist/supplement the pilot in attaining the desired flight profile

Navigate
- Monitor the aircraft states
- Select and tune applicable navigation modes
- Monitor compliance with the applicable required navigation performance
- Process and resolve any identified conflicts

Communicate and coordinate with the ATCO and the ground flight crew
- Monitor compliance with the applicable required communication performance

Manage
- Monitor the system state
- Monitor compliance with the applicable required surveillance performance
- Detect system failures
- Process declarations of emergency

- Reduce the workload of the on-board pilot in extreme situations (e.g., turbulent weather conditions)
- Support the approach phase in case of blind landing
- Take over the on-board pilot's responsibility in the event of incapacitation

RPAS operations

Mission planning and accomplishment
- Check all relevant information to create and validate the mission plan
- Check for any deviations in the executed mission plan and perform corrections
- Monitor compliance with the applicable required navigation performance

Communicate and coordinate with ATCO
- Monitor compliance with the applicable required communication performance

Manage
- Monitor the system state
- Monitor compliance with the applicable required surveillance performance
- Detect system failures
- Manage the aircraft to avoid any sensor degradations or losses based on predictive cautions

Fig. 7 Piloting tasks analysis [16]

A different research classifies different tasks according to their difficulty. The mental workload in flight driving was obtained by simulating different flight conditions and analyzing the participant’s EEG signals while performing them. Each participant performed three simulated flight experiments containing three mental flight tasks. *Table 2* shows the breakdown of all tasks for their respective mental load flight conditions. The tasks mainly consisted of reading information that was displayed on the flight interface. Subjects were then required to monitor the dial information and respond accurately and quickly to the questions presented. [17]

Table 2 Pilot's tasks classified by mental load flight conditions [17]

High workload	**Low workload**
Indicated airspeed meter	Indicated pitch meter
Pitching scale information	Pitching scale information
Altimeter information	Altimeter information
Directional angle information	-
Rolling corners information	-
Status of steering conditions	-
Landing gear conditions	-
Engine status	-

2.2.2 General Aviation Accidents

Aviation accidents resulting from pilot workload represent a significant concern in aviation safety. Pilot workload refers to the mental and physical demands imposed on pilots while operating an aircraft, encompassing factors such as task complexity, time pressure, and situational awareness. When pilots experience high workload levels, their cognitive resources may become overwhelmed, potentially leading to errors in decision making, situational evaluation, and task performance, which can increase the risk of accidents.

Several factors can contribute to elevated pilot workload, including adverse weather conditions, air traffic congestion, equipment malfunctions, and communication challenges. In dynamic and demanding situations, pilots may struggle to manage multiple tasks simultaneously, prioritize critical actions, and maintain adequate attention to key flight parameters, all of which can compromise flight safety. [18]

Furthermore, human factors such as fatigue, stress, and individual differences in cognitive abilities can increase pilot workload and its associated risks. Fatigue, whether due to extended duty hours, or inadequate rest periods, can impair cognitive functioning and degrade pilot performance, causing errors and accidents.

To mitigate the risks associated with pilot workload, aviation regulators, aircraft manufacturers, and airlines employ various strategies and technologies. These include the development of cockpit automation systems to assist pilots in managing tasks and reducing cognitive load, implementation of standardized procedures and checklists to streamline operations. Another recent priority is the enhancement of crew resource management (CRM) training to improve teamwork, communication, and decision-making skills among flight crews.

The establishment of effective fatigue risk management systems (FRMS) and scheduling practices to optimize pilot rest and duty periods are critical for combating fatigue related workload issues. Additionally, ongoing research and data analysis efforts contribute to a deeper understanding of the factors influencing pilot workload and the development of evidence-based interventions to enhance aviation safety. [18]

A high cognitive load may negatively affect the job performance of a pilot. This has been proven since research concluded that most aviation accidents occur in flights where the pilot’s mental workload is too high. That's why new instruments for information processing are required during flying tasks for faster and more accurate decision making. Hence the importance of addressing human factors and workload management in aviation operations.

2.3 Workload evaluation

This subchapter focuses on the workload evaluation in aviation related tasks. It introduces the Multi Attribute Task Battery (MATB) as a standardized tool for measuring mental workload. Furthermore, it explores different methods in which the workload can be evaluated in flight simulators.

2.3.1 Workload measurement with MATB

Multi Attribute Task Battery (MATB) is software that simulates aviation related tasks including system monitoring, tracking, communication and resource management. Its main purpose is to evaluate the user’s performance and workload levels. What makes this software ideal for workload evaluation is its customization, since it lets the user configure the test duration, number of active sub tasks, event rates and response times. The sub tasks may be simultaneous or with a determined order of succession. This creates the possibility of increasing the task's difficulty and overall workload of the user. [19]

The Multiple resource theory asserts that people have a limited set of resources available for mental processes, in particular during high mental workload. This theory explains how difficult single tasks can run into processing difficulties and how dual task performance is more likely to be hampered by performing similar tasks rather than dissimilar tasks (Wickens, 2002; 2008).

With respect to the prediction of mental workload levels on MATB task performance, the mental workload shows a significant inverted U-shaped effect during the tasks, with both very low and high mental workload being associated with a decrease in performance. This performance variability with the increasing mental workload can be properly analyzed in *figure 8*. “High workload levels could lead to errors and accidents, while low workload could lead to frustration and lack of attention” (Hancock et al., 1995).

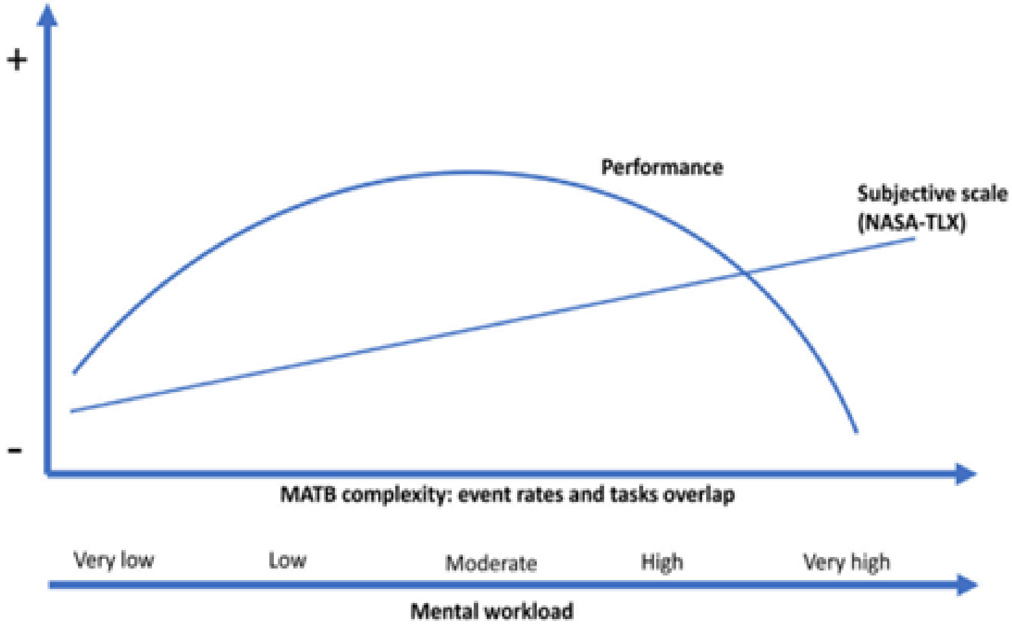


Fig. 8 Performance variability with increasing mental workload. [19]

A study conducted by Sauvet F. (2024) reviewed 19 articles which analyzed the effects of workload and its relation with different MATB configurations. A limitation for this study is that with the chosen systematic review method, the number of articles is not enough to conduct a meta-analysis. Still, most results show a degradation in performance due to an increase in the workload levels of the user. This effect has been identified as predominant in sub tasks related to tracking. With an increasing tendency in the study of human machine interaction, MATB can be used as a simulation tool to study the effect of mental workload in aircraft pilots, with different operational and psychological constraints.

Types of subtasks

System Monitoring: This task tests the attention and impulsivity of the user. It is formed by four moving scales and two warning lights. The performance metrics for this subtask are the average reaction time, accuracy (percentage of correct response), and false alarm (percentage of omission). These metrics are obtained for each button and scale and furtherly analyzed for the overall task performance.

Tracking: This subtask is related to manual control. For this task the user must use a joystick and keep the target at the center of the window. The performance measurement is done with the root mean square distance between the target and the cursor. With this sub task it is possible to measure the subject's objective performance decrease due to the increase in mental workload.

Communications: For this task, prerecorded voice messages will announce signals with a given radio channel and frequency. Once these announcements are played on the computer's speaker, the participants must discriminate which messages are relevant by answering them and choosing the proper channel and frequency. Responses are valid when the participant matches the radio channel and frequency to the corresponding target. This type of task tests auditory discrimination and memory skills.

Resource Management: The purpose of this task is to simulate in flight fuel management. Participants must maintain the fuel levels in tanks A and B to a previously defined value. This is achieved by interacting with the pumps. This task tests strategy and planning. Within the fuel management scenario with failing pumps, participants must achieve the lowest normalized mean value of deviation from the target fuel level.

In aviation, human multitasking performance is always evident. Pilots are constantly monitoring the information generated by different subsystems of the plane, such as hydraulic, pneumatic and electric. Pilots interpret all the acquired data simultaneously and with the available tools, make the appropriate decision based on situational awareness. Monitoring mental workload is necessary to properly develop and implement adaptive support systems for human machine systems. With the recent growth in automation technologies, pilots are dealing with more complex systems, where multitasking performance is gaining more relevance. [19]

2.3.2 Evaluating workload in aviation task environments

As aircraft automation and information technology have advanced, pilots face increasing multitasking demands, leading to concerns about mental workload, which arises from cognitive, rather than physical, demands. Understanding mental workload is critical for aviation, where cognitive demands directly impact performance and safety. The study conducted by Li et al. (2022) investigates the effects of multitasking on mental workload in a simulated flight. It uses a combination of physiological, subjective, and performance variables. The study aimed to identify different ways in which cognitive overload negatively impacted performance, providing valuable data for workload optimization. [20]

The study consisted of a flight simulator, where participants were asked to perform multiple aviation related activities. Several tools were used to evaluate their mental workload, including physiological indicators such as electroencephalography (EEG) and heart rate variability (HRV). For the subjective evaluation a NASA Task Load Index (NASA-TLX) was filled by the participants after every task. The combination of these methodologies led to an evaluation of cognitive load during different tasks difficulties. [20]

Participants were divided into groups with varying multitasking conditions to determine how workload intensity changed with increasing cognitive demands. The simulated flight experiment is composed of four subtasks: the flight target tracking task, the meter monitoring task, the emergencies handling task, and the residual capacity task. Among these, the first three were designated as primary tasks, while the residual capacity task served as a secondary task. These tasks were integrated into a program operated via computer, with a structured interface.

This study shows that as mental workload increases during multitasking, task performance declines, with time pressure identified as the main contributing factor. NASA-TLX scores increased with task difficulty, particularly in mental and temporal demand, indicating increasing workload due to time pressure rather than operational difficulty. Performance declined in the tracking and secondary tasks as task load increased, suggesting that limited cognitive resources were being redistributed. fNIRS data showed increased prefrontal cortex (PFC) activation with higher task loads, relating it with the multiple resources theory in multitasking. [20]

Both heart rate and prefrontal cortex (PFC) activation proved to be effective in detecting changes in mental workload under simulated flight conditions. The research conducted by Li et al. (2022) provides an analysis of mental workload during multitasking in a simulated flight environment. By employing a multiple phase approach, the study successfully demonstrates the impact of increasing task difficulty on cognitive load, performance, and physiological stress. These findings emphasize the need for continuous advancements in training methodologies and cockpit interface design to ensure optimal pilot performance and aviation safety. Future research should explore adaptive workload management systems and real world applications of cognitive workload monitoring to improve operational efficiency. [20]

3. Flight Simulator Development

3.1 MATLAB code description

This MATLAB code defines a function called ‘labview_interface’, which establishes a bidirectional communication interface between MATLAB and LabVIEW using TCP/IP protocol. The function accepts a variable number of input arguments, allowing users to connect to a LabVIEW server by specifying the remote host, port, and timeout values. If no arguments are provided, it defaults to connecting to the local host on port 2000. The function also validates the remote host input to ensure it is a valid string. Upon successfully establishing a TCP/IP connection using MATLAB’s ‘tcpip’ function, it sets various parameters such as the network role as a client and the byte order as Big Endian. If the connection is successful, the function creates a timer that calls a nested function, dotransfer’, every 0.01 seconds to manage data transfer between MATLAB and LabVIEW. The ‘dotransfer’ function reads a specified number of bytes from the LabVIEW interface, processes the received data, and prepares it for sending back to LabVIEW, ensuring proper byte order. Additionally, the code includes error handling to display messages for connection issues and to stop any existing timers in case of disconnections. Overall, this code enables seamless data communication between MATLAB and LabVIEW, allowing for the exchange of information in real time.

This MATLAB code prompts the user with an input dialog box asking for two pieces of information: the remote LabVIEW host address and the timeout duration in seconds. The dialog title is set to "LabVIEW Interface," and it provides default values of "localhost" for the host and "0.5" for the timeout. If the user provides input, the code calls the ‘labview_interface’ function, passing the specified remote host address, a fixed port number of 10,000, and the timeout value converted from a string to a number. Finally, the code clears the variable storing the user's input to free up memory and maintain a clean workspace. Overall, this code serves as a simple user interface for setting up a connection to a LabVIEW server from MATLAB.

This MATLAB code defines a function called ‘lvread’, which retrieves data from the LabVIEW interface. The function allows variable output arguments and starts by accessing the user data stored in a timer object that is tagged as ‘labview_interface’. It uses the ‘get’ function to retrieve this user data, which is presumably the data exchanged between MATLAB and the LabVIEW server. The retrieved data is then assigned to the output variable ‘varargout’, allowing it to be returned to the caller. Overall, this function serves as a simple method for reading data that has been made available by the LabVIEW interface during communication.

This MATLAB code defines a function called ‘lvwrite’, which is designed to send data to the LabVIEW interface. The function accepts a variable number of input arguments and begins by attempting to validate the first argument using the ‘validateattributes’ function. It checks whether the input is a numeric vector of type double and real numbers. If the validation is successful, the code sets the user data of the LabVIEW interface to the provided input data, ensuring the data is formatted as a row vector. In the event of an error during validation, an error message is displayed, indicating that only numeric classes are allowed. Finally, the function assigns an empty output variable ‘varargout’ before completing its execution. Overall, this function facilitates the transfer of numeric data from MATLAB to LabVIEW, while ensuring that the input meets specified validation criteria.

3.2 Introduction to LabVIEW

To properly make an immersive aircraft dashboard to make the simulation more realistic, graphical programming software will be used. This software is often used for designing tests, measurements and control systems. LabView is ideal software for the design of the interface, since it uses a flow chart method for graphical data flow. This way is possible to display randomly generated values from a desired range at a periodic time. Each block diagram will have its respective indicator in the front panel to visualize the desired data. This way is possible to use gauge numeric indicators to replicate the manifold pressure the air temperature gauges. Additional styles may be added to adjust the shape and size of the visual indicators to make them seem as realistic as possible. An example of a base gauge numerical indicator can be seen in *figure 9*.

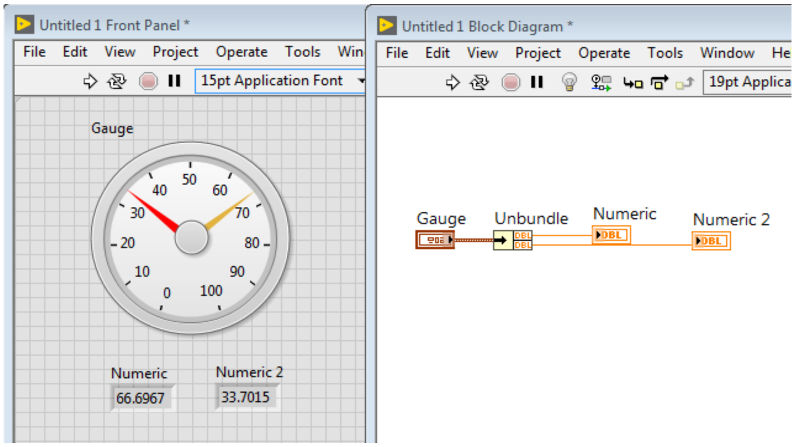


Fig. 9 Example of LabVIEW program with gauge numerical indicator [22]

As for how the program works. The block diagram is where the graphical code resides. They can be seen as programmatic functions that include boolean, string and numeric blocks. By linking the blocks between each other it is possible to move data from one item to the next one. A “timing” block can create a delay for the iterations of a loop if needed. This way the entire process may be automated, reducing the tasks during the evaluation and letting the researcher pay the proper attention to the experiment realization. [21]

3.3 LabVIEW program description

The LabVIEW program is designed to facilitate TCP communication between a host computer and a remote system via TCP/IP. The entire system operates within a while loop, allowing continuous execution until a stop condition, controlled by a Boolean stop button, is met. The program begins by initializing a TCP listener, configured to wait for an incoming connection, using specified port and timeout values. Once a connection is established, the program reads data from the client through a TCP Read block and processes this data as a double-precision floating-point value. An internal conditional check ensures error handling for connection failures or timeouts. The program then proceeds to send data back to the client through a TCP Write block, using the same floating-point format. Conditional logic within the loop ensures that only valid data is transmitted, while error handling ensures termination if necessary. Once the stop condition is triggered, either due to an error or user intervention, the TCP connection is gracefully closed using a TCP Close block. Status indicators are incorporated throughout the program to monitor the connection status, detect errors, and control the stop condition, ensuring reliable operation of the TCP communication process. [23]

A cluster is connected directly to the TCP block to unbundle the data and convert them into array type. From here it is then possible to select which specific values are assigned to the X and Y axis of the plot. These chosen append arrays are bundled once again and finally the output of the cluster is directly connected to the XY Graph block. [24]

Case structures in LabVIEW are the equivalent to the commonly used statements if then, else which are used to selectively execute lines of code in other programming environments. Cases structures in LabVIEW follow the same principle but employ a graphical structure instead. The case structure block will have multiple cases with its respective code and will use the selector input to determine which piece of graphical code to execute. LabVIEW automatically selects and runs the matching case, while a Default Case can handle any unmatched values. Only one case is executed at a time, and the structure ensures no overlap between cases. This graphical construct simplifies decision-making by visually organizing conditional logic in the block diagram. [25]

The background image in the LabVIEW program ensures immersion close to a flight simulation, is part of a simulator for the Dornier 228NG aircraft family built by EUROMEC. [26]

Internet Protocol (IP), User Datagram Protocol (UDP), and Transmission Control Protocol (TCP) are the basic tools for network communication, with TCP/IP being widely used for communicating over single or interconnected networks, including the internet. TCP/IP communication in LabVIEW provides a user-friendly interface for reliable network communication, hiding underlying complexities. LabVIEW’s TCP/IP functions, located in the “Data Communication” palette, facilitate the process of opening connections, reading/writing data, and closing connections, similar to DAQ or File I/O processes. A computer can act as either a client or a server in TCP/IP communication, and LabVIEW allows users to develop custom applications to manage both roles. The server listens to remote connections and responds accordingly. Server access control can be implemented by verifying client permissions using the remote address output of the “TCP Listen VI”. [27]

In LabVIEW, users can monitor input devices such as keyboards, mice, and joysticks using the built-in Input Device VIs, which are located in the Connectivity, Input Device Control palette. These VIs allow the monitoring of attached devices by providing various data related to their operations. For example, when reading from a joystick, the process begins by calling Initialize Joystick.vi, where the device index input identifies the device to be monitored. This index is a numeric value assigned by the operating system.

The device ID output from Initialize Joystick.vi is connected to the device ID input of Acquire Input Data.vi. For a joystick, this VI returns data such as axis information, button status, and directional inputs that reflect user operations. The axis information provides rotating input data, while button status is given through a cluster of Boolean controls representing each button on the joystick, although the mapping may not always match the physical button layout. Directional information, known as the Point of View Hat input, indicates the joystick’s direction.

Finally, after completing the data acquisition, Close Input Device.vi is used to close the reference to the device. This is done by wiring the device ID from Acquire Input Data.vi to the device ID input of Close Input Device.vi to properly terminate the connection to the input device. [28] The “Reinitialize to Default” method was used on each control to restart it by employing an invoke node. [29]

3.3.1 Parallel Tasks

Inspired by the MATB paper, a system monitoring sub task was added to the main tracking activity of the simulator. These tasks are analogous to activities that aircraft crew members perform in flight, and by using them it is intended to detect and measure the performance degradation due to increased workload.

The system monitoring task will occur simultaneously to the tracking task. Two gauges with continuously moving scales will be placed next to the tracking window. Participants will be asked to interpret and regulate the information displayed in the gauges, which represent both temperature and pressure respectively. The system monitoring task tests impulsiveness and attention by asking the participant to restart a gauge once it enters a certain range of values. For that, the subject must press a retractive switch, which is part of the joystick device. Each gauge measures a different variable, and it has assigned a correspondent switch. The position of the retractive switches correspond to the on-screen placement of the gauges for both right and left handed people. The intention of this placement is to make it more intuitive for the subject and reduce the overall cognitive load of the sub-task.

The participants will need to restart the gauges until they stop moving, this will happen once the button is pressed a certain number of times previously defined. This task's performance will be measured with the average error between the valid range of values and the registered value of the gauge at the instant when the button was pressed. As the program is running, there is no performance feedback for the system monitoring sub task, just a meter that registers how many times the button has been pressed.

3.2.2 Write Data from LabVIEW to Excel

As part of the parallel task, it is a requirement to keep track of the instants in which the user presses the retractive switch to trigger the correspondent gauge response. Since the purpose of this sub task is to measure the impulsivity and attention of the participant, it is compulsory to properly evaluate the timing and precision of the response.

As previously described, the participant must restart a gauge once it enters a certain range of values. These ranges will be set with anticipation and kept the same during the entire session. Since the values will be constant, it is possible to measure the precision of the response in function of how close to the desired value was to the actual gauge placement.

To ensure that all the values are properly saved for further analysis, a section of the LabVIEW code was created with the sole purpose of writing all the data from LabVIEW and exporting it into an excel file. The first step was to create an invoke node value from the gauge that was measuring both the pressure and the temperature and set it as control. Since both gauges have their respective individual loop, the entire procedure was done for each gauge. Since it's just one signal measured at the time for each case, it is not necessary to merge the signals. A case structure is employed, so that the data saving is controlled by a boolean button. This means that the selected data within the case structure will be saved and exported to excel only when the boolean is set to be true.

A LabVIEW block named *Set Attributes* is used to configure and set the dynamic attributes. Here is where the signal name and values are established to further appear as the Excel column headline. The invoke node value from the gauge is set as signals in. A signal index value of zero must be set as well in this step. Lastly, the signal name, which will be the same as the respective gauge description. It's necessary to set only one dynamic data attribute since there will be only one input to write on excel for each gauge.

A second LabVIEW block named *Write Meas Files* is used to configure the write to measurement file. This basically consists of choosing a file name and its location in the local files, the type of action for saving the file, what to do in case a file with the same name already exists and the file format. To keep the same format for all the measured variables, the chosen format was to save the data in a .xlsx file, with only one header per segment and one X value (time) column. All the data have a tabulator for delimiter.

Finally, to have a correct time format for the data written in excel, a block named *Get Date/Time in seconds* is used to have the same timestamp as the computer. This ensures that all the measurements are made without a delay and with the correct format and time zone.

Since this segment of the code is connected to the two continuously moving gauges scales, the boolean button will be set to be true as the gauge increases their scale. When the participant presses the retractive switch, it will automatically switch off the boolean, turning the case structure to false. This will restart the gauge scale and stop writing the data to excel. Immediately afterwards, the Boolean will automatically turn true once again, restarting the loop. The new data will be appended to the same excel file keeping on with the loop, until it is stopped.

To automate the entire process, a small segment was created within LabVIEW, to keep track of how many times the retractive switch was pressed, restarting its respective gauge. For this purpose, a while loop structure was created with a shift register on both ends. A numeric constant of zero was placed to the left end, to initialize the count. With a numeric *add* block, a new condition is set so that every time the boolean value connected to the retractive switch is changed from 0 to 1, a numeric indicator will increase its magnitude by one. We can subsequently use this indicator to stop the loops once it reaches a certain value.

Each gauge has its respective indicator that keeps the count of how many times it's been restarted. By placing a time delay on each sub task, it is possible to calculate the average time that it will take to finish it and how many times the loop must repeat itself in order to reach the desired duration. By knowing the total duration of the task, it is possible to set a corresponding number of loops that better fit the purpose of each sub task. All the numeric indicators have their corresponding global variable set to zero, so that every time the LabVIEW VI program is restarted, all the numerical variables return to their initial value.

3.4 Interface design

The interface presented in the VisualSupport.vi file, developed in LabVIEW, represents an environment intended for aviation simulation tasks. It combines intuitive task segmentation and aviation standard visual cues. This user interface integrates both a tracking task and a system monitoring task, adapted into a realistic background inspired by the flight deck of a Do-228 simulator. [26]

Device Index 1

1

axes total 0
buttons total 0
pov total 0
device name

axes total 0
buttons total 0
pov total 0
device name

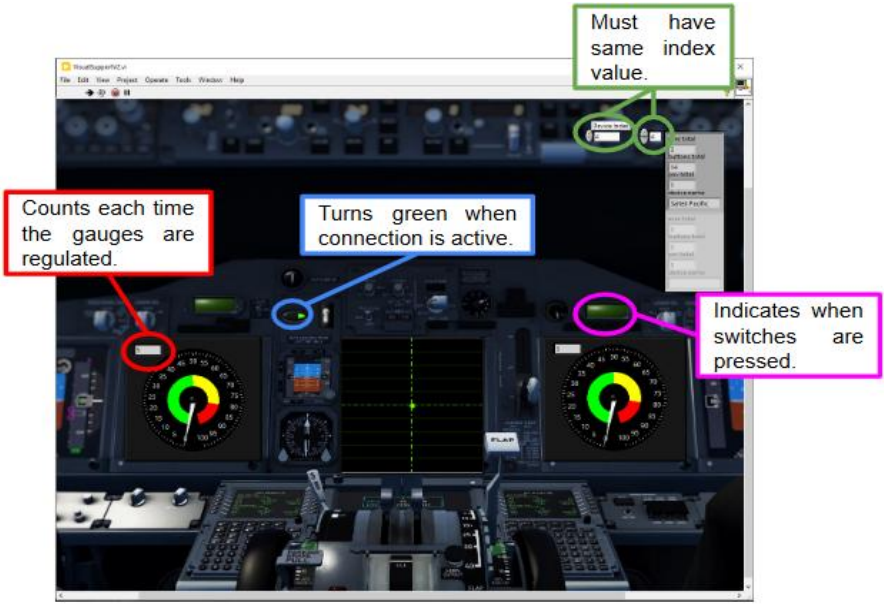
Fig. 10 Flight simulator interface *VisualSupport.vi*

As can be seen in *figure 10*, at the center of the interface is the tracking task, visualized through a centered square panel containing green gridlines and crosshairs. The purpose of this panel is to facilitate real time tracking activities, such as maintaining a reference point or following a moving target. The decision to utilize a black background with bright green lines provides high contrast, thus ensuring good visibility under varying conditions while minimizing user fatigue. The simplicity of the panel supports focused attention, which is critical for precision tasks often encountered in flight operations.

Besides the tracking panel there are two circular gauges placed next to it, which are dedicated to the system monitoring task. The left gauge simulates to indicate air temperature, while the right gauge represents the manifold pressure gauge. Both gauges have an intuitive, color segmented display, wherein green indicates a safe operational range, yellow signals caution, and red warns of critical conditions. The analog style of these gauges aligns well with traditional aviation instrumentation, allowing users (particularly those with piloting experience), to interpret the systems status. The inclusion of numeric displays within the gauges serves as guidance for the participant. With them it is possible to keep track of the times the gauge has been restarted, which is essential for the competition of the parallel task.

Positioned above the main task areas are several hardware and communication controls, including the settings for the device index. This parameter facilitates the connection and monitoring of the joystick’s status. Additional panels on the right side of the interface provide real time information on connected input devices, including axis counts, button states, and device name.

With respect to the boolean buttons in the interface, there are a total of three of them. The two rectangular green lights placed on top of both circular gauges are directly linked to the gauge’s controls. Every time they turn on, it indicates that their respective joystick switch is being pressed. The third and last boolean button is placed in the left top corner of the tracking panel. This oval shaped button, half black and half green, will turn on when the TLC connection with MATLAB is active. This description is accompanied with a graphical guide in *figure 11*.



Must have same index value.

Counts each time the gauges are regulated.

Turns green when connection is active.

Indicates when switches are pressed.

Fig. 11 LabVIEW usage and features

3.4.1 Selection of indicators

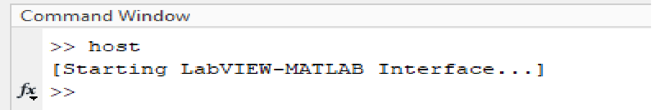
After conducting brief research into the pilot’s workload and its main tasks, it was concluded that the best alternative is to consider only tasks that mainly consist of reading information that is being displayed on the flight dashboard interface. Additionally, the tasks should be adding a low mental load to prioritize the accuracy of the system instead of the pilot performance. This is due to the fact that the participants are not intended to be trained professional pilots with extensive experience. The chosen tasks require the participant to monitor the interface information and properly interpret it. For that, gauge type indicators will display information that is often reviewed by the pilots directly from the avionics while airborne. The simulated indicators are air temperature gauges and the manifold pressure gauge, commonly found in aircraft.

The chosen indicators are part of the avionics of any modern plane and are often accompanied by engine controls, flight control systems, navigation, and communication among other systems that carry out flight management tasks. [30] Generally, the bigger the aircraft, the more electronic systems it has. The reason why these indicators were chosen is due to their simplicity. They consist of at least one pointer that measures either the air temperature or the pressure inside the intake manifold. The main advantage is that even if reading them does not represent a challenge itself, doing it under a high visual workload may increase its difficulty. That's why including them in the experiment protocol will demonstrate if tactile guidance may indeed relieve visual workload and assist the pilot with navigation tasks at the same time.

3.5 Simulator TCP connection instruction

To properly use the LabVIEW simulator with all its features, a list of steps must be followed in a certain order. They are listed below:

1.- Switch on USB + Power supply for Arduino.
2.- Initiate LabVIEW.
3.- Open Existing: VisualSupportV2.vi.
4.- Run program and match device index with Saitek Pacific.
5.- Restart VisualSupportV2.vi
6.- Keep LabVIEW as a background program. Once the connection is made, interaction must be exclusively with MATLAB.
7.- Initiate MATLAB.
8.- Run host.m, ensuring labview_interface.m, lvread.m and lvwrite.m are in the same folder. (*Figure 12* shows how it appears on the command window once the connection is made)
9.- Accept predetermined values for timeout and local host.
10.- Wait until the connection between the programs is made.
11.- Run experiment.m and follow the MATLAB command window.
12.- To initiate a parallel task in LabVIEW, turn the main red switch to A, and return it to OFF.



```
>> host
[Starting LabVIEW-MATLAB Interface...]
>>
```

Fig. 12 TLC Connection confirmation

A note to consider is that Switches T1 & T5 are linked to the left gauge, while switches T3 & T7 to the right one. For a clear overview of the Saitek Joystick, *figure 13* shows its switches layout.

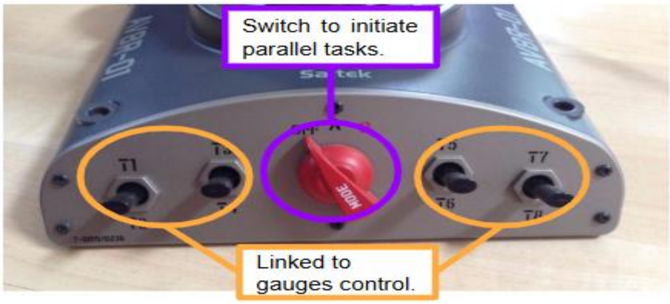


Fig. 13 Saitek Joystick switches layout

4. Experimental methodology

4.1 Equipment and materials

Holdens research published relevant discoveries related to the disadvantages of using consumer grade technology for scientific research due to its inherent delays and variables when it comes to reaction time. His work concluded that such delays and variables are associated with the hardware equipment, and that it often results in significant errors when measuring reaction time. This phenomenon was proved to affect different types of stimuli, for both tactile and visual configurations. Results showed how reaction time tests conducted by consumer grade technology often depend on the CPU processing speed, which can vary depending not only on the number of cores, but also on the RAM memory of the device, and the programs running in the background. The tactile guidance joystick is considered as a USB peripheral connected to the computer and will also introduce latency due to the communication protocols. Another consideration is the refresh rate of the monitor screen used, which mainly affects the stimulus delivery time.

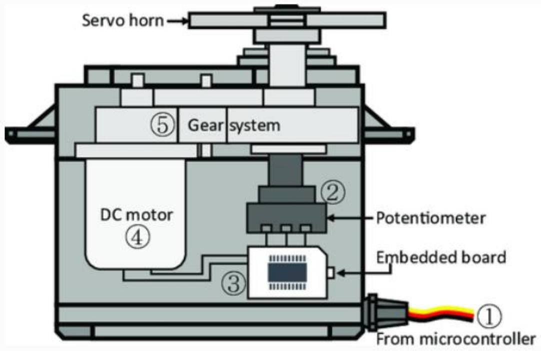


Fig. 14 Schematic of a SG90 servo motor [31]

The individual electronic components that are part of the tactile guidance device also have a corresponding delay that must be considered. The delay associated with the used micro servo SG90 refers to the time it takes for the servo motor to respond to a change in its control signal. This delay can vary depending on factors such as the load on the servo, the quality of the power supply, and the specific characteristics of the servo itself. Generally, SG90 micro servos have relatively low latency, meaning they respond quickly to changes in their input signal. A schematic of the servo SG90 can be seen in *figure 14*. It is assumed that the best way to define a delay is with respect to the pulse width modulation. The angle of the servo is determined by the duration of a pulse that is applied to the control wire and is expecting a new signal every 20 milliseconds. [31]

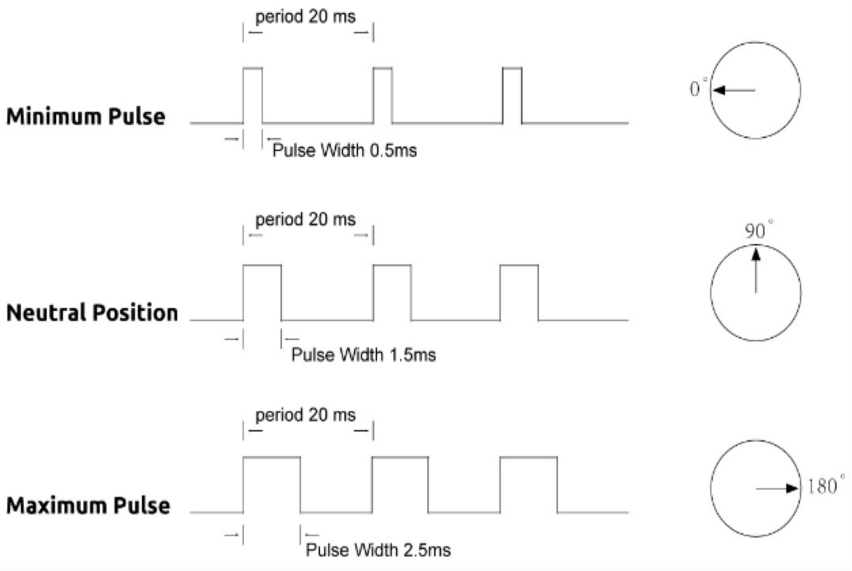


Fig. 15 Pulse width modulation for a SG90 servo motor [31]

Figure 15 shows how the device makes pulse readings every 20 milliseconds independently of the pulse width, hence the average delay due to the servo motor is assumed to be 20 milliseconds without considering external factors such as the design of the system and the specific operating conditions.

To reduce the influence that hardware may have on the tested performance parameters, all the experiments must be run under the same conditions. The same computer must be used to store data for all the sessions in order to reduce the variability of the reaction time measurements. Peripherals and monitors used during the experimental stages must remain constant as well. A standard average delay of 15 msec is often introduced by computer related variables [8], plus the 20 msec attributed to the servo motor [31]. This same delay of 35 msec will be considered for all experiment sessions. Additionally, to reduce even more the variability for each condition, a couple of trials before the testing will be set to reduce the variability effect over the experiment results.

4.2 Participants

Based on the conclusions driven from the Sensorial Stimuli Reaction Time experiment, a criterion to choose the participants must be followed to reduce the influence that diverse factors such as age, education level and time spent on the computer may have on reaction times. After a certain number of successful evaluations, it was concluded that for each type of stimulus modality, reaction time varies considerably between people from groups with diverse characteristics. In order to ensure accurate results and to have control over the reaction time inconsistency, the requirements of *table 3* were established for the participants selection.

Table 3: Participants requirements

Requirement	**Desired range**
Age	21 - 30 years
Educational level	Tertiary education
Daily Time Spend on computer	2 - 4 hours
Availability (Per week)	1 - 2 hours

By choosing participants with similar characteristics, the inconsistency of reaction time will be considerably reduced. This offers another benefit, which is accessibility to the ideal participants. Bachelor and master's students could be the best alternative, since they all share a similar age range, educational level and often spend the same on the computer due to their responsibilities. They should also be related to the software used and hardware, which will ease the protocol follow-up. For these experiments a total of 12 undergraduate student volunteers that meet the requirements must be recruited as subjects for data collection. The number of participants must be divisible by all factors. [32] There are two segments with three configurations each, which give 6 different procedures for each task type. And for two task types it makes a total of 12 configurations.

All the participants should be in good health, with normal or corrected vision. Since the joystick handle is symmetric, there is no preference over the subject's dominant hand. Before starting with any procedure, the selected participants must sign an informed consent about the experiment. They must agree with the publication of their test results and be compromised to attend all the required sessions.

4.2.1 Participant information and consent form

During the first experimental session, a document was given to each participant consisting of a brief introduction to the study "Haptic and visual joystick guidance". This document included all the relevant information about the study scope and purpose. It included details about the participants' criteria, required preparation, involved risks, withdrawal procedure and also regarding the data storage. Attached to the document was the contact information of the student responsible for the experiment and the supervisor.

If there was any comment or doubt during the experimental period, they could reach any of the responsible at any moment.

After being introduced to the study, a consent form was handed over, and participants were requested to fill in the appropriate boxes to claim that they agreed with the given information and the experiment methodology. It consisted of four different sections. *Section A: General Agreement* (research goals, participant tasks and voluntary participation), *Section B: Potential Risks of Participating* (Including data protection), *Section C: Research Publication, Dissemination and Application*, and *Section D: Long Term Data Storage, Access and Reuse*.

As part of the general agreement on the consent form, it was informed to all the participants that as compensation for their participation in the experimental evaluation, an aircraft LEGO set was going to be given to them after finalizing the required number of individual sessions. One week after the last participant concluded his last session, an email was sent to all participants at the same time, indicating that they could pass by the laboratory and pick their reward. There was no preference given, and they could choose their reward depending on their arrival order.

Fig. 16 Reward LEGO sets

4.3 Test Procedure

The experiment will consist of two different segments, with three different configurations for each of them. There will be one dedicated setup for each type of guidance, and a third one for a multisensory configuration using both the tactile and visual guidance simultaneously. This is purposely designed to compare visual and tactile guidance under different scenarios in which it is possible to evaluate the benefits of both. The classification is inspired by the meta-analysis conducted by M. Prewett which was previously discussed.

The first segment will be a direct comparison between the visual, tactile and the combination of both guidance stimuli, and will follow a similar methodology as the one used to measure the effect on the learning effect. The experiment will be focused on measuring the accuracy of each system and its reaction time without the influence of additional workload. Participants will be asked to guide the joystick to 20 generated positions with front and back travel range. Different variants for the course of joystick movements will be generated for each individual task. It is planned for all participants to go through each variant as the sessions progress, but in a different order. The order will be prescribed by the Latin square method to avoid the order effect. For this method to work, the number of participants should be divisible by the number of procedures.

A Latin square in analytical experiments is a design method structured into rows and columns, where each treatment level appears exactly once in each row and column. It's employed to control nuisance variables like spatial or temporal variation by systematically assigning treatments across the rows and columns. This design helps ensure that the effects observed are more likely due to the treatments themselves rather than external factors. Latin squares are particularly useful when resources or space are limited, as they allow for efficient experimentation while still controlling for sources of variability. [32]

This first segment consists of a task with a duration of 60 seconds. It will have a uniform duration between 3 and 6 seconds for each generated position. This procedure will be repeated three times for configurations A, B and C respectively. Configuration (A) will be assigned to the tactile guidance with no visual feedback. As for the visual guidance (B), the tactile features of the joystick will be deactivated, and the participant will rely only on visual feedback to follow the desired patterns. Lastly, multisensorial configuration (C) will consist of having both the visual feedback and the tactile features of the joystick active. By measuring the time to reach the target position (TRTP), the average error (AE) and the workload for each type of cue, it is possible to evaluate which system has a better performance when no workload influence is considered. Measurements must be regularly monitored as the data is stored to reveal outliers during the experiment and not once it is concluded.

What differentiates the first second segment from the first part is the effect on an additional cognitive load, which will be considered in the evaluation. This segment will follow the same methodology, with the same parameters to evaluate, task and configurations used for the first part. It is intended to increase the visual workload of the participants by asking them to accurately monitor and interpret the indicated values in the air temperature gauge, and the manifold pressure gauge. This side task will take place at the same time as the main evaluation and will evaluate which system has a better performance with a parallel task influence considered. Having an independent evaluation that considers workload will show relevant results of the benefits of multisensorial information processing. All the test procedures are described as a flow chart in *figure 17*.

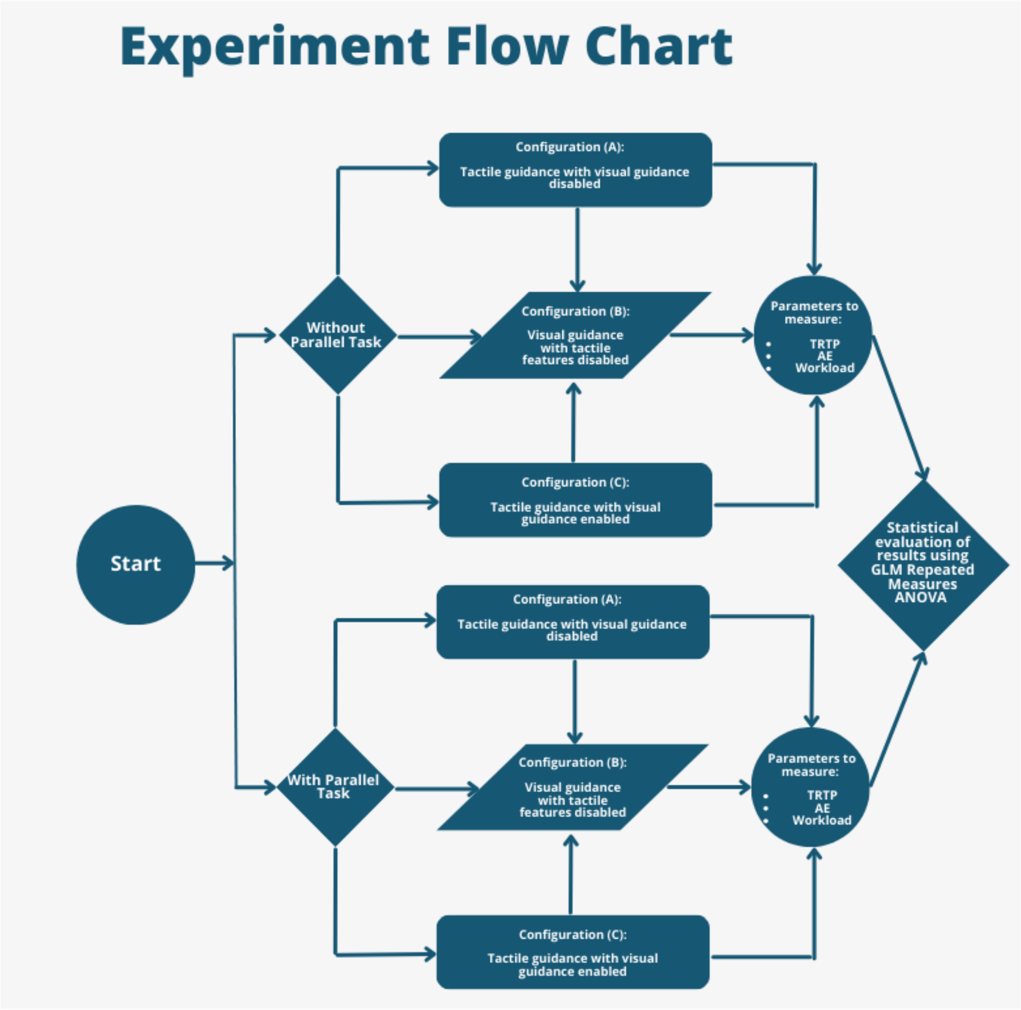


Fig. 17 Experiment flow chart

Similar to the original methodology proposed by P. Zikmund [2], after each session is concluded, participants will be asked to fill in a questionnaire to evaluate their workload for each configuration of the second segment using the Bedford scale. This is intended to measure in the first instance if the haptic guidance and the multisensorial configuration have a direct influence on the participant's workload perception. Even if other parameters will mathematically evaluate the benefits on the device in terms of performance enhancement, it is desired to have complementary feedback coming directly from participants.

Before the evaluation of the participants begins, the entire procedure will be tested. It is indispensable to ensure that the equipment is calibrated and properly connected. The program will be run several times by the tester, ensuring that the joystick works smoothly and that the data is being stored properly. Each experimental session, including both segments with their respective configurations and the time disposed for the calibration and preparation of the hardware, is expected to last around 24 minutes per participant to be successfully finalized. The post-questionnaire is already considered in this time approximation.

The entire experiment test procedure described in *figure 17* will be repeated for a secondary type of navigation task. The reason for this is that participants will perform two types of joystick tasks, one with a fixed target position and another where the target position changes continuously over time. These task types are not considered factors in the primary analysis but are examined separately to maintain clarity in interpreting its effects. In total, every participant will go through 12 different tasks, considering all the guidance configurations, influence of parallel tasks and type of target position.

4.4 Sessions criteria

When it comes to the number of sessions that each participant will go through until the stored data is valid enough to have relevant results, it is important to take as reference the conclusive results from Pavel's investigation about the influence that learning effect has in tactile guidance. This work explores the influence that learning effect has during pilot aircraft interactions due to tactile guidance methods. Since pilots rely on muscle memory, the performance of haptic feedback relies on the learning phases of it. In order to reduce as much as possible, the results variants.

From this investigation it was concluded that to properly conduct a comparison between haptic and visual guidance for aircraft control that will demonstrate if there is a benefit from introducing haptic feedback, it is necessary to pay special attention to the intervals between each session. Intervals may be flexible with the condition that sessions must be planned with anticipation and consistently followed with the only condition of having at least one night between them to provide sufficient time for knowledge consolidation and to prevent overtraining. After 12 successfully tested sessions with a mean time between sessions of 45 hours, results showed that the most significant learning effect was tracked to take place up to session 6. Afterwards there is a significant improvement.

With the retrieved information it is possible to determine how many sessions are ideal to properly evaluate the performance of the systems by reducing as much as possible the results variants due to the learning curve. A total of five sessions will be assigned to eliminate the impact that the learning effect has over the results. These first training sessions have the purpose of letting the participants understand and get used to the device. Muscle memory will take place and the interactions with the device will be enhanced. Sessions will take place weekly, and participants will have the flexibility of choosing the day on which they want to be evaluated. Sessions will be preferably carried out during the morning, to avoid possible fatigue from daily tasks and activities that may affect the performance of the participants.

Once the learning effect is reduced enough not to be considered, two additional sessions will take place with the main purpose of measuring the efficiency of each stimulus modality. Same intervals as for the training sessions will be followed. With this evaluation strategy it is expected that the parameters of TRTP, AE and workload will have less variance and a lower standard deviation for each configuration on every segment of the methodology. With constant values, the true impact of the haptic guidance enhancement to navigation performance will be properly measured. The same applies to the evaluation of the multisensory impact in the reduction of cognitive overload.

Another reason for having two definitive sessions after eliminating the learning effect is that a total of 12 tasks must be completed by each participant. This is the required number of sessions to satisfy all the possible combinations of guidance methods, cognitive load and task type. Diving them in two days will reduce the fatigue in participants and reduce the possibility of a random effect or error occurring during the final measurements. Below in *table 4* the registration date of each participant can be seen for their respective sessions.

Table 4: Session dates for all participants

	Training					**Measurement**	
Name	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Guillermo	28.3.	1.4.	15.4.	22.4.	23.4.	24.4.	30.4.
Honza	28.3.	31.3.	1.4.	15.4.	16.4.	23.4.	24.4.
Lucia	1.4.	15.4.	22.4.	28.4.	29.4.	30.4.	2.5.
Javier	1.4.	15.4.	17.4.	22.4.	28.4.	30.4.	2.5.
Brandon	11.4.	17.4.	23.4.	24.4.	25.4.	30.4.	2.5.
Mateo	15.4.	16.4.	17.4.	22.4.	23.4.	24.4.	25.4.
Nichole	15.4.	16.4.	22.4.	23.4.	24.4.	25.4.	30.4.
Min	16.4.	17.4.	23.4.	24.4.	25.4.	30.4.	2.5.
Filip	17.4.	23.4.	28.4.	29.4.	5.5.	7.5.	7.5.
Lukas	15.4.	16.4.	23.4.	24.4.	5.5.	6.5.	7.5.
Petr	15.4.	16.4.	22.4.	23.4.	29.4.	30.4.	5.5.
Carlos	15.4.	16.4.	17.4.	22.4.	28.4.	29.4.	30.4.

4.5 Configuration for evaluating mental workloads

Several factors were previously defined to keep track of the workload levels. These factors include the overall duration of each session, the duration of each individual segment with its respective setup configuration, and lastly the number of sub tasks for the segments that involve parallel tasks.

Level duration

The average duration for each session is 24 minutes. The whole session consists of 6 segments with different setups and an approximate duration of 4 minutes each. Only 3 out of the 6 sessions involve parallel tasks, which are meant to increase the workload level of the participants. The duration of each session is equal for both low and high mental workload tasks. A study conducted by Bowers in 2014 showed how there is an increase in false alarm reactions for high work level tasks that last less than 3 minutes. This effect is stabilized after 4 minutes of testing. [19]

Number of tasks

Several studies have proven that the divided attention required for the multitasking supervision of a system demands a greater workload on humans, than a sustained attention required to fulfill an individual task (McDowd, 2007). [19] With this into account, each 24 minute session of the experiment will be divided by two 2 additional categories with 3 segments each. For the first category, only one tracking task will be employed. These tracking tasks will have slightly different setup configurations regarding the type of stimulus given to the participant to fulfill the task (haptic, visual or both). The second category on the other hand will have an additional parallel task of system monitoring on each of the 3 segments. While the system monitoring task will remain constant, the tracking task will have the same variations in its setup configuration as the first category. Additionally, the tasks could have a fixed target position or a target position that changes continuously over time.

Event rates

The main purpose of the experiments is to compare and measure the participant’s efficiency to the fulfillment of tracking tasks with different types of stimuli, and the impact that an increase of workload has on it. Therefore, it is possible to maintain the same levels of workload for the simultaneous tasks. For the comparison, low mental workload segments won’t involve additional sub tasks at all, and the high mental workload segments will employ the same tasks, thus it is not necessary to modify at all the events rates.

Feedback

Studies have shown that feedback in the form of success or failure of the tasks, can affect performance. [19] During the training sessions, MATLAB generates a plot of the root mean square distance between the desired generated positions and the measured joystick position during the tracking task. Since this plot is automatically generated after each segment of the experiment, the participants may observe their performance and adjust their responses.

Training

To familiarize participants with the haptic device (its sensitivity and function), several training sessions are introduced in the experiment. The justification is described in another section of the paper. The practice sessions will have a fixed duration of 24 minutes and have the same structure of 6 segments and the same amount of low and high mental workload tasks. Several variations of the experiment will be generated, so that the parallel task is present in half of the segments, independently of its type of stimulus (haptic, visual or both). The variations of the experiment will differ for each participant.

Subjective measurements

In aviation, since human performance can directly impact safety and operational success, monitoring mental workload from the pilot's perspective is critical. The Bedford Scale seen in *figure 18* is a subjective but effective tool widely used to measure perceived mental workload. While originally developed to evaluate general fatigue and drowsiness, the scale has found significant application in aviation, particularly in the evaluation of pilot workload during various phases of flight operations and simulation studies. [34]

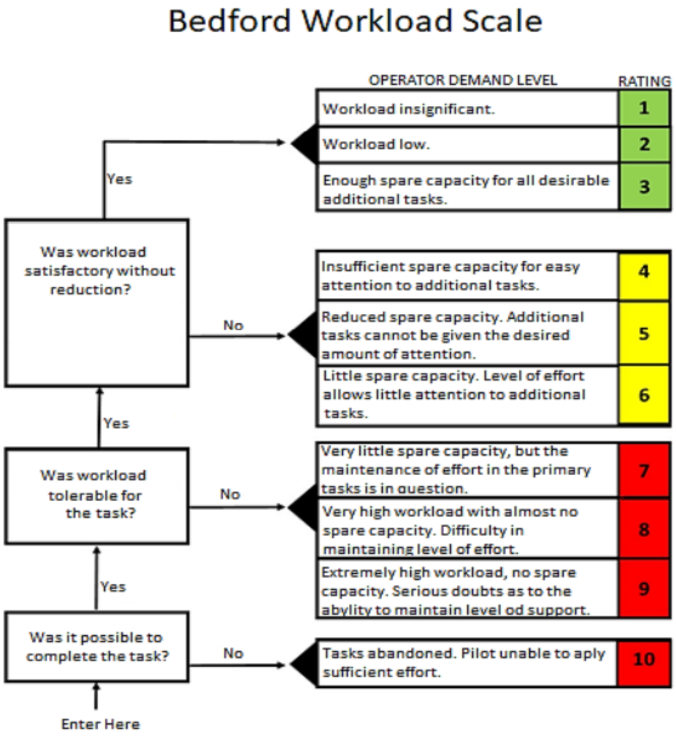


Fig. 18 Bedford Workload Scale [34]

In the aviation context, the Bedford Scale functions as a tool for understanding how pilots experience cognitive load under different flight conditions. The Bedford Scale enables researchers to quantify this subjective experience, offering insights that can lead to cockpit design, crew scheduling, fatigue management strategies, and training programs.

Moreover, the Bedford Scale has the advantage of not being intrusive. Unlike physiological or evaluations based on performance, which may require specialized equipment or disrupt normal operations, the Bedford Scale relies solely on the pilot's self-evaluations. This makes it highly adaptable for both studies based on simulations and real flight evaluations. Its simplicity makes it a key element in human factors research, particularly in studies examining the interaction between workload, fatigue, and flight performance. [34]

It is also important to acknowledge the limitations of the Bedford Scale. As a subjective measure, it depends on the pilot's self-awareness and willingness to report honestly. Since this may introduce certain variability in the results, it is necessary to support it with objective performance metrics, such as reaction time tests and error rates. [20] The Bedford Scale is a ten point subjective rating system that allows individuals to report the demand level of the task. Each participant will be asked to evaluate their own workload after the competition of each segment. This will be done for each segment of the entire session, so that it is possible to observe the impact of each training session on the participants' perspective.

Environmental constraints

Regarding the environmental constraints, all the experimental sessions take place in the same room, where the simulator is located. To reduce the variability of the psychological stress experienced by each participant due to environmental conditions such as the temperature exposure, noise and oxygen levels, the room is kept at ambient conditions. Participants will listen to white noise during the experiment to isolate the external noises that might compromise the task and to eliminate the servo's noises from the haptic guidance device. These noises could be interpreted as auditory stimuli, and the effect must be eliminated.

5. Statistical Analysis

5.1 ANOVA Overview

ANOVA works as a hypothesis test. It tests whether there are statistically significant differences between three or more groups. For the comparison there must be a measurable variable in common with the determined groups. The procedure is similar to a T-test, with the difference that ANOVA can be used to compare three or more groups. There are basically two possible hypotheses with an ANOVA test. Either the mean values of the groups are equal, or there is a difference between them. [35]

There are some assumptions that must be satisfied before performing an ANOVA calculation. Data is required to perform an ANOVA test. This data is obtained from a sample of a specific population that is being analyzed. From this sample of subjects, an additional division is made, and then it is possible to compare these formed groups. Once having the groups, a measurable variable must be determined, so that data can be obtained for the statistical analysis. [35]

The first assumption is the level of measurement of the variable, which must be appropriate for the ANOVA test. The independent variable should be nominal, and the dependent variable must be metric like. The nominal variable will define the groups, and the metric variable will be the measurable value. The second assumption is the independence between the measurements, which means that the measured values of each group should not affect the other’s group values. Third assumption refers to normal distribution, which means that data within each group should be normally distributed, especially for small groups. As the size of the sampled group increases, this assumption becomes less relevant. Fourth assumption is the homogeneity of variances. It means that the variance in each group should be similar. Small and large variances must be avoided to keep them similar between groups.

The ANOVA test uses the ratio of the variance between groups and the variance within groups to calculate if there is a difference. The variance between groups uses the deviation of the mean values, while the variance within groups uses the individual data points.

Mean of all data:

$$G = \frac{\Sigma x}{N}$$

$$\Sigma x \ = \ Sum\ of\ all\ the\ values$$

$$N \ = \ Number\ of\ values$$

Mean of groups:

$$Mi = \frac{\Sigma^{Group} x}{n_{Group}}$$

Sum of squares between the groups:

$$SS_{btw} = \Sigma n_i (M_i - G)^2$$

$$n_i = Number\ of\ values\ on\ each\ group$$

Sum of squares within the groups:

$$SS_{wi} = \Sigma\Sigma(x_{mi} - M_i)^2$$

$$x_i = Individual\ values$$

Mean squares between the groups:

$$MS_{btw} = \frac{SS_{btw}}{df_{btw}}$$

$$df_{btw} = Number\ of\ degrees\ of\ freedom\ between\ the\ groups$$

Mean squares within the groups:

$$MS_{wi} = \frac{SS_{wi}}{df_{wi}}$$

$$df_{wi} = Number\ of\ degrees\ of\ freedom\ within\ the\ groups$$

F value:

$$F = \frac{MS_{btw}}{MS_{wi}}$$

The p-value is obtained with an F-distribution, which is the result of the quotient of two chi-square distributions which are divided by the respective degrees of freedom. The required values for this calculation are the F-value and both df_{btw} & df_{wi}. To determine a statistically significant difference, p-value must be smaller than 0.05. [35]

Variations of ANOVA

There are different extensions to the one way ANOVA, which are employed for more complex cases and analysis. A two way ANOVA for example, tests the effect of two independent variables on a dependent variable. The independent variables are also known as factors and are basically a nominal variable. In this type of ANOVA, the test will study if any of the two factors have an effect on a determined metric variable. It will also determine if there is any interaction between these factors. Same assumptions apply to this test. The significance of the interaction effect is also determined by the p-value. [35]

In a two way ANOVA the total variance of the dependent variable is determined by the variance explained by the individual factors, its interaction and the variance error. The variance error is the proportion that can't be explained.

ANOVA with repeated measurements is used when the same subjects are measured multiple times under different conditions or at different time points. It shows if there is any variability within individuals. This analysis determines if there is a statistical significance between two or more dependent samples (conditions or time points). There will be one different measurement for each condition and each person involved. [35]

Post-hoc test

This test is employed when there is a significant difference between groups determined by the ANOVA test. It is necessary since ANOVA can only determine the existence of a statistically significant difference, but it doesn't specify which groups are different from each other. A post-hoc test will point out which groups differ from each other and by how much, by performing a pair-wise comparison between the groups. This test can also show that no individual group has a significant difference, even if this was previously determined by the initial ANOVA test. [35]

5.2 ANOVA Model Selection

The statistical evaluation analyzes two factors, the type of guidance and the presence or absence of a secondary parallel task. The guidance modality has three configurations, which could be either visual guidance, haptic guidance, or a combination of both. Each of these is tested under two cognitive load conditions, either with or without an additional parallel task, resulting in a 3 × 2 factorial structure. Additionally, participants perform two types of joystick tasks, one with fixed target position and another where the target position changes continuously over time. These task types are not considered factors in the

primary factorial analysis but are examined separately to maintain clarity in interpreting its effects. To evaluate the influence of guidance modality and parallel task on task performance, several dependent variables are analyzed. These variables are the time to reach the target position (TRTP), the average error (AE) between the target and the joystick position, and the workload perceived by the participant. AE and Workload will be evaluated on both task 1 and task 2, and TRTP is going to be evaluated only in the case of task 1 (with fixed target position), and both
Rewriting this in terms of ANOVA test, the independent variables, also known as factors, must be nominal. The independent factors for this study are the guidance modality and the cognitive load conditions. The dependent variables should be metric like, and that's why TRTP, AE & workload are chosen. Each dependent variable will have the same structure defined in *table 5* to organize the raw data. For this study, the effect or interaction between the dependent variables won't be analyzed. Instead, the ANOVA test will be performed individually for each dependent variable.

Table 5: Structure of the experiment for each dependent variable.

Participant No.	Haptic w/o Parallel	Haptic + Visual w/o Parallel	Visual w/o Parallel	Haptic w/ Parallel	Haptic + Visual w/ Parallel	Visual w/ Haptic
1						
2						
↓						
12						

In order to determine which type of ANOVA fits better the experiment characteristics, two aspects must be considered. The first consideration is that it involves 12 participants, and each of them must complete each configuration. This is the reason why repeated-measures ANOVA should be employed, since the key requirement for this type of test is that the same participants experience all the conditions. The second important consideration to be taken is that the two independent factors are within subjects, meaning the same subjects are measured multiple times. With these two considerations together, the two-way repeated-measures ANOVA is the best choice, since the same participants will be measured across all combinations of two independent variables (type of guidance and the cognitive load condition).

In this study, the objective is to examine how different types of guidance methods (tactile, visual, and a combination of both) under varying cognitive load conditions (high workload and low workload) affect participants' performance on three dependent variables: time to reach the target position (TRTP), average error (AE), and workload. A two-way repeated-measures ANOVA will be conducted separately for each dependent variable. The first null hypothesis states that there will be no significant differences in performance across the different guidance methods, while the alternative hypothesis proposes that at least one guidance method will significantly differ from the others. The second null hypothesis states that there will be no significant difference in performance between the high and low cognitive load conditions, whereas the alternative hypothesis suggests a significant difference between the two levels of cognitive load. Lastly, the null hypothesis for the interaction effect posits that there will be no interaction between guidance method and cognitive load, meaning the effect of the guidance method will be consistent across cognitive load levels. The alternative hypothesis asserts that an interaction exists, such that the impact of the guidance method on performance will vary depending on the cognitive load condition. All these hypotheses are summarized and organized by their effect type in *table 6*.

Table 6: Hypothesis summary

Effect Type	**Null Hypothesis (H0)**	**Alternative Hypothesis (H1)**
Main effect of Guidance Method	No difference between guidance methods	At least one guidance method differs
Main effect of Cognitive Load	No difference between high and low workload	High and low workload differ
Interaction of Guidance & Load	No interaction	There is an interaction

The proposed statistical evaluation is ideal for the experiment structure and chosen factors. The analysis evaluates both the independent and interactive effects of guidance modality and cognitive load. This approach ensures that the findings are both statistically robust and relevant to understanding how different guidance methods support human performance in multitasking environments. The dependent variables also help to determine which method has the best performance under several simulated flight conditions.

5.3 Statistical analysis procedure

To perform a proper statistical analysis, it is essential to organize the data correctly. For the proposed ANOVA test, a .xls file with all the experimental data must be created. In this file, each participant has one row, and each repeated measurement has a separate column. The columns may also represent each level combination of the two within-subjects' factors. The order of the columns is important, since it will serve as reference for the structure of the within-subjects' factors later on.

Once the file is properly created with all the data collected from the experimental sessions, it is possible to begin the statistical procedure. The first step is to launch the chosen software and load the created data file. Due to license limitations, StatotSoft's STATISTICA version 14.0 ended up being the best alternative. The software version can be seen once it's initiated as it appears in *figure 19*. After choosing the file, the selected excel sheet must be imported as a spreadsheet. The data range in terms of the columns and rows is determined automatically, the names of the variables are taken from the first row, and the case names are taken from the first column.

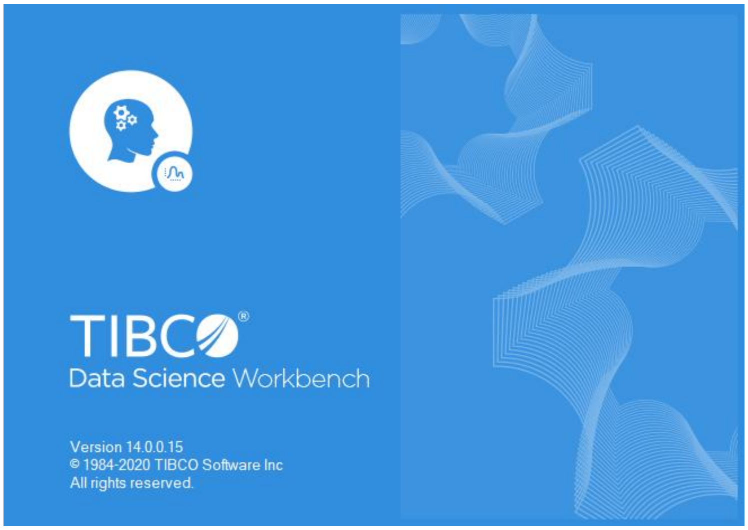
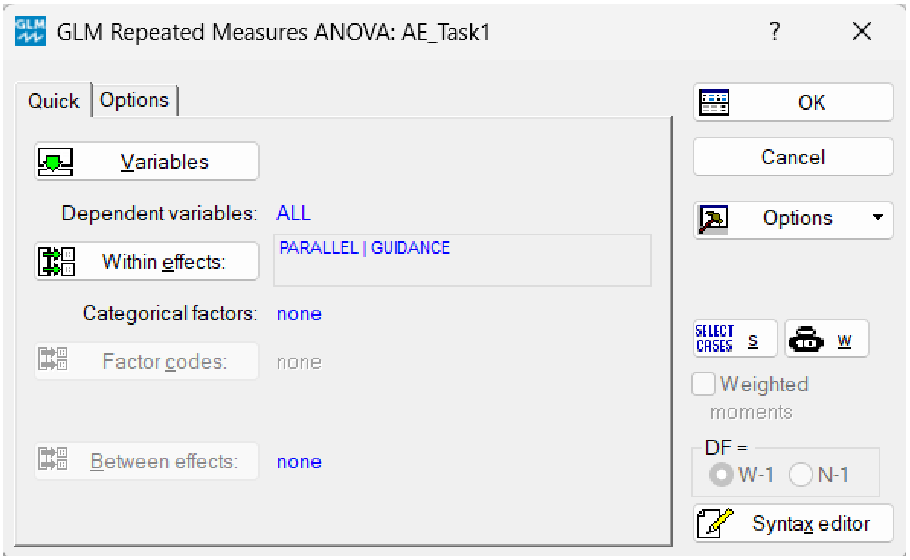


Fig. 19 STATISTICA version 14.0.0.15 initiation

Within STATISTICA, there are two different types of modules. With ANOVA module, it is only possible to specify one within-subject (repeated measures) factor. With the General Linear Models module on the order hand, it is possible to specify multiple repeated measures factors. When specifying the within-subject factor and the respective number of levels, the variables in the dependent variable list are assigned to levels for the repeated measures factor. The specified factor should always account for all the selected dependent variables.

By taking this into account, due to the requirements of the proposed statistical evaluation, a repeated measures ANOVA from the General Linear Models modules has to be chosen. This type of test uses general linear models to analyze designs with any combination of categorical independent variables, continuous predictor variables, or repeated measures. Due to the specification method, multiple dependent variables can be specified for any type of analysis. Both univariate and multivariate results are available when multiple dependent variables are specified. With respect to the specification method, the quick specs dialog consists of multiple dependent variables that can be specified for any kind of analysis.



GLM Repeated Measures ANOVA: AE_Task1

Dependent variables: ALL

Within effects: PARALLEL | GUIDANCE

Categorical factors: none

Factor codes: none

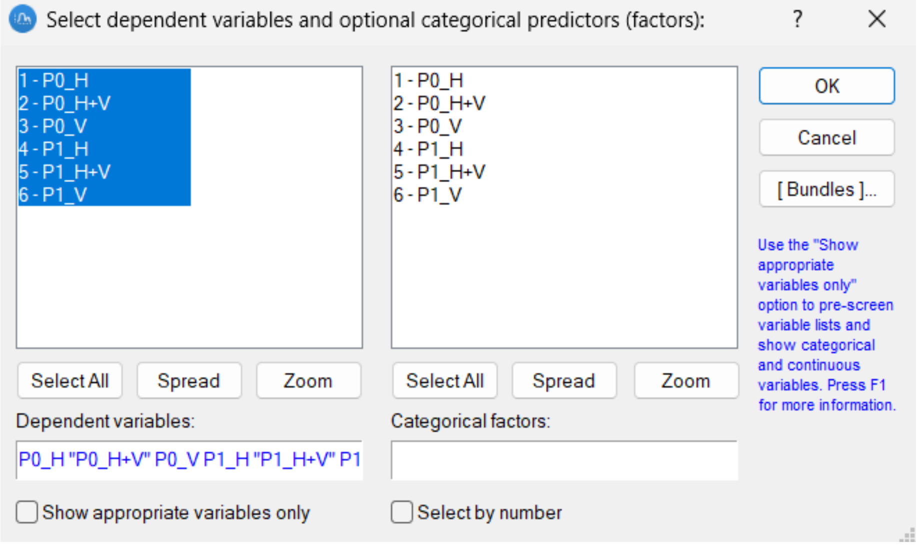
Between effects: none

☐ Weighted moments

DF = ● W-1 ○ N-1

Fig. 20 General linear model repeated measures ANOVA configuration

To perform this analysis, two additional steps must be followed for the GLM repeated measures ANOVA as seen in *figure 20*. The first one is to select dependent variables as it appears in *figure 21*. Here it is necessary to select all the dependent variables in the same order as they appear in the excel file with all the collected data. The order is important because this version of STATISTICA uses the order of the variables (columns selected as dependent variables) to map them automatically to the within-subjects factor combinations. Within-subjects factors consist of adjacent variables from the order in which they were specified.



Select dependent variables and optional categorical predictors (factors):

1 - P0_H
2 - P0_H+V
3 - P0_V
4 - P1_H
5 - P1_H+V
6 - P1_V

1 - P0_H
2 - P0_H+V
3 - P0_V
4 - P1_H
5 - P1_H+V
6 - P1_V

Use the "Show appropriate variables only" option to pre-screen variable lists and show categorical and continuous variables. Press F1 for more information.

Dependent variables: P0_H "P0_H+V" P0_V P1_H "P1_H+V" P1

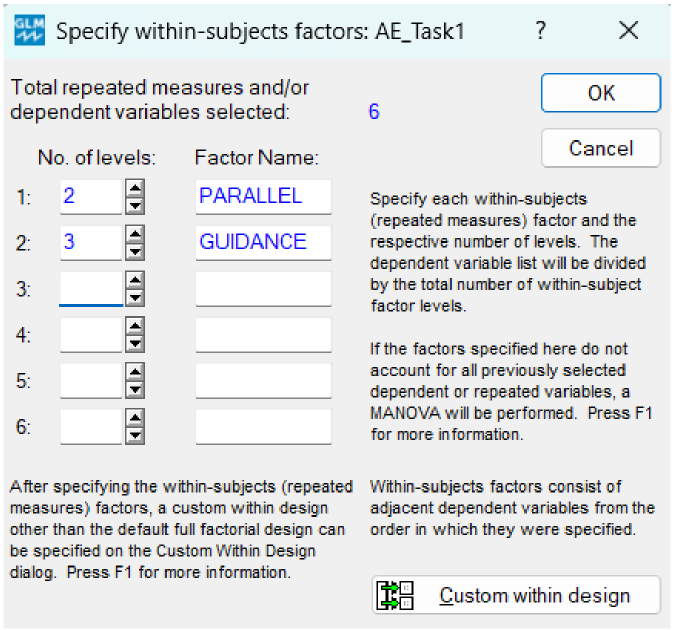
Categorical factors:

☐ Show appropriate variables only

☐ Select by number

Fig. 21 Dependent variables selection

After choosing the dependent variables, the next step is to specify the within-subjects' factors. In this case, the type of guidance factor will have three levels, and the cognitive load condition factor, two. This can be corroborated in *figure 22*. Each factor will have its respective name in order to distinguish them, and the total number of levels should match the dependent variables selected. For the simplicity of the analysis, a default full-factorial design of the within-subjects' factors will be used.



Specify within-subjects factors: AE_Task1

Total repeated measures and/or dependent variables selected: 6

	No. of levels:	Factor Name:
1:	2	PARALLEL
2:	3	GUIDANCE
3:		
4:		
5:		
6:		

Specify each within-subjects (repeated measures) factor and the respective number of levels. The dependent variable list will be divided by the total number of within-subject factor levels.

If the factors specified here do not account for all previously selected dependent or repeated variables, a MANOVA will be performed. Press F1 for more information.

After specifying the within-subjects (repeated measures) factors, a custom within design other than the default full factorial design can be specified on the Custom Within Design dialog. Press F1 for more information.

Within-subjects factors consist of adjacent dependent variables from the order in which they were specified.

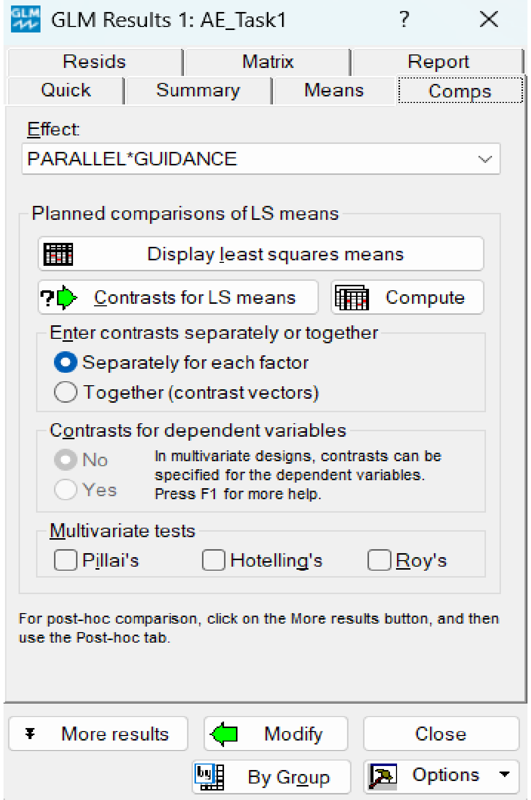
Fig. 22 Within-subjects factors specification

It is possible to verify if each variable column was assigned to the correct combination of factor levels by opening the *marginal tables, observed unweighted* from within the *plot or show means*. This must be done after running the analysis. *Table 7* has the correct order of within subjects' factors for each condition.

Table 7: Within-subjects factors for each condition

Parallel Factor	**Guidance Factor**	**Condition**
1	1	P0_H
1	2	P0_H+V
1	3	P0_V
2	1	P1_H
2	2	P1_H+V
2	3	P1_V

After properly defining the variables and the within-subject factors, STATISTICA will compute the main effects for each within-subject factor and the interaction between them. It will also make a Mauchyl's test for sphericity, with Greenhouse-geisser corrections if needed. Lastly it will include multivariate and univariate tests. The reason why there are no results of the Mauchly's Test of Sphericity presented is that one of the factors has only two levels of within subject's factor. The result tab appears as in *figure 23*.



Effect
PARALLEL*GUIDANCE

Enter contrasts separately or together
● Separately for each factor
○ Together (contrast vectors)

Contrasts for dependent variables
● No
○ Yes
In multivariate designs, contrasts can be specified for the dependent variables. Press F1 for more help.

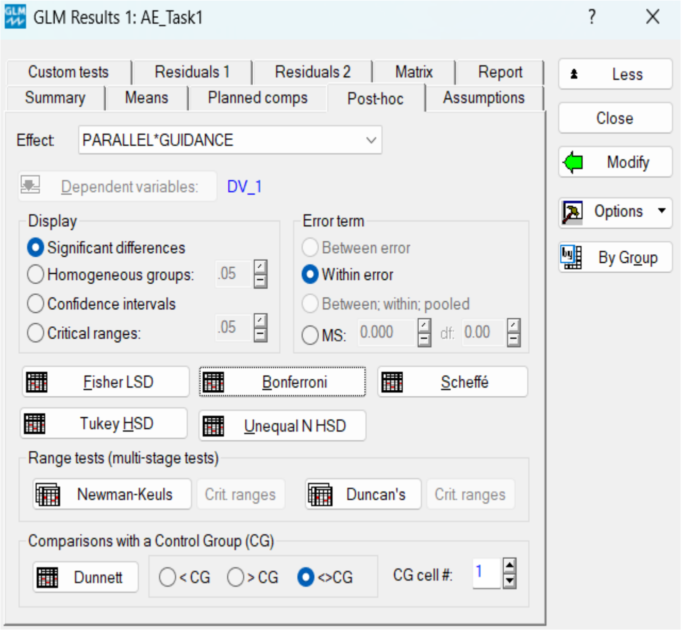
Multivariate tests
☐ Pillai's ☐ Hotelling's ☐ Roy's

For post-hoc comparison, click on the More results button, and then use the Post-hoc tab.

Fig. 23 General linear model repeated measures ANOVA results

For the factor to be considered statistically significant, the p-value or significance value must be less than 0.05. The meaning behind this value is that there is less of a 5% probability that the observed differences between groups means values are due to random chances. With a p-value lower than the minimum of 5%, the null hypothesis is rejected, and it is concluded that there is a significant difference between at least two of the factors. For all the presented STATISTICA results, when the p-values are colored red, it means that they are statistically significant. This can be seen either on the *tests of within-subjects effects* tables, or on the *post-hoc analysis results* tables. This is a key element for data interpretation and analysis. It determines if the ANOVA hypothesis is fulfilled or not.

In the tests of within-subjects effects tables it is possible to analyze the main effect of each individual factor. If they have a statistically significant effect, then it's necessary to compute a post-hoc comparison also known as post-hoc analysis for each of them. In STATISTICA for a post-hoc comparison, it is necessary to press on the more result button, and then use the post-hoc tab. Within this tab it is possible to choose a method to display the significant differences between variables. The alternatives are Fisher LSD, Bonferroni, Scheffe, TukeyHSD and Unequal N HSD. The post-hoc tab where is possible to select the test method appears as in *figure 24*.



GLM Results 1: AE_Task1

Effect: PARALLEL*GUIDANCE

Dependent variables: DV_1

Display
● Significant differences
○ Homogeneous groups: .05
○ Confidence intervals
○ Critical ranges: .05

Error term
○ Between error
● Within error
○ Between; within; pooled
○ MS: 0.000 df: 0.00

Comparisons with a Control Group (CG)
○ < CG ○ > CG ● <>CG
CG cell #: 1

Fig. 24 Post-hoc test method selection

Post-hoc comparison methods are used following a significant ANOVA to determine which group means differ. These methods vary in how they balance control of Type I error and statistical power. Fisher’s LSD is a test with high power but weak control over Type I error, making it suitable only for a small number of comparisons. Bonferroni correction is highly conservative, reducing the risk of false positives but often at the cost of statistical power, especially when many comparisons are made. Scheffé’s test is the most conservative method, appropriate for testing complex or non pairwise contrasts, but it has the lowest power. Tukey’s HSD provides a good balance between Type I error control and power for all pairwise comparisons, assuming equal group sizes. When sample sizes are unequal, the Tukey-Kramer method is preferred as it adjusts for this imbalance while maintaining solid control over Type I error. Overall, Tukey’s methods are generally the best choice for post-hoc pairwise comparisons due to their statistical balance, while Bonferroni and Scheffé are more appropriate when stricter error control is necessary. Fisher LSD, Scheffe, TukeyHSD and Unequal N HSD, are commonly used for between-subjects designs.

When it comes to two-way repeated measures ANOVA with two within-subjects’ factors (factor A: 3 levels, factor B: 2 levels), post-hoc testing depends on whether the significant effects lie on the main effects or the interaction. Traditional methods like Tukey HSD, Fisher LSD or Scheffé are not appropriate for repeated measures designs because they assume independent samples. In other words, these methods will ignore the within-subjects’ correlations, which will inflate the type 1 error, leading to invalid results. With this into consideration, to be able to keep a standard statistical procedure, Bonferroni was determined to be the best alternative for the post-hoc comparisons of the experiment, since it doesn't assume independent groups and is suitable for repeated measures. [36]

After selecting the Bonferroni post-hoc method and running the pairwise comparison, it is possible to analyze the paired samples test. Here once again the p-value will determine if there is a statistically significant result on the dependent variables within the factor. All dependent variables must be statistically significant to determine if there is a main effect on the factor.

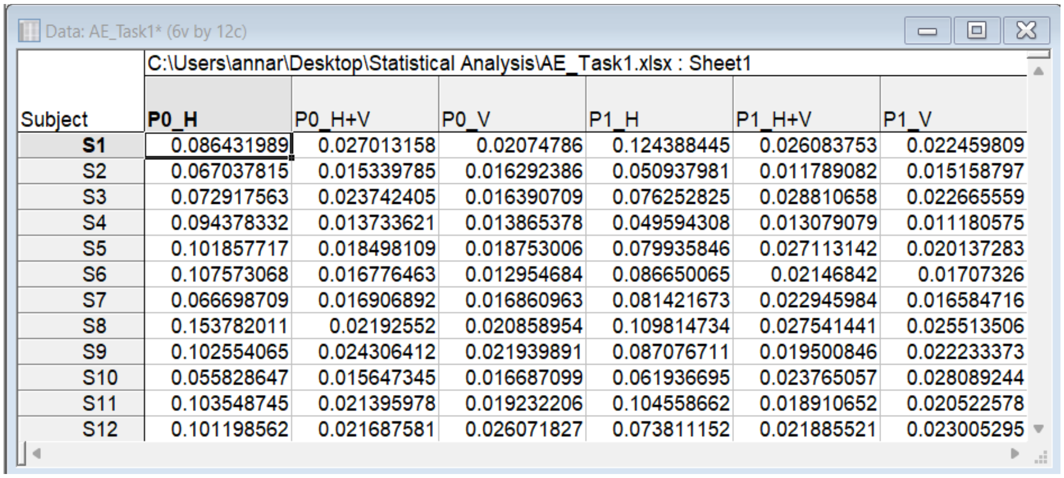
As previously mentioned, Bonferroni is a very conservative method, which can lead to large, adjusted p-values of 1, or small adjusted p-values of 0. The reason for the p-values of 1 is that sometimes when the unadjusted p-value is relatively large, even after the Bonferroni correction, it will still exceed 1. Since p-values cannot be greater than 1, it is capped and appears as 1 in the post-hoc result. This suggests that there is no statistical evidence for a difference in the comparison.

In case of a p-value of 0, it typically results from very small unadjusted p-values. If after the Bonferroni correction it is a small number still, then it may be rounded to 0 in the output, which does not mean that is absolute zero, but that is very close to it. This indicates very strong evidence against the null hypothesis. [37]

Optional graphs also help to visualize the effects. The profile plot of the results may be generated by STATISTICA. In this chart it is possible to observe the effect of both factors on the estimated marginal means of the analyzed variable. On this graph, each curve represents a different guidance method (guidance1: haptic, guidance 2: visual + haptic, guidance 3: visual). The x-axis of the plot is for the cognitive load, 1 refers to when the parallel task is inactive, and 2 for when it is active. The y-axis is for the estimated marginal means of the measured dependent variable, and it appears as DV_1.

6. Results

Average error task 1



C:\Users\annar\Desktop\Statistical Analysis\AE_Task1.xlsx : Sheet1

Subject	**P0_H**	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
S1	0.086431989	0.027013158	0.02074786	0.124388445	0.026083753	0.022459809
S2	0.067037815	0.015339785	0.016292386	0.050937981	0.011789082	0.015158797
S3	0.072917563	0.023742405	0.016390709	0.076252825	0.028810658	0.022665559
S4	0.094378332	0.013733621	0.013865378	0.049594308	0.013079079	0.011180575
S5	0.101857717	0.018498109	0.018753006	0.079935846	0.027113142	0.020137283
S6	0.107573068	0.016776463	0.012954684	0.086650065	0.02146842	0.01707326
S7	0.066698709	0.016906892	0.016860963	0.081421673	0.022945984	0.016584716
S8	0.153782011	0.02192552	0.020858954	0.109814734	0.027541441	0.025513506
S9	0.102554065	0.024306412	0.021939891	0.087076711	0.019500846	0.022233373
S10	0.055828647	0.015647345	0.016687099	0.061936695	0.023765057	0.028089244
S11	0.103548745	0.021395978	0.019232206	0.104558662	0.018910652	0.020522578
S12	0.101198562	0.021687581	0.026071827	0.073811152	0.021885521	0.023005295

Fig. 25 Average error task 1 measurements

Table 8: Tests of within-subjects' effects for average error task 1

Effect	**Type III Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F-value**	**p-value**
PARALLEL	0.000084	1	0.000084	0.7085	0.417868
Error	0.001296	11	0.000118		
GUIDANCE	0.072710	2	0.036355	141.1728	0.000000
Error	0.005665	22	0.000258		
PARALLEL* GUIDANCE	0.000645	2	0.000323	3.3043	0.055614
Error	0.002147	22	0.000098		

All the average error task 1 measurements are displayed in *figure 25*. *Table 8* indicates that only the guidance factor is statically significant for the AV variable in task 1. The p-value for the guidance factor is rounded to 0. The parallel factor and the interaction effect between both factors have a p-value higher to 0.05. They have a p-value of 0.417868 and 0.055614 respectively. To know exactly which dependent variables differ from the rest and for how much, it is necessary to perform a post hoc analysis. *PARALLEL*GUIDANCE* is the interaction effect between both factors.

Table 9: Post-hoc results for average error task 1

	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
P0_H		0.000000	0.000000	0.227922	0.000000	0.000000
P0_H+V	0.000000		1.000000	0.000000	1.000000	1.000000
P0_V	0.000000	1.000000		0.000000	1.000000	1.000000
P1_H	0.227922	0.000000	0.000000		0.000000	0.000000
P1_H+V	0.000000	1.000000	1.000000	0.000000		1.000000
P1_V	0.000000	1.000000	1.000000	0.000000	1.000000	

The Bonferroni post-hoc comparison in *table 9* reveals several statistically significant differences between experimental conditions. Most notably, the conditions P0_H and P1_H show significant differences when compared to other conditions, including P0_H+V, P0_V, P1_H+V, and P1_V. This indicates that the effects of haptic guidance are meaningfully distinct from these other groups regardless of the cognitive load. In contrast, comparisons among P0_H+V, P0_V, P1_H+V, and P1_V all show p-values of 1, suggesting that these conditions do not differ significantly from one another after adjusting for multiple comparisons. Additionally, the comparison between P0_H and P1_H has a moderate adjusted p-value of 0.227922, which is not statistically significant, implying that these two conditions are relatively similar. Overall, the pattern of results shows that P0_H and P1_H have the biggest deviation in terms of its statistical profile, while the remaining conditions tend to form a homogeneous group with no significant differences among them after Bonferroni correction.

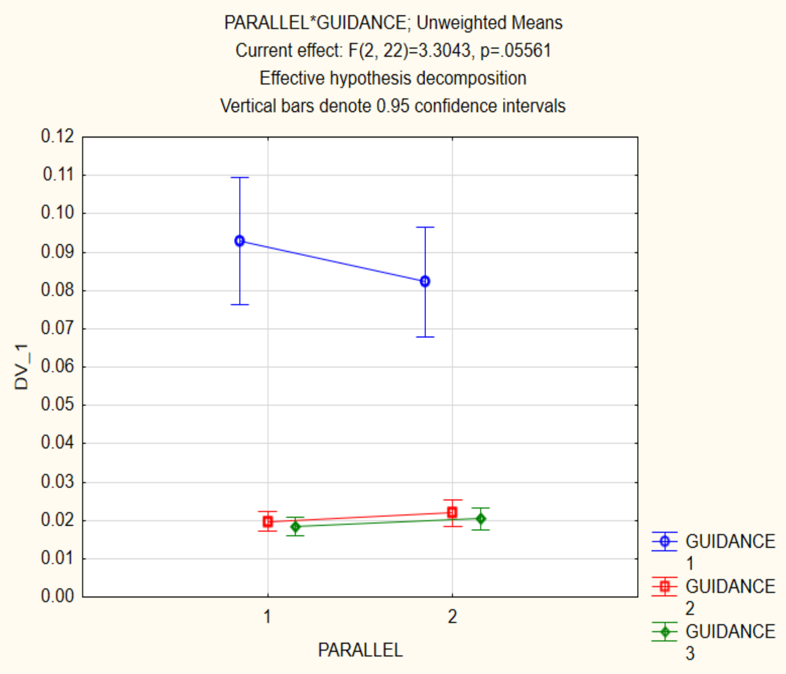


Fig. 26 Profile plot for average error task 1

By analyzing the profile plot in *figure 26*, the biggest change in the mean value of AE happens for the haptic guidance method. The AE marginal means drops from approximately 0.092 to 0.08 when the parallel task is initiated. For the other two guidance methods the effect of the parallel task seems to be minimal. The haptic guidance has significantly higher mean values of AE.

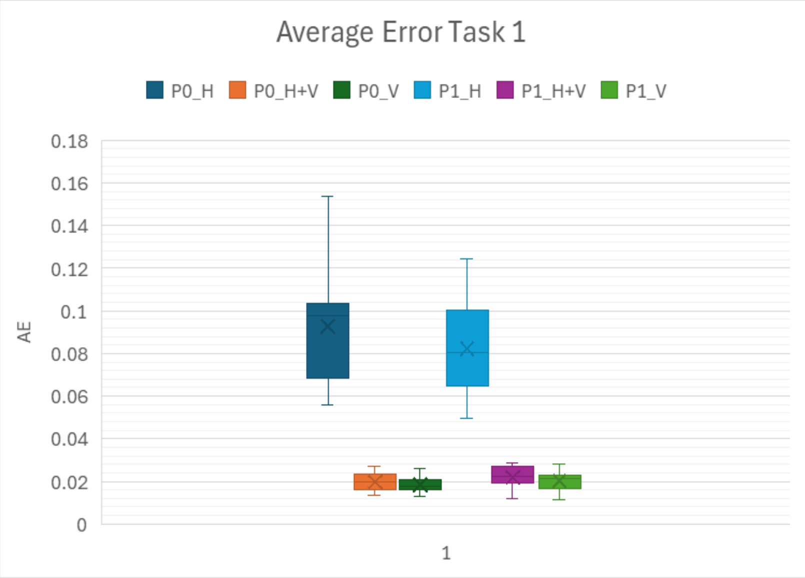
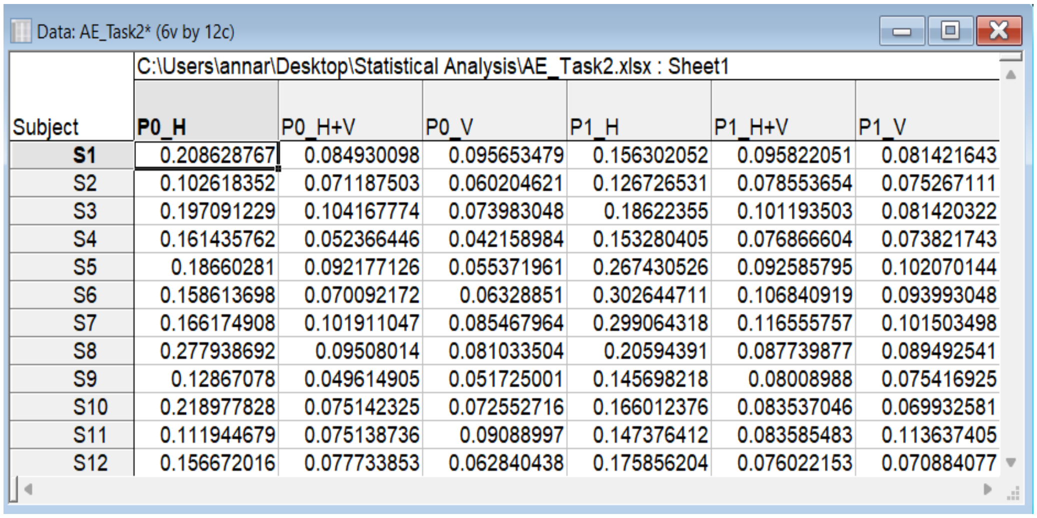


Fig. 27 Box plot for average error task 1

The box plot in *figure 27* also shows how the average error varies more for haptic guidance, independently of the parallel task condition. The ranges of values for between guidance conditions P0_H and P1_H are the largest. Additionally, the mean AE for both conditions is high compared to the rest, which explains the statistical significance. Most of the AE measurements for these conditions are kept within the same range, with the exception of P0_H maximum value. This value is not considered to be an outlier, but explains the AE marginal mean difference seen in the profile plot for the haptic guidance method between P0_H and P1_H.

Average error task 2



C:\Users\annar\Desktop\Statistical Analysis\AE_Task2.xlsx : Sheet1

Subject	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
S1	0.208628767	0.084930098	0.095653479	0.156302052	0.095822051	0.081421643
S2	0.102618352	0.071187503	0.060204621	0.126726531	0.078553654	0.075267111
S3	0.197091229	0.104167774	0.073983048	0.18622355	0.101193503	0.081420322
S4	0.161435762	0.052366446	0.042158984	0.153280405	0.076866604	0.073821743
S5	0.18660281	0.092177126	0.055371961	0.267430526	0.092585795	0.102070144
S6	0.158613698	0.070092172	0.06328851	0.302644711	0.106840919	0.093993048
S7	0.166174908	0.101911047	0.085467964	0.299064318	0.116555757	0.101503498
S8	0.277938692	0.09508014	0.081033504	0.20594391	0.087739877	0.089492541
S9	0.12867078	0.049614905	0.051725001	0.145698218	0.08008988	0.075416925
S10	0.218977828	0.075142325	0.072552716	0.166012376	0.083537046	0.069932581
S11	0.111944679	0.075138736	0.09088997	0.147376412	0.083585483	0.113637405
S12	0.156672016	0.077733853	0.062840438	0.175856204	0.076022153	0.070884077

Fig. 28 Average error task 2 measurements

Table 10: Tests of within-subjects' effects for average error task 2

Effect	**Type III Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F-value**	**p-value**
PARALLEL	0.004684	1	0.004684	3.6472	0.082575
Error	0.014127	11	0.001284		
GUIDANCE	0.168864	2	0.084432	82.7224	0.000000
Error	0.022455	22	0.001021		
PARALLEL* GUIDANCE	0.000338	2	0.000169	0.2510	0.780221
Error	0.014805	22	0.000673		

Table 10 shows that guidance is the only statistically significant factor with a p-value rounded to 0. Parallel factor has a p-value of 0.082575 and the interaction of both guidance and parallel factors has a p-value of 0.780221. The measurements for this dependent variable can be seen in *figure 28*.

Table 11: Post-hoc results for average error task 2

	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
P0_H		0.000000	0.000000	0.829632	0.000001	0.000001
P0_H+V	0.000000		1.000000	0.000000	1.000000	1.000000
P0_V	0.000000	1.000000		0.000000	1.000000	1.000000
P1_H	0.829632	0.000000	0.000000		0.000000	0.000000
P1_H+V	0.000001	1.000000	1.000000	0.000000		1.000000
P1_V	0.000001	1.000000	1.000000	0.000000	1.000000	

Table 11 reveal several statistically significant differences between the conditions. Haptic conditions P0_H and P1_H have a significant difference with a p-value of 0 when compared to P0_H+V, P0_V, P1_H+V, and P1_V. These results demonstrate that the addition of visual or combined stimuli has a meaningful effect on the outcome measures, distinguishing them significantly from the haptic-only conditions. However, no significant difference was found between P0_H and P1_H, indicating consistency in responses across time or phases for the haptic condition. The p-value between both haptic conditions is of 0.829632.

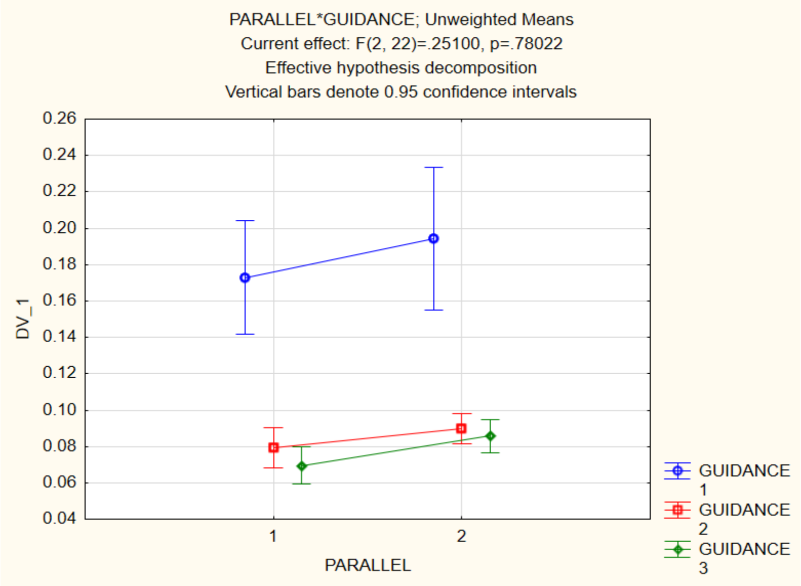


Fig. 29 Profile plot for average error task 2

Figure 29 consists of a profile plot, where it is possible to graphically observe the differences between the three different guidance types and the effect of the parallel task. All three guidance methods seem to have the same reaction to the parallel task, since the average error increases with the condition P1 in all cases. Still, the haptic guidance curve has the most significant growth on the AE marginal means compared to the other two, with a difference of 0.02 points between P0_H and P1_H. Visual guidance increases its AE approximately by 0.01 points, and the combination of visual and haptic roughly increases 0.005. Also, haptic guidance shows higher values of AE overall.

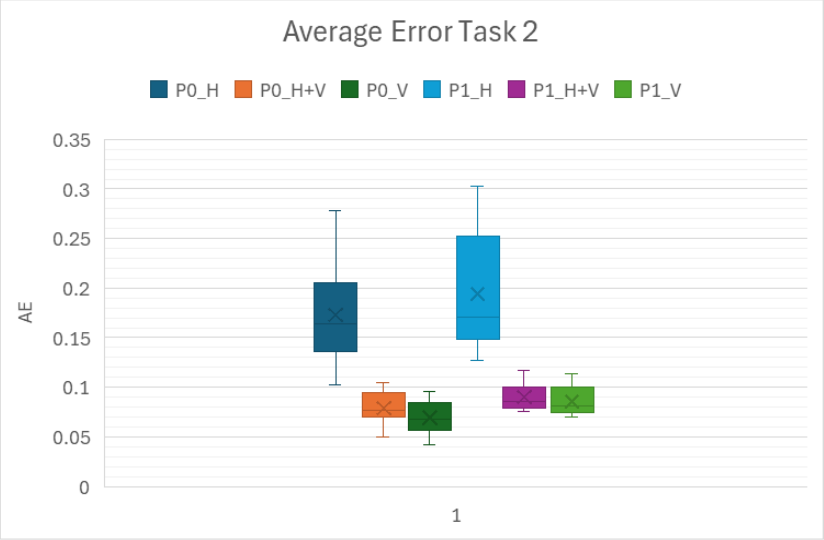
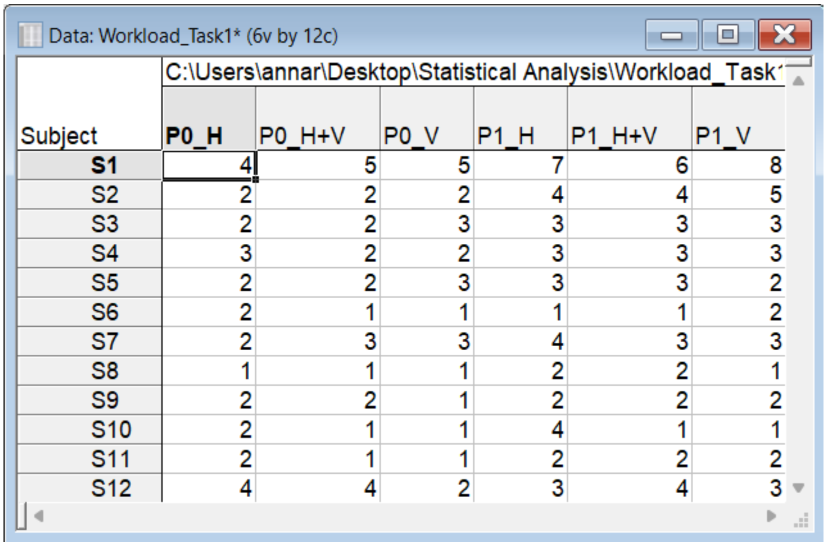


Fig. 30 Box plot for average error task 2

The box plot in *figure 30* illustrates the Average Error (AE) for Task 2 under six experimental conditions, varying by sensory input and cognitive load. The highest errors and variability are observed in the haptic guidance conditions (P0_H and P1_H), indicating that relying solely on haptic input leads to less accurate and less consistent performance, particularly under cognitive load (P1_H). In contrast, conditions involving visual input (P0_V, P0_H+V, P1_V, P1_H+V) show markedly lower average errors and reduced variability.

Workload task 1



Subject	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
S1	4	5	5	7	6	8
S2	2	2	2	4	4	5
S3	2	2	3	3	3	3
S4	3	2	2	3	3	3
S5	2	2	3	3	3	2
S6	2	1	1	1	1	2
S7	2	3	3	4	3	3
S8	1	1	1	2	2	1
S9	2	2	1	2	2	2
S10	2	1	1	4	1	1
S11	2	1	1	2	2	2
S12	4	4	2	3	4	3

Fig. 31 Workload task 1 measurements

Table 12: Tests of within-subjects’ effects for workload task 1

Effect	**Type III Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F-value**	**p-value**
PARALLEL	10.8889	1	10.8889	12.25000	0.004970
Error	9.7778	11	0.8889		
GUIDANCE	1.0000	2	0.5000	1.06452	0.362002
Error	10.3333	22	0.4697		
PARALLEL* GUIDANCE	0.1111	2	0.0556	0.13253	0.876571
Error	9.2222	22	0.4192		

It can be seen in *table 12* that for the workload variable in task 1, the only statistically significant factor is the parallel effect. Parallel factors have a p-value of 0.004970, while guidance factor and the interaction effect between factors both show a p-value of 0.362002 and 0.876571, respectively. To be considered statistically significant, the p-value must be higher than 0.05. The workload measurements for task 1 are summarized in *figure 31*.

Table 13: Post-hoc results for workload task 1

	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
P0_H		1.000000	1.000000	0.069268	1.000000	0.570839
P0_H+V	1.000000		1.000000	0.015325	0.291032	0.143742
P0_V	1.000000	1.000000		0.007114	0.143742	0.069268
P1_H	0.069268	0.015325	0.007114		1.000000	1.000000
P1_H+V	1.000000	0.291032	0.143742	1.000000		1.000000
P1_V	0.570839	0.143742	0.069268	1.000000	1.000000	

The Bonferroni post hoc results from this GLM repeated measures ANOVA in *table 13* reveal less statistically significant differences compared to the previous tables. Since red values represent significant p-values ($p < 0.05$), the data show that the P0_H+V and P0_V conditions differ significantly from the P1_H condition, with p-values of 0.015325 and 0.007114, respectively. These findings demonstrate that the haptic condition with parallel tasks active (P1_H) has significant changes in participant responses when compared to conditions without additional cognitive load that include visual stimuli. No pairwise comparisons between the same guidance type with a different cognitive load reached statistical significance after Bonferroni adjustment. This indicates that there is no major effect due to the transitions from phase 0 to phase 1 exclusively.

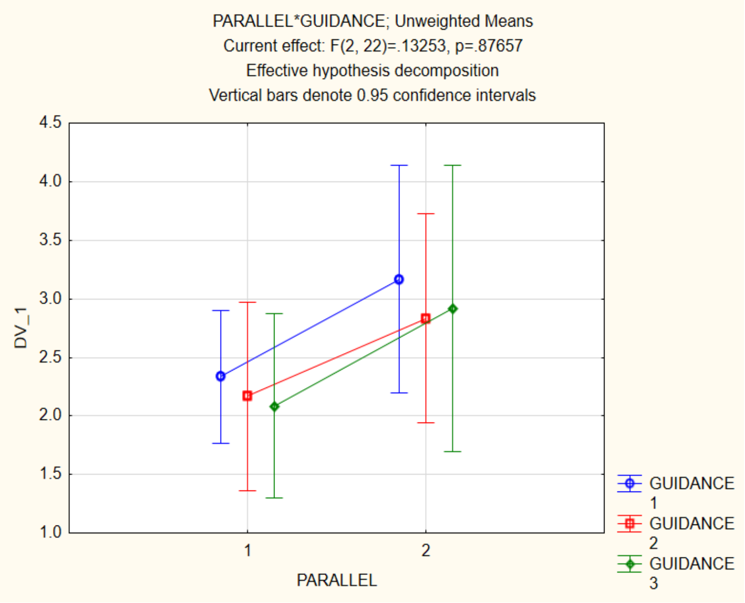


Fig. 32 Profile plot for workload task 1

For this variable, the profile plot in *figure 32* shows a similar behavior of all guidance types when transitioning from phase 0 to phase 1. The growth in the workload marginal means is similar for all guidance types, with an increase of approximately 1 point after the transition. The main difference resides in the growth of the visual guidance, which starts with the lowest workload with approximately 2 points and ends up as the second highest with the most drastic increase. By comparing the overall workload marginal means, haptic guidance shows slightly higher values. Visual guidance and the combination of haptic and visual are similar to each other.

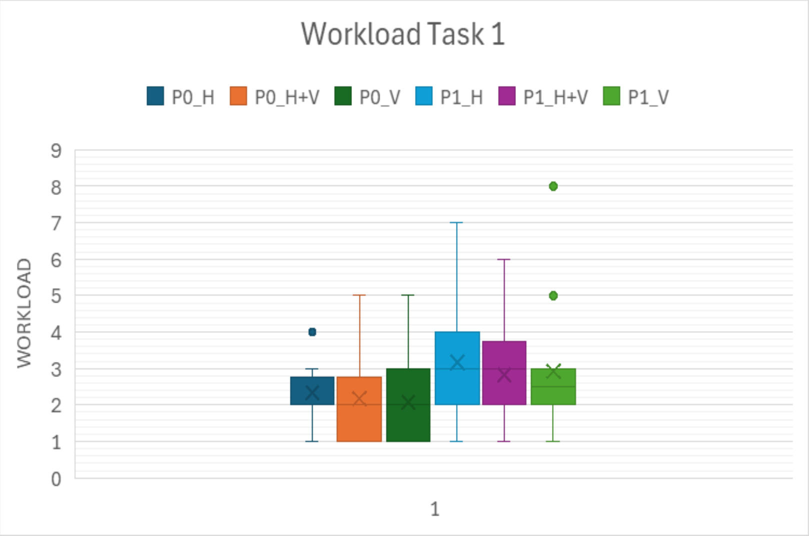
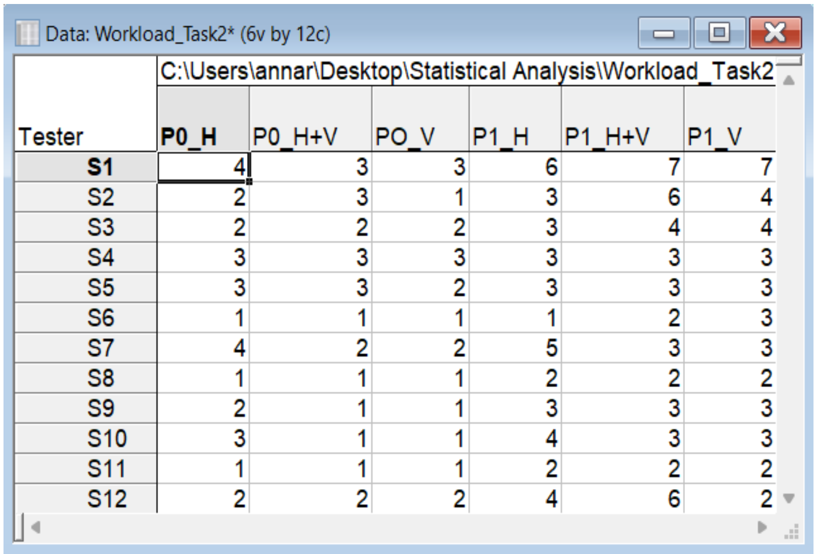


Fig. 33 Box plot for workload task 1

Figure 33 shows the workload ratings for Task 1 across the same six experimental conditions, with workload measured on a scale from 0 to 9. Overall, participants reported moderate perceived workload, with median values ranging between approximately 2 and 4. The lowest median workload appears in the P0_H with approximately 2.5 points and P0_H+V with approximately 2.2. Similarly, P0_V and P1_V show medians around 2.5 to 3. Conditions involving a parallel task (P1_H, P1_H+V, P1_V) show slightly higher mean and spread, with P1_H showing the highest mean of 3.6 and widest range, suggesting increased perceived workload when only haptic feedback is available under dual-task conditions. Notably, P1_H and P1_H+V each include high outliers above 5, with one point in P1_H+V reaching 8. These outliers represent individual participants who experienced significantly more strain than others, possibly due to difficulty managing both the primary and parallel tasks simultaneously, even when visual input was present.

Workload task 2



Tester	P0_H	P0_H+V	PO_V	P1_H	P1_H+V	P1_V
S1	4	3	3	6	7	7
S2	2	3	1	3	6	4
S3	2	2	2	3	4	4
S4	3	3	3	3	3	3
S5	3	3	2	3	3	3
S6	1	1	1	1	2	3
S7	4	2	2	5	3	3
S8	1	1	1	2	2	2
S9	2	1	1	3	3	3
S10	3	1	1	4	3	3
S11	1	1	1	2	2	2
S12	2	2	2	4	6	2

Fig. 34 Workload task 2 measurements

Table 14: Tests of within-subjects' effects for workload task 2

Effect	**Type III Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F-value**	**p-value**
PARALLEL	36.1250	1	36.1250	29.71028	0.000201
Error	13.3750	11	1.2159		
GUIDANCE	1.7778	2	0.8889	1.09317	0.352678
Error	17.8889	22	0.8131		
PARALLEL* GUIDANCE	2.3333	2	1.1667	3.85000	0.036841
Error	6.6667	22	0.3030		

Workload in task 2 is the first dependent variable that shows two statistically significant factors in the tests of within-subject effects table. As seen in *table 14*, the parallel factor has a p-value of 0.000201, and the interaction effect between guidance and parallel factor has a p-value of 0.036841. With both p-values being lower than 0.05, they are considered to be significant. On the other hand, the remaining guidance factor has a p-value of 0.352678. For further analysis, *figure 34* contains the measurements of the workload for task 2.

Table 15: Post-hoc results for workload task 2

	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
P0_H		1.000000	0.106897	0.007463	0.000085	0.007463
P0_H+V	1.000000		1.000000	0.000085	0.000001	0.000085
P0_V	0.106897	1.000000		0.000007	0.000000	0.000007
P1_H	0.007463	0.000085	0.000007		1.000000	1.000000
P1_H+V	0.000085	0.000001	0.000000	1.000000		1.000000
P1_V	0.007463	0.000085	0.000007	1.000000	1.000000	

The Bonferroni post hoc results in *table 15* reveal a strong and consistent pattern of statistically significant differences across nearly all condition comparisons involving haptic (H), visual (V), and combined (H+V) sensory modalities across the two phases (P0 and

P1). Statistical significance is present in nearly every pairwise comparison between phase 0 and phase 1 conditions. Specifically, P1_H, P1_H+V, and P1_V all differ significantly from each of the phase 0 conditions: P0_H, P0_H+V, and P0_V. For example, P0_H vs P1_H (p = 0.007463), P0_H+V vs P1_H+V (p = 0.000001), and P0_V vs P1_V (p = 0.000007) are all highly significant. Even within the same sensory modality, the phase transition leads to statistically different outcomes, indicating a robust phase effect regardless of modality. These results demonstrate that participants' responses changed significantly when the parallel task was active across all sensory input conditions. The only pairwise comparisons that do not show significance are those within the same phase.

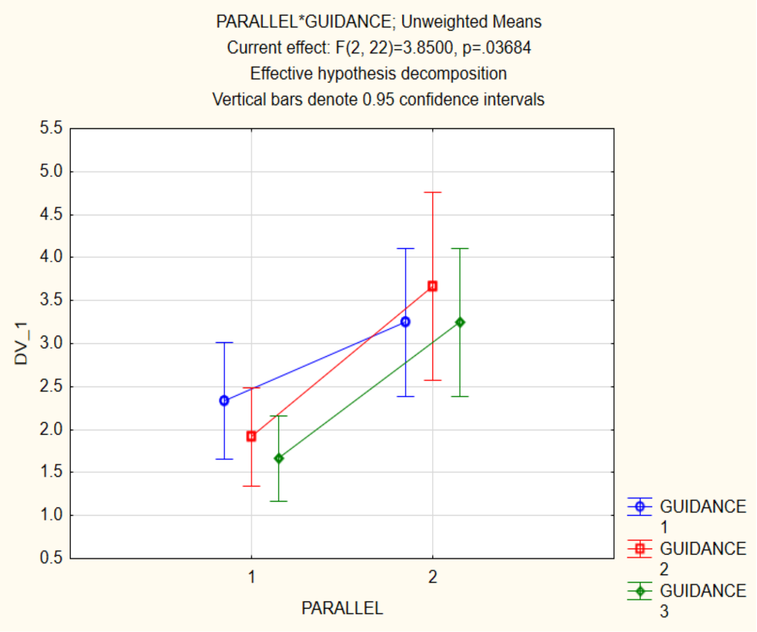


Fig. 35 Profile plot for workload task 2

Figure 35 profile plot displays the interaction between guidance type and the presence of a parallel task on subjective workload. All guidance types show an increase in workload when a parallel task is introduced, but the growth varies. The most pronounced rise occurs in the haptic + visual condition, which jumps from a workload level of approximately 2 points under no parallel task to the highest with the task present of 3.7. In contrast, the visual condition starts with the lowest workload at 1.8 points and exhibits a more moderate increase to 3.2. The haptic condition shows a relatively steady and moderate rise. It starts with the highest workload of 2.4 points and ends up with just 3.2.

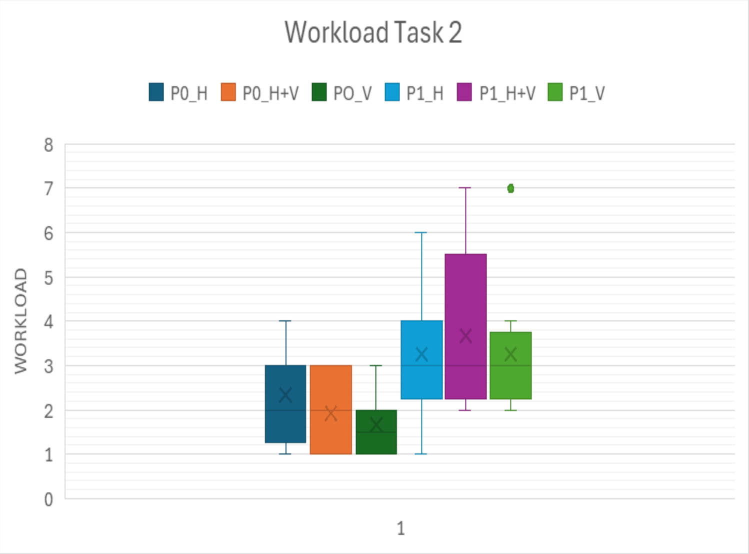
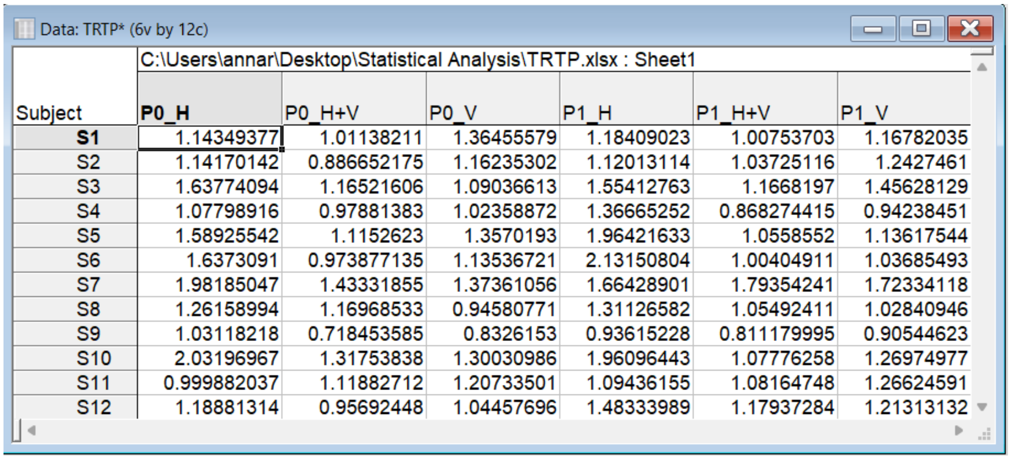


Fig. 36 Box plot for workload task 2

The box plot in *figure 36* presents the workload ratings for Task 2 across the same six experimental conditions, using a 0-9 scale. The lowest median workload is observed in the P0_V and P0_H+V conditions, both around 2, indicating that when no parallel task is present, participants found the task easier to fulfill with visual or multimodal feedback. As expected, conditions involving a parallel task (P1) generally show higher workload, with P1_H+V standing out with the highest median of 4.5 and largest variability, including a maximum value of 7. An outlier can also be identified for P1_V condition at 7. P1_H and P1_V also show elevated medians between 3.5 and 4 points, confirming that the presence of a parallel task consistently raises workload regardless of sensory modality.

Time to reach target position task 1



C:\Users\annar\Desktop\Statistical Analysis\TRTP.xlsx : Sheet1

Subject	**P0_H**	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
S1	1.14349377	1.01138211	1.36455579	1.18409023	1.00753703	1.16782035
S2	1.14170142	0.886652175	1.16235302	1.12013114	1.03725116	1.2427461
S3	1.63774094	1.16521606	1.09036613	1.55412763	1.1668197	1.45628129
S4	1.07798916	0.97881383	1.02358872	1.36665252	0.868274415	0.94238451
S5	1.58925542	1.1152623	1.3570193	1.96421633	1.0558552	1.13617544
S6	1.6373091	0.973877135	1.13536721	2.13150804	1.00404911	1.03685493
S7	1.98185047	1.43331855	1.37361056	1.66428901	1.79354241	1.72334118
S8	1.26158994	1.16968533	0.94580771	1.31126582	1.05492411	1.02840946
S9	1.03118218	0.718453585	0.8326153	0.93615228	0.811179995	0.90544623
S10	2.03196967	1.31753838	1.30030986	1.96096443	1.07776258	1.26974977
S11	0.999882037	1.11882712	1.20733501	1.09436155	1.08164748	1.26624591
S12	1.18881314	0.95692448	1.04457696	1.48333989	1.17937284	1.21313132

Fig. 37 Time to reach target position task 1

Table 16: Tests of within-subjects' effects for time to reach target position task 1

Effect	**Type III Sum of Squares**	**Degree of Freedom**	**Mean Square**	**F-value**	**p-value**
PARALLEL	0.0497	1	0.0497	4.0457	0.069436
Error	0.1351	11	0.0123		
GUIDANCE	1.6212	2	0.8106	12.7228	0.000213
Error	1.4017	22	0.0637		
PARALLEL* GUIDANCE	0.0123	2	0.0062	0.2689	0.766708
Error	0.5033	22	0.0229		

In *table 16* it can be seen how for the TRTP dependent variable there is only one factor that is statistically significant. Guidance factor has a p-value of 0.000213, being the only one lower than 0.05. The parallel effect has a p-value of 0.069436 and the interaction effect between both factors is of 0.766708. All the recorded times to reach target position in task 1 can be seen in *figure 37*.

Table 17: Post-hoc results for time to reach target position task 1

	P0_H	P0_H+V	P0_V	P1_H	P1_H+V	P1_V
P0_H		0.000452	0.011720	1.000000	0.001172	0.069712
P0_H+V	0.000452		1.000000	0.000017	1.000000	0.738049
P0_V	0.011720	1.000000		0.000376	1.000000	1.000000
P1_H	1.000000	0.000017	0.000376		0.000041	0.002278
P1_H+V	0.001172	1.000000	1.000000	0.000041		1.000000
P1_V	0.069712	0.738049	1.000000	0.002278	1.000000	

The Bonferroni-adjusted post hoc results for TRTP variable seen in *table 17*, reveal several significant differences between all sensory modalities (haptic, visual and the combination of both) across the two phases (P0 and P1). Notably, significant differences were observed between P0_H and both P0_H+V and P0_V, indicating that at the initial stage, performance varied depending on whether the input was of a single sense or multisensory. Similarly, P0_H differed significantly from P1_H+V, suggesting a combined effect of modality and parallel task presence. Within the P1 phase, P1_H showed significant differences compared to P0_H+V, P0_V, P1_H+V, and P1_V, highlighting considerable changes in performance with a higher cognitive load, especially for the haptic condition. Additionally, significant differences between P1_H and P1_H+V and between P1_H and P1_V point to distinct modality effects when there is an active parallel task.

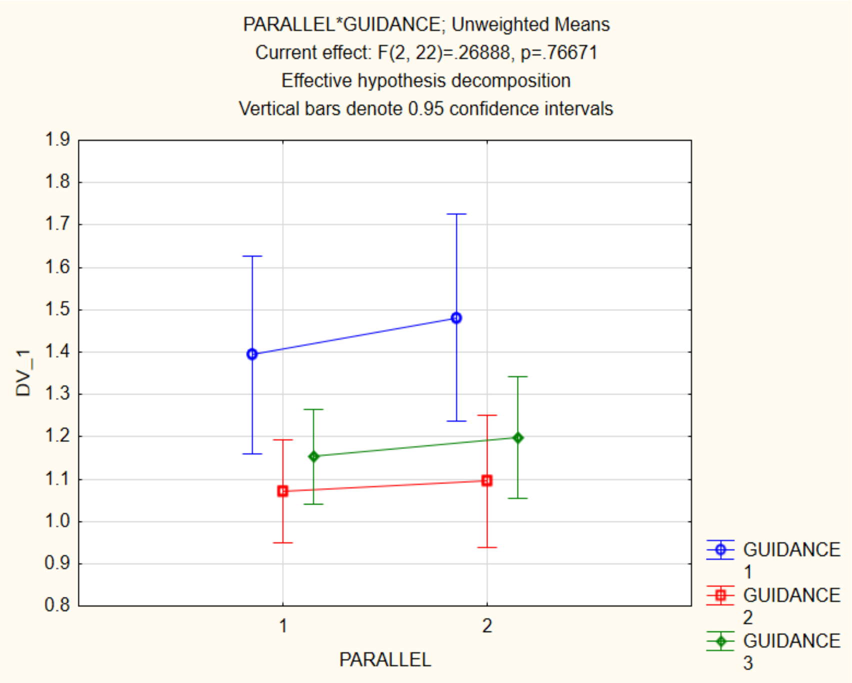


Fig. 38 Profile plot for time to reach target position task 1

The profile plot of dependent variable TRTP can be seen in *figure 38* and shows that across all guidance conditions, the presence of a parallel task produces only minimal changes in performance. The haptic + visual guidance condition shows the lowest mean time to reach the target in both parallel task conditions, followed by visual and then haptic alone. Overall, the time to reach the target appears relatively stable across all combinations of guidance type and parallel task presence. However, the haptic guidance method has higher recorded values compared to the other two methods, which explains the significance difference from the tests within-subjects effects.

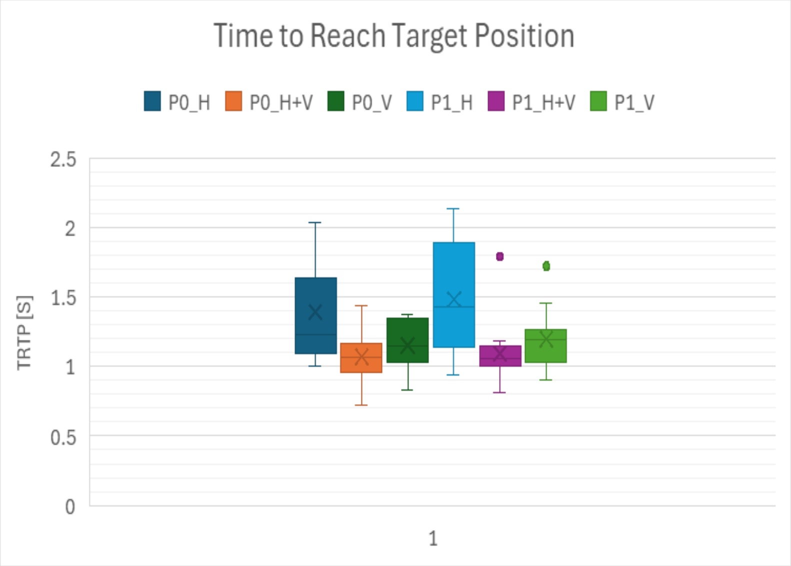


Fig. 39 Box plot for time to reach target position task 1

Figure 39 illustrates that when only haptic guidance (H) is used, TRTP increases and becomes more variable. A higher variability and mean value are observed under parallel tasking (P1_H). In contrast, visual guidance (V) alone results in relatively short and consistent TRTP in both P0_V and P1_V conditions, showing greater resilience to multitasking. The combination of haptic and visual guidance (H+V) shows the best overall performance. Specifically, the P0_H+V condition demonstrates the shortest and most consistent TRTP. Overall, the addition of visual guidance significantly enhances task performance and mitigates the negative impact of multitasking, particularly when used in combination with haptic guidance.

7. Discussion

The results of this study highlight meaningful differences in participants' performance across various guidance methods and cognitive load conditions, as measured by average error (AE), time to reach the target position (TRTP), and subjective workload. Each dependent variable exhibited distinct response patterns, suggesting that the type of guidance and presence of a parallel task affect performance outcomes in different ways.

For Average Error, which reflects task accuracy, haptic guidance consistently produced the highest error rates across both tasks. This was particularly evident when participants were multitasking (P1), indicating that haptic input alone is susceptible to misinterpretation or is harder to follow under high cognitive load. In contrast, both visual and combined guidance modalities led to significantly lower error rates. Importantly, there were no major differences between visual only and combined modalities, suggesting that the visual component is the reason behind the accuracy benefits. Adding haptics stimuli does not significantly improve or impair this performance. This implies that visual guidance is more precise and intuitive, especially under high cognitive load, whereas haptic guidance alone is less reliable in maintaining accuracy.

Looking at Time to Reach the Target Position (TRTP), a similar effect is observed. Participants took the longest and most inconsistent times to reach the target when using haptic guidance alone. Visual guidance, on the other hand, enabled faster and more consistent task completion. The combined guidance condition performed similarly well, sometimes even better, indicating that adding haptic feedback to visual input does not slow down performance but improves it. However, the differences in TRTP across phases (with and without a parallel task) were not as pronounced, suggesting that timing is less affected by multitasking than accuracy or perceived workload. Nonetheless, the data reinforces the idea that visual input enhances task efficiency, and that haptic guidance, while slower, becomes especially inefficient under increased cognitive demands.

The Workload data offers insight into how demanding each guidance method is on the participant. Across both tasks, the presence of a parallel task (P1) significantly increased perceived workload across all conditions. However, haptic guidance was consistently rated as more mentally demanding than the visual and combined guidance types. While visual guidance started with the lowest workload ratings, the combined guidance method (H+V) showed the highest increase in perceived workload in Task 2 when multitasking was introduced. This indicates that while combined guidance supports better performance in terms of accuracy and speed, it may come at a cognitive cost due to the need to integrate information from multiple sensory channels. Haptic guidance, on the other hand, appeared to demand higher mental effort regardless of task phase, possibly because interpreting haptic signals without visual support is more challenging.

Overall, these findings demonstrate that visual guidance is the most effective modality for multitasking scenarios, offering high accuracy, fast task completion, and low mental workload. Combined guidance provides slightly higher performance benefits but may increase cognitive load due to the complexity of dual sensory input. Meanwhile, haptic guidance is the least favorable, associated with higher error, slower responses, and greater perceived workload, especially under cognitive stress.

Based on the hypotheses and the interpretation of the results for each dependent variable, it is possible to evaluate which hypotheses were supported and under which tasks. The guidance method had a significant main effect on the AE for both tasks, and on the TRTP for task 1. Hence, the alternative hypothesis (H1) for guidance factor is supported by these two variables only. Parallel factor had a clear effect only on workload variables, with statistically significant p-values on the workload for both tasks 1 and 2. Lastly, the interaction effects show a null hypothesis (H0) for almost all variables, with the exception of workload in task 2. All the other dependent variables show no statistical significance. This is summarized in *table 18*.

Table 18: Supported hypothesis for the GLM repeated measures ANOVA results

Dependent Variable	**Main Effect: Guidance Method**	**Main Effect: Cognitive Load**	**Interaction Effect**
AE Task 1	Alternative Hypothesis (H1)	Null Hypothesis (H0)	Null Hypothesis (H0)
AE Task 2	Alternative Hypothesis (H1)	Null Hypothesis (H0)	Null Hypothesis (H0)
Workload Task 1	Null Hypothesis (H0)	Alternative Hypothesis (H1)	Null Hypothesis (H0)
Workload Task 2	Null Hypothesis (H0)	Alternative Hypothesis (H1)	Alternative Hypothesis (H1)
TRTP Task 1	Alternative Hypothesis (H1)	Null Hypothesis (H0)	Null Hypothesis (H0)

8. Conclusion

This thesis explores how different guidance affects task performance and perceived workload in simulated aviation tasks under varying levels of cognitive demand. To be precise, the guidance modalities could be haptic, visual, and combined stimuli. With background research of the topic, the study analyzes the relationship between tactile stimuli and reaction times, cognitive load in aviation, and the use of multisensory signs for task enhancement. The investigation was focused on the context of pilot workload and multitasking, areas that are critical to aviation safety and performance. Workload directly affects pilot's decision making, reaction times, and overall flight safety. Excessive cognitive load can result in errors, delayed responses, and reduced situational awareness, which represent significant risks in aviation operations. Understanding the limits of human cognitive capabilities allows for the design of better training programs and workload distribution strategies that prevent overload conditions.

A flight simulator was developed using MATLAB and LabVIEW, with carefully designed interface elements that resemble real cockpit scenarios. This allowed for controlled testing conditions. The experimental methodology involved a sample of participants who underwent sessions designed to simulate single task and multitasking conditions, while their performance was measured across three key dependent variables: average error (AE), time to reach the target position (TRTP), and subjective workload. Workload evaluation was implemented using a structured mental workload evaluation configuration based on aviation tools.

Statistical analysis using ANOVA revealed clear effects of guidance modality and cognitive load. Visual guidance consistently led to better outcomes showing lower AE, faster TRTP, and reduced perceived workload. These results were obtained independently of the cognitive load. Combined guidance improved performance in some cases but was also associated with increased workload, likely due to the cognitive cost of integrating dual sensory inputs. Haptic guidance consistently resulted in the poorest performance, with higher error, slower task completion, and elevated workload ratings, particularly when multitasking was introduced.

This study contributes to the growing research on multisensory interface design and workload management in aviation. It highlights the critical role of visual input in ensuring task efficiency and cognitive load management. This is essential for maintaining safe flight operations. It also suggests that while multisensorial modalities may offer performance gains, they must be balanced with their cognitive demands. The findings also highlight the limitations of relying on haptic feedback alone in critical tasks or multitasking environments.

Some limitations were encountered during the study. Hardware constraints, such as the use of consumer grade technology, may have introduced signal delays and reduced the accuracy of given feedback. These delays affect both tactile and visual stimuli and are influenced by factors such as CPU speed, RAM, background processes, USB peripheral latency, and monitor refresh rates. Such variability affects the reliability of reaction time measurements in scientific settings. Future work should aim to refine haptic feedback systems using high precision hardware.

With respect to the flight simulator, implementing the tactile guidance device directly on a simulated flight should be the best alternative to properly introduce aircraft dynamics. Unfortunately, in current conditions it is not possible to implement aircraft dynamics into the testing due to hardware limitations. Nevertheless, to be able to analyze to which degree the prototype assists the pilot in bringing the aircraft back to a stable position with forces and moments involved could have relevant results for further development. It should be considered for future aerospace engineering research if the resources are available for it.

Additionally, while the main objective of the experiment was to compare the efficiency of different guidance methods in navigation tasks under varying cognitive loads, workload still played a crucial role in interpreting performance outcomes. Although continuous workload monitoring was not implemented, future studies could benefit from integrating physiological methods such as heart rate variability (HRV) and functional near-infrared spectroscopy (fNIRS) to gain more accurate insights. These techniques offer reliable, noninvasive ways to measure mental workload but were excluded in this study due to budget and equipment limitations. Incorporating them in future research could enhance the understanding of cognitive demands, especially in multitasking scenarios.

9. References

[1] Wickens, Christopher. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science. 3. 159-177. 10.1080/14639220210123806.

[2] Zikmund, Pavel & Horpatzká, Michaela & Macik, Miroslav. (2024). Learning Effect in Joystick Tactile Guidance. IEEE transactions on haptics. PP. 10.1109/TOH.2024.3368663.

[3] P. Zikmund, M. Macik, L. Dubnicky, and M. Horpatzska, "Comparison of joystick guidance methods," in 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom). IEEE, 2019, pp. 265–270.

[4] Johansson, S.R. (1978). Tactile sensibility in the human hand: receptive field characteristics of mechanoreceptive units in the glabrous skin area. Journal of Physiology, 281, 101-123.

[5] Erp, J & Carter, Jim & Andrew, I. (2006). Iso's work on tactile and haptic interaction guidelines. Proceedings of Eurohaptics. 467-470.

[6] D. Voudouris and K. Fiehler, "Enhancement and suppression of tactile signals during reaching." Journal of Experimental Psychology: Human Perception and Performance, vol. 43, no. 6, p. 1238, 2017

[7] G. Juravle and D. Spence, "Juggling reveals a decisional component to tactile suppression, "Experimental Brain Research, vol. 213, no. 1, pp. 87-97, 2011.

[8] Chan, Alan & Ng, Annie. (2012). Finger Response Times to Visual, Auditory and Tactile Modality Stimuli. Lecture Notes in Engineering and Computer Science. 2196. 1449-1454.

[9] Holden, Jameson & Francisco, Eric & Tommerdahl, Anna & Lensch, Rachel & Kirsch, Bryan & Zai, Laila & Pearce, Alan & Favorov, Oleg & Dennis, Robert & Tommerdahl, Mark. (2020). Methodological Problems with Online Concussion Testing. Frontiers in Human Neuroscience. 10.3389/fnhum.2020.509091.

[10] Prewett, Matthew & Elliott, Linda & Walvoord, Ashley & Coovert, Michael. (2012). A Meta-Analysis of Vibrotactile and Visual Information Displays for Improving Task Performance. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on. 42. 123-132. 10.1109/TSMCC.2010.2103057.

[11] J. B. F. van Erp, Tactile Displays for Navigation and Orientation: Perception and Behavior. Leiden, The Netherlands: Mostert and Van Onderen, 2007.

[12] G. de Rooij, D. Van Baelen, C. Borst, M. M. van Paaseen, and M. Mulder, “Supplementing haptic feedback through the visual display of flight envelope boundaries,” in AIAA Scitech 2020 Forum, 2020, p. 0373
[13] W. Wang, J. Yang, Y. Yu, Q. Wu, J. Yu, S. Takahashi, Y. Ejima, and J. Wu, “Tactile angle discriminability improvement: roles of training time intervals and different types of training tasks,” Journal of neurophysiology, vol. 122, no. 5, pp. 1928-1927, 2019.

[14] L.R. Elliot J. van Erp, E. S. Redden, and M. Duistermaat, “Field-based validation of a tactile navigation device,” IEEE transactions on haptics, vol. 3, no. 2, pp. 78-87, 2010

[15] K. Fellah and M. Guiatni, “Tactile display design for flight envelope protection and situational awareness,” IEEE transactions on haptics, vol. 12, no. 1m pp. 87-98, 2018.

[16] Lim, Yixiang & Bassien-Capsa, Vincent & Ramasamy, Subramanian & Liu, Jing & Sabatini, Roberto. (2017). Commercial Airline Single-Pilot Operations: System Design and Pathways to Certification. IEEE Aerospace and Electronic Systems Magazine. 32. 4-21. 10.1109/MAES.2017.160175.

[17] Xue, Jinna & Wang, Changyuan. (2023). Research on Pilots’ Mental Workload Classification in Simulated Flight. International Journal of Advanced Network, Monitoring and Controls. 8. 75-82. 10.2478/ijanmc-2023-0048.

[18] Lee, Kyongsun, "Effects of Flight Factors on Pilot Performance, Workload, And Stress at Final Approach to Landing Phase of Flight" (2010). Electronic Theses and Dissertations. 1628. https://stars.library.ucf.edu/etd/1628

[19] Pontiggia, A., Gomez-Merino, D., Quiquempoix, M., Beauchamps, V., Boffet, A., Fabries, P., Chennaoui, M., & Sauvet, F. (2024). MATB for assessing different mental workload levels. Frontiers in Physiology, 15, 1408242. https://doi.org/10.3389/fphys.2024.1408242

[20] Li, W., Li, R., Xie, X., & Chang, Y. (2022). Evaluating mental workload during multitasking in simulated flight. Brain and Behavior, 12, e32489. https://doi.org/10.1002/brb3.2489

[21] NI Apps. (2010, July 30). Writing your first labview program. YouTube. https://youtu.be/ZHNlKyYzrPE?si=TuRGY8pjgpBDBPkk

[22] EMERSON. (2023, November 6). Create a labview indicator that displays data in 360 degrees. NI. https://knowledge.ni.com/KnowledgeArticleDetails?id=kA00Z000000kEu0SAE&l=cs-CZ

[23] Juan Camilo Gómez Cadavid (2024). LabVIEW-MATLAB Interface [Official] (https://www.mathworks.com/matlabcentral/fileexchange/61208-labview-matlab-interface-official
), MATLAB Central File Exchange. Retrieved October 12, 2024.

[24] LabVIEW Para Todos. (2022, June 19). JOYSTICK CON LABVIEW Y ARDUINO [Video]. YouTube. https://youtu.be/3s_H_MXaJyQ

[25] NI Apps. (2010, July 30). Using LabVIEW Case Structures [Video]. YouTube. https://www.youtube.com/watch?v=cgS3u0nupfs

[26] Aviation Business Article: Aviation Business Middle East. (2023, October 9). *GA ATS orders Do-228 flight simulator*. https://www.aviationbusinessme.com/news/23887-ga-ats-orders-do-228-flight-simulator

[27] NI Support Document: National Instruments. (2024, July 9). *Basic TCP/IP communication in LabVIEW.* NI. https://www.ni.com/en/support/documentation/supplemental/06/basic-tcp-ip-communication-in-labview.html

[28] NI Knowledge Base Article: National Instruments. (2025, February 10). H*ow Can I Monitor a Joystick, Keyboard, or Mouse in LabVIEW?* NI Knowledge Base. https://knowledge.ni.com/KnowledgeArticleDetails?id=kA03q000000YIEOCA4&l=cs-CZ

[29] NI Forums Post: National Instruments. (2010, January 12). *How do I get my numerical inputs to reset back to 0 once the program stops?* NI Community. https://forums.ni.com/t5/LabVIEW/How-do-I-get-my-numerical-inputs-to-reset-back-to-0-once-the/td-p/1050507

[30] FEDERAL AVIATION ADMINISTRATION. (2001). Instrument Flying Handbook. FAA-H-8083-15. https://www.sheppardair.com/download/faa-h-8083-15.pdf

[31] SunFounder. (n.d.). Servo Motor (SG90). Servo Motor (SG90) - SunFounder Ultimate Sensor Kit documentation. https://docs.sunfounder.com/projects/ultimate-sensor-kit/en/latest/components_basic/27-component_servo.html

[32] Richardson, J. T. E. (2018) ‘The use of latin-square designs in educational and psychological research’, Educational Research Review, 24, 84-97.

[33] Roscoe, A. H., & Ellis, G. A. (1990). A subjective rating scale for assessing pilot workload in flight: A decade of practical use. Royal Aerospace Establishment Technical Report 90019.

[34] Berlik, Marcin. (2024). The usefulness of subjective task load assessment methods for predicting pilot task load in general aviation organizations. Scientific Papers of Silesian University of Technology Organization and Management Series. 2024. 19-32. 10.29119/1641-3466.2024.200.2.

[35] DATAtab Team. (2025). ANOVA Tutorial: Analysis of Variance Made Simple. DATAtab e.U. https://datatab.net/tutorial/anova

[36] IBM Corporation. (2021). One-way ANOVA: Post hoc tests. IBM Documentation. Retrieved May 22, 2025, from https://www.ibm.com/docs/en/spss-statistics/beta?topic=anova-one-way-post-hoc-tests

[37] Bland, J. M., & Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. BMJ, 310(6973), 170. https://doi.org/10.1136/bmj.310.6973.170

10. Figures

11. Tables

12. Appendix

A1 - Flight simulator interface VisualSupport.vi – LabVIEW Project