

Does folding improve the usability of interactive surfaces in future airliner cockpits? An evaluation under turbulent conditions and varying cognitive load

Apport de la physicalité et du changement de forme pour pallier les faiblesses de l'interaction tactile dans les cockpits d'avions

Alma Cantu
alma.cantu@ncl.ac.uk
ISAE-SUPAERO, University of
Toulouse
Toulouse, France

Jean-Luc Vinot
Catherine Letondal
Sylvain Pauchet
catherine.letondal@enac.fr
sylvain.pauchet@enac.fr
ENAC, University of Toulouse
Toulouse, France

Mickaël Causse
mickael.causse@isae-supero.fr
ISAE-SUPAERO, University of
Toulouse
Toulouse, France

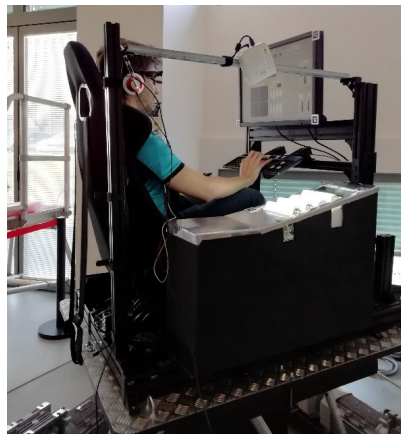


Figure 1: The motion platform mounted on a 6-axes simulator. It is equipped with a main screen and a keyboard (in front of the user), an area dedicated to the flexible surface (at the right of the user), and a joystick (at the left of the user, invisible on this picture). The display onto the flexible surface is performed by a projector.

ABSTRACT

In recent years, many aircraft manufacturers have proposed innovative cockpit concepts based on touchscreens. Although having a large number of advantages, this type of solution suffers from severe limitations in operational use, in particular, eyes-free interaction is nearly impossible and touchscreens are extremely complex to use during turbulent conditions. We examined the contribution of physicality to overcome these weaknesses by introducing a shape-changing touchscreen that offers folds on which the user hand can rest. This surface has been assessed in simulator during piloting conditions that varied in terms of turbulences and mental workload. Results showed that the folds helped reducing physical effort by stabilizing the arm and hand. This physicality was also associated with better performance in the piloting task, as well as with better situational awareness of the state of the aircraft's systems, most certainly because the shapes offered by the folds had better visual properties (saliency), making their monitoring less expensive in terms of attentional resources.

CCS CONCEPTS

• **Human-centered computing** → **Usability testing**; **Haptic devices**.

KEYWORDS

Shape-changing interfaces, Touchscreens, Safety, Human factors, Situation awareness, Turbulence, Aeronautics

RÉSUMÉ

Ces dernières années, de nombreux constructeurs aériens ont proposé des concepts de cockpit basés sur des écrans tactiles. Bien que très avantageux, ce type de solution souffre de sévères limitations en utilisation opérationnelle, notamment en raison d'une utilisation sans les yeux presque impossible et parce que les écrans tactiles sont difficilement utilisables durant des conditions turbulentes. Nous avons cherché à examiner l'apport de la physicalité pour pallier à ces faiblesses en introduisant une surface tactile pouvant changer

de forme et offrant des plis sur lesquels les mains peuvent se reposer. Cette surface a été évaluée en simulateur au cours de conditions de pilotages variables en termes de turbulences et de charge mentale. Les résultats ont montré que le fait de disposer de plis permettait une réduction de l'effort physique en stabilisant le bras et la main. Cette physicalité était également associée à de meilleures performances de pilotage, ainsi qu'à une meilleure conscience de l'état des systèmes de l'avion, certainement parce que les formes offertes par les plis présentaient de meilleures propriétés visuelles.

MOTS-CLÉS

Interfaces à changement de forme, Interaction tactile, Sécurité aérienne, Facteurs humain, Conscience de la situation, Turbulence, Aéronautique

1 INTRODUCTION

In airliner cockpits, the pilot crew activity is divided into four hierarchical tasks: piloting (controlling the aircraft's attitude), navigating (following a flight plan), communicating (maintaining communication with the ground) and managing systems (supervising the operation of aircraft systems) [30]. To conduct these complex activities, pilots interact with aircraft systems via specialized interfaces, grouped into functional units and displayed on different screens dedicated to each of the crew's main activities. They operate these systems with physical controllers: buttons, switches, pulls, sidestick etc. [33]. During the flight, pilots are responsible for monitoring, extracting, analyzing, and processing a lot of information. They carry out a multitude of operations and coordinate all activities within the crew, but also with external entities working on the flight [24]. To ensure their task management in the cockpit [14], pilots must divide their attention to numerous channels of information and develop skills resulting from constantly repeating procedures, postures and gestures. Many external factors constantly interrupt pilots and exacerbate the complexity of performing these multiple tasks, making pilot crews particularly vulnerable to errors [24].

Recently, many aircraft manufacturers and suppliers have proposed innovative cockpit concepts based on the use of touchscreens [2]. The challenge for industry is to respond to the increasing complexity of systems with greater flexibility and lower costs. The touch cockpit concept allows manufacturers to offer high-performance, adaptive, and generic product lines to address civil or military avionics needs. It also brings important benefits for pilots, such as direct information manipulation [3, 5], visualization of aggregate data on large display surfaces [15], and improvement of cognitive resource management [19]. To date, only a few manufacturers offer operational touch solutions, such as Rockwell Collins' Pro Line Fusion product line or avionics systems for Honeywell's Gulfstream G500 and G600 models. This rather slow deployment of touchscreens in modern cockpits is partly due to their severe limitations in operational settings. Contrary to current physical interactors, whose perception and manipulation are also promoted via the sense of touch and proprioception, they place a high demand on the visual channel to adjust the gesture, thus eyes-free interaction is difficult [34]. They are also extremely complex to use during turbulent conditions [10], the pilot's ability to interact with them may be substantially impaired or eliminated during periods of heavy cockpit

vibration. In addition, compared to full use of physical space, touchscreen surface interaction does not promote situational awareness and crew collaboration [8, 11]. Finally, perception of the information can be difficult because the screen can get dirty or because of the presence of smoke in the cockpit [33], and their usability is markedly reduced by stress or cognitive overload [7]. Thus, the generalization of touchscreens in a safety-critical domain like aeronautics would be possible only if their usage, especially in degraded context, is at least as efficient and reliable than classical interactors.

2 RELATED WORK

As opposed to approaches aiming at enhancing the design of touched-based interactions to improve robustness and visual dimensions in the cockpit [20], a possible avenue can be to combine the advantages of touchscreen interfaces and physical controllers [12]. Introducing a graspable area in touchscreens would allow stabilizing the interacting arm and a reduction of pointing errors during turbulences. Cockburn et al. established that the possibility to stabilize the interacting arm improved the user performance with touchscreens during turbulent conditions [10]. In this respect, graspable elements within the surface are likely to reduce physical effort during the interaction. In addition to an improvement of the physical interaction, relying on the physicality of shape and movement could also improve the perceptual dimension of the touchscreens. This physicality of "tangible" interactors, based on changing the shape of a flat surface, is indeed a promising property for efficiently conveying information. Studies from the physical information field support the idea that physicality improves users' efficiency to perceive information [21] without complex symbolic representations [16, 18] and provides a more creative and mindful way to look at the data [31, 32]. Physical objects convey haptic perception, have a high visual salience, and their borders and edges structure the information space [28]. In this sense, the physical discontinuity of tangible interactors, originating from the emergence of an object on a flat surface, might also increase information salience [6] and the building and maintenance of the situational awareness about the system states in the cockpit (e.g., the landing gear is fully extended). Pauchet et al. [27] showed that digital non flat interactors could reduce visual and mental demands during a piloting task, allowing to focus more on other cockpit activities. In addition, when set in motion, interactors based on shape change might prove to be very efficient to drive the attention, enhancing event detection and further contributing to situational awareness. Physiologically, the perception of movement is different from the perception of immobile object due to the different properties between the magnocellular visual pathway and the parvocellular visual pathway. The former detects quick changes in the position of an object and have a higher temporal resolution than the latter [23]. As such, a dynamical aspect (movement) of tangible interactors may allow stimulating magnocellular cells, allowing a faster information perception of some critical information. Existing example of cockpits systems that use their physicality to enhance perception and awareness are the auto-trim and the throttle levers. Those two systems, respectively aiming at controlling pitch and speed, can be moved intentionally by the pilots or can move automatically, for example according to the autopilot actions. Their physicality and

movements improve the situational awareness of the crew about the evolution of their states.

3 OBJECTIVES AND HYPOTHESES

This paper presents the human factor evaluation of a flexible surface designed to address the above-mentioned limitations of touchscreens in airliner cockpits. Indeed, we aimed to assess whether a shape changing touch-based surface could better cope with turbulences or high mental workload than traditional touchscreens, by providing graspable interactors and an enhanced perception of the system states.

To compare the performance of flattened mode vs folding mode of a surface, sixteen participants were placed in a moving platform where they experienced a flying scenario with two different levels of difficulty and three levels of turbulences. A navigation and a surveillance task were performed with the flexible surface to characterize its pro and cons in terms of interaction robustness and situational awareness. The participants were equipped with an electromyogram (EMG) to quantify activity in the deltoid muscle. The flexible surface was used in either folded or flattened mode. We hypothesized that H1) turbulent conditions should affect interaction performances [10]. Following insights gained on a qualitative preliminary evaluation of the flexible surface [28], we also hypothesized that H2) folded vs. flattened mode should improve interaction performance and reduce effort in the shoulder, especially under high turbulences. We finally hypothesized that H3) the physical dimension (shape, movement) of the folded mode should allow a better situational awareness in comparison to the flattened mode [33].

4 MATERIAL AND METHOD

4.1 The flexible surface

Multi-plied are accordion-fold interactive displays integrating shape changing design principles [28]. Following numerous participatory workshops with professional pilots and a deep analysis of their professional activity [22], the Multi-plied prototypes were designed to address some of the operational limitations of touch-based interaction in airliner cockpits. This previous exploration highlighted the ability of two concepts to support various in-flight scenarios use-cases with a high level of safety and efficiency, while promoting crew situational awareness and collaboration. The two functional prototypes that emerged from the workshops were an articulated display series and a pleated display surface, called respectively multi-screen and flexible surface. In this paper, we focused on the evaluation of the flexible surface. This prototype is assumed to make touch interaction more efficient and reliable and to promote the pilot's awareness about the different aircraft system states (e.g., strobe state). The flexible surface is composed of a unique surface measuring 15 cm by 33 cm, foldable in four parts (see Figure 2). The thin 3D printed flexible surface can be deformed thanks to four mechanical arms. The display is generated via a projector (see Figure 1) whose image is dynamically mapped onto the surface. Some parts of this prototype are made of a touch sensitive surface, based on five sensitive rectangles in the rear part of the two bottom folds and the four edges made by the folds.

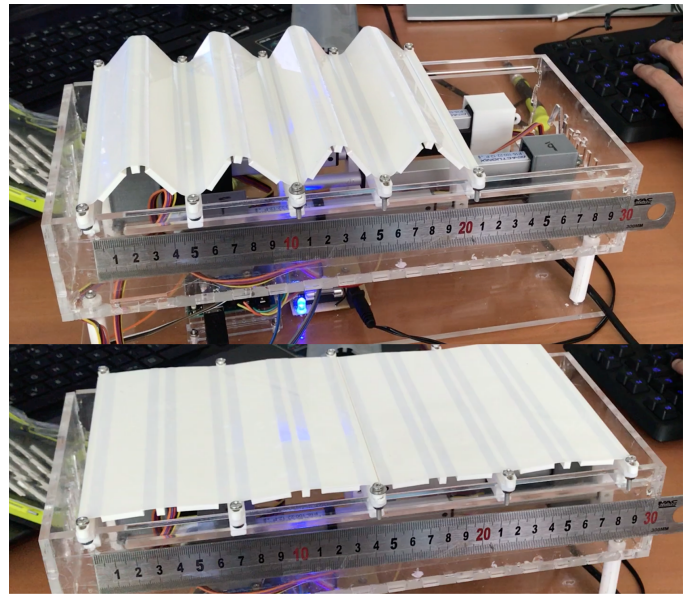


Figure 2: The flexible surface fully folded (top) and fully flattened (bottom).

4.2 Participants

Sixteen participants were enrolled in the experiment. Their age ranged from 20 to 45 years (mean = 27 years). Four participants (25%) were private pilots or future pilots (as a hobby or on training to be professional airliners) and the others were familiar with the aeronautics field and had previous experience with flight simulators. There were two left-handed persons. It should be noted that even left-handed professional pilots must interact with systems with their non-dominant hand (i.e., their right hand), for example when flying in the left seat (captain position) and interacting with the pedestal. All participants had a visual acuity allowing them to see clearly without glasses in a 2m sphere (size of the experimental platform). The tasks performed by the participants (abstract piloting task, tasks with the prototypes) were fully accessible to novice participants. The aim of the study was not to reproduce a fully realistic piloting situation but rather to evaluate the fundamental properties offered by our flexible surface in terms of interaction and situational awareness during simple and synthetic tasks. This choice also allowed us to recruit "normal" participants, airline pilots being a population that is more difficult to access.

4.3 Motion platform and turbulences

The experiments took place in a motion platform with six degrees of freedom (pitch, yaw, and roll) to recreate turbulent conditions (see Figure 1). Participants were installed in a driving seat mounted on the platform. A front screen was installed approximately 70 cm in front of the participant's chest to display the abstract piloting task and some other information related to the surveillance task. A joystick was placed at the left of the seat and a customized keyboard was fixed below the front screen. In order to ease the interaction with the keyboard, it contained only the key required to perform the

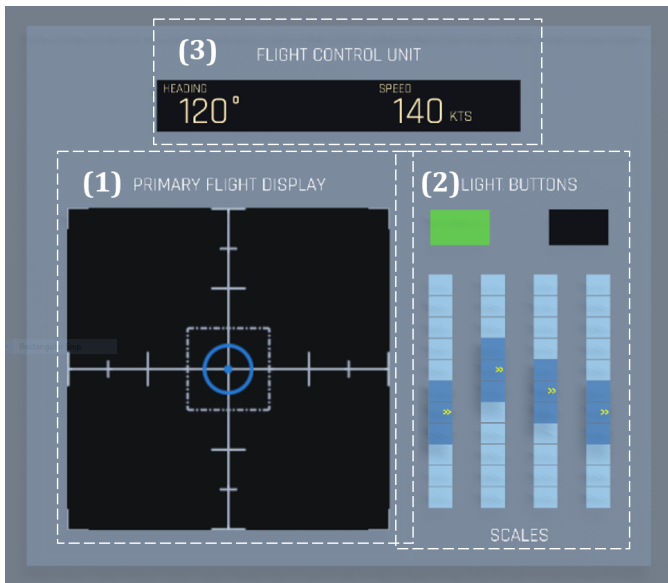


Figure 3: Illustration of the display of the front screen. The abstract piloting task (replicating the MATB-II) consisted of a tracking task (1) and a system monitoring task (2). This display was also used during the navigation task to display the selected heading and speed values (3). Finally, this screen was used during the surveillance task to display a series of questions to assess the situational awareness (not illustrated on this picture).

task and those were distributed among the keyboard to avoid erroneous keystrokes. The flexible surface was placed at the right of the participants following the recommendations for touchscreen use in cockpits [4]. The joystick was meant to be used with the left hand and both the keyboard and the prototype were meant to be used with the right hand. This configuration reproduces the flight captain configuration, where the prototype is used as the ECAM panel. The participants were wearing a headphone that was used to give them instructions during the experimentation (e.g., heading instructions). This headphone had a noise-cancelling function to improve vocal instructions perception and to avoid potential biases due to the mechanical noise of the flexible surface. In line with Hourlier et al. experimental settings [20], we reproduced two levels of turbulences, corresponding to weak and strong turbulences according to the aeronautical classification [1]. The turbulence signal originated from a record made during a real flight (embedded accelerometer and gyroscope). The signal was linearly transformed in frequency and intensity to match aeronautical classification parameters.

4.4 Abstract piloting task

The abstract piloting task was used to reproduce a piloting situation and to generate a sustained mental workload while performing the tasks with the flexible surface. It is a reproduction of the widely used (e.g. [35]) MATB-II [29]. It consists of a set of tasks designed to reproduce piloting activity while being accessible to novice participants. The MATB-II makes possible to control the difficulty by

adjusting different parameters (cursor speeds, range direction, indicators' rates, etc.). In our experiment, the abstract piloting task was performed in parallel to the tasks with the flexible surface. This situation does not reproduce the actual real-life pilot's activity, in particular considering the navigation task. Pilots would not hand-fly and do data entry in the autopilot at the same time. This decision was made to recreate a highly demanding environment and to force the participants to divide their attention to several tasks (tracking, system monitoring and tasks with the flexible surface). However, in order to keep the tasks feasible, we selected only two of the four tasks of the original MATB-II: the tracking task and the system monitoring task. The tracking task and the system monitoring task were displayed on the front screen (see Figure 3.1 and Figure 3.2) and required the use of the keyboard and the joystick. The tracking task reproduces the manual control of the aircraft trajectory. It consists in continuously keeping a cursor in the center of a target thanks to the joystick. The system monitoring task consists of visually monitoring six indicators. Each of these six indicators is associated with a keyboard key that the participant must press as fast as possible to return the indicator to its nominal state (two light buttons and four moving scales). The level of difficulty of the abstract piloting task depended both on the speed/amplitude of the erratic cursor movements of the tracking task and the rate at which each indicators of the system monitoring task leave their nominal state. Performance measurement was focused on the mean distance between the cursor and the center of the target. The participants' performances to this task were focused on the tracking task and measured by the mean distance of the cursor from the center of the target.

4.5 Tasks with the flexible surface

The two tasks performed with the flexible surface were the core of this study. The navigation task addressed the "interaction dimension" of the flexible surface and the surveillance task was designed to address the "situational awareness dimension" of the flexible surface.

4.5.1 Navigation. The navigation task reproduces the manual input of flight parameters in the autopilot. Every 15 seconds, a vocal instruction announced a flight parameter (heading or speed) and an associated value as following "set parameter to value". The participant was asked to modify alternatively the heading and the speed. The required value was randomly chosen among realistic values that could be reached during the time allowed for each instruction. There were eight voice instructions and task duration was about two minutes. The participant was required to input the value of the parameter using the appropriate area of the flexible surface.

During this task, the upper area (area 1) of the flexible surface displayed the values of the two flight parameters. The area just below (area 2) remained empty. Area 3 and 4 were touch-sensitive and served to edit the parameters, i.e to select the heading and speed values, respectively (see Figure 4.1). In addition of providing a touch-based area, these parameter editing areas also displayed the parameter names. Heading selection (area 3) was done using four buttons to add or subtract 1 or 10 degrees to the current heading value (with a possibility of long presses to scroll the value). Speed selection (area 4) was done using a slider, where one had to keep

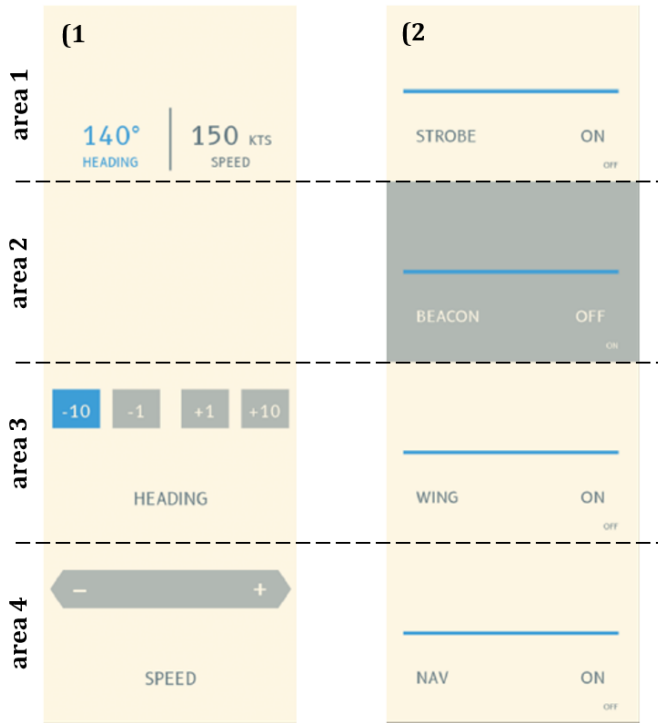


Figure 4: Illustrations of the flexible surface display, divided in 4 areas, during the navigation (1) and the surveillance (2) tasks (in flattened mode). On the navigation task illustration (1), the button increasing the heading value by ten units is pressed, as indicated by the blue color.

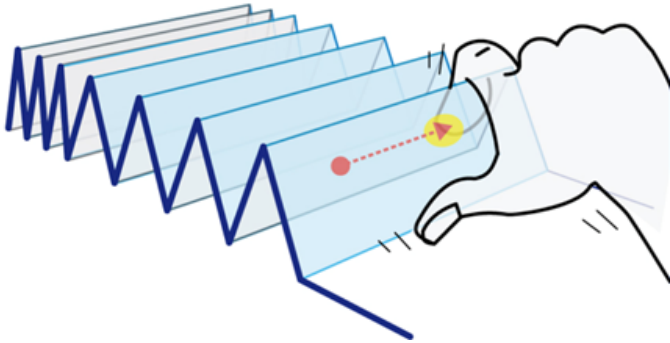


Figure 5: "Slide" interaction with a fold during the navigation task. The finger behind the fold is sliding over the fold to modify the speed parameter.

the finger pressed on the area and to move it to the left or the right for decreasing or increasing the value, respectively. The farther the finger was moved from the center, the faster the speed decreased or increased (see Figure 5). The difference between the folded and the flattened mode condition was that the area 1, 3, and 4 were folded. The area 1 was folded to allow reading the parameters value while the area 3 and 4 were folded to be used to put the palm of the hand and to stabilize the arm, possibly improving the interaction

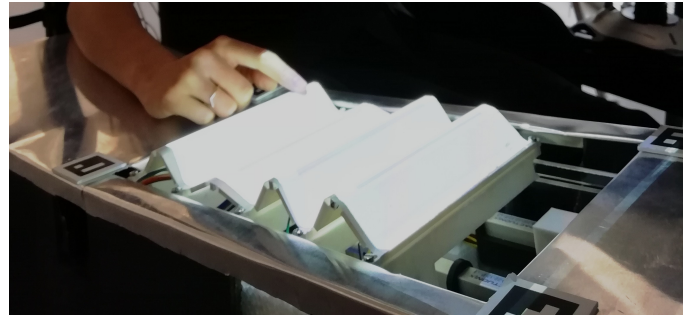


Figure 6: The flexible surface during the surveillance task. On the picture, the user is tapping the last fold to modify the corresponding parameter's state.

robustness during the turbulent conditions. In order to promote eye-free interaction, the current speed and heading values were also displayed on the front screen (see Figure 3.3).

The participants had to be as fast and accurate as possible when selecting the flight parameter values. During this task the performance was measured by the time taken to set the right value. This task was performed in several conditions by participants. To ensure the same level of complexity between participants, flight parameter values to input were the same for each participant. However, to avoid bias, these values were randomly distributed among the conditions.

4.5.2 Surveillance task. The surveillance task reproduces the permanent implicit consciousness that the pilots must maintain about the aircraft system states. As in an actual flight situation, these system states evolve as a result of pilot actions and are also subject to changes, even without explicit actions from them. Each of the four selected parameters ("strobe", "beacon", "wing" or "nav") were displayed on one of the four areas of the flexible surface, with the name of the parameter and its current state ("on" or "off"). The parameter states changed after an explicit action from the participant following a vocal instruction asking to modify a parameter state, or regularly without actions (see Figure 6). Depending on whether the flexible surface was in flattened or folded mode, the states of the parameters were respectively represented by coloration (see Figure 4.2) or by the shape of the fold. A flattened fold or a dark area indicated that the parameter was "off"; a raised fold or a bright area indicated that the parameter was "on". To modify a parameter state, the participants had to tap the area (in flattened mode) or the fold (in folded mode) corresponding to the parameter; subsequently, the state switched from "on" to "off" or from "off" to "on". Regularly the task was interrupted to test the awareness of the participants about the states of the parameters: A series of questionnaires was displayed on the front screen and the participants used the keyboard to recall each of the four parameter states. This method was inspired by the SAGAT, a widely used technique to measure situation awareness [13]. It consists in interrupting the experimentation at some random point in time and asking the participant to answer questions about the state of the systems. In the current experimentation, the questions were focused on the four modifiable parameters (e.g., no question about the system monitoring task).

There was a total of four questionnaires and three parameters modifications (with a modification every 6.5 seconds in average, randomly via explicit modifications from the participant or by automatic system modifications) between each questionnaire. It resulted in a total of nine parameters modifications such that the task duration was approximately two minutes (including questionnaires). Participants were asked not to over-focus on the flexible surface and to perform the abstract piloting task as correctly as possible. During this task, the performance was measured by the number of errors when answering the questionnaires (when participants did not report properly the parameter state) and by the average total time taken to fulfill each series of questions. This task was performed in several conditions by participants. The states of the 4 folders (on/off) were the same for each participant. The values were randomly distributed among the conditions.

4.6 Questionnaire

In order to identify the subjective usability of the flexible surface, we asked the participants to fulfill SUS questionnaires. SUS is a usability test widely used in the field of human computer interaction [9]. It consists of a 10-item questionnaire with five response options. It provides a score on a 0-100 scale. Only one questionnaire was added in order not to overwhelm the participants. The SUS questionnaire is a bit dated but it is still widely used, especially in related studies, allowing us to compare our results.

4.7 Electromyographic measurements of the deltoid muscle

We measured the muscular effort in the shoulder during the navigation task as it required to interact with the flexible surface. The participants were equipped with an EMG to record muscular effort of the shoulder, more specifically, the electric potential generated by the medium deltoid activity. Two electrodes were placed between the acromion to the lateral epicondyle of the elbow and one electrode was placed on the 7th cervical to serve as a reference [36]. The raw signal was rectified by applying a root mean square function [25] (see Figure 7). All EMG data has been normalized with respect to an EMG value taken before each experimental scenario, during a maximum isometric voluntary contraction (MIVC) test for normalizing deltoid muscle EMG. Meticulous visual inspection of the signal was conducted to ensure that no motion artefacts were present, which can happen when an electrode is moved or the skin is stretched [26].

4.8 Procedure

The participants were informed about the risks of the procedure, the fact that they could interrupt the experimentation at any moment, and about the fact that recorded data were anonymous and will be used only for scientific purposes. Then, the participants were given instructions describing the experiment and they were installed on the driving seat mounted on the platform. They were secured with a safety harness attached to the platform structure and were equipped with the EMG and the headphone with noise-cancelling function.

They first performed the navigation task under the six following conditions:

- Two interaction modes: “flattened mode” and “folded” mode;

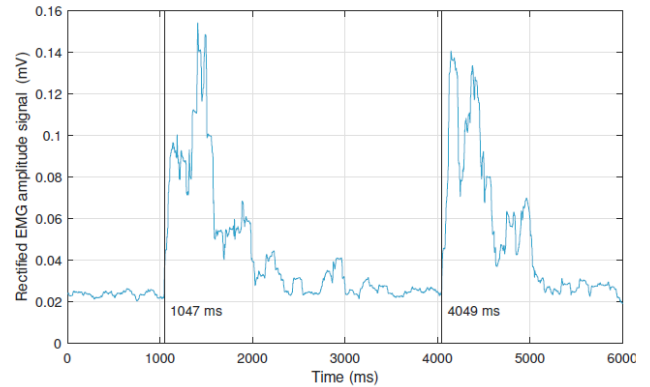


Figure 7: Illustration of the onsets of two arm movements on the flexible surface (vertical black lines) during the navigation task and its effect on the medium deltoid activity (blue line).

- Three levels of turbulence: none, weak and strong.

To perform the navigation task under the six possible conditions, the participants were asked to accomplish the task (and thus the parallel abstract piloting task) six times. The order of the six different conditions was randomized across the participants.

Then the participants performed the surveillance task under the four following conditions:

- Two interaction modes: “flattened mode” and “folded” mode;
- Two levels of cognitive demand: low and high (depending on the cursor movement’s speed/amplitude plus the rate at which the states of the indicators of the system monitoring task change).

To perform the task under the four possible conditions, the participants were asked to accomplish the surveillance task (and thus the parallel abstract piloting task) four times. The order of the four conditions was randomized across the participants.

Before each task with the flexible surface, the participants performed a short training session to ensure that they fully understood the tasks rules. It also helped ensuring that they get familiar with the abstract piloting task. After each experimental condition of the task with the flexible surface, participants fulfilled the SUS questionnaire. The whole experiment lasted approximately two hours.

4.9 Statistical analysis and experimental variables

Results were analyzed using STATISTICA software version 10 (StatSoft Inc). We performed 2 x 3 repeated measures analysis of variance (ANOVA) for each dependent variable to assess the effects of our experimental factors (independent variables). More precisely, during the two tasks, dependent variables included the subjective usability (SUS score) and the performance to the piloting task (mean distance of the cursor from the center of the target). In addition, during the navigation task, performance with the flexible surface was assessed via the time taken to input the parameters. During

the surveillance task, the performance with the flexible surface was investigated via the number of errors to the questionnaires. Finally, the effort in the shoulder (EMG signal amplitude of the deltoid muscle) was also measured during the navigation task. The effects of three within-subject factors were considered during both tasks: The two flexible surface modes (folded, flat), the three levels of turbulences (no turbulence, weak turbulence, strong turbulence), and the two levels of task difficulty (low, high). Post hoc tests were conducted using the Tukey HSD test (this test corrects for multiple pairwise comparisons). We used Mauchly's test to test the sphericity assumption for the within-subject (repeated measures) factor with more than two levels (or more than a single degree of freedom, i.e., the main effect of turbulence with three levels, which would have 2 degrees of freedom). In this case the univariate F statistic reported in the ANOVA table is based on the assumption that the hypotheses associated with each single degree of freedom are orthogonal. The Greenhouse-Geisser correction was applied if the assumption of sphericity was violated (i.e., significant Mauchly's test). The Greenhouse-Geisser correction computes reduced degrees of freedom to reflect the fact that the within-subject hypotheses are correlated, and not independent. We also used the Kolmogorov-Smirnov test to test the normality of all variables. They all met the assumptions of normality (Kolmogorov-Smirnov test significance was always > 0.05). Effect sizes were reported using partial eta-squared (η^2). Due to unexpected events during the experiment (failure to record the data), some participants data are missing for two analyzes.

5 RESULTS

This section detail the following results. Findings show that the interaction is affected by turbulence, but that the folded screen does not differ from the control condition during navigation tasks. However, the folded mode results is less muscle strain (sensed using an EMG sensor) and better situational awareness.

5.1 Navigation task

5.1.1 Subjective usability of the flexible surface. A two-way ANOVA (14 valid participants) showed a significant main effect of turbulences on the subjective usability, $F(2, 26) = 10.74, p < 0.001, \eta^2 = 0.45$, with a lower perceived usability of the flexible surface (irrespective of the flexible surface mode) during strong turbulences ($M = 59.03, SE = 5.03$) vs. weak turbulences ($M = 63.97, SE = 3.96$), HSD, $p < 0.001$, and during weak turbulences vs. no turbulence ($M = 70.80, SE = 4.61$), HSD, $p = 0.032$). The main effect of the flexible surface mode and the interaction effect were not significant ($p > 0.05$).

5.1.2 Performance to the abstract piloting task. A two-way ANOVA (14 valid participants) showed a significant effect of turbulences on the performance to the abstract piloting task, Greenhouse-Geisser corrected: $F(1.06, 13.85) = 44.76, p < 0.001, \eta^2 = 0.77$ (see Figure 8), with a monotonic decrease of performance with increased turbulences. Performance to the abstract piloting task (cursor distance to the center) were lower in strong turbulences ($M = 0.28, SE = 0.03$) vs. weak turbulences ($M = 0.20, SE = 0.01$), HSD, $p < 0.001$, and lower during weak turbulences vs. no turbulence ($M = 0.14, SE =$

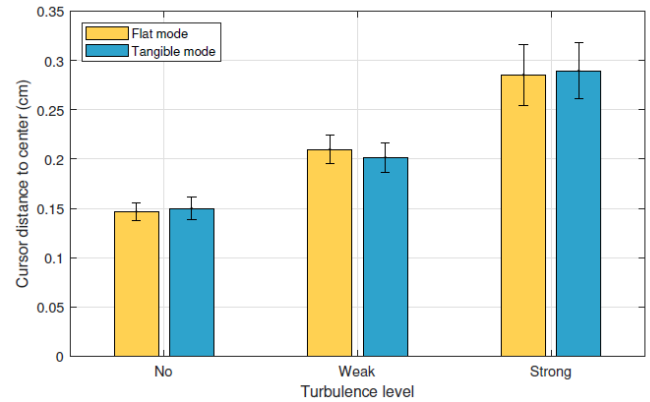


Figure 8: Mean performance to the abstract piloting task during the navigation task across the three levels of turbulences and the flexible surface mode. Errors bars represent the standard error mean.

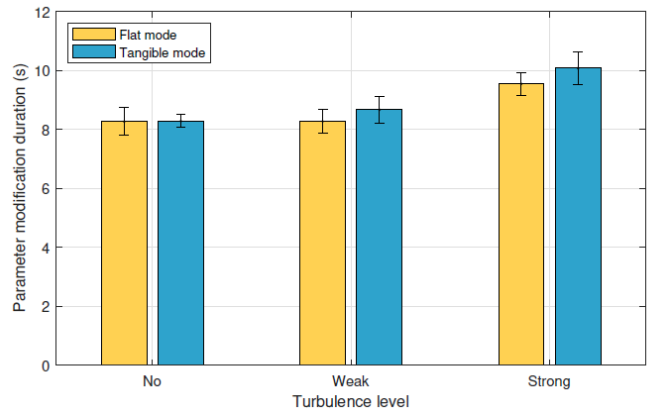


Figure 9: Time taken to input the parameters during the navigation task across the three levels of turbulences and the flexible surface mode. Errors bars represent the standard error mean.

0.01), HSD, $p = 0.001$. The main effect of the flexible surface mode and the interaction effect were not significant ($p > 0.05$).

5.1.3 Performance with the flexible surface. A two-way ANOVA (16 valid participants) showed a significant main effect of turbulences on the time taken to input the parameters, $F(2, 30) = 17.06, p < 0.001, \eta^2 = 0.53$ (see Figure 9), with an increased time under strong turbulence ($M = 9.81, SE = 0.36$) vs. weak turbulence ($M = 8.47, SE = 0.38$) and vs. no turbulence ($M = 8.27, SE = 0.26$), HSD, $p < 0.001$ in both comparisons. The main effect of the flexible surface mode and the interaction effect were not significant ($p > 0.05$).

5.1.4 Effort in the right shoulder while interacting with the flexible surface. A two-way ANOVA (16 valid participants) showed a significant main effect of turbulences on the EMG signal amplitude, Greenhouse-Geisser corrected: $F(1.16, 14.01) = 8.36, p = 0.009, \eta^2$

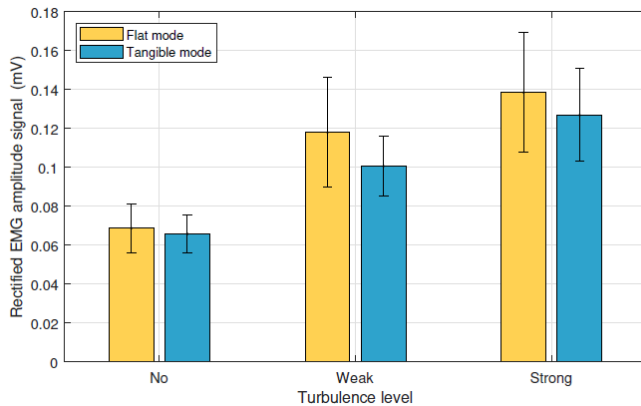


Figure 10: Mean rectified EMG signal amplitude during the navigation task across the three levels of turbulence and the flexible surface mode. Errors bars represent the standard error mean.

= 0.41 (see Figure 10). The intensity of shoulder muscles activity was higher during strong turbulence ($M = 0.12$, $SE = 0.03$) vs. weak turbulence ($M = 0.10$, $SE = 0.02$), HSD , $p = 0.001$, and higher during weak turbulence vs. no turbulence ($M = 0.06$, $SE = 0.01$), HSD , $p = 0.001$. We also found a significant main effect of the flexible surface mode, $F(1, 12) = 4.97$, $p = 0.045$, $\eta p^2 = 0.29$ (see Figure 10), with a lower shoulder muscles activity in the folded mode ($M = 0.09$, $SE = 0.02$) vs. flattened mode ($M = 0.10$, $SE = 0.03$). The interaction effect was not significant ($p > 0.05$).

5.2 Surveillance task

5.2.1 Subjective usability of the flexible surface. A two-way ANOVA (16 valid participants) showed no significant main nor interaction effects of the flexible surface mode and the difficulty ($p > 0.05$ for all analyzes).

5.2.2 Performance to the abstract piloting task. A two-way ANOVA (16 valid participants) showed a significant main effect of the flexible surface mode on the performance to the piloting task, $F(1, 15) = 13.03$, $p = 0.002$, $\eta p^2 = 0.46$, with better piloting performance when using the flexible surface in folded ($M = 0.16$, $SE = 0.01$) mode vs. flattened mode ($M = 0.17$, $SE = 0.01$). We also found a highly expected main effect of the difficulty of the piloting task to the performance to this same task, with lower performance when difficulty was high ($M = 0.22$, $SE = 0.01$) vs. low ($M = 0.11$, $SE = 0.01$), $F(1, 15) = 89.64$, $p < 0.001$, $\eta p^2 = 0.86$. The interaction effect was not significant ($p > 0.05$).

5.2.3 Performance with the flexible surface. A two-way ANOVA (16 valid participants) showed a significant main effect of the flexible surface mode on the number of errors to the questionnaires, $F(1, 15) = 4.83$, $p = 0.043$, $\eta p^2 = 0.24$ (see Figure 11), with more errors committed when using the flexible surface in flattened mode ($M = 4.28$, $SE = 1.15$) vs. folded mode ($M = 3.03$, $SE = 1.08$). The main effect of difficulty and the interaction effect were not significant ($p > 0.05$).

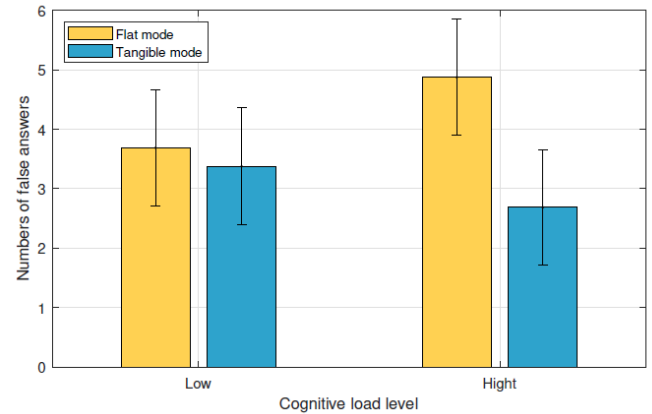


Figure 11: Mean number of errors during the surveillance task across the two levels of cognitive load and the flexible surface mode. Errors bars represent the standard error mean.

6 DISCUSSION

In this paper, we showed that providing folds (in folded mode) allowed a general reduction of the physical effort because it helped to stabilize the user arm. The physicality of the flexible surface in this mode was also associated with higher performance to the parallel abstract piloting task and with a higher situational awareness of the aircraft system states, most likely because this interaction had better visual properties and could free some critical attentional resources.

6.1 Navigation task

Our results supported H1, a context of high turbulences was highly deleterious to the different tasks in the cockpit. The turbulent conditions provoked a diminution of the performance to the abstract piloting task as well as to the navigation task performed with the flexible surface (irrespective of the surface mode). Consistently, high turbulences had a negative impact on the subjective usability of the flexible surface and increased the muscular effort in the shoulder when interacting with it. Taken together, these results confirmed that turbulence must be carefully considered in the design of future touchscreen cockpit displays [1, 10]. Partially failing to validate H2, we found no significant performance advantages for the folded mode over the flattened mode to cope with the turbulence during the navigation task. Also, the folded mode of the flexible surface was not significantly perceived as more usable than the flattened mode.

Interviews with the participants raised several issues during the navigation task that might be corrected in the future. For example, participants reported that interacting without the eyes was not really possible, even in the folded mode. A participant explained that: "the fold was convenient because it allows [her] to grab the surface but then [her] hand was masking the interaction area. [She] tried to interact without looking at it, but the touchscreen interaction was not easy enough". Another explanation was that the tactile area behind the folds was not entirely interactive. As a reminder, it

consisted in five separate square tactile zones. Consequently, the interaction required too much accuracy and eyes-free interaction was not really possible due to lack of guidance on the tactile area. However, despite this lack of observable performance advantages with the folded mode, it generated a lower muscular effort in the shoulder than the flattened mode, likely reflecting the benefits of being able to grasp the folds to stabilize the hand.

6.2 Surveillance task

The results confirmed that the folded mode allowed a better situation awareness (less error to the questionnaires) in comparison to the flattened mode, validating H3. First, the number of errors was lower when reporting the system state with the folded mode. This outcome tends to support that the physical emergence of the shape on a flat surface acted as an efficient external cue of the system state, in a similar fashion than the landing lever and the auto-throttle position act as an external reminder to know the system states. Second, folded mode allowed higher performance to the abstract piloting task than the flattened mode. It supports the idea that the folded mode helped maintaining attention toward the system states while performing other concurrent tasks and without reducing the attention paid to them. We assumed that the participants could more easily monitor the system states in peripheral vision thanks to the better salience of the folds. A participant reported that: "It [was] way easier with the folds". This result indicated that the emergence of the shape and the movement of the flexible surface was an efficient way to maintain the user attention about the system states (i.e., on/off). It tends to support the supposed advantages of physical information to convey efficiently information [16, 18].

6.3 Limits

Some limits should be mentioned. First, our results suggested that the folded mode allowed a better situation awareness (less error to the questionnaires) thanks to the physical emergence of shape, in comparison to the flattened mode, validating H3. It would be interesting to investigate whether this outcome can also partially result from the orientation of the surface in relation to the pilot. One can argue that the folded mode allowed a better visibility in peripheral vision due to the position of the folds. A flat but tilted surface (with the same angulation than the folds) could be also efficient. However, most participants reported that they had to turn their head to look at the flexible surface (i.e., the surface was not really visible in peripheral vision in none of the two modes), independently from its mode, suggesting that shape changing property was also important for improving their situational awareness, in addition to the simple higher visibility of the information. In addition, several parts of our experiment did not reproduce an ecologically valid situation. The majority of the participants were not licensed pilots. However, it is important to note that the tasks performed by the participants (abstract piloting task, tasks with the prototypes) were fully accomplishable to novice participants. The aim of the study was not to reproduce a piloting situation as realistically as possible but rather to evaluate the fundamental properties offered by our flexible surface in terms of interaction and situational awareness during simple and synthetic tasks. Given the nature of the tasks

employed, we believe that the impact of our participant profile was very low.

In addition, the flexible surface used during this experiment was an early prototype. The prototyping of tangible operational devices is known to be complex [17], and this work confirms it. Some improvements could be done regarding the tactile interaction as it raised concerns from several participants. Firstly, it would be interesting to increase the resolution of the tactile area to improve the reliability and accuracy of the tactile area. Secondly, a better guidance of the fingers in folded mode would be desirable in the future. This could be done by adding some texture to the surface. Finally, it would be interesting to move the touch area to the other side of the fold (the front). We discovered that the way the fold was meant to be grabbed (relying on the thumb and using the other finger behind the fold, see Figure 5) was not the most convenient way to interact (despite being tested during a pilot study). It would be interesting to allow the thumb to interact with the front of the fold and to rely on the other fingers to stabilize the position.

Finally, regarding the number of participants, 16 participants is above the bottom limit to apply a one-way ANOVA. This study has to be considered as a pilot study investigating what results could be expected. It leads to another one, more consequent, that requires more participants. For example, this evaluation aimed at being independent from expertise such that despite a small number of participants they come from different backgrounds. Yet it does not mean that expertise does not have an impact. Further studies could multiply the number of participants according to different levels of expertise to apply a two-way ANOVA regarding expertise, and as such, measure its effect.

7 CONCLUSION

This study, by examining the advantages of physicality and shape changing properties for touchscreen interfaces, identified several principles that may lead to a better integration of touchscreen interfaces in airliner cockpits. First, in accordance with existing studies, we confirmed the strong negative impact of turbulence on their subjective and actual usability: turbulence indeed affected touchscreen interaction as well as muscular effort. Second, the folded mode required a lower muscular effort than when the task was performed in flattened mode. The grabbing possibility offered by the folds explains this latter result. Another interesting result concerned the impact of folding on situation awareness. We found that the flexible surface in folded mode promoted a better building and maintenance of the situational awareness about the state of the aircraft system. Thirdly, we believe that a few improvements on the design of the prototype can be made to better support hypothesis H2, (only partially confirmed) e.g. by increasing the resolution of the tactile area and moving it to the front of the fold (see Section Limits). Fourthly, our results also confirmed that a high workload context tends to reduce subjective usability of touchscreens. Future research might investigate the factors that can explain this specific issue. A final contribution was the development of new interaction widgets on the flexible surface to select target headings using incremental buttons and target speed using a slider. It is worthwhile noting that these are novel interaction mechanisms in aviation, as pilots mostly

use rotating knobs located on the glare shield to select navigation target.

REFERENCES

- [1] Federal Aviation Administration. 2019. Aeronautical information manual: Official guide to basic flight information and ATC procedures.
- [2] Alexandre Alapetite, Rune Fogh, David Zammit-Mangion, Christian Zammit, Ian Agius, Marco Fabbri, Marco Pregnolato, and Loïc Becouran. 2012. Direct tactile manipulation of the flight plan in a modern aircraft cockpit. In *Proceedings of International Conference on Human-Computer Interaction in Aerospace (HCI Aero 2012)*. Association for Computing Machinery, New York, NY, United States, 2–5.
- [3] Huseyin Avsar, Joel E Fischer, and Tom Rodden. 2016. Designing touch screen user interfaces for future flight deck operations. In *Proceedings of IEEE/AIAA 35th Digital Avionics Systems Conference (DASC 2016)*. IEEE, Sacramento, CA, United States, 1–9.
- [4] Jérôme Barbé, Régis Mollard, and Marion Wolff. 2014. Ergonomic approaches to integrate touch screen in future aircraft cockpits. *Journal Européen des Systèmes Automatisés* 48, 4-5-6 (2014), 303–318.
- [5] David Barstow. 2012. The aviation iPad revolution. *Journal of Air Traffic Control* 54, 2 (2012).
- [6] Jacques Bertin. 1983. *Semiology of graphics: diagrams networks maps*. Technical Report.
- [7] Guy A Boy. 2012. What do we mean by Human-Centered Design of Life-Critical Systems? *Work* 41, Supplement 1 (2012), 4503–4513.
- [8] Scott Brave, Hiroshi Ishii, and Andrew Dahley. 1998. Tangible interfaces for remote collaboration and communication.. In *CSCW*, Vol. 98. 169–178.
- [9] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [10] Andy Cockburn, Carl Gutwin, Philippe Palanque, Yannick Deleris, Catherine Trask, Ashley Coveney, Marcus Yung, and Karon MacLean. 2017. Turbulent Touch: Touchscreen Input for Cockpit Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 6742–6753. <https://doi.org/10.1145/3025453.3025584>
- [11] Philip R Cohen and David R McGee. 2004. Tangible multimodal interfaces for safety-critical applications. *Commun. ACM* 47, 1 (2004), 41–46.
- [12] Juan Angel Lorenzo del Castillo and Nadine Couture. 2016. The aircraft of the future. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '16*. ACM Press, New York, New York, USA, 1–8. <https://doi.org/10.1145/2950112.2964582>
- [13] Mica R Endsley. 1988. Situation awareness global assessment technique (SAGAT). In *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference*. IEEE, 789–795. <https://doi.org/10.1109/NAECON.1988.195097>
- [14] Ken Funk. 1991. Cockpit task management: Preliminary definitions, normative theory, error taxonomy, and design recommendations. *The International Journal of Aviation Psychology* 1, 4 (1991), 271–285.
- [15] Miguel A Garcia-Ruiz. 2013. *Cases on Usability Engineering: Design and Development of Digital Products: Design and Development of Digital Products*. IGI Global.
- [16] James J Gibson. 1979. The theory of affordances. The ecological approach to visual perception.
- [17] Jeffrey C F Ho. 2019. ClothSurface: Exploring a Low-Cost Prototyping Tool to Support Ideation for Shape Displays. *Multimodal Technologies and Interaction* 3, 3 (2019).
- [18] Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM Press, 437–446.
- [19] Sylvain Hourlier, Sandra Guérard, Jean Luc Barou, and Xavier Servantie. 2015. Testing Touch Screens in Realistic Aeronautic Turbulent Conditions (Light to Severe). *SAE International Journal of Aerospace* 8, 2 (sep 2015), 2015–01–2532. <https://doi.org/10.4271/2015-01-2532>
- [20] Sylvain Hourlier, Sandra Guerard, and Xavier Servantie. 2019. Chapter 3: Avionics touch screen in turbulence: simulator design and selected human-machine interface metrics. In *Advances in Aviation Psychology, Volume 3, Improving Aviation Performance through Applying Engineering Psychology* (taylor & francis ed.), Michael Vidulich and Pamela Tsang (Eds.).
- [21] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the efficiency of physical visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press, 2593–2602.
- [22] Catherine Letondal, Jean-Luc Vinot, Sylvain Pauchet, Caroline Boussiron, Stéphanie Rey, Valentin Becquet, and Claire Lavenir. 2018. Being in the Sky: Framing Tangible and Embodied Interaction for Future Airliner Cockpits. In *Proceedings of the 12th International Conference on Tangible, Embedded, and Embodied Interaction - TEI '18*, Vol. 2018-Janua. ACM Press, New York, New York, USA, 656–666. <https://doi.org/10.1145/3173225.3173229>
- [23] Margaret Livingstone and David Hubel. 1988. Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science* 240, 4853 (1988), 740–749. <https://doi.org/10.1126/science.3283936>
- [24] Loukia D Loukopoulou, RK Dismukes, and I Barshi. 2001. Cockpit interruptions and distractions: A line observation study. In *Proceedings of the 11th international symposium on aviation psychology*. Ohio State University Press Columbus, OH, 1–6.
- [25] K Mahaphonchaikul, D Sueaseenak, C Pintavirooj, M Sangworasil, and S Tungjitsukulmun. 2010. EMG signal feature extraction based on wavelet transform. In *ECTI-CON2010: The 2010 ECTI International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, 327–331.
- [26] Andrea Merlo and Isabella Campanini. 2010. Technical aspects of surface electromyography for clinicians. *The open rehabilitation journal* 3, 1 (2010).
- [27] Sylvain Pauchet, Catherine Letondal, Jean-Luc Vinot, Mickaël Causse, Mathieu Cousy, Valentin Becquet, and Guillaume Crouzet. 2018. GazeForm. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*. ACM Press, New York, New York, USA, 1193–1205. <https://doi.org/10.1145/3196709.3196712>
- [28] Sylvain Pauchet, Jean Luc Vinot, Catherine Letondal, Alexandre Lemort, Claire Lavenir, Timothée Lecomte, Stéphanie Rey, Valentin Becquet, and Guillaume Crouzet. 2019. *Multi-plié: A linear foldable and flattenable interactive display to support efficiency, safety and collaboration exploring and ironing out design complexities with airliner pilots*. Technical Report. <https://doi.org/10.1145/3290605.3300384>
- [29] Yamira Santiago-Espada, Robert R Myer, Kara A Latorella, and James R Comstock Jr. 2011. The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user’s guide. (2011).
- [30] Paul C Schutte and Anna C Trujillo. 1996. Flight crew task management in non-normal situations. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 40. SAGE Publications Sage CA: Los Angeles, CA, 244–248.
- [31] Simon Stusak. 2015. Exploring the Potential of Physical Visualizations. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*. ACM Press, New York, New York, USA, 437–440. <https://doi.org/10.1145/2677199.2691604>
- [32] Faisal Taher, John Hardy, Abhijit Karnik, Christian Weichel, Yvonne Jansen, Kasper Hornbæk, and Jason Alexander. 2015. Exploring Interactions with Physically Dynamic Bar Charts. <https://doi.org/10.1145/2702123.2702604>
- [33] Jean-Luc Vinot, Catherine Letondal, Sylvain Pauchet, and Stéphane Chatty. 2016. Could tangibility improve the safety of touch-based interaction?. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '16*. ACM Press, New York, New York, USA, 1–8. <https://doi.org/10.1145/2950112.2964581>
- [34] Simon Voelker, Kjell Ivar Øvergård, Chat Wacharamanatham, and Jan Borchers. 2015. Knobology revisited: A comparison of user performance between tangible and virtual rotary knobs. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*. ACM Press, 35–38.
- [35] Ziheng Wang, Ryan M Hope, Zuoguan Wang, Qiang Ji, and Wayne D Gray. 2011. An EEG workload classifier for multiple subjects. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 6534–6537.
- [36] P Zipp. 1982. Recommendations for the standardization of lead positions in surface electromyography. *European Journal of Applied Physiology and Occupational Physiology* 50, 1 (1982), 41–54.