Investigating Multimodal Augmentations Contribution to Remote Control Tower Contexts for Air Traffic Management

Maxime Reynal, Pietro Aricò, Jean-Paul Imbert, Christophe Hurter, Gianluca Borghini, Gianluca Di Flumeri, Nicolina Sciaraffa, Antonio Di Florio, Michela Terenzi, Ana Ferreira, et al.

Investigating Multimodal Augmentations Contribution to Remote Control Tower Contexts for Air Traffic Management

Maxime Reynal[1], Pietro Arico[2], Jean-Paul Imbert[1], Christophe Hurter[1], Gianluca Borghini[3], Gianluca Di Flumeri[2], Nicolina Sciaraffa[2], Antonio Di Florio[4], Michela Terenzi[3], Ana Ferreira[3], Simone Pozzi[3], Viviana Betti[2], Matteo Marucci[2] and Fabio Babiloni[2]

[1]*French Civil Aviation University (ENAC), Toulouse, France*

[2]*Sapienza University, Roma, Italy*

[3]*Deep Blue, Roma, Italy*

[4]*Fondazione Santa Lucia, Roma, Italy*

Keywords: Multimodal Stimuli, Interactive Sound, Vibrotactile Feedback, Remote Control Towers.

Abstract: The present study aims at investigating the contribution of multimodal modalities to the context of Remote Towers. Interactive spatial sound and vibrotactile feedback were used to design 4 different types of interaction and feedback, responding to 4 typical Air Traffic Control use cases. The experiment involved 16 professional Air Traffic Controllers, who have been called to manage 4 different ATC scenarios into ecological experimental conditions. In two of the scenarios, participants had to control only one airport (i.e., Single Remote Tower context), while in the other two scenarios participants had to control simultaneously two airports (i.e., Multiple Remote Tower context). The augmentation modalities were activated or not in a balanced way. Behavioral results highlighted a significant increase in overall participants performance when the augmentation modalities were activated in Single Remote Tower context. This work demonstrates that some types of augmentation modalities can be used into Remote Tower context.

1 INTRODUCTION

Due to its important cost, Air Traffic Services can sometimes be difficult to provide in very low traffic density airports (approximately one or two flight per day). Some solutions have already been proposed to tackle this issue (Nene, 2016). One could be to bring those services into places where human resources could be brought together and then easier to manage. These remote control rooms would not be located on the controlled airport areas but centralized in more accessible and more densely populated places. Air Navigation Authorities and laboratories have already issued information and recommendations for further developments in this context (Calvo, 2009; Braathen, 2011; Fürstenau et al., 2009). The present study falls in this research field with the purpose of increasing human performances in Remote Tower context by using Human-Machine Interaction (HMI) concepts, and therefore, increase safety.

A Remote Control Tower (RT) is a control tower that is not located in the airports area. The equipment that can be found in a RT is similar to the one used in a regular control tower. However, the external view on the airport vicinity is streamed via cameras that are located on the controlled airport. Since the airport environment is only visible through screens, it opens the way for multiple visual augmentations. Night vision, zooming, driving the field of view, and even holographic projections (Schmidt et al., 2006), are some of the augmentation modalities that have already been studied in existing RT solutions. A multiple RT is a RT used to control not one but two or more airports simultaneously. This concept is more and more studied due to the fact that it may allow to greatly reduce the costs related to control towers and may also facilitate the management of human ressources. However, controlling at the same time several aircraft located at the vicinity of several airports may have an impact on the work of the Air Traffic Controllers (ATCos). In the case the systems are not properly designed, ATCos workload would be likely to increase, which may

Reynal, M., Arico, P., Imbert, J., Hurter, C., Borghini, G., Di Flumeri, G., Sciaraffa, N., Di Florio, A., Terenzi, M., Ferreira, A., Pozzi, S., Betti, V., Marucci, M. and Babiloni, F.
Investigating Multimodal Augmentations Contribution to Remote Control Tower Contexts for Air Traffic Management.
DOI: 10.5220/0007400300500061
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 50-61
ISBN: 978-989-758-354-4

also increase the risk of error, and then safety. In this paper, we investigate the contribution of several new augmentation modalities in the context of single and multiple RT in terms of behavioral and subjective measurements.

As many possible visual augmentations have already been tested, the present study focuses on auditory and tactile augmentation modalities. Moreover, it is also possible to change the way specific information is provided to the operators using sensory channels other than the visual one. This may enable to relieve the visual channel that tends to be overloaded during the Air Traffic Control (ATC) task. In this study, four augmentation modalities were designed and tested for the RT framework: an interaction form based on spatial sound, a spatialized sound feedback, and two sorts of vibrotactile feedback. These augmentation modalities were used to address specific issues that are presented in detail in the form of use cases in the next section.

The present paper can be read in continuation of a previous study (Arico et al., 2018). It is structured as following: section 2 proposes a brief state of the art on the technologies that have been used in the present study, then the different modalities that have been designed and tested are described in section 3. In section 4, the participants population, the experimental design, protocol and setup, and the metrics used are presented. Section 5 relates the results that are discussed in section 6 before concluding the study in section 7.

2 RELATED WORKS

In this section, we focus on the state of the art of Remote Towers, interactive systems acting on sound, and haptics, more precisely vibrotactile feedback used to communicate information to the user. Several studies have been led to try to understand more deeply the ATCos' working task, the sensory channels and mental processes involved or even their emotional states (Pfeiffer et al., 2015). Vision is the most studied human sense for Single Remote Towers (Cordeil et al., 2016; Hurter et al., 2012; Van Schaik et al., 2010). Some studies related to hearing channel for RTs have been led, like for example an innovative method of sound spatialization using binaural stereo aiming at discriminating enroute ATCos communication (Guldenschuh and Sontacchi, 2009). Multiple RT context has been little studied in scientific literature due to the fact that it is still a rather recent subject. A first study was published in 2010 to demonstrate the feasibility of multiple RT, at least to control two small airports simultaneously (Moehlenbrink and Papenfuss, 2011; Papenfuss and Friedrich, 2016).

In a physical control tower, ATCos are used to hear different types of sounds: aircraft engine sounds starting from the parking area in front of the tower, engine tests at the end of the runway, communications with pilots, or the sound of the wind around the tower, for example. They already have discrete spatialized sound sources to deal with. Therefore, sound can become a way to provide additional information to them. However, too many spatial sound sources may make it difficult to dissociate from each others. A solution to this problem may be to make these sound sources interactive. Interactive systems acting on sound described in literature are mostly based on eye gaze information and not on the head orientation. Firstly, we can cite the work of (Bolt, 1981), who imagined in the early 1980s a system aiming at focusing the user attention on sounds while facing a wall of screens operating simultaneously and broadcasting different images (about 20 simultaneous images and sounds). The sounds played by these televisions were amplified according to the users gaze. Since this study, the will to amplify the sound towards the user has been identified several times in recent literature. In particular, the OverHear system (Smith et al., 2005) extended GAZE and GAZE-2 (Vertegaal, 1999; Vertegaal et al., 2003) studies by providing a method for remote sound amplification based on the user gaze direction using directional microphones. We can also mention the AuraMirror tool (Skaburskis et al., 2003) that informs the user graphically of his attention by superimposing on his vision a particular shape (e.g. colored "bubbles") around the concerned interlocutors in a multi-speaker situation. Other interactive systems acting on sound were presented by (Savidis et al., 1996), which is similar in some ways to the present work. Using this system, the user is surrounded by interactive sound sources organized in a "ring" topology. They can select specific sound sources thanks to 3D pointing, gestures and speech recognition inputs. The goal of this system is to provide the user means to explore the auditory scene with the use of direct manipulation (Hutchins et al., 1985) via a simple metaphor mapping of a structured environment.

Vibrations are also important in the ATCos environment. For example, the wind can cause the tower to oscillate, giving the ATCos an indication of its direction and speed, or vibrations induced by engine tests give them a feedback on some of the immediate actions that pilots can perform. Various studies presented different systems allowing communicating spatial information to users via vibrotactile feedback. Some contributions to vibrotactile feedback in aeronautics field can be found. Van Erp and his collegues developed a tactile display system to indicate navigation information (waypoints) to pilots through a belt (Erp et al., 2005; Van Erp et al., 2004). They found a way to present direction and distance information through vibrotactile feedbacks and their results were significant concerning direction information. Another system developped by Raj et al. provided tactile cues to helicopter pilots aiming at helping them to perform hover maneuvers (Raj et al., 2000). The results showed that this kind of information can significantly enhance pilots performance. Also, in the context of car driving, multiple studies used vibrotactile patterns to indicate directions or obstacles to be avoid to the driver (Petermeijer et al., 2017; Schwalk et al., 2015; Meng et al., 2015; Gray et al., 2014; Jensen et al., 2011; Ho et al., 2005). More generally, others use vibrotactile feedback to manage the allocation of user attention (Sklar and Sarter, 1999).

3 AUGMENTATION MODALITIES DESIGN FOR SPECIFIC ATM USE CASES

Following discussions with ATCos and integrating their experience into the design process, several potentially unsafe situations were identified which allowed us to imagine specific use cases in Single and Multiple RT contexts. For each of these use cases, HMI solutions have been developed to enable ATCos to respond better and/or more quickly to the potential problems generated. The main goal of the experiment reported in this paper was to investigate these contributions in terms of behavioral and subjective measurements for Single and Multiple RT. In this section, we developp these use cases, and introduce the augmentations modalities associated to, what they rely on in the literature and what we expected from them.

3.1 Interactive Spatial Sound to Improve Sound Sources Location

A common situation in ATC is the impossibility to see the aircraft. In case of heavy fog or simply during the night time, some airports still continue their activity. ATCos often look for visual contact with the planes they are controlling but this is no longer possible in such specific circumstances. The Audio Focus modality have been designed to overcome this issue by no longer relying on the sense of sight to locate aircraft, but on the sense of hearing. It relies on aircraft engine sounds (which cannot be heard normally in a physical tower). For this purpose, we proposed to add these spatial engine sounds coming from aircraft to the normal environment of the control tower. In this first use case, since the sense of sight is no longer available, the auditory one becomes the primary direct perception output modality. Savidis and his colleagues developped in the late 90's an environment which provided the user a hierarchical navigation dialogue based on 3D audio, 3D pointing and gestures (Savidis et al., 1996). The Audio Focus principle, which is based on the high correlation between head orientation and visual attention (Stiefelhagen et al., 2001), is quite similar to this one and is based on the increase of engine sounds that are located along the participant head sagittal axis.

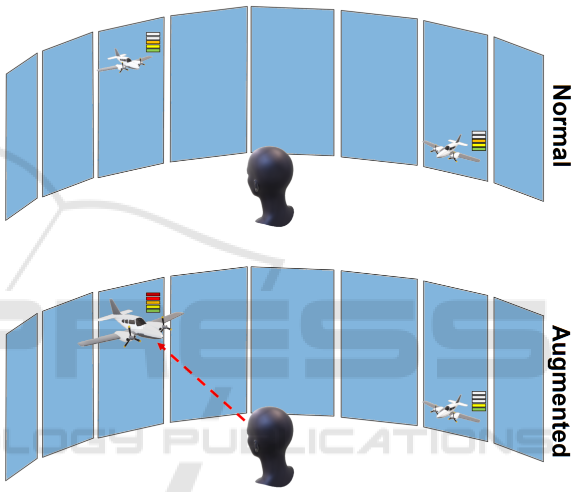

Figure 1: Audio Focus modality principle. When not activated (top), aircraft are associated to engine sounds but does not interact with user's movements. When activated (bottom), the interaction modality will enhance sound volume of aircraft which are aligned with the user's head.

Aircraft are linked to synthetic engine sounds that are spatialized. When the visibility is poor and does not allow the ATCo to see the aircraft, he or she can move the head to make these engines sounds volume varying. When a sound appears to be much louder than the other ones, an aircraft is in front of him (Figure 1). Also, the distance between the concerned aircraft and the user point of view is mapped into the gain of the sound sources (louder when the aircraft is close to the tower). This interaction form is designed to help the user to locate the aircraft in the airport vicinity when there is no visibility, instead of having an impossible visual contact with it.

This concept is also similar to the one exposed by (Bolt, 1981) in that it allows the user to play with sounds to have a better comprehension of her or him direct environment. We can tell that this concept belongs to the category of *enactive interfaces* because it allows the user to have a specific form of knowledge by moving her or him body (*act of doing*) (Loftin, 2003). This first modality was the subject of a full-fledged study. The first results were published in (Arico et al., 2018).

3.2 Spatial Sound Alert to Improve Abnormal Events Location

The second situation that have been identified is when pilots execute some actions without having previously been authorized by the ATCo to perform. Typically, it can happen that pilots start their engine or even start to move on the parking area without prior dialogue with the tower. The augmentation modality that have been designed to assist the ATCo during this type of event is called Spatial Alert. A spatial sound alert (typically a high frequency ”bip” sound) is raised when an unauthorized movement on ground is detected by the systems, along the azimuth of this event (Figure 2). The aim is to catch user’s attention toward the related event. The sound alert will stop by the time the user’s head is aligned with the event. This type of auditory displays have already been studied in the literrature (Simpson et al., 2005) and often provided better accuracy and response times in the location of events, especially in degraded visibility conditions (Simpson et al., 2005). Unauthorized movements are a common use case in ATC. Spatial Alert is here used to warn the ATCo that an abnormal situation, potentially dangerous, has been detected. For the purpose of this experiement, this situation can be an unauthorized movement on ground of an aircraft on the apron or a runway incursion.

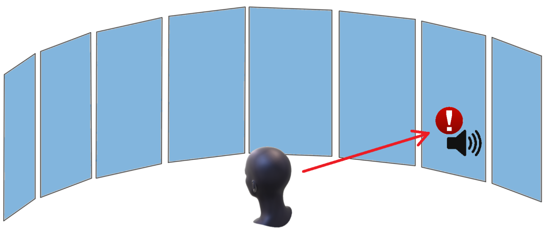

Figure 2: Spatial Sound alert modality principle. An event that requires ATCo attention occurs on an azimuth that is not currently monitored. A spatial sound alert is raised on this azimuth to attract the ATCo’s attention.

3.3 Runway Incursion Alert to Improve Unauthorized Movements on Ground

One of the most dangerous situations in airports is the runway incursion. This situation appears when an aircraft has crossed the holding point to go on the runway, while another one is ready to land (Figure 3). In the past, this situation as spawn several crashes including the deadliest in history (Weick, 1990), that is why the related augmentation modality could be very important and must be disruptive. Some airports are already equipped with systems aiming at informing the tower that such a situation is happening. The augmentation modality that is related to this type of use case is the Runway Incursion alert. During the experiment, ATCos were seated on a wooden chair which vibrates continuously when this type of event occurred. The ATCo had to acknowledge this event by clicking on a button located on the radar HMI in front of her.him, then the vibrations stopped. Spatial Sound alert is also coupled with this modality (a spatial sound alert is always raised in the direction of the holding point when a runway incursion situation is detected).

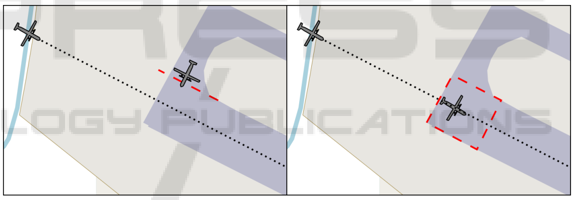

Figure 3: A schematic view on a runway incursion situation. The aircraft on the left of the two images is about to land in the next seconds. While this aircraft run through the final leg, another one crosses the line between taxi routes and runway.

3.4 Feedback to Distinguish Multiple Radio Calls

Multiple RT Tower concepts led to several problematic and one of the most recurrent is the radio frequency management. Airports commonly used different radio frequencies to communicate with aircraft. In the case of more than one airport controlled at the same time, the radio frequencies of each of the concerned airports will deliver radio messages simultaneously. This use case is potentially confusing for the ATCo, who should answer rapidly and precisely to pilots. To try to avoid this kind of confusion, the Call from Secondary Airport modality have been designed and tested during this experiment. Its principle is simply to add vibrations behind the back of the seat on which the ATCo is seated when messages from secondary airport are emitted. The vibrating patterns used for this feedback is different from the one used for Runway Incurion modality. First, this one is located behind the back of the seat, when Runway Incursion modality vibrations are located under the seat. Secondly, when the Runway Incursion vibrations are continuous, these ones are composed of a succession of up and down signals.

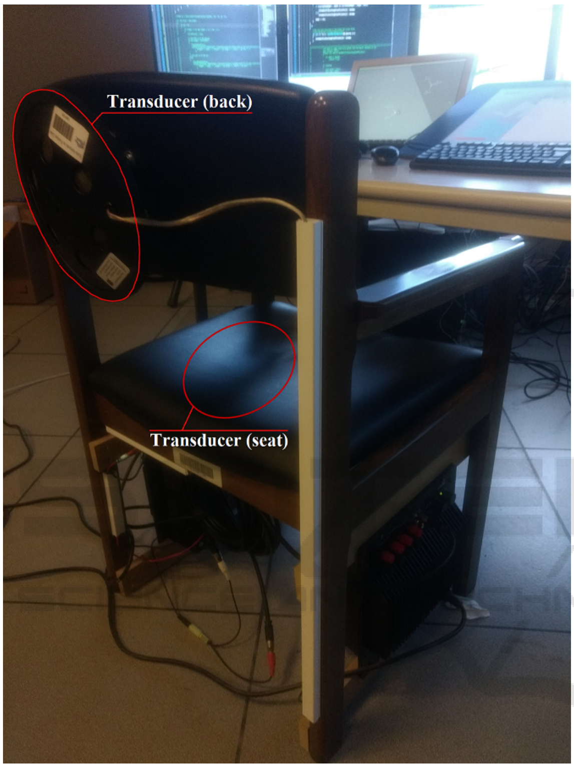

Figure 4: The wooden chair equipped with two transducers used for Runway Incursion and Call from Secondary Airport alerts.

4 MATERIALS AND METHOD

4.1 Participants

A total of sixteen French professional ATCos (8 females and 8 males, Mean age: 39.4 years; $SD = 7$) took part to the experiment. None of them had auditory problems. Their mean experience in years in control tower was 10 years ($SD = 6.8$). Each participant filled an informed consent after a detailed explanation of the study, which was conformed to the revised Declaration of Helsinki.

4.2 Nature of the Task

The participants were asked to seat on the haptic chair in front of a panoramic wall of screens. The setup was comparable to the one we can found in an operational RT facilities room. The participants could use radars to avoid being out of their elements. Another screen was used to have a minimalist view on the secondary airport, in the case of Multiple RT context. Two pseudo-pilots took place in another room to pilot the aircraft which were visible on the scene and to speak with the participant in aeronautical language, following the scenario scripts. These scripts had to be both plausible for the ATCos and comparable to each other to further be able to conduct statistical analyses. They were designed to get as close as possible to the actual working conditions in a real airport. Typically, participants had to deal with common events such as pilots asking to start their aircraft, order to reach the holding point, taking off, landings or circling the runway. They also had to manage abnormal situations such as unauthorized events on the parking area or runway incursions. All the scenarios have been written with the help of ATM expert and were conducted in poor visibility conditions where no aircraft were visible.

4.3 Experimental Procedure

The experimental protocol was designed to quantify the contribution in terms of subjective, behavioral and performance data of Spatial Alert, Runway Incursion and Call from Secondary Airport modalities. The first phase was a presentation to the participant of the experimental protocol and all the modalities that they had to use. The training exercise was firstly required for all the participants to make them familiar with the experimental platform and employed technologies. During this exercise, they executed a dedicated training scenario. They were asked to come before the experiment to get trained and be familiar with the platform facilities and each of the modalities. Afterwards, they were called to come later to follow the entire experiment protocol. The sequence of activities which were performed started with a welcome to the participant, a short reminder of the platform facilities and an introduction to the experiment itself. Then the core experiment began: for each of the 4 different scenarios, the participant were briefed, then they had to play the scenarios before finishing with a post-run questionnaire. The entire experiment lasted 2 hours per participant. Antoher questionnaire was given to them at the end of the experiment (post-experiment questionnaire).

4.4 Experimental Conditions

There were 4 different configurations: Single RT context without any augmentation, Single RT context with augmentations, Multiple RT context without augmentation, and Multiple RT context with augmentations. In order to decrease the learning effect, 4 different scripts were designed for a total of 8 scenarios.

Two experimental conditions were tested: *Context* [Single, Multiple] and *Augmented* [No, yes]. Single and Multiple scripts were made similar using comparable traffic complexity and events to raise. Fog was added to the visuals to make poor visibility conditions all along the experiment (Figure 5). The scenarios order has been randomized across the experimental conditions for all the participants. Hence, each participant had to come through 4 distinct scenarios in a different order.

Multiple Remote Tower aspects were managed using a screen in order to display a view on the secondary airport on the right of the ATCo position. These events were raised during all scenarios. Augmentations were activated only during augmented scenarios. There were 3 types of events related to each modality tested: *Spatial* event linked to Spatial Sound alert modality, *Runway Incursion* event linked to Runway Incursion alert modality, and *Call from Secondary Airport* event linked to Call from Secondary Airport modality. Audio Focus modality was always activated during the augmented remote tower condition.

4.5 Scenarios

To allow the measurement of behavioral and subjective values, we designed 8 different scenarios in total, called SRT 1 and 2, SART 1 and 2, MRT 1 and 2, and MART 1 and 2. These scenarios were developed considering 4 different scripts, designated by Single 1 for SRT 1 and SART 1 scenarios, Single 2 for SRT 2 and SART 2, Multiple 1 for MRT 1 and MART 1, and Multiple 2 for MRT 2 and MART2. Single and Multiple scenarios were different from each other but included equivalent operational events, with the goal to decrease potential learning effects. During the passes, 4 different scenarios were randomly presented to each participant, whose were divided into two groups: the first was composed of SRT 1, SART 2, MRT 1 and MART 2 scenarios, and the second was composed of SRT 2, SART 1, MRT 2 and MART 1 scenarios.

The 4 scripts were followed by the pseudo-pilots during the core experimental phase to create the different situations in which they could raise the events we wanted to test. During augmented scenarios, each event was linked to the appropriate augmentation modality. For the scenarios without augmentations, these events were raised in the same way than for augmented scenarios, to make the comparison between these two types of scenario feasible. The only difference was that for scenario without augmentations, the augmentation modalities were not activated.

4.6 Objectives and Hypothesis

The experimental platform was composed of hardware and software, which, in conjunction, provided realistic environment that was as much suitable as possible to immerse the ATCos in the context of their work as if they were in a real tower. The overall objective of the experiment was to promote user immersion and to increase performance while reducing workload. In this perspective, the working hypothesis can be formulated as following: "*User performances, in terms of reaction times and perceived workload, are improved when all the augmentation modalities are activated (i.e. Audio Focus interaction, Spatial Sound Alert, Runway Incursion alert, and Call from Secondary Airport feedback*".

4.7 Experimental Setup

The setup used for this experiment was quite substantial to get as close as possible to the real working conditions of a control tower. It was composed of the core position in which the participants took place to do the experiment, and two pseudo-pilot positions located in another room.

Experimental setup was composed of 8 UHD Iiyama Prolite X4071 screens for the panoramic display. Secondary airport was displayed with a 40 inches Eizo Flexscan screen using Flight Gear open flight simulator visuals. The ground radar view was displayed with a Wacom 24 inches high definition tablet, and the air radar view with a 19 inches Iiyama display. Radio communications were made using two Griffin PowerMate button used as push-to-talk actuators, and two microphones (one for each airport). Participants were seated on the haptic chair on which two Clark Synthesis T239 Gold tactile transducers have been attached (Fontana et al., 2016) for vibrotactile feedback (Figure 4). Spatial sound was relayed using Iiyama screen speakers to provide a physical spatial sound due to the physical positions of the eight speakers. User's head orientation has been retrieved using a Microsoft HoloLens mixed reality headset (visual augmentation facilities were not used here and the participants were asked to use it with the glasses raised upon their head). The 3D environment was

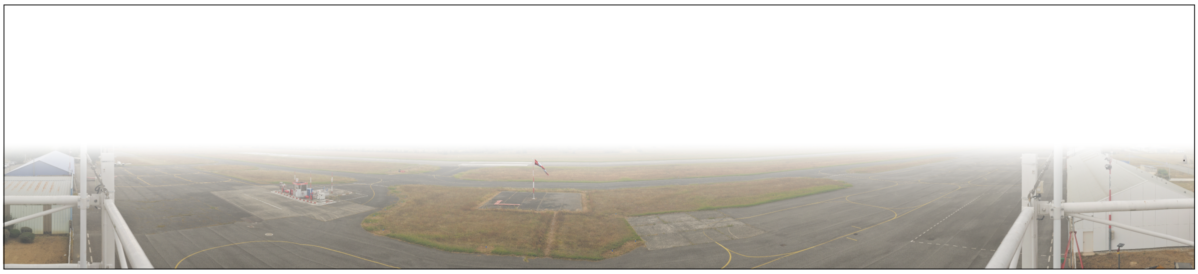

Figure 5: A screen capture of the visual conditions displayed during the experiment. The fog was set to completely cover the Muret airport runway and its environment. No aircraft were visible when they were not on the parking areas.

made using real photographs of Muret airport (south of France) mapped onto a 3D scene that was developed with Unity editor. The different software modules for the augmentations (Audio Focus, Runway Incursion vibrations, Call from Secondary Airport vibrations and Spatial Sound alert) were written in C# language using Microsoft .Net framework 4.6 and Direct Sound library. Network communications were developed using ENAC Ivy bus technologies which provides a high level mean of communication using string messages and a regular expression binding mechanism.

The only difference between the two pseudo-pilots positions was a supplementary screen for one of them in order to monitor the overall exercise. A position was composed of a Iiyama 40 inches Prolite X4071UHSU screen, a Wacom 24 inches tablet for the ground radar, them same push-to-talk buttons used for participants' position for radio communications, and a Corsair H2100 headset.

4.8 Metrics

4.8.1 Behavioral Measurements

Behavioral data were acquired by automatically computing response times during the experiment: since an event was raised, a timer was started until it was taken into account by the user. For the Spatial Sound alert, the timer was started when an aircraft moves on the ground without authorization, until the moment when the participants' head was aligned with the azimuth of the event. For the Runway Incursion alert, the timer was started when an aircraft crosses the holding point while another one was on the final leg, ready to land, until the moment when the participants presses the corresponding button on the radar HMI in front of him to tell that they had taken into account the alert. Finally, for the Call from Secondary Airport event, the timer was started when a message comes from the secondary airport, until the moment when the participants answered to this message by clicking on the secondary airport radio communication button end starting to speak to the pilot.

Since all the scripts contained at minimum one of each event, average values were computed from each type of event, for Single and Multiple contexts, and when the augmentations were activated or not. In total, there were 4 mean values per participant for Spatial Sound and Runway Incursion events: one for Single RT context without augmentations, one for Single RT context with augmentations, one for Multiple RT context without augmentations, and one for Multiple RT context with augmentations. Call from Secondary airport events could be tested only in Multiple RT context. Therefore, for this particular event, 2 average values were computed for each participant: one in Multiple RT context without augmentations, and another one in Multiple RT context with augmentations.

4.8.2 Subjective Measurements: Cognitive Workload and Questionnaires

After each scenario (i.e. run), participants were asked to fill a post-run questionnaire. The first part was the NASA-TLX (Hart and Staveland, 1988), which was used to estimate the cognitive workload. It uses six 100-points range subscales to assess mental workload (Mental , Physical and Temporal demands, Performance, Effort and Frustration). At the the end of the experiment the participants were asked to be part of a guided interview with a questionnaire. In this questionnaire they were asked to rate each of the augmented solutions they tested in terms of contribution for the Usefulness, Accuracy, Situation Awareness, Sense of presence and Cognitive workload. In the second part of the guided interview (supported by a custom made questionnaire) they were asked to rate the suitability of each solution in different operational contexts.

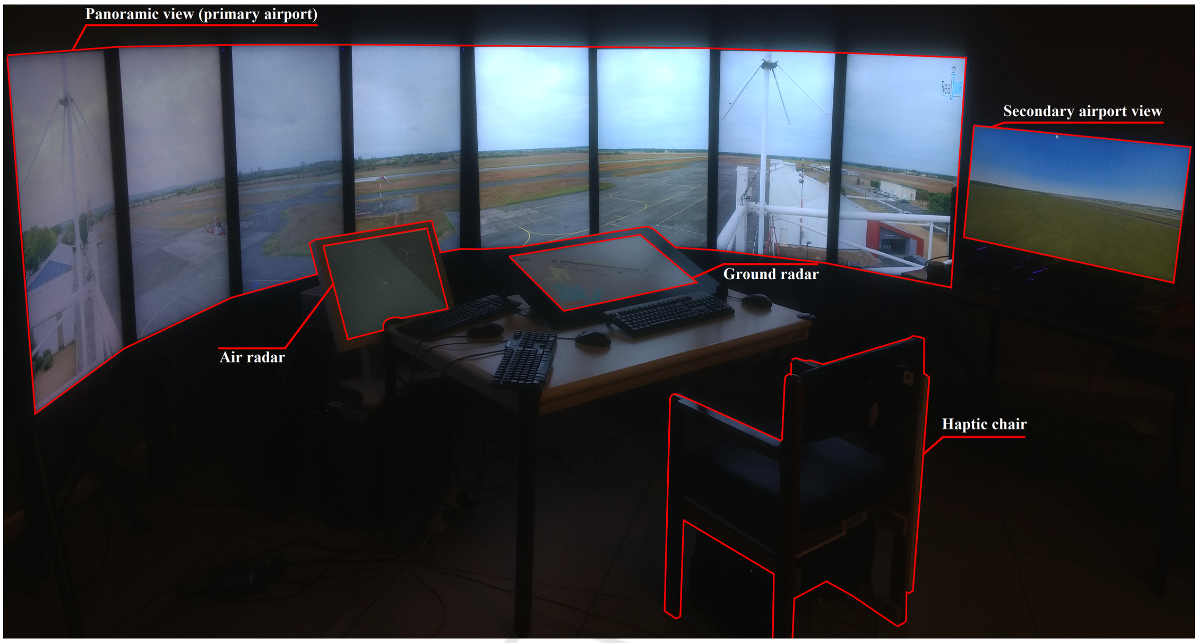

Figure 6: A photo of the setup used for the experiment. The main panaramic view displays Muret airport vicinity to the participants, whose were seated on the haptic chair. To be as ecological as possible compared to a real setup, they also had two radar views. The view on Lasbordes secondary airport was proposed through the screen on the righ.

4.8.3 Subject Matter Expert Ratings

Direct and non-intrusive evaluation was also carried out by a Subject Matter Expert (SME) during the experimental session. This SME was a professional ATCo who contributed to the experiment. A dedicated post-run questionnaire was designed to collect his insights about each participant. This SME questionnaire consisted rating the overall performance and workload during the run. For each point, the SME provided a rate from 0 (Very Low) to 10 (Very High). In addition, the SME filled simultaneously during the experiment the performance reached by the participants, by using a tablet with a simple software showing a slider that he could move from 0 to 10, in order to make a real-time rating.

5 RESULTS

5.1 Performance Results

Firstly, to analyze response times values related to Spaial Sound Alert modality, a two-way 2 × 2 ANOVAs (CI = .95) with repeated measures (Context [Single, Multiple] × Augmented [No, Yes]) was conducted and Tukey's HSD was used for post-hoc analysis. The Analysis revealed a main effect on Augmented factor and a Context × Augmented interaction.

No significant differences was found between Single and Multiple contexts [$F(1,14) = .85$; $p = .37$; $\eta_p^2 = .06$]. However, there was a significant difference between response times when the augmentations were activated or not [$F(1,14) = 7.27$; $p < .05$; $\eta_p^2 = .34$]. Tukey's post-hoc analysis revealed that participants were significantly faster to resolve the unauthorized movement on ground when the Spatial Sound Alert modality was activated ($M = 18455,54$; $SD = 4217,22$) than when it was not ($M = 41438,31$; $SD = 7733,1$). Finally, a significant Context × Augmented interaction was also found [$F(1,14) = 6.92$; $p < .05$; $\eta_p^2 = .33$]. Tukey's post-hoc analysis revealed that participants were significantly faster to resolve the unauthorized event when they were in Single Remote Tower configuration with the augmentations activated ($M = 9820,64$; $SD = 1276,44$) than without augmentations ($M = 58103,2$; $SD = 14813,18$). No significant result was found in Multiple Remote Tower configuration. Others two-ways 2 × 2 ANOVAs (Context [Single, Multiple] × Augmented [No, Yes]) were conducted over response times values for Runway Incursion and Call from Secondary Airport modalities. Regarding these two modalities, no significant differences have been highlighted between Context and Augmented factors.

After the behavioral measurements analyses, another analysis was conducted to confront (1) the perceived performance by each participant

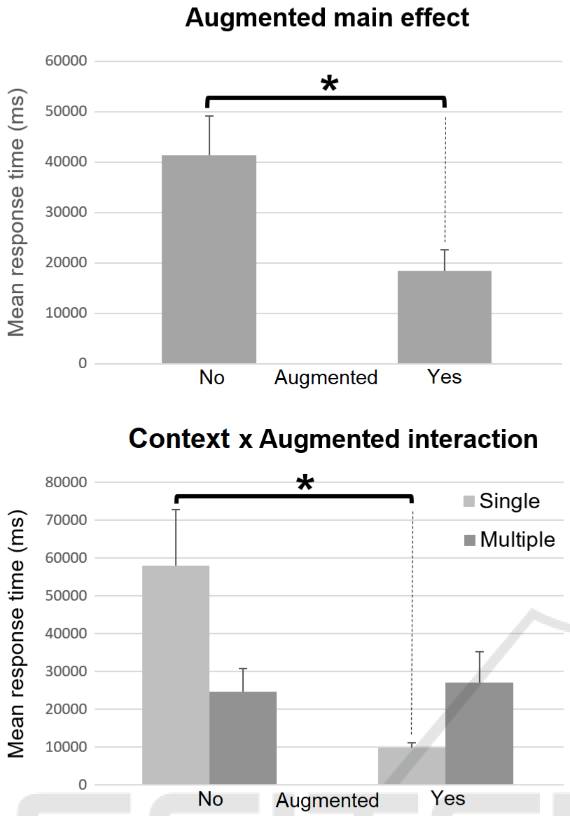

Figure 7: Results from inferential analysis on Response time behavioral data for Spatial Sound Alert modality only. Top diagram is Augmented main effect. Bottom diagram is for Context × Augmented interaction. Error bars are standard errors.

($Perf_{NASA-TLX}$), by using NASA-TLX factors, and (2) the performance rated by the SME, both during each scenario ($Perf_{SME}$) and after each scenario ($Perf_{SMEpost}$). For each of these variables a two-way ANOVAs 2 × 2 (CI = .95) with repeated measures (Context [Single, Multiple] × Augmented [No, Yes]) on the normalized performance values has been conducted using Arcsin transform (Wilson et al., 2013). In particular, the transformation consists in calculating the Arcsin of the square root of the related value between 0 and 1. From a subjective point of view, results did not show any significant effect across conditions and modalities [$F = .00034$; $p = .98$]. Also for the post-run performance rating given by the SME, the test did not highlight any significant trend among conditions and modalities [$F = 1.53$; $p = .02$]. On the contrary, the rating $Perf_{SME}$ showed a significant trend highlighting a decrease of performances when no augmentation modalities were activated. In details, the test showed a main effect [$F = 5.98$; $p < .05$], and in particular, the Duncan post-hoc test showed that for Multiple RT context, the augmentation modalities induced a significant decrement in performance ($p < .05$). No significant trends have been highlighted for the Single RT context.

To sum up, the results highlighted an advantage in using the proposed augmentation solutions to enhance operators performances (i.e. reducing response times), especially during Single RT context. Although participants did not report any difference in terms of perceived performance across the different experimental conditions, SME highlighted a significant decreasing in performance when augmentation modalities were activated during the Multiple RT context.

5.2 Perceived Workload Results

The perceived workload was evaluated considering both the NASA-TLX based workload scores (i.e. directly filled by the participants, $WL_{NASA-TLX}$), and post-run SME ratings (WL_{SME}). For both values, it was performed a two-way ANOVAs 2 × 2 (CI = .95) with repeated measures (Context [Single, Multiple] × Augmented [No, Yes]) on the normalized workload values, by using the Arcsin transformation (Wilson et al., 2013). ANOVA results did not show any significant main effect [$WL_{NASA-TLX}$: $F = .84$; $p = .037$; WL_{SME}: $F = .0003$; $p = .99$]. Anyhow, a significant effect has been highlighted for each of the factors. In other words, the workload was lower if the augmentation modalities were disabled, and higher during Multiple RT context with respect to Single one. In this regard, both subjective and SME scores showed the same significant trend ($p < .05$), except for the NASA-TLX assessment in which the comparison between Multiple and Single RT contexts showed a strong trend, but not significant ($p = .1$).

Therefore, cognitive workload measurements showed that (1) Multiple RT context induced higher workload with respect to the Single one, and (2) when activated, the augmentation modalities that were developed induced an overall higher workload perception with respect to when they are not activated. On the contrary, no significant differences have been highlighted during the other operational events.

5.3 Results from Post-experiment Interview

At the end of the experiment, each participant had to filled a post-experiment questionnaire in which they had to rate the 4 augmentation modalities regarding 5 different assertions on a scale going from 1 (Strongly Disagree) to 5 (Strongly Agree): *Perceived* *contribution to usefulness* ("The current augmentation modality is a useful aid for RT operations"), *Perceived contribution to accuracy* ("The current augmentation modality is accurate enough to support you during the RT operations"), *Perceived contribution to situation awareness improvement* ("The current augmentation modality improves your situation awareness in RT operations"), *Perceived contribution to the sense of immersion* ("The current augmentation modality improves your sense of immersion in RT operations"), and *Perceived contribution to workload reduction* ("The current augmentation modality does not have a negative impact on your workload in RT operations").

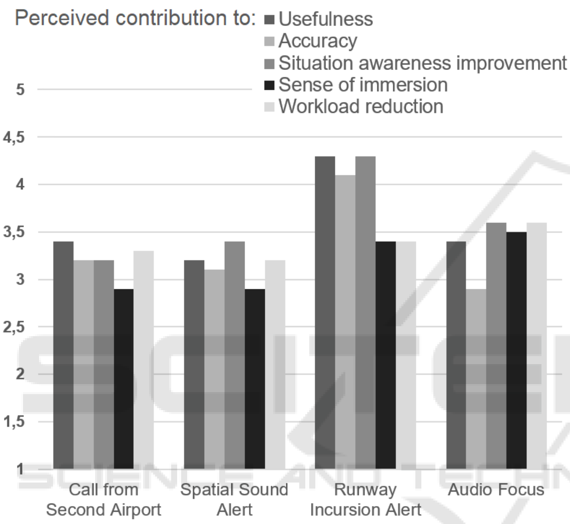

Figure 8: Results from post-experiment interview. The scale goes from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*).

The general feeling that emerges from the post-experiment results was that the augmentation modalities were considered more useful, accurate and to better support ATCo's situation awareness was the Runway Incursion alert (Figure 8). In general, the augmentation modalities were not perceived to contribute greatly to workload reduction and to immersion, but Audio Focus modality followed by Runway Incursion alert were the aumgnentation modalities that had the best scores for these two points.

6 DISCUSSION

The purpose of the present study was to investigate the impact of specific augmentation modalities in RT context, in terms of behavioral and subjective measurements. Multimodal interactions and feedback have been tested not only into Single RT context, but also in Multiple RT context. As the most important sense in ATC is the sight (Wickens et al., 1997), the augmentation modalities selected for this experiment were specifically focusing on both auditory and touch sensory channels, in order to prevent overloading or impairing the visual channel. This choice was supported by previous experimental activities (Arico et al., 2018).

The results of this experiment highlighted a relatively clear advantage for augmentation modalities when applied on specific operational events (especially for Spatial Sound alert), since the related response time values were significantly shortened (3 times less on average) once the augmentation modalities were activated. Anyhow, the perception of the overall performance across the whole conditions and modalities by both the participants and the SME sides was significantly lowered when the augmentation modalities were activated. The same trend was confirmed by the cognitive workload measurements, suggesting an increase of experienced workload when augmentation modalities were activated. Such surprising behavior could be interpreted as a need for a major familiarization of participants in using such solutions. Indeed, feedback with ATCos showed that audio and vibrotactile modalities are not, at least, consciously used in actual RT context. Nowadays, most control towers are sound proof, and vibrations are generally not felt by the controllers. As highlighted by the state of the art on remote towers, more care is taken on visual sensory channel than on audio and vibrotactile ones. Even though the use of these two sensory channels (i.e. hearing and touch) is seen as natural, the novelty of these interaction forms and the information provided need a longer appropriation time, implying a specific care in learning sessions. In other words, although ATCos have been well trained before the experiment to use in a proper way a setup that could be useful in specific operational situations (i.e. Runway Incursion and Spatial Sound alerts), it could also be too much distracting, inducing in some way a decreasing in their overall performances and increase in experienced workload.

This is consistent with the feedback from some participants on the post-experiment interview where they confirmed that vibrotactile feedback was a bit distracting due to the fact that they are not used to this type of distraction in current tower operations. Of course, it has to be stressed that this is one of the possible explanations on why the augmentation solutions increased the participants workload. Anyhow, we could assume that such increase in workload did not decreased the overall performances. Rather, the performance during the Spatial Sound alert was even enhanced by the proposed solutions.

7 CONCLUSION

Regarding these results, the assumption that the overall performance of the participants is improved when the augmentation modalities are activated is partially validated. In other words, the augmentation modalities did not induce any overload situation (i.e. excessive workload, and lower performance), but although the workload was higher in Multiple RT context, it did not negatively affect the performance. This study is a first step and other ones have to be performed in a more targeted way to understand the single contributions of the different multimodal augmentations in RT context.

FUNDINGS AND ACKNOWLEDGMENT

This work was co-financed by the European Commission by Horizon2020 project Sesar-06-2015 the embodied reMOte Tower, MOTO, GA n. 699379. We would like to thank all the air traffic controllers who took part in this experiment as well as all the people involved in the MOTO project.

REFERENCES

Arico, P., Reynal, M., Imbert, J., Hurter, C., Borghini, G., Flumeri, G. D., Sciaraffa, N., Florio, A. D., Terenzi, M., Ferreira, A., Pozzi, S., Betti, V., Marucci, M., Pavone, E., Telea, A. C., and Babiloni, F. (2018). Human-machine interaction assessment by neurophysiological measures: A study on professional air traffic controllers. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4619–4622.

Bolt, R. A. (1981). Gaze-orchestrated dynamic windows. *SIGGRAPH Comput. Graph.*, 15(3):109–119.

Braathen, S. (2011). Air transport services in remote regions. International Transport Forum Discussion Paper.

Calvo, J. (2009). Sesar solution regulatory overview – single airport remote tower. Technical report.

Cordeil, M., Dwyer, T., and Hurter, C. (2016). Immersive solutions for future air traffic control and management. In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, ISS Companion '16, pages 25–31, New York, NY, USA. ACM.

Erp, J. B. F. V., Veen, H. A. H. C. V., Jansen, C., and Dobbins, T. (2005). Waypoint navigation with a vibrotactile waist belt. *ACM Trans. Appl. Percept.*, 2(2):106–117.

Fontana, F., Camponogara, I., Cesari, P., Vallicella, M., and Ruzzenente, M. (2016). An exploration on whole-body and foot-based vibrotactile sensitivity to melodic consonance. *Proceedings of SMC*.

Fürstenau, N., Schmidt, M., Rudolph, M., Möhlenbrink, C., Papenfuß, A., and Kaltenhäuser, S. (2009). Steps towards the virtual tower: remote airport traffic control center (raice). *Reconstruction*, 1(2):14.

Gray, R., Ho, C., and Spence, C. (2014). A comparison of different informative vibrotactile forward collision warnings: does the warning need to be linked to the collision event? *PloS one*, 9(1):e87070.

Guldenschuh, M. and Sontacchi, A. (2009). Transaural stereo in a beamforming approach. In *Proc. DAFx*, volume 9, pages 1–6.

Hart, S. G. and Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In Hancock, P. A. and Meshkati, N., editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 139 – 183. North-Holland.

Ho, C., Tan, H. Z., and Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6):397 – 412.

Hurter, C., Lesbordes, R., Letondal, C., Vinot, J.-L., and Conversy, S. (2012). Strip'tic: Exploring augmented paper strips for air traffic controllers. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 225–232, New York, NY, USA. ACM.

Hutchins, E. L., Hollan, J. D., and Norman, D. A. (1985). Direct manipulation interfaces. *Hum.-Comput. Interact.*, 1(4):311–338.

Jensen, M. J., Tolbert, A. M., Wagner, J. R., Switzer, F. S., and Finn, J. W. (2011). A customizable automotive steering system with a haptic feedback control strategy for obstacle avoidance notification. *IEEE Transactions on Vehicular Technology*, 60(9):4208–4216.

Loftin, R. B. (2003). Multisensory perception: beyond the visual in visualization. *Computing in Science Engineering*, 5(4):56–58.

Meng, F., Gray, R., Ho, C., Ahtamad, M., and Spence, C. (2015). Dynamic vibrotactile signals for forward collision avoidance warning systems. *Human factors*, 57(2):329–346.

Moehlenbrink, C. and Papenfuss, A. (2011). Atc-monitoring when one controller operates two airports: Research for remote tower centres. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 55, pages 76–80. Sage Publications Sage CA: Los Angeles, CA.

Nene, V. (2016). Remote tower research in the united states. In *Virtual and Remote Control Tower*, pages 279–312. Springer.

Papenfuss, A. and Friedrich, M. (2016). Head up only – a design concept to enable multiple remote tower operations. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, pages 1–10.

Petermeijer, S., Cieler, S., and de Winter, J. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis and Prevention*, 99:218 – 227.

Pfeiffer, L., Valtin, G., Müller, N. H., and Rosenthal, P. (2015). The mental organization of air traffic and its implications to an emotion sensitive assistance system. *International Journal on Advances in Life Sciences*, 8:164–174.

Raj, A. K., Kass, S. J., and Perry, J. F. (2000). Vibrotactile displays for improving spatial awareness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 181–184. SAGE Publications Sage CA: Los Angeles, CA.

Savidis, A., Stephanidis, C., Korte, A., Crispien, K., and Fellbaum, K. (1996). A generic direct-manipulation 3d-auditory environment for hierarchical navigation in non-visual interaction. In *Proceedings of the Second Annual ACM Conference on Assistive Technologies*, Assets '96, pages 117–123, New York, NY, USA. ACM.

Schmidt, M., Rudolph, M., Werther, B., and Fürstenau, N. (2006). Remote airport tower operation with augmented vision video panorama hmi. In *2nd International Conference Research in Air Transportation*, pages 221–230. Citeseer.

Schwalk, M., Kalogerakis, N., and Maier, T. (2015). Driver support by a vibrotactile seat matrix recognition, adequacy and workload of tactile patterns in take-over scenarios during automated driving. *Procedia Manufacturing*, 3:2466 – 2473. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.

Simpson, B. D., Brungart, D. S., Gilkey, R. H., and McKinley, R. L. (2005). Spatial audio displays for improving safety and enhancing situation awareness in general aviation environments. Technical report, Wright State University, Department of Psychology, Dayton, OH.

Skaburskis, A. W., Shell, J. S., Vertegaal, R., and Dickie, C. (2003). Auramirror: Artistically visualizing attention. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pages 946–947, New York, NY, USA. ACM.

Sklar, A. E. and Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human factors*, 41(4):543–552.

Smith, D., Donald, M., Chen, D., Cheng, D., Sohn, C., Mamuji, A., Holman, D., and Vertegaal, R. (2005). Overhear: Augmenting attention in remote social gatherings through computer-mediated hearing. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1801–1804, New York, NY, USA. ACM.

Stiefelhagen, R., Yang, J., and Waibel, A. (2001). Estimating focus of attention based on gaze and sound. In *Proceedings of the 2001 workshop on Perceptive user interfaces*, pages 1–9. ACM.

Van Erp, J., Jansen, C., Dobbins, T., and Van Veen, H. (2004). Vibrotactile waypoint navigation at sea and in the air: two case studies. In *Proceedings of EuroHaptics*, pages 166–173.

Van Schaik, F., Roessingh, J., Lindqvist, G., and Fält, K. (2010). Assessment of visual cues by tower controllers, with implications for a remote tower control centre.

Vertegaal, R. (1999). The gaze groupware system: Mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 294–301, New York, NY, USA. ACM.

Vertegaal, R., Weevers, I., Sohn, C., and Cheung, C. (2003). Gaze-2: Conveying eye contact in group video conferencing using eye-controlled camera direction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 521–528, New York, NY, USA. ACM.

Weick, K. E. (1990). The vulnerable system: An analysis of the tenerife air disaster. *Journal of management*, 16(3):571–593.

Wickens, C. D., Mavor, A. S., McGee, J., Council, N. R., et al. (1997). Panel on human factors in air traffic control automation. *Flight to the future: human factors in air traffic control*.

Wilson, E., Underwood, M., Puckrin, O., Letto, K., Doyle, R., Caravan, H., Camus, S., and Bassett, K. (2013). The arcsine transformation: has the time come for retirement.