FlyBrate: Evaluating Vibrotactile Cues for Simulated Flight

Luis Lutnyk, David Rudi, Emanuel Meier, Peter Kiefer, and Martin Raubal

Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland

ABSTRACT
Contemporary aircraft cockpits rely mostly on audiovisual information propagation which can overwhelm particularly novice pilots. The introduction of tactile feedback, as a less taxed modality, can improve the usability in this case. As part of a within-subject simulator study, 22 participants are asked to fly a visual-flight-rule scenario along a predefined route and identify objects in the outside world that serve as waypoints. Participants fly two similar scenarios with and without a tactile belt that indicates the route. Results show that with the belt, participants perform better in identifying objects, have higher usability and user experience ratings, and a lower perceived cognitive workload, while showing no improvement in spatial awareness. Moreover, 86% of the participants state that they prefer flying with the tactile belt. These results suggest that a tactile belt provides pilots with an unobtrusive mode of assistance for tasks that require orientation using cues from the outside world.

1. Introduction

Flying an airplane is considered one of the most cognitively challenging tasks (Wickens, 2008a). Even though experience and increasing automation can make some parts of a pilot’s job routine procedures, pilots are still required to constantly adapt and react to a changing environment and suddenly occurring events (Vidulich & Tsang, 2012). Signals and alarms may require a sudden and unexpected shift in priorities, while requiring pilots to uphold communication with other crew members as well as personnel on the ground (Helmreich & Foushee, 2010). On top of these additional tasks that pilots have to attend to, the aircraft must be navigated to the intended destination in a stable and safe state (Wickens, 2002). Considering this high workload, especially trainees and rookie pilots who lack experience in handling challenging and unfamiliar situations may benefit from further assistance.

The most common way of presenting information to users is via the visual sense (Burke et al., 2006). The cockpit of an aircraft is no exception to this (Liggett, 2009). For example, a multitude of displays and instruments are installed and deployed alongside digital and paper maps to guide the pilots. The auditory modality is also used extensively in aviation, for example by alarm sounds, automated call-outs from the radio altimeter or air traffic control (ATC) communication. This poses challenges for the pilot, in particular in high workload scenarios, as demonstrated by Dehais et al. (2014), who showed that pilots missed auditory alerts when a workload increase by a windshear event was introduced in a simulated landing.

When different information is presented using the same modality or sensory channel, the mental demand for perceiving and processing the information can exceed the capacity of a pilot’s mental resources, and lead to a decrease in performance (Basil, 2012; Wickens, 2008b). Conversely, providing information across different modalities can lower the interfering demand for the mental resources required for processing the information provided on the single sensory channel. Thus allowing pilots to absorb the information more efficiently at levels of residual capacity in mental resources (Wickens, 2008b).

Given this high usage of the auditory and visual modes in aviation, there seems to be untapped potential in providing further assistance and guidance using the tactile mode (McGuirl & Sarter, 2001). The efficiency of tactile equipment for such assistance has already been demonstrated in a variety of other use cases, such as car driving (Hogema et al., 2009; Van Erp & Van Veen, 2004), navigating through terrain on foot (Elliott et al., 2010), or as an assistive device for the blind (Velázquez, 2010).

This article investigates the potential of the tactile mode for guidance in a general aviation context. We contribute to the existing body of work by reporting on a study, in which a tactile belt was utilized to support pilots in a typical task they face when flying under visual-flight-rules. In the presented flight simulator study, participants used the tactile belt’s vibrations coming from different directions to find specific objects in the simulated environment while flying maneuvers according to verbal instructions. Eye tracking technology, which allows capturing and displaying a person’s
gaze (Duchowski, 2017), was used in the study to evaluate the participants’ performance and progression during the flights. Their individual responses to questionnaires and in an open interview provided insights into their perceived workload and user experience, as well as the tactile system’s usability.

2. Related work

This article studies the value of vibrotactile cues for pilots’ performance, perceived cognitive workload, user experience, and spatial awareness. Therefore, we first discuss related work on tactile displays in general and in the aviation context, before elaborating on its use to enhance spatial awareness.

2.1. Tactile displays

Technology intended to assist people can often overwhelm them instead (Burke et al., 2006). According to the multiple resource theory (Wickens, 2008b), this can in some cases be related to an overload of the senses (Basil, 2012) – a circumstance which can be avoided by better distributing the information load across different modalities (e.g., audio, visual, tactile) when one modality, such as the visual modality, is heavily used already. The modality that has been described as best suited for a multi-modal interface is the tactile modality as it is often less overloaded or has been neglected so far (Burke et al., 2006; Van Veen & Van Erp, 2000).

Tactile displays have been, amongst others, built into car seats to improve users’ localization accuracy and response times (de Vries et al., 2009), and to provide the visually impaired with important information about the direction, distance, and height of objects (Durá-Gil et al., 2017; van Erp et al., 2017). Moreover, Tsukada and Yasumura (2004) explored the encoding of distance and directional information through multiple vibration motors on a belt and found that users can intuitively recognize directions to their destinations. This has led to successful application of tactile belts for pedestrian wayfinding (Heuten et al., 2008), indicating that users could pay more attention to the environment when receiving tactile feedback from a belt (Pielot & Boll, 2010). It has also been shown that spatial vibrotactile cues can effectively direct users’ visual attention to a given direction which can increase safety in driving (C. Ho et al., 2005).

Rosa et al. (2020) found that vibrotactile warnings can be effective in conditions of increasing mental workload and in the residual capacity region for effectively attending a secondary task. Multi-modal interfaces combining, e.g., audible and tactile modalities for a single alarm have been found to increase perceived urgency, but with increasing workload, the effect disappears as demonstrated by Biondi et al. (2017) for a take-over request in a semi-autonomous driving scenario. Additionally, Murphy and Dalton (2016) describe how high visual load can lead to inattentive numbness, i.e., reduced sensitivity to tactile stimuli. A study by Gallace et al. (2006) also found an analogue of visual change blindness in the tactile domain, depending on the intervals between different vibration patterns, which implies that no long intervals should be introduced between vibrations when changing pattern or location should be detected.

The design of tactile displays with regard to the main parameters, frequency, intensity, locus, and duration, has been explored in numerous studies and is covered extensively by Jones and Sarter (2008). Hence, we will adhere primarily to the recommendations provided by this work in the scope of this study (see Section 3.2.1).

In the aviation context, tactile feedback is a modality that has not been utilized intensively compared to other modalities. Van Erp and Self (2008) provide a comprehensive overview of the different applications of tactile displays in air, as well as for sea and land environments, but focus on military use cases. The authors highlight the capabilities of the technology to improve users’ spatial orientation, and that they had a positive effect on performance, especially in circumstances of degraded vision or high workload. The reviewed applications in military aviation included guidance for flying maneuvers, holding altitude, controlling drift while hovering a helicopter and landing unmanned aerial vehicles from a remote ground control station. In the civil aviation domain tactile displays have been evaluated in different scenarios. Van Baelen et al. (2021) investigated the use of tactile feedback for flight envelope protection in a comparison study between force feedback on the control stick, vibrotactile alerts and a combination of both. They conclude that for communicating when and in what direction the flight envelope protection is active, force feedback is valuable during operation and vibrotactile feedback can be added during training. C.-Y. Ho et al. (2001) studied how air-traffic-controllers handled interruption tasks coming from different modalities and found that the presentation of concurrent tasks via different sensory channels led to improved time-sharing performance, possibly due to reduced resource competition. Prinet et al. (2016) assessed the effectiveness of tactile spatial alerts for notifying pilots of an aircraft intruding on their flight path. They found that 100% of tactile alerts were detected and that tactile displays are effective for conveying spatial information which can help to prevent collisions. In a comparison study, pilots were warned about location and severity of icing on their airplane through visual cues and tactile cues (McGuirl & Sarter, 2001). It was found that performance was equally high for both types of cues and that with tactile cues the pilots were better able to recall and compare the previous and current icing condition. Ziat et al. (2016) found that tactile feedback improves the landing performance of pilots in a highly simplified experimental setup with a stylus pen controlling an abstract aircraft representation on a display. Moreover, tactile feedback is also already used in aircraft in the form of so called stick shakers that vibrate when pilots are about to stall (Trant, 1955).

However, little work has been done to assess tactile displays as part of the orientation procedure in general aviation that is mainly based on the visual orientation using objects in the environment (flying under visual flight rules (Eurocontrol, 2009)). This article intends to address this.
2.2. Eye tracking

Tactile signals may also be combined with eye tracking technology for validation of directional tactile cues or for gaze based interactions. In general, eye tracking can provide an unobtrusive way to observe, analyze and utilize a person’s visual attention (Duchowski, 2017). Rantalä et al. (2020) studied the temporal and spatial mechanisms between the combination of the gaze and tactile modalities, and found that tactile feedback performed equally well as both, visual and auditory feedback, to provide users with signals in an unobtrusive way. Moreover, Gkonos et al. (2017) combined a tactile belt with an eye tracker to provide pedestrians with navigation instructions. In the aviation context, eye tracking has been used for a variety of purposes (Di Stasi & Díaz-Piedra, 2019; Peißl et al., 2018; Ziv, 2016), such as assistance for instructors (Rudi et al., 2020). However, there is little work that combined eye tracking with tactile displays. Our research utilizes eye tracking primarily to analyze participants’ attention distribution and to enable the experimenter to trigger the next vibrotactile cues at the exact moment when the participants identified the current target object.

2.3. Spatial awareness and sketch maps

Spatial awareness is defined as a concept in conjunction with situation awareness. In the aviation context, situation awareness is defined as the pilots’ need to perceive their aircraft and environment, understand the significance of the perceived information, and anticipate the effect that environmental elements may have on the aircraft (Endsley et al., 1998). These three aspects are commonly referred to as different levels of situation awareness, with perception being level 1, understanding being level 2, and anticipation level 3. Moreover, situation awareness can be divided into correlated subcategories (Endsley, 1995), out of which Wickens (2002) defines automation awareness, task awareness, and spatial awareness to be the most important in the aviation context.

Spatial awareness in aviation may be defined as the perception, understanding, and projection of the position and extent of one’s own aircraft as well as the position and extent of objects in the aircraft’s environment (Bolton & Bass, 2009). The importance of pilots’ spatial awareness has been previously discussed, but there still exists only little research on designing an appropriate measure (Rudi, 2019). Bolton and Bass (2009) were among the first to explore appropriate measures and suggested to exploit existing measures of situation awareness for this purpose.

Hence, one more recent approach was presented by Stähli et al. (2018), who suggested to utilize measures from the Geographic Information Science that have a longstanding history of dealing with topics related to spatial cognition (Montello & Raubal, 2013). More precisely, the authors suggested to utilize sketch maps as a measure for spatial awareness – an approach that we will also follow in this study.

Sketch mapping is a method that is often used for studies in which participants have perceived or traversed a spatial scene and allows for an externalization of the participant’s spatial knowledge of that scene. Participants are provided with an empty drawing canvas, either on paper or on a computer, and given an instruction such as “please sketch the street network and any landmarks you remember.” Thus, sketch maps allow for the assessment of cognitive maps (Blades, 1990; Portugali, 1996) in real and virtual environments (Billinghurst & Weghorst, 1995). Cognitive maps are individuals’ mental representations of the properties and locations of objects in spatial environments (Tolman, 1948), and are related to an individual’s sense of direction (Kozlowski & Bryant, 1977).

For interpretation of the drawings, a variety of different sketch map analysis tools have been proposed throughout the years, such as the spatial-query-by-sketch system (Egenhofer, 1997) or CogSketch (Forbus et al., 2008). More recent solutions are the Gardony Map Drawing Analyzer (GMDA) that can be utilized for quantitative sketch map analysis (Gardony et al., 2016) and the more qualitative SketchMapia introduced by Schwering et al. (2014). In this article we chose to utilize the more recent and quantitative GMDA tool.

3. Methodology

3.1. Study design

The goal of this research is to determine whether a tactile belt allows general aviation pilots to improve their target identification performance and spatial awareness, while lowering their perceived cognitive workload and enhancing their user experience.

For this purpose, we present a simulator study, as part of which the participants had to fly a series of maneuvers and try to identify building structures on the ground. This task emulates that pilots are flying under visual flight rules (VFR), which require good visibility (Eurocontrol, 2009) and generally rely on a visual orientation using objects in the environment which can easily be identified. The study procedure was inspired by the study conducted by Tannen et al. (2004) who tasked fighter pilots with identifying targets from their cockpit while using different configurations of head-coupled visual and spatial audio displays.

The experiment was set up as a counterbalanced within-subject design. Every participant completed one flight for each of the following conditions:

- Condition NoBelt: Finding the objects without any additional assistance. (Baseline)
- Condition Belt: Finding the objects with the assistance from the tactile belt, which indicates the relative direction to the next object.

To avoid a learning effect, pilots flew two possible flight scenarios (A and B) in randomized order. As shown in Figure 1, the flown routes in the two scenarios were a mirror-image of each other and each scenario used 5 comparable but different objects. This prevented pilots to learn the sequence of the required banking maneuvers and the appearance of the objects. Participants either flew Belt-first
or NoBelt-first with one of two flight scenarios (A or B) assigned. Table 1 shows an overview of the possible configurations for the 4 groups participants were assigned to.

Based on previous research, indicating an advantage of tactile feedback for spatial allocation, the following hypotheses are proposed and examined in the remainder of the article:

H1 – Performance Participants identify more objects correctly and require significantly less time to do so in the Belt condition compared to the NoBelt condition.

H2 – Spatial Awareness Participants are able to draw significantly more accurate sketch maps of their flight route in the Belt condition compared to the NoBelt condition.

H3 – Cognitive Workload Participants perceive a significantly lower cognitive workload in the Belt condition compared to the NoBelt condition.

H4 – Usability Participants consider the Belt condition to have a significantly better usability than the NoBelt condition.

H5 – User Experience Participants consider the Belt condition to have a significantly better user experience than the NoBelt condition.

The independent variable is the applied assistance with Belt or NoBelt. The dependent variables are performance (measured by number of correct identifications, time to identification, and missed objects), spatial awareness (measured by accuracy of sketch maps), perceived cognitive workload (measured by score in the standardized NASA-TLX questionnaire (Hart, 2006)), and user experience (measured by score in the standardized UEQ questionnaire (Schrepp et al., 2014)).

Further, the pilots’ gaze behaviour in the Belt and NoBelt conditions is recorded with an eye tracker to examine the influence of the tactile feedback on pilots’ visual search behaviour. Then, three common eye movement metrics for the analysis of visual search are calculated from the eye tracking data to explore quantitative differences between conditions: fixation duration, saccade amplitude and fixation saccade ratio (Holmqvist et al., 2011).

3.2. Study setup

As part of the study setup, participants wore a tactile belt and a mobile eye tracker, while controlling a flight simulator. To comply with COVID-19 restrictions, the experimenter sat in a separate room at a controller station. The study setup (participant and experimenter) is depicted in Figure 3.

3.2.1. Tactile belt

The tactile belt was built on the basis of a motorcycle kidney belt and enhanced by adding an Arduino Lilypad microcontroller1, which is designed for usage in wearables and e-textiles. Figure 2 (right) shows a picture of the belt. The microcontroller was sewn into the belt’s fabric and connected with conductive threads to eight Lilypad vibe boards that were attached equally spaced across the inner surface of the tactile belt, similar in design to Gkonos et al. (2017) and Tsukada and Yasumura (2004). The vibe boards use Precision Microdrives 310 vibration motors2 with a vibration frequency of around 200 Hz, which provides good sensitivity on the abdomen as shown in Jones and Sarter (2008). The authors also conclude that greatest localization accuracy is achieved with eight to twelve motors. Hence, it was chosen to use the maximum limit of eight motors allowed by the microcontroller.

The vibrations were controlled with a custom system that allowed to activate the eight vibe boards (i.e., initiate a vibration) using the experimenter PC’s numpad. These commands were sent to the microcontroller using the serial
connection. The microcontroller would then pulse the output that is connected to the corresponding vibe board (see Figure 2, left). The vibration duration was set to 500 ms for the on and off phases of the vibration pattern which is in the range of durations presented in Jones and Sarter (2008) and the pulse durations were kept constant throughout the study.

### 3.2.2. Eye tracker

The eye tracker used as part of the study setup was a second generation mobile eye tracker by SMI running at 120 Hz. The eye tracker was provided by the manufacturer and was used for calibration, as well as for recording and monitoring the participants’ eye movements at the experimenter station.

### 3.2.3. Simulator

The flight simulator was provided by Lufthansa Aviation Training and is primarily used as part of exhibitions. The simulator (see Figure 3) included a steering yoke in front of the participants, a driving seat, a metal chassis (for stability), three monitors mounted at eye level, as well as levers for throttle, propeller and mixture, an instrument board, pedals, and various buttons on the yoke itself. All these components were connected to a high-performance PC.

However, to make the study as accessible to laymen participants as possible, the majority of the flight simulator’s control inputs and instruments were disabled. Additionally, the autopilot was configured so that it would maintain a predefined altitude above sea level. Pulling out or pushing in the yoke, which usually leads to altering the altitude of the aircraft, would only lead to minimal changes in the aircraft’s pitch. The autopilot would quickly counteract this and maintain the predefined altitude. This means that crashing the aircraft was impossible under normal circumstances, barring extreme maneuvers such as repeated barrel rolls. Participants were made aware of this and discouraged from flying such maneuvers. For the off-chance of a crash occurring, the damage settings were also disabled.

The simulator software used for the experiment was Prepar3D v3. The aircraft model chosen for the flight...
simulator sessions was the Mooney Acclaim which includes a Garmin G1000 avionics suite that provides a glass cockpit layout (Sweet, 1995) as found on contemporary general aviation aircraft.

The monitor in front of the participant featured the bank angle indicator in a prominent position that made it easily visible for participants. This indicator was one of the components participants had to monitor in order to complete the maneuvers they were asked to fly (see Section 3.3). Unlike the experimenter, participants did not see the navigation map to avoid them simply following the ownership symbol and the route depicted there.

3.2.4. Experimenter station
On the instructor side the setup included a laptop that allowed surveying the eye tracker and controlling the tactile belt, as well as a monitor displaying some of the key flight parameters shown to the participants. In particular, the monitor displayed copies of the three views displayed to the participant, and an additional view showed the aircraft’s electronic navigation map which contains the location of the aircraft, the flight route, and the locations of the target objects. In conjunction with the ownership symbol displayed on the map, the experimenter could judge where the participant was in relation to their next target object and decide which vibe board to activate if the flight was conducted with the belt. If, e.g., the next target object was 90 degrees to the right of the ownership symbol, the experimenter would trigger the vibe board placed at 90 degrees to the right on the tactile belt. This “Wizard of Oz” approach could potentially be replaced with a piece of software that interfaces to the simulator’s API and activates the vibe boards automatically. However, the manual activation worked quite well and so the additional development work for an automatic activation was deemed unnecessary. Instead, for enabling consistent activation of the vibration cues, a fixed protocol for the activation was defined and the whole study was conducted by the same experimenter to minimize a potential link between participants’ performance and the activation by the experimenter.

3.3. Procedure
To evaluate our hypotheses and to acquire the necessary data, a flight simulator experiment was conducted. The participants were tasked with flying along a predefined route by following the verbal instructions given by the experimenter. While doing so, they were asked to identify specific objects in the simulated environment. During the flight simulator sessions, participants wore an eye tracker, allowing the experimenter to determine whether the pilots identified the correct target. After they had completed a flight, participants were asked to draw a sketch map of their flight route and fill in the post-study questionnaires (see Figure 4).

3.3.1. Participants
The study was conducted during the COVID-19 pandemic and the number of available participants was limited. A total of 22 participants (7 female, 15 male) could be recruited. While the majority of participants were laymen with no experience in handling a real aircraft, five participants stated that they had accrued at least 10 real-life flight hours, with the maximum being 190 h. 11 participants had at least some previous experience with flight simulation, including games. 14 participants stated that they had at least some experience with handling a yoke or a joystick.

Three participants had to wear their glasses during the experiment to be able to recognize objects on the simulator monitors, which is why the eye tracker feed could not recognize and display their gaze, as wearing glasses would not allow the device to correctly track the pupils. Instead, these users were asked to point at the objects upon identification.

3.3.2. Pre-study
A pre-study was conducted ahead of the main study involving two participants aged 33 and 36. The most important findings of the pre-study were that the belt was too small for one of the participants and that the difficulty of the target identification was too high. To improve on these limitations, an optional belt extension was created, and the objects were made more noticeable by placing them in more exposed locations and enlarging them. It has to be noted that the front motors were not shifted by the belt extension, only the two back motors and their location was still on the back of the extended belt. Even more, the shift of the back motors turned out to be irrelevant in the main study because no full turns were required for the guidance to the target objects as shown later. It was also suggested to stress the importance of the sketch map task before the flights as both pre-study participants forgot to also focus on remembering the route until they were asked to draw the maps. The pre-study also allowed the experimenter to train a consistent activation of the vibration cues as per the protocol.
3.3.3. Main study

The main study consisted of four phases: an introductory and tutorial phase, a first flight phase followed by a debriefing, a second flight phase followed by a debriefing and a concluding interview phase. Each of the two flight phases included a test flight, a main flight, and a post-flight questionnaire, plus drawing the sketch maps.

3.3.3.1. Introductory and tutorial phase. After arriving at the location of the experiment, participants were asked to wear a surgical face mask for the duration of the experiment to adhere to prevailing safety rules. Next, they read a description of the study and filled out a consent form and a demographics questionnaire. The demographics questionnaire also contained questions about their condition and their experience with real or simulated flying.

Following the introduction, participants were given a tutorial for the flight simulator by the experimenter. This included an overview of the general handling of the aircraft and some experiment specific advice, such as the advice to use the 30° bank angle indicator when flying turns to improve consistency across participants. Participants were also shown multiple ways to judge the changing heading angle during turns, such as the on-board compass, the track display or the usage of the environment for orientation. They were then given as much time as needed to get used to flying turns and to familiarize themselves with the flight simulator.

If their first flight was planned to be the Belt condition, participants would then be equipped with the tactile belt. A trial run of the vibrations was performed to familiarize the participants with the tactile feedback, in which the participants had to judge what direction the vibration was coming from.

Independently of whether they were wearing the belt for the flight or not, the participants were then equipped with the eye tracker, and a 1-point calibration procedure as per manufacturer instructions was completed.

3.3.3.2. First flight phase. The first flight phase with the selected condition included a test flight and a main flight. In the more simple test flight, the participants were tasked with flying two maneuvers and finding two water towers in the environment. Pictures of the water towers were shown to them before the test flight started. If they would later have their first main flight in Scenario A, which starts with a left turn, their test flight would start with a right turn and vice versa for Scenario B.

To initiate a maneuver, the experimenter would first call out the direction and relative heading angle of the maneuver, followed by the target that the participants had to look for. An example of such an instruction would be “Please turn 90 degrees to the left, and look for the red water tower.”

In case they wore the belt, the test flight also allowed the participants to familiarize themselves with the changing vibration signals. The more the participants steered towards their target, the further towards the front and center the vibration wandered. Once activated, the vibration was repeated continuously in a loop (500 ms vibration on, 500 ms off, etc.). Meanwhile, the experimenter could follow their performance on the eye tracker feed and further reassure the quality of the calibration.

Upon successful completion of the test flight, the participants were presented with the five objects they had to identify in their first main flight. They were also reminded of the upcoming sketch map task after the flight to make them less likely to forget about paying attention to their route and overall environment.

The first main flight was guided by the same kind of instructions and call-outs as the test flight. But, due to the increased duration of the flight, the instructions announced by the experimenter were repeated right after they were first issued, which made it easier for participants to remember both the maneuver they had to fly and the object they were tasked with finding.

To identify a target, participants were asked to focus their gaze on the appropriate object. The experimenter would see this by using the eye tracker video feed. Additionally, they had to either give a verbal sign of confirmation (“I can see the target.”) or point at the simulator monitor where the object was located. The start of either of those two actions served as the point in time where the identification attempt was officially initiated and the timestamp for it was created. Upon successful identification of a target, participants received a confirmation from the experimenter. The instructor then told them to either hold the current direction or to adjust the trajectory to correct for previous deviations from the planned route.

If the participant did not identify the target correctly, the experimenter would notify them accordingly and record the incorrect call out. In case they flew past an object or had no realistic way of finding it, participants received a “strike” for the current target object, though they were not informed of this. Instead, the experimenter gave them new instructions that were aimed at getting them back to flying towards the target object. This usually involved flying a circular pattern until the object was in the field of view again. In case a participant amassed three strikes on the same object, they were given instructions to find the next object.

Once the participant had identified the last object, the first main flight was finished, and the recording of the eye tracker and the flight simulator were stopped. The participants were asked to take off the eye tracker and the belt if they had worn it during the flight, as neither was required for the post-flight tasks.

In the subsequent debriefing, the participants were tasked with sketching the route they just had flown and the target objects they had identified on a blank sheet of paper. To make the location and order of the objects unambiguous, the participants were asked to mark them with a cross symbol and number them in chronological order of identification. There were no further restrictions on the style and format of the sketch maps. Examples of drawn sketch maps are depicted in Figure 5.
Upon completion of the second set of sketch maps, the participants were about the participants share their subjective opinions in a semi-structured interview. While the semi-open nature of the interview allowed for any behavior. Before leaving the room, participants were offered a reward in the form of some sweets of their choosing.

### 3.3.3.3. Second flight phase

Once they had completed the post-study questionnaires, participants were led back into the simulator room for the flights with the second condition, starting with the second test flight. If they had not used the belt for the previous flight, they were equipped with it at this point. If they already had used the belt, they conducted the second flight without it. The rest of the flight followed the same procedure as detailed for the first test flight.

After the test flight, the second main flight followed the same procedure as the first main flight, except for the altered condition, but with the same set of questionnaires and sketch maps.

### 3.3.3.4. Interview phase

Upon completion of the second set of post-flight questionnaires, the participants were invited to share their subjective opinions in a semi-structured interview. While the semi-open nature of the interview allowed for any form of feedback, specific questions were asked as well. These were about the participants’ preference regarding flying with or without the belt, their perceived advantages and disadvantages that the tactile belt had to offer, the quality of the targets to identify, and if there was a change in their search behavior. Before leaving the room, participants were offered a reward in the form of some sweets of their choosing.

## 4. Results

This section details the results gained from the data collected during the flight simulator experiment. It is subdivided into the different topics that were investigated, i.e., performance, spatial awareness, cognitive workload, and user experience.

Given the within-subject study design, the statistical analysis is conducted considering the results as paired groups. A 95% confidence interval was used for all cases.

### 4.1. Performance

To compare the performance of participants between the two conditions (Belt=B and No Belt=NB), four aspects were considered. First, the total number of incorrect identifications participants made were counted. Second, the number of strikes they accrued from flying past an object and third, the number of objects they missed altogether after getting a third strike on a single target were evaluated. Fourth, the overall time spent looking for targets was computed as the sum of all the five times from starting a leg of the route until the successful identification of the announced target (in seconds). Table 2 summarizes the results while separating between the first three aspects reported as absolute numbers and the last two reported as averaged numbers.

With regard to the number of incorrect identifications, the B condition had lower numbers for all three categories, participants did not get a single strike or missed a single target, and the number of incorrect identifications was reduced compared to the NB condition. A Shapiro–Wilk test showed that none of these data sets are normally distributed, which is why a Wilcoxon signed rank test was used to test for significance. The number of strikes participants received was significantly different ($p < 0.05$) between the conditions. There were no significant differences for incorrect identifications and misses. For the five participants that had real-life flight hours, no strikes or misses were recorded in either of the conditions and one incorrect identification in the NB condition was logged.

Participants spent significantly less time (in seconds) searching for objects in the B condition (mean = 273, SD = 84) compared to the NB condition (mean = 350, SD = 171). For 15 out of 22 participants, their average time searching was lower when using the belt. A Shapiro–Wilk test showed that the data was not normally distributed, hence, a Wilcoxon signed rank test was used. The test showed that the differences in average time searching are significant ($p < 0.05$). Note that 3 of the total 44 flights had incomplete log files but the required timestamps to determine the average time searching could be extracted from the audio and front camera video recordings of the eye tracker through manual analysis for these participants. One

### Table 2. Number of incorrect identifications, strikes and misses per condition, as well as the average time searching for objects per condition; with their Z and p values for a Wilcoxon signed rank test.

<table>
<thead>
<tr>
<th></th>
<th>NoBelt condition</th>
<th>Belt condition</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect identifications</td>
<td>7</td>
<td>2</td>
<td>-1.414</td>
<td>0.157</td>
</tr>
<tr>
<td>Strikes</td>
<td>10</td>
<td>0</td>
<td>-2.271</td>
<td>0.023</td>
</tr>
<tr>
<td>Misses</td>
<td>1</td>
<td>0</td>
<td>-1.000</td>
<td>0.317</td>
</tr>
<tr>
<td>Average time searching (s)</td>
<td>350</td>
<td>273</td>
<td>-2.289</td>
<td>0.022</td>
</tr>
</tbody>
</table>

Significant difference in **bold** ($p < 0.05$).
participant was not able to identify one of the objects and accrued three strikes. Since no identification timestamp could be registered, the time when the last strike was given acted as the cutoff point for that leg of their route.

### 4.2. Spatial awareness

The sketch maps drawn by the participants were evaluated as an indicator for determining whether the tactile belt helped to enhance their spatial awareness. One participant did not provide the sketch maps and stated that they did not remember enough to do so. The participant was excluded from the following analysis. With the two conditions, a total of 42 sketch maps were considered for analysis.

The scenery files edited in Prepar3D’s SimDirector were used to define the reference coordinates of the object locations for both scenarios. The sketch maps were scanned, digitized and then imported into the Gardony Map Drawing Analyzer (GMDA) tool (Gardony et al., 2016). The GMDA tool allowed a comparison of the relative and absolute directional differences between the drawn and actual object position to calculate different measures, indicating how closely a participant’s sketch map resembles the reference map. The outcome of these comparisons is presented in Table 3.

TheSQRT(CanOrg) measure results in a score between 0 and 1, where a higher score indicates a better accuracy of the drawn configuration. Sketch maps drawn after flights of the NB condition resulted in higher scores regarding SQRT(CanOrg) compared to the sketch maps from the B condition. A Shapiro–Wilks test determined that neither data set is normally distributed. A Wilcoxon rank sign test showed no significant differences.

The two metrical measures of distance and angle accuracy also result in a score between 0 and 1, with a higher score indicating a better representation of distance and angles. Similarly to the SQRT(CanOrg) measure, both NB condition’s distance and angle accuracy showed higher scores than B condition’s counterparts. Shapiro–Wilk tests showed the data is not normally distributed and a Wilcoxon signed rank test indicated no significant differences for either of the two measures.

A fourth measure that was considered is the bidimensional regression (BDR) (Friedman & Kohler, 2003), which provides us with the correlation coefficient r that reflects the resemblance between the configuration of points in both maps on a scale from 0 to 1. The sketch maps from the NB condition scored a higher r value than from the B condition. A Shapiro–Wilks test showed no normal distribution and the Wilcoxon signed rank test showed no significant differences.

### 4.3. Cognitive workload

The perceived cognitive workload was measured using the NASA-TLX questionnaire (Hart, 2006) for both conditions post-flight. A Shapiro–Wilk test showed a normal distribution for the overall NASA-TLX (i.e., the sum of all categories per participant), so a paired samples t-test was employed. The overall NASA-TLX was significantly (p < 0.05) higher for the NB condition (mean = 48.77, SD = 20.433) than for the B condition (mean = 37.23, SD = 15.611).

The six categories of the NASA-TLX were non-normally distributed as per a Shapiro–Wilk test and analyzed using a Wilcoxon signed rank test with a Benjamini-Hochberg correction for false discovery rate (Thissen et al., 2002) including the overall NASA-TLX. All of the six individual categories had lower mean values for the B condition, while only mental demand (p < 0.05) and performance (p < 0.05) were significantly lower. Physical demand (p = 0.086, Z = –1.866), temporal demand (p = 0.150, Z = –1.440), effort (p = 0.078, Z = –2.011) and frustration (p = 0.150, Z = –1.506) showed no significant differences. A summary of the analysis is given in Figure 6.

### 4.4. Usability

Participants judged the tactile belt’s usability using the System Usability Scale (SUS) (Brooke, 1996). A Shapiro–Wilk test showed that the SUS scores of the participants follow a normal distribution. A paired samples t-test showed that the SUS score of the B condition (mean = 84.77, SD = 9.093) was significantly higher (p < 0.005) than for the NB condition (mean = 73.18, SD = 14.438).

### 4.5. User experience

Participants also filled in the User Experience Questionnaire (UEQ) (Schrepp et al., 2014). Shapiro–Wilk tests showed that the data sets of the B condition for the categories perspicuity, efficiency, and novelty do not follow a normal distribution, while all other data sets do for both conditions. Since not all of the data sets are normally distributed, a Wilcoxon rank sign test with a Benjamini-Hochberg correction for false discovery rate was used to determine whether the differences are significant.

The scores of all user experience categories of the B condition were significantly higher than their counterparts of the NB condition, i.e., attractiveness (p < 0.05), perspicuity (p < 0.01), efficiency (p < 0.005), dependability (p < 0.05), stimulation (p < 0.01) and novelty (p < 0.001). A summary of the analysis is given in Figure 7.

### 4.6. Interview

A semi-structured interview was conducted at the end of the study. Participants were asked a set of questions involving their preference regarding flying with or without the belt, their perceived advantages and disadvantages the tactile belt
had to offer, the quality of the targets to identify, and if there was a change in their search behavior.

### 4.6.1. Preference

Nineteen of the 22 participants preferred performing the task while wearing the tactile belt. One participant preferred to not wear the belt, while two had no clear preference.

### 4.6.2. Advantages

The most commonly named advantage when using the tactile belt was its assistance in determining when a flight maneuver was completed and thus when the aircraft should be leveled out again. This was mentioned by 14 participants, who also noted that without the belt, they had a harder time judging the required heading change. 11 participants responded that they used the belt’s vibrations to infer where to look for the next target object. Moreover, eight participants perceived the strength of the vibration signals to be just right, while five participants found them too weak and four too strong.

### 4.6.3. Disadvantages

Five participants noted that the resolution of the belt was too low. With only eight available vibration motors, judging the exact direction to the next object was found to be difficult. For two participants, the frequency of the vibrations was too high. Four participants noted that while they found the vibration signals to be fine, they could see them getting annoying or distracting if they were used “as is” for longer flights.

### 4.6.4. Target identification

When asked about the two sets of target objects they had identified during the flights (see Figure 1, right), 10 participants found these groups to be balanced and did not perceive one to be more difficult than the other. For four participants, the objects in scenario A were more difficult to identify, while three participants found the objects in scenario B to be more difficult to spot. The remaining participants provided statements concerning the targets also found the two groups to be balanced in terms of difficulty, but named single objects they found to be more difficult.

---

**Figure 6.** Results of the “raw” NASA-TLX questionnaires for the six categories and overall value for both conditions. Whiskers indicate standard error. Statistical significance marked by: *$p < 0.05$.

**Figure 7.** UEQ benchmark results for the six categories and both conditions. Whiskers indicate confidence intervals. Statistical significance marked by: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
than the others: Three objects were mentioned four times, those being the antenna, the windmill and the ruins. Moreover, two of the participants answered that they found the difficulty of the tasks too easy and would have enjoyed a higher difficulty.

4.6.5. Search behaviour
Fifteen participants noted that when using the tactile belt, their search behavior was drastically different compared to the flight without the belt. In the latter case, they used a wider scanning search pattern, checking all three available monitors with regularity. When assisted by the belt, they mainly focused on the central monitor once the vibration motor in the front was active. This effect did also show up in the visual analysis of the eye tracker recordings, as will be explained in the following section.

4.7. Gaze behaviour
The changes in visual search behaviour reported in the interview (see previous section) were also supported by an exploratory analysis of the eye tracking data.

A first visual inspection of the participants’ scanpaths between conditions revealed a more focused scan on the middle display, i.e., in the direction of flight, in the B condition for 14 participants. The scanned area was smaller and less visual attention was directed at the left and right display. Two exemplary scanpaths of one participant demonstrating the described effect of a changed scanning strategy are shown in Figure 8. Four participants’ scanpaths did not show the more focused scan in the B condition and the recordings from four participants did not contain enough gaze data for an analysis.

Further exploratory analysis was performed using common eye tracking metrics. The quantitative analysis was based on the fixation and saccade export files of SMI’s BeGaze software. Three eye tracking metrics which are commonly used for the analysis of visual search were computed from the data based on the definitions by Holmqvist et al. (2011).

The metrics, fixation duration, saccade amplitude and fixation saccade ratio, were computed per participant, per condition and a joint dataset was created for each metric. Boxplots of the resulting datasets are depicted in Figure 9. Two far outliers were removed from the saccade amplitude data.

The mean fixation duration was slightly lower in the NB condition (mean = 333.4, SD = 92.09) than in the B condition (mean = 357.3, SD = 82.71), and the mean saccade amplitude was slightly higher in the NB condition (mean = 6.57, SD = 1.35) than in the B condition (mean = 6.11, SD = 0.87). The fixation saccade ratio was slightly lower in the NB condition (mean = 7.33, SD = 3.35) than in the B condition (mean = 8.45, SD = 3.42). The boxplots also reveal a smaller interquartile range for the B condition compared to the NB condition, particularly for saccade amplitude.

The eye tracking metrics datasets were non-normally distributed as per a Shapiro–Wilk test and a Wilcoxon signed rank test was performed on the three datasets. There was no statistically significant difference in mean fixation duration ($p = 0.349$, $Z = 0.936$), mean saccade amplitude ($p = 0.586$, $Z = -0.936$) or fixation saccade ratio ($p = 0.199$, $Z = -1.285$).

5. Discussion
Based on these results, the initially proposed five hypotheses for evaluating the tactile belt as additional assistance in simulated flight are discussed.

5.1. H1 – Performance
The performance indicators of accrued strikes, misses and incorrect identifications show a clear advantage for the B condition: The tactile belt increases the confidence of the participants in their decision making, as it directly indicates the direction to the next target. This observation was reflected by the majority of responses in the interview, where the above functionality was the most named

---

**Figure 8.** Exemplary scanpaths of one participant. A wider area was scanned when not assisted by the belt (left), whereas the scanning was more focused in the center of the middle display without the belt (right).
A significant difference was the lack of strikes issued for the B condition compared to the ten strikes of the NB condition. Note, that no strikes and misses for the B condition also means that the missing motors in the back when using a belt extension was no issue. These results indicate that the belt benefits the overall safety of the flight: Failing to identify landmarks in a visual approach or similar real-life scenario may lead to problems, such as missed approaches, loss of orientation or complications with other air traffic. While there were still two incorrect identification attempts in the B condition, this was another improvement over the seven incorrect identifications in the NB condition.

The differences in the times until objects were successfully identified were found to be significantly lower for the B condition. This result should be evaluated carefully though, as the mean total identification time for the NB condition was influenced by its overall worse performance in terms of strikes and misses. The strikes and misses are evidently highly correlated with longer total time for identification, as the duration of the turnaround maneuver after flying past the target is added to the total time for identification. When eliminating all flights with disproportional delays due to strikes and misses, there was no significant difference in terms of the time until identification between the two conditions. Nonetheless, this does not diminish the expressiveness of the average time searching measure as flying past a target would result in a longer search time in the real world as well.

An explanation for the overall longer time searching for the NB condition could be seen in potential risk aversion from the participants. Some of them stated that even though they were fairly confident that they had identified an object, they wanted to make sure they were correct and waited with officially calling it out until they were closer to it. This however also implies that participants were more confident when wearing the belt.

The B condition performed significantly better with regard to the number of strikes and the average time searching, though the difference in incorrect identifications and misses was not found to be significant. Hypothesis H1 can thus be partially accepted.

In line with previous applications of tactile feedback in the aviation domain (Van Baelen et al., 2021; Van Erp & Self, 2008) and in other domains such as wayfinding (Heuten et al., 2008; Tsukada & Yasumura, 2004), the directional tactile cues were effective in conveying spatial information to the pilot. This confirms that at levels of residual capacity in mental resources (Wickens, 2008b), performance in a target search and navigation task can be improved by using tactile cues for communicating direction.

5.2. H2 – Spatial awareness

The spatial awareness aspect of the study was examined using the sketch maps that were drawn by the participants after each flight. The locations of the drawn objects were compared with a reference map from the Prepar3D flight.
simulator. Results from the GMDA analysis revealed no significant differences between the maps for the two conditions.

Many participants noted how they had a harder time remembering the route during the first flight, sometimes even forgetting about the sketch map task completely until they were asked to draw the map. Being used to the procedure, they found drawing the second sketch map easier as they made sure to remember their routes more carefully. Based on this feedback, the GMDA results were analyzed with regard to differences between the first and second sketch map of each participant. However, these results did not bring up any significant differences either.

Another explanation for the results between the two conditions not showing significant differences was explained by some participants as follows: While the belt allowed them to focus on the object finding task, it also made them not pay as much attention to their surroundings. Since in the B condition the belt guided them well through the entire flight they did not observe the environment as consciously as they had to for the NB condition, searching for the objects in a more active way. This possible over-reliance on a navigation or guidance system can impact spatial awareness if no specific training is given to uphold sufficient scanning.

Since no significant differences regarding the sketch maps were observed between the study conditions, hypothesis H2 has to be rejected.

The fact that some participants directed less attention to their surroundings when following the directional cues from the tactile belt can be compared to a degradation in spatial knowledge acquisition due to an overreliance on navigation systems (Parush et al., 2007). It also relates to the issue of overtrust in automation. Wickens (1995) has to be rejected.

Our findings seem to contradict the capability of tactile feedback to improve spatial orientation as highlighted by by Van Erp and Self (2008). However, the participants could also have prioritized the navigation task and finding the objects over the spatial knowledge acquisition and sketch map task in both conditions, especially in their first flight.

5.3. H3 – Cognitive workload

The results of the NASA-TLX indicate that participants experienced a significantly lower overall cognitive workload in the B condition. More specifically, with regard to the questionnaire’s six subcategories participants experienced significantly lower mental demand while at the same time perceiving their performances to be significantly better.

The lower perceived mental demand, as well as the higher perceived performance could be tied to the overall difficulty of judging the heading angle change. Participants noted that they were not always sure if they had completed the maneuver they were tasked with the help of the outside view and compass instrument. The tactile feedback made this very easy, i.e., as soon as the belt would vibrate in the front the pilots knew that they were flying straight towards the target. This notion was further supported by participants’ remarks on how the belt gave them an additional sense of safety when identifying targets.

With regard to the physical demand, frustration, and temporal demand the B condition received a lower score for all individual aspects but none was significantly different. The outcome concerning the physical demand is not surprising as there is no noticeable difference between the two conditions except that participants wore a lightweight belt. The results regarding frustration indicate that the belt did not add any frustration but also could not decrease it. The finding that the perceived temporal demand was not significantly lower could be explained by the fact that the belt’s assistance did alleviate some of the temporal demand, but its effect was dampened by the constant vibration cycles which some participants noted to be worth turning off after the front motor is vibrating (i.e., they are on course) until the course has to be adjusted again.

Both the overall cognitive workload and three of its six subcategories were shown to be significantly lower and thus better for the B condition. The third hypothesis H3 can thus be partially accepted.

The reduced cognitive workload observed in our study relates to the findings of previous studies on vibrotactile feedback in the aviation domain that studied the use of tactile feedback at different levels of cognitive workload (C. Ho et al., 2005; Rosa et al., 2020; Van Baelen et al., 2021; Van Erp & Self, 2008). Our results extend these findings in that perceived cognitive workload can decrease when vibrotactile cues provide assistance to a navigation and target search task.

5.4. H4 – Usability

The usability of both conditions achieved results above the threshold of 68, which is regarded as above the average for an overall system score (Lewis & Sauro, 2018). More specifically, the mean score of the B condition gets an A+ on a curved grading scale while the mean score of the NB condition gets a B- (Lewis & Sauro, 2018). The grades indicate that participants felt they were able to complete their tasks safely, effectively, and efficiently. The simplified study design and implementation along with the straightforward tasks may have been contributing factors to these good results.

The B condition was scored significantly higher than the NB condition, which was in accordance with the majority of the responses from the interview. The positive influence of the belt can be explained with the wearable providing valuable assistance that made it easier for participants to find the objects. In essence, they had to align their gaze with the direction signaled by the equipment. This can also be observed in the eye movement data of participants such as the one depicted in Figure 8. The belt also allowed pilots to focus on the object finding task by taking away some difficulty from the flying task as mentioned above.

Given the significantly higher SUS score for the B condition hypothesis H4 can be accepted.
5.5. H5 – User experience

All six categories of the UEQ were significantly better for the B condition compared to the NB condition. Users thus enjoyed their experience during the study significantly more when they had the tactile belt assisting them. This can again be directly seen as a confirmation of the generally stated preference of flying with rather than without the belt.

The higher novelty score is not surprising considering that none of the participants had experience flying with a tactile belt. Similarly, the higher perceived efficiency is easily explained by the fact that the belt indicated the correct direction and participants had to scan less. Moreover, since the efficiency category also indicates satisfaction with how fast the system reacts, it appears that the changes regarding which motors were vibrating were performed well by the experimenter.

The high perspicuity score can be explained with participants’ feedback that the belt’s signals and assistance were intuitive and easy to learn, with only one participant feeling they would need more training to use the belt to its full potential. The good stimulation rating indicates that participants felt motivated to use the tactile feedback and perceived it to be fun to use.

Attractiveness measures the overall impression of the product, so the higher score of the B condition can be seen in direct analogy to the majority of participants regarding it as preferable. Dependability is a measure for how much the user feels in control during the task and how predictable and secure the interactions are. The higher score for the B condition can be explained by participants trusting the belt’s signals quickly and fully relying on them in most cases. Participants felt that following the vibrations was a very safe and easy way to complete the task.

Concerning the significantly higher scores across all six UEQ categories for the B condition hypothesis H5 can be accepted.

Additionally, the overall positive rating of the tactile feedback by the participants also aligns with the positive findings from previous work about its application inside and outside the aviation domain (Burke et al., 2006; C. Ho et al., 2005; Van Baelen et al., 2021; Van Erp & Self, 2008; Van Veen & Van Erp, 2000). An effect of inattentional numbness as described by Murphy and Dalton (2016) has not been observed during the study which indicates that the task difficulty needs to be higher than the difficulty of the target search task in our study for the described effects to materialize.

5.6. Limitations

The following section lists the limitations of our study in terms of methodology and findings from the post-study analysis.

The vibration signals were not issued automatically based on the position of objects and the aircraft but through the “Wizard of Oz” methodology. Consistency by the experimenter was ensured through a manageable number of motors and thus relatively coarse granularity of the signals, the limited complexity of the routes, the assistance of the electronic navigation map and a strict adherence to the study protocol. Moreover, while not a specific question in the interview, most participants were surprised to find out that the vibrations were not handled automatically, which indicates that we successfully kept the experimenter effects small.

Regarding the user experience, the statistically significant difference between the conditions in the novelty score of the UEQ raises the question of a possible novelty effect for the B condition. However, the use of test flights before the main flights (see Figure 4) should reduce such influence. Also, it is clear that long term effects of using the tactile belt on user experience scores were not a focus of our study design.

An additional measure of performance that could have been employed is the deviation from an optimal flight path, e.g., a flight path made up of straight lines between the objects. However, limitations of the simulator would have made this quite complex given the possibility for the pilots to fly back towards the target object after missing it and receiving a strike.

While the study was designed to be accessible to laymen to be able to recruit enough participants (particularly considering that the study was executed during the COVID-19 pandemic), it also led to simplifications and generalizations of the tasks, controls, and the simulated environment. The results from the study should therefore transfer well to novice pilots, but might provide less validity for expert pilots. A noticeable number of participants remarked the low difficulty of identifying the visible objects. This was a problem introduced by Prepar3D’s rendering priorities, as additionally placed objects would always be rendered earlier than randomly generated objects. It seems like participants with either experience in flight simulators or other digital environments such as video games were able to notice the rendering differences. The target objects were also better noticeable due to aliasing effects, which was especially apparent for the thinner objects such as the antenna or the windmill. However, these effects were present in both conditions and therefore did not give either condition an inadvertent advantage.

Some of the participants had real-life or simulated flight experience which may have influenced the study outcome concerning the performance and cognitive workload differences. For example, one participant always inspected the called out direction before initiating the flight maneuver and explained that due to the extended piloting experience there was an automatism to first making sure that no incoming traffic or other obstacles were impeding their maneuvers. Similarly, participants more experienced in maneuvering aircraft may have been less stressed overall. Given our within-subject design these experience differences should not have biased our study results as it affected both conditions.

Finally, the lab study may be considered a limitation, as one participant with piloting experience noted that she was not sure whether the current implementation of the belt would work well in a general aviation aircraft. Her reasoning was that, due to the compact design of such an aircraft,
vibrations from the engine would make the vibrations of the belt less noticeable. Another participant with a piloting license did not agree with this thought though, as she expected the belt’s vibrations to work as intended in a smaller aircraft as well.

6. Conclusion

This article assessed the benefits of utilizing a tactile belt with eight vibration motors for conveying directional information to pilots. The motivation behind this approach is based on the overall heavy reliance on audiovisual information in the cockpit. Especially novice pilots may quickly become overwhelmed by the amount of information in need of processing. Thus, following the multi resource theory (Basil, 2012), additional supportive information is instead conveyed using the tactile modality. The tactile modality is assumed to not interfere with the cognitive resources required to process the audiovisual information and is not fully exploited as part of the human-computer interaction in aircraft today.

The effects of the tactile belt were assessed with regard to pilots’ object identification performance, user experience, spatial awareness, and cognitive workload as part of a counterbalanced within-subject simulator study. 22 participants were asked to fly along a predefined route and look for specific objects in the simulated environment. While doing so, the participants were wearing an eye tracker that allowed the experimenter to evaluate their progress in the study and adjust the belt’s vibrations following a wizard-of-oz principle.

The results showed that while wearing the belt, the participants had a significantly lower number of missed targets, required significantly less time to identify the objects, while reporting significantly lower perceived cognitive workload compared to not wearing a belt. The usability and user experience of the system with the tactile belt were also rated significantly higher compared to the system without the belt. The spatial awareness of participants measured by the sketch maps they drew after each flight (with and without the belt) did not show any significant differences between the two. However, the interviews carried out at the end of each study supported the overall positive findings, with 86% of participants preferring the flight with the belt. A commonly mentioned advantage of using this wearable were the help it offered in judging the direction to the next target object. This reduced the area participants had to monitor on the simulator screen, increasing their efficiency in the object identification task. However, the reduced monitored area in the outside view presumably had a contradictory effect and reduced their spatial awareness of the outside world.

The study results show that a tactile belt is a promising way of communication and relaying information in the cockpit. It significantly supports pilots in identifying objects in the environment in an unobtrusive way, which is essential when flying under VFR and executing a landing maneuver. This shows that vibrotactile feedback provides effective assistance for target search and orientation in simulated flight and that such complex tasks can in fact be significantly enhanced as postulated by amongst others Burke et al. (2006) and McGuirl and Sarter (2001). The use of the tactile modality to convey spatial information can also be viewed as an extension of the multiple resource theory as described by Wickens (2008b). The tactile interaction adds another level to the modalities dimension in the 4-D multiple resource model, as previously defined by Wickens.

7. Outlook

The outcome of this study shows that exploring tactile feedback as a way to enhance the human computer interface in cockpits, particularly in the context of VFR and for novice pilots is worth further research. Future research could include novice pilots of a flight school flying under VFR in a real-world study setup or a higher fidelity simulator. Moreover, the scenarios could be chosen to reflect more complex environments and increase the workload by introducing additional factors such as wind shears.

Further use cases for vibrotactile feedback in aviation can be explored. A collision alert similar to the ground-proximity warning system (ground collision) or the traffic collision avoidance system (mid-air collision) could be implemented. Other examples may include warnings for entering restricted or controlled airspace or a landing assistance which conveys the distance towards the landing strip or indicate the aircraft’s relative position through tactile feedback.

Additionally, other types of tactile displays can be assessed as part of future work, e.g., Jansen et al. (2008) already demonstrated an implementation of a tactile vest for helicopter landing applications instead of a belt. A similar tactile vest could provide higher resolution and more directional information to the user, e.g., about the vertical location of an object relative to the own aircraft. Additionally, distance and speed of the object could be encoded in the vibration pattern or by adapting frequency, duration, amplitude and locus of the vibration.

Other extensions of the presented study could introduce a different role to the eye tracking. In this article eye tracking was primarily used for analysis, but applications could be similar to what was introduced by Anagnostopoulos et al. (2017) who tracked the user’s gaze and provided guidance if users continuously dwelled over an area of interest. The system could also use vibration signals to direct the attention to an instrument or monitor.

Notes

1. https://docs.arduino.cc/retired/boards/lilypad-arduino-main-board
2. https://www.precisionmicrodrives.com/product/310-101-10mm-vibration-motor-3mm-type
5. https://www.prepare3d.com/SDKv2/LearningCenter/getting_started/mission_creation/simdirector_overview.html
Acknowledgements

We thank all our participants who took part in our study, despite the circumstances of the COVID-19 pandemic. Moreover, we thank Charalampos Gkonos for providing us with the tactile belt to adapt and use for our work. Finally, we thank Lufthansa Aviation Training Switzerland for providing us with the flight simulator.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Luis Lutnyk http://orcid.org/0000-0002-9458-8575
Peter Kiefer http://orcid.org/0000-0003-4457-0438
Martin Raubal http://orcid.org/0000-0001-5951-6835

References


About the Authors

Luis Lutnyk is a doctoral student at the Chair of Geoinformation Engineering of ETH Zurich. His research focuses on human-computer interaction, especially in aviation. He is currently exploring the use of eye tracking in the cockpit and how novel interactions can help pilots and improve pilot training.

David Rudi is currently employed as senior consultant. He previously conducted his post-doctoral research at the Chair of Geoinformation Engineering of ETH Zurich with focus on the aviation domain. He completed his doctoral studies with focus on enhancing spatial awareness of commercial aviation pilots.

Emanuel Meier is currently employed as a GIS specialist and product analyst, focusing on obstacles and terrain in an aviation context. He concluded his master’s programme in Geomatics at ETH Zurich, specializing in the GIS and navigation fields.

Peter Kiefer is a Senior Scientist, leading the geoGAZElab at the Chair of Geoinformation Engineering at ETH Zurich. His research is concerned with assisting humans in spatial decision situations by the use of eye tracking. Peter Kiefer’s research is interdisciplinary, bridging Geographic Information Science, Eye Tracking, and Human-Computer Interaction.

Martin Raubal is professor of geoinformation engineering at ETH Zurich. His research interests focus on spatial decision-making for sustainability, more specifically he concentrates on analyzing spatio-temporal aspects of human mobility, spatial cognitive engineering, and mobile eye-tracking to investigate visual attention while interacting with geoinformation and in spatial decision situations.