Towards adaptive AR interfaces for passengers of autonomous urban air mobility vehicles: analyzing the impact of fight phases and visibility conditions on user experience through

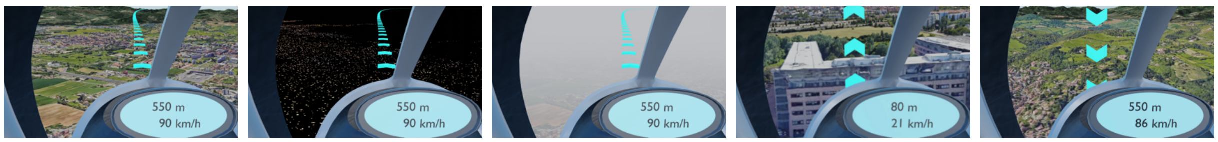
Towards Adaptive AR Interfaces for Passengers of Autonomous Urban Air Mobility Vehicles: Analyzing the Impact of Flight Phases and Visibility Conditions on User Experience Through Simulation

Lorenzo Valente*
Politecnico di Torino, Italy

Filippo Gabriele Pratticò†
Politecnico di Torino, Italy

Marco Nobile‡
Politecnico di Torino, Italy

Fabrizio Lamberti§
Politecnico di Torino, Italy



(a) cruise - clear daylight (b) cruise - night (c) cruise - foggy (d) takeoff - clear daylight (e) landing - clear daylight

Figure 1: Example of flight with one of the interface designs studied in this work (namely, Zip Line) over different flight phases and visibility conditions.

ABSTRACT

This work compares four possible designs of Augmented Reality (AR) interfaces for passengers of an Autonomous Aerial Vehicle (AAV) envisioned as an air taxi in the Urban Air Mobility (UAM) context. The four designs were evaluated and compared through a video-based study considering two potentially influential factors: flight phases (namely takeoff, cruise, and landing) and visibility conditions (i.e. clear daylight, night, and foggy). Dimensions included in the analysis were perceived safety, anxiety, situational awareness, cognitive workload, trust, predictability, and preference. The results showed that preferred interface by the passengers may vary depending on the considered combination of the two factors.

Keywords: advanced air mobility; autonomous transportation systems; augmented reality

Index Terms: Computing methodologies—Modeling and simulation—Simulation types and techniques—Interactive simulation; Computing methodologies—Computer graphics—Graphics systems and interfaces—Mixed / augmented reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

1 INTRODUCTION

The technological advancements in the fields of automation and aircraft manufacturing have been accelerating dramatically in recent years, pushing industries and academies to discover and redefine the concept of transportation. The field of Autonomous Aerial Vehicles (AAVs), in particular, is playing and will play a key role in this ongoing revolution. The aforementioned context is leading to the realization of Urban Air Mobility (UAM), aiming at the use of self-operating air vehicles in various fields, including emergency management, life-saving operations, weather monitoring, and passenger transportation [2]. The many opportunities and potentials that the UAM will be able to bring into people's everyday lives are already pushing the market itself to take an interest in it and find great investment [26], as testified by various companies like, e.g., Volocopter, which are committed to releasing Vertical Take-Off and Landing (VTOL) vehicles – VTOLSs – by 2030 [16]. Through a well-regulated use of these air vehicles, it will be possible to aim for the concept of sustainable mobility, leading to a drop in global city traffic congestion and a reduction in air pollution [23]. To this end, to reduce emissions from these aircraft is to target the use of electrical VTOLs [14].

Building on the general concept of Autonomous Systems (ASs), studies related to UAM are still few, whereas studies related to Autonomous Vehicles (AVs) have been multiplying over the years; particular attention has been paid, among others, to their interfaces, with the aim of improving public trust and acceptance [21]. The attitude of distrust towards ASs arises from several factors: first and foremost from poor, and in some cases absent, education in the field of ASs themselves, especially considering the field of self-operating vehicles, and from a misguided and misleading redistribution of information related to the latest innovations and advances in the field, often aimed at increasing the anxiety already inherent in the population [5]. It is plausible that many individuals may feel some anxiety about relying completely on an automated guidance system without the presence of a human driver or pilot at the helm. The absence of a physical leading person brings up doubts about safety, control, and the ability to handle unforeseen situations, generally decreasing passengers' trust [12]. Particularly in this field, it has been studied and shown that a lack of confidence in AAVs driving abilities causes a lack of confidence in UAM in general [1].

A further element to be considered to reduce people's anxiety and, consequently, increase confidence is to provide simulated experiences that can attract curiosity, society's interest, and conduct behavioral studies. Studies in the literature present various works, more or less interactive, aimed at familiarizing people with AAVs and, more generally, with UAM: it turns out to be of considerable importance how to design simulations of unmanned air vehicles to maximize public acceptance [30].

From this body of literature some guidelines and general trends for ASs have emerged, such as using Head Up Display (HUD) or Augmented Reality (AR) interfaces to visualize contextual information to the passenger. Indeed, it is well known how AR interfaces in AVs help to increase road safety and improve the passenger experience in general, managing to increase trust and acceptance, resulting in an expanding field of research [27].

Nevertheless, although the results obtained so far seem promising, it is not possible to take "as is" for good what has been done in these

*e-mail: lorenzo.valente@polito.it
†e-mail: filippogabriele.prattico@polito.it
‡e-mail: marco.nobile@studenti.polito.it
§e-mail: fabrizio.lamberti@polito.it

fields and apply it directly to AAVs because:

- The third dimension of motion, introduced by flight, presents a non-negligible factor for ordinary people, since they are not used to dealing with changes in height [4].
- Ordinary people are unfamiliar with transportation scenarios involving self-operating vehicles (especially air vehicles), and this causes them to perceive a higher risk [9].

Moreover, in the context of UAM, there is a paucity of studies that investigated AR interfaces. A few studies can be found with pilots in mind (e.g., to improve situational awareness in low visibility conditions), but the application to passengers is under-explored yet.

An early study that investigated the effect of trust of Human-Machine Interfaces (HMIs) for the passengers of AAV is represented by the pioneering work of Colley et al. [4]. The promising results produced by the study were obtained using a flight simulation in the cruising phase with a clear sky during daylight. Nevertheless, no variations were considered in flight conditions during the simulation: they have not investigated whether the trust augmented by HMIs would vary during more stressful flight conditions for passengers such as different phases of flight, e.g., vertical takeoffs or landings, or during low or reduced visibility conditions; hence, it is not possible to claim with certainty that the results obtained would remain the same under other conditions.

Subsequently, Meinhardt et al. [19] presented another work, again related to increasing passengers' acceptance through the use of a properly designed HMI; in this case, different visibility conditions (clear daylight, at night with only city lights, and foggy) are leveraged to stress passengers (and HMI) under worst-case circumstances that may be encountered during flight. Unfortunately, the work just presented the simulation system and the scenario enabling potential user studies, without performing an actual experimental evaluation. Moreover, the authors expected to analyze one HMI design across the different visibility conditions, without considering the flight phases factor.

In this paper, we attempt to overcome the aforementioned limitations by comparing, via an online video-based study, four different HMIs along three different phases of the flight (takeoff, cruise, and landing) and with different visibility conditions (clear daylight, at night with only city lights, and foggy).

2 Related Work

This paper builds on previous work related to the fields of AV interface design, AR interfaces used by aircraft pilots, and interfaces aimed exclusively at passengers in AAVs.

2.1 Interfaces for Increasing Trust in Autonomous Transportation Systems

Raats et al. [25] mentioned that the concept of trust has been identified as extremely relevant to the successful design of smart technologies such as AV, primarily questioning what the key points were in the HCI methodology on which the research itself is based. In this context, Morra et al. [21] proposed a methodology to validate the effectiveness of overlaid interfaces: the methodology consists of comparing two interfaces, based on AR-HUD, by varying the amount of information shown to the user. Through a Virtual Reality (VR) based driving simulation, Morra et al. were able to observe that the interface that shows only the most relevant information increases the sense of confidence and decreases cognitive load.

Remaining in the context of AVs, Sawitzky et al [35] studied how the use of AR User Interfaces (UIs) in a driving simulation via VR could change the levels of trust in passengers: namely, it was found that the interface that provided information about the status of the system via driving path visualization significantly increased trust in AVs.

In support of these arguments, Ruijten et al. [28] argued through their study that it is possible to increase the sense of trust by trying to raise the level of anthropomorphism of the interfaces used.

2.2 AR Interfaces for Pilots

Even before developing targeted AR interfaces, the idea of augmenting pilots' vision with additional information about, e.g., altitude, directly in the helmet was explored: this intention led to the realization of helmet-mounted displays, or head-mounted displays (HMDs). Over the years, these displays have developed more and more taking into account several construction characteristics, the most important of which have been weight and overall Field Of View (FOV). However, the focus of the early HMDs was limited to the projection of real-world symbols and images, without deriving additional information [18].

After increasing the visual information directly on the helmets, Ernst et al. [8] thought of further increasing the information available to the pilots through a system that could allow the complete replacement of the external view with a virtual screen: the images made available to the pilots were created through a combination of external sensors. Through this technology, the overall weight could be drastically reduced and the pilots' view, which was no longer limited by the helmet, could be increased.

The latter design uses an approach in which information is conveyed through the use of virtual cockpits: an alternative method involves the use of see-through cockpits onto which the additional information is projected, aiming to reduce pilots' heavy workload [11]. A step forward in this direction, proposed by Tran et al. [33], is to use AR glasses, specifically the Microsoft HoloLens 2 [20] HMD, to add artificial visual guidance to the crew, with the aim to enhance situation awareness and confidence in taking time-critical decisions without the support of a second crew member.

A further study aimed to decrease pilots' workload through the use of proper interfaces was performed by Katins et al. [13]: in their study, they designed a mock-up of Mixed Reality (MR) interfaces, using Microsoft HoloLens 2 [20], inside a full-sized flight simulator. The study additionally showed how the MR interfaces positively affected participants' situational awareness and overall landing routine efficiency.

2.3 Interface for Passengers of AAVs

The body of literature targeting UAM passengers is quite limited yet. An early study that sought to combine the research of HMI in the field of AVs to UAM, also focusing on the interaction between the passenger and the vehicle itself, is that of Otte et al. [22]: according to their results, the design of HMI should be as simple and minimalistic as possible, permitting both a technology-affiliated and an aversive passenger to be able to better understand the target to be reached. In addition, HMIs should always include general flight data such as altitude and speed to increase trust.

Edwards et al. [7], classified the major passengers' concerns and hypotheses for potential mitigations into six general categories: perceived safety, noise and vibration, availability and access, well-being, concern for the environment, and vehicle motion. Building on these hypotheses, Colley et al. [4] proposed seven interfaces to find the best-performing visualization in communicating path information to passengers by simulating the cockpit of an AAV: through a video-based online within-subject questionnaire, they relied on various specific measures, such as trust in automation, cognitive workload, perceived safety, situational awareness and a ranking for preference. After finding the best representation, they sought to understand the extent to which it helped passengers in increasing trust and confidence. As previously remarked, the results produced by their study were conducted only during the cruising phase, with a clear sky during daylight. The study did not analyze the potential impact on the interface preference of the flight phases, such as vertical takeoffs or landings, or of different visibility conditions, such as during low or reduced visibility.

Meinhardt et al. [19] proposed to use the best-performing interface derived from the study by Colley et al. [4] to analyze the variation of acceptance in passengers when visibility conditions change (clear daylight, at night with only city lights, and foggy) to stress passengers. However, as anticipated, the work was limited to only presenting the simulation system and scenario during various visibility conditions, without performing an actual experimental evaluation, and without considering the flight phases factor.

3 Materials and Methods

This section describes the realized flight simulation. It leverages the three-dimensional reconstruction of the cockpit of an AAV, which has been implemented using the Blender modeling and animation suite [10]. Specifically, Blender was used to render the forty-five videos of the simulated flight to study the impact of four variants (plus one baseline) of potential AR interface mock-ups. The videos depict all available combinations of experiences during the flight simulation, allowing the analysis of the interface behaviors during the three flight phases over the 3D reconstruction of the city of Bologna under three different visibility conditions.

3.1 Interfaces

Regarding the mock-up AR interfaces, it was decided to represent them as AR-HUDs, simulating their projection directly onto the glass windows that compose the cockpit of the AAV.

For this study, a neutral turquoise color was chosen for all the proposed interfaces, following the style already widely used for AVs-related studies [37] and adopted in [4] as well. The interfaces, together with the rationale for their choice, were as follows:

- *Arrow (A):* This interface (Fig. 2a) is inspired by the concept of “Gaze” [31] through the design and implementation of a three-dimensional arrow placed in front of the vehicle and has already been examined in the study by Colley et al [4]. Based on orientation and altitude change, the arrow changes accordingly by adjusting its orientation with 2 Degrees Of Freedom (DOF), indicating the next direction the AAV will take. The rationale for choosing to include this interface is primarily practicality in providing the direction information to the passenger, respecting the concept of simplicity and minimalism of the previously mentioned study [22].

- *Compass (C):* Building on the classical concept of a compass, this interface (Fig. 2b) involves the projection around the vehicle of a three-dimensional, spherical compass within which is a cross-shaped symbol that indicates the next direction (as for the Arrow interface). This interface is not present in the study by Colley et al. [4], but it was decided to introduce it ex novo because the compass is a conventional tool used in cockpits for pilots to help understand the three-dimensional motion of the vehicle and is expected to provide a reference frame for the passenger also in low visibility conditions.

- *Zip Line (ZL):* Based on the Chevron Zip Line [36] and found as the most promising in [4], this interface involves an orderly succession of two-dimensional arrows at the distance of three m, depicting the entire trajectory the vehicle will traverse, resembling the pattern of a roller coaster rail (Fig. 2c). According to previous results, it is expected that this representation will result in the least cognitive load for users and was also rated the best for trust in automation.

- *Tunnel (T):* The tunnel interface (Fig. 2d) involves the use of a reticulated tunnel that encapsulates the occupied airspace throughout the flight. Similarly to the previous interface, the rationale for including the tunnel is that it was found to be the second most popular interface in the study by Colley et al [4], and it allows the visualization of the direction combined with the size of the vehicle throughout the flight. Regarding the implementation details of the tunnel, it was decided to change the design from the original version, adopting a design more frequently found in the literature (for pilots) through the use of circular sections, departing from the previous rectangular choice, 11.5 m in size (slightly larger than the size of the Volocity’s rotor diameter [34]) and located at every 10 m to ensure proper visual fluidity. With such implementation, it is like an intermediate interface between the Compass and the Zip Line.

To respect the concept regarding the presence of general flight data from the previously mentioned study [22], it was decided to keep the AAV’s altitude and speed information just on the on-board display panel and do not show them on the AR-HUDs, for all variants of the interfaces.

To establish a weighted judgment against other interfaces, it was decided to use a modality with no interfaces (NI) as depicted in Fig. 2e.

3.2 Simulated Scenario and Flight Phases

To recreate the city of Bologna, it was chosen to rely on a plugin called *Blosm for Blender* [24]. Through the tool, leaning on APIs provided by Google, it is possible to choose a specific section of a map inside Google Maps and import it into a Blender scene.

Concerning the AAV, it was decided to use the Volocopter Volocity [34] as reference aircraft, as done also in [4]. Its flight features were used to define the path, and the 3D model of the interior was slightly adapted to suit the aim of the study (removing flight control sticks and instrumentation).

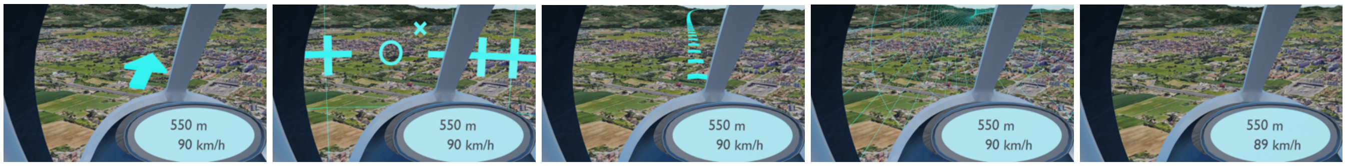
Regarding the flight experience, it was decided to consider the three flight phases separately, focusing on (and highlighting) the substantial differences. The takeoff phase (Fig. 1d) involves an altitude change of about 550 m, starting from a designated take-off vertiport at FICO Eataly World Center [6]. The average take-off speed was set at 45 km/h, thus establishing the total ascent time of about 40 seconds. For the cruise phase (Fig. 1a-c), a route (Fig. 3) of about five minutes was planned at an average speed of 90 km/h. The cruise path involves directional changes adopted from [4], including both an ascending/descending altitude change of about 100 m, and both a left and right directional change over Bologna. Regarding the landing phase (Fig. 1e), it was decided to symmetrically replicate the behavior chosen for the take-off phase: it starts from an altitude of about 550 m, descends at an average speed of 45 km/h, and reaches a second vertiport in Piazza Maggiore, Bologna. With the aim to stress the interfaces in the takeoff and landing an upright vertical motion was devised since considered the worst-case trajectory for these phases.

3.3 Flight Visibility Conditions

A key element of the study was the fruition of the overall experience at different levels of visibility. For this reason, three different scenarios were designed that were capable of providing different experiences characterized by important variations in visibility conditions. The three different flight conditions implemented were as follows:

- *Clear Daylight condition:* Distinguished by clear skies in daylight (Fig. 1a).

- *Night condition:* Distinguished by the darkness of night (no moonlight) with lights from the urban area as only reference (Fig. 1b).



(a) Arrow interface (b) Compass interface (c) Zip–Line interface (d) Tunnel interface (e) No interface.

Figure 2: Interfaces included in the study (framed during cruise phase).

Figure 3: Devised path of the AAV over the city of Bologna, adapted from [4].

- *Foggy condition:* Distinguished by poor visibility in daylight by simulating a fog with 15m of range of sight (Fig. 1c). The fog thickness was kept constant at all simulated altitudes to not introduce bias.

4 Experiment

This section presents the exploratory study conducted to compare the interface described under the different phases and conditions.

4.1 Experiment Design

The study was arranged by following a $3 \times 3 \times 5$ (flight phases, conditions, and interfaces respectively) within-subject factorial design. The study was arranged with an online questionnaire that embedded the videos of the simulated experience, which was distributed to participants involved by a network of contacts.

4.1.1 Sample

A Before-Experience Questionnaire (BEQ) was administered before starting the experiment, which included items about demographics, experience flying on commercial flight and helicopter (piloted by someone else), previous knowledge and expertise with technologies related to AR, individual proclivity to sense of presence when watching a video/movie/TV series, and individual propensity to trust autonomous transportation systems vehicles. According to data collected, the sample was made of 17 individuals aged between 22 and 34 ($\tilde{x}$ = 26.35 y.o., *s.d.* = 3.20 y.o.); 47% were females, 53% males. Out of them, 59% were moderately to very familiar with the use of AR interfaces, AR devices, or AR applications/games, and 41% were little to no familiar with it. Moreover, 82% of the sample reported a high proclivity to get immersed when watching a video/film/TV series. Only 12% have been on a helicopter at least once while 76% have been on a commercial flight (plane) during the last year. Regarding the propensity to trust AS, 53% think we should be careful with unfamiliar AS, only 29% think they could trust a system than they mistrust it, and 47% think AS generally works well. Finally, 47% would be willing to ride an air taxi if they had the opportunity.

4.1.2 Measures

Subjective feedback was collected through a questionnaire, which was administered after exposure to the flight phase and condition (AFCQ) that was structured as follows.

Situational Awareness. Was collected using the SART questionnaire [32] (one item, related to understanding flight simulation). According to the literature it is expected that higher levels of situational awareness lead to reduced anxiety in passengers and lower cognitive load.

Cognitive Workload. Was assessed using the single item scale from Paas [17]. According to previous studies lower cognitive workload levels are associated with higher trust in AS [29].

Perceived Safety. Was measured with two items from [3, 4] for perceived safety and anxiety subscales respectively.

Trust. Was investigated using the trust subscale of the trust in automation questionnaire (one item) [15].

Predictability. Was investigated using the predictability subscale of trust in automation questionnaire (one item) [15].

4.1.3 Protocol

The protocol of the experiment was structured as follows: for each flight phase, and for each visibility condition, the five videos (for the four interfaces plus the one with no interface) associated with the given phase and condition were shown (video duration was on average 30s); hereafter, the participant was administered with the corresponding AFCQ and had to indicate the ranking for the explicit preference of the interfaces (considering phase and condition). After the three conditions (i.e., once for each phase) the participants were requested to rank the interfaces for the given phase. At the end of all phases, a rank for overall preference was collected as well.

On average, a session lasted about 45 min. The order of the flight phases, visibility conditions, and videos were presented in randomized order to prevent exposure bias.

The questionnaire (including videos) used for the experimental protocol (but displaying a fixed order of phase and conditions) can be accessed at `https://forms.gle/qKK5Mky1wptqMUZdA`

4.2 Results

The statistical significance of the data collected was analyzed, with the RealStatistics tool (v7.3), using the three-way repeated-measures ANOVA test for the AFCQ data, with two-factor or one-factor repeated measure ANOVA as follow-up (with Bonferroni correction) when appropriate. The statistical significance of the results about the preference rank sampled after flight phase was analyzed using the two-way repeated-measures ANOVA test, with one factor repeated measure ANOVA as follow-up (with Bonferroni correction), and lastly, the one-way repeated-measures ANOVA test was used to calculate the statistical significance regarding the overall preference (with Bonferroni correction). Adopted Bonferroni corrected alpha were $\alpha_{corr} \leq 0.0167$ (for the flight phase and visibility condition factors) and $\alpha_{corr} \leq 0.0050$ (for interface factor). The normality assumption was verified with the D'Agostino-Pearson test.

Situational Awareness - Understanding Regarding situational awareness, particularly concerning flight situation understanding, statistical significance was found for interfaces (p-value $< .001$) whereas no significant differences were found neither for flight phases or for conditions. A significant interaction effect was found for visibility conditions on interfaces (p-value $= .003$) whereas no significant differences were found for the other interaction effects.

Regarding takeoff (Fig. 4a), ZL was significantly greater compared to NI for all visibility conditions, T was greater than NI during night and foggy while A was greater compared to NI during night. Finally, ZL and T were greater than C during night. Regarding cruise (Fig. 4b), T and ZL were higher compared to NI for all visibility conditions, and A was higher than NI during clear daylight. Ultimately, T and ZL were greater compared to C during clear daylight and night. Finally, for landing (Fig. 4c) T and ZL were greater than NI for all visibility conditions, A was greater than Ni for clear daylight and night, and C was greater than NI only at night. Lastly, ZL and T were greater than C for clear daylight and night.

Regarding the effects of visibility conditions on interfaces, in terms of understanding the situation, considering takeoff, it was observed a significant drop from night to foggy for ZL. During the cruise, A generally decreased from clear daylight to night and fog, and T and ZL decreased from clear daylight to fog. There were no significant effects during the landing.

Cognitive Workload Concerning cognitive workload, statistical significance was found for the interfaces (p-value $< .001$) whereas no significant differences were found neither for flight phases or for conditions. No significant interaction effects were found.

Regarding takeoff (Fig. 5a), C was significantly greater than ZL during night. Regarding cruise (Fig. 5b), C was greater than ZL for all visibility conditions while C was greater than T for clear daylight and night. Finally, for landing (Fig. 5c) C was greater than T and ZL during night.

Perceived Safety - Safety Concerning safety, statistical significance was found for the interfaces (p-value $< .001$) whereas no significant differences were found neither for flight phases or for conditions. A significant interaction effect was found for visibility conditions on interfaces (p-value $= .045$) whereas no significant differences were found for the other interactions. In the different flight phases, the interface that increased passengers' sense of safety the most was T, followed steadily by ZL. The general trend of safety, as visibility decreased, dropped for all interfaces and flight phases.

Regarding takeoff (Fig. 6a), T and ZL were significantly greater compared to NI for all visibility conditions and A was greater compared to NI during night. Concerning cruising (Fig. 6b), A, T, and ZL were superior to NI during all visibility conditions, C was greater than NI during foggy and T and ZL were superior to C and A during clear daylight and night. Finally, for landing (Fig. 6c), T and ZL were superior compared to NI for all visibility conditions, A was superior compared to NI for clear daylight and night, T was superior compared to C during clear daylight and night and ZL was superior to C at night.

Regarding the effects of visibility conditions on interfaces, during takeoff, NI decreased significantly from clear daylight to night and foggy. During the cruise, NI decreased from clear daylight to foggy, T decreased from clear daylight and night to foggy and ZL decreased from clear daylight to foggy. Finally, during landing, C and A decreased from clear daylight to foggy.

Perceived Safety - Anxiety Concerning anxiety, statistical significance was found for the flight phases (p-value $= .035$) and for interfaces (p-value $< .001$) whereas no significant difference was found for visibility conditions. No significant interaction effects were found.

Regarding the effects of flight phases on interfaces, concerning clear daylight (Fig. 7a), NI was significantly greater compared to A, T, and ZL for cruise and landing, C was greater than T for cruise and landing while C was greater than ZL for cruise only. Concerning night (Fig. 7b), NI was greater than T and ZL for all flight phases, NI was greater than A for cruise and landing, C was greater than T for cruise and landing while C was greater than ZL for cruise. Finally, for foggy (Fig. 7c), NI was greater than T and ZL for all flight phases, NI was greater than A for cruise and landing, NI was greater than C for cruise, C was greater than T and ZL for cruise and A was greater than ZL for cruise.

Regarding clear daylight (Fig. 7a), NI increased from takeoff to cruise and landing, T decreased from takeoff to cruise, and ZL increased from cruise to landing. Regarding night (Fig. 7b), NI increased from takeoff to cruise. Finally, regarding foggy (Fig. 7c), ZL decreased from takeoff to cruise but increased from cruise to landing.

Trust Concerning trust, statistical significance was found for the interfaces (p-value $< .001$) and for the visibility conditions (p-value $= .046$) whereas no significant differences were found for flight phases. No significant interaction effects were found.

Regarding takeoff (Fig. 8a), T and ZL were greater compared to NI for all visibility conditions and A was greater than NI at night. Concerning cruise (Fig. 8b), A, T, and ZL were greater than NI for all visibility conditions, C was greater than NI during foggy and T and ZL were greater than A and C during clear daylight and night. Finally, on landing (Fig. 8c) A, T, and ZL were greater than NI for all visibility conditions, T was greater than C during clear daylight and night and ZL was greater than C during night.

Regarding the effects of visibility conditions on interfaces, during takeoff, NI decreased from clear daylight to night and foggy. During cruise, NI and ZL decreased from clear daylight to foggy and T decreased from clear daylight and night to foggy. Finally, during landing, A and C decreased from clear daylight to night.

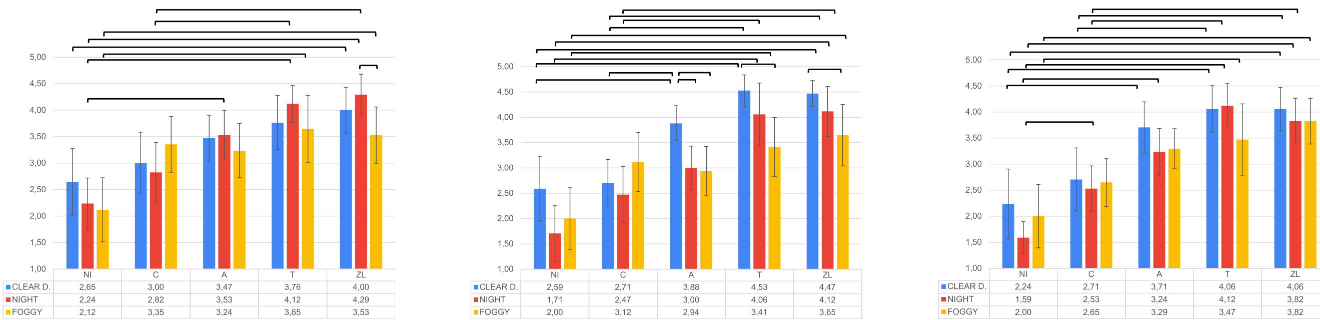
Predictability Concerning predictability, statistical significance was found for visibility conditions (p-value $= .035$) and for interfaces (p-value $< .001$) whereas no significant difference was found for flight phases. A significant interaction effect was found for flight phases on interfaces (p-value $= .045$) whereas no significant differences were found for the other interactions.

Regarding the effects of flight phases on interfaces, concerning clear daylight (Fig. 9a), all interfaces were greater than NI for all flight phases, T and ZL were greater than C for cruise and landing and T and ZL were greater than A for cruise. Concerning night (Fig. 9b), A, T and ZL were greater than NI for all flight phases, C was greater than NI for takeoff and cruise, T was greater than C for all flight phases, ZL was greater than C for takeoff and cruise and T and ZL were greater than A for cruise. Finally, for foggy (Fig. 9c), T and ZL were greater than NI for all flight phases, A was greater than NI for takeoff and cruise and C was greater than NI for takeoff.

Regarding clear daylight (Fig. 9a), NI and C decreased from takeoff to cruise, and T and ZL increased from takeoff to cruise and decreased from cruise to landing. Regarding night (Fig. 9b), NI and C decreased from takeoff to cruise and ZL decreased from cruise to landing. There were no significant effects regarding foggy.

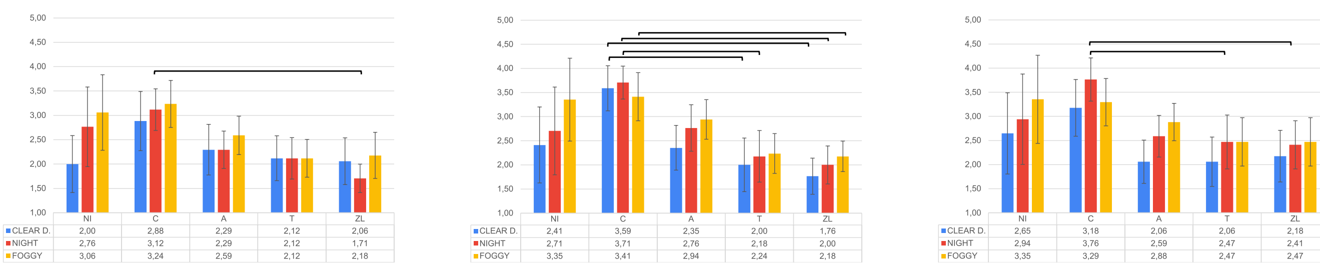
Preference (AFCQ) Regarding preference associated with a specific phase and a specific visibility condition, statistical significance (Table 1) was found for interfaces (p-value $< .001$) whereas no significant differences were found neither for flight phases nor for conditions. A significant interaction effect was found for visibility conditions on interfaces (p-value $< .001$), whereas no significant differences were found for the other interactions.

Regarding takeoff (Table 2), T was greater compared to NI concerning all visibility conditions, ZL was greater compared to NI for clear daylight and night, and T was greater than C during night.



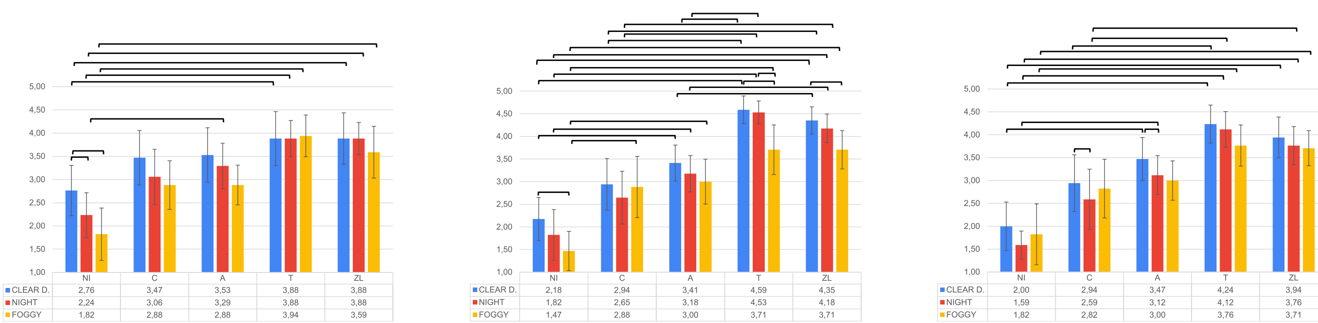
(a) Takeoff (b) Cruise (c) Landing

Figure 4: SART - Understanding.



(a) Takeoff (b) Cruise (c) Landing

Figure 5: Cognitive Workload (the lower the better).



(a) Takeoff (b) Cruise (c) Landing

Figure 6: Perceived Safety.

Regarding cruise (Table 2), T and ZL were greater compared to NI for all visibility conditions, A and C were greater compared to NI during night and T and ZL were greater compared to A and C during clear daylight and night. Finally, at landing (Table 2), T and ZL were greater compared to NI for all visibility conditions, A was greater than NI for clear daylight and night, C was greater compared to NI during night and T was greater than C during night.

Regarding the effects of visibility conditions on interfaces, during takeoff, C increased significantly from clear daylight and night to foggy. During cruise, NI decreased from clear daylight to night and C increased from clear daylight to night and foggy. Finally, during landing, T improved from clear daylight to night.

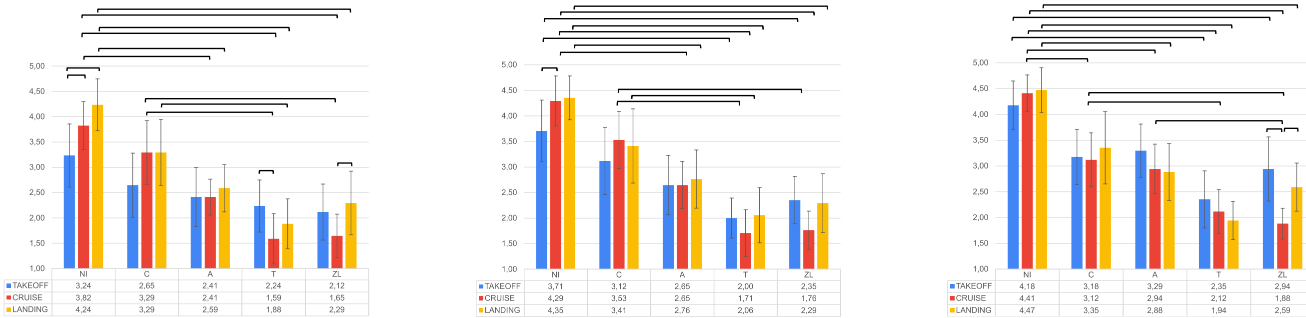
Ranking (After flight phases (FP) and overall) Regarding preference for different flight phases, statistical significance (Table 1) was found for interfaces (p-value $< .001$) but no significant differences were found for flight phases and for the interaction effect of flight phases on interfaces.

Regarding all flight phases (Table 3), A, T, and ZL were higher than NI and C and T and ZL were higher than A.

Regarding overall (Table 3), all interfaces were higher than NI, T and ZL were higher than C and T was higher than A.

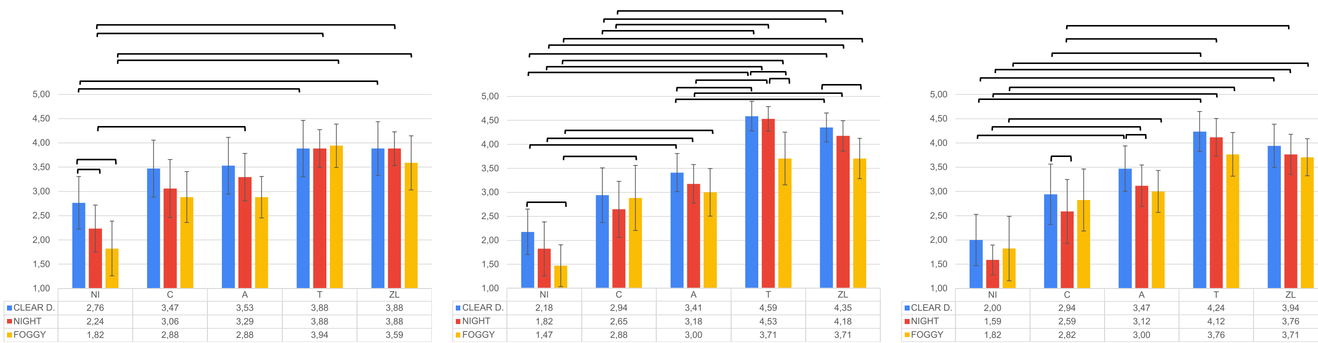
4.3 Discussion

Situational Awareness - Understanding Regarding situational awareness, particularly understanding of the flight situation, the results were influenced by the interfaces and the interaction effect for visibility conditions on the interfaces. Understanding was highest for T and ZL and lowest for NI. Understanding scores were lower in the vertical flight phases (i.e., takeoff and landing) than in cruise, and for A, T, and ZL, the scores were lowered by reducing visibility (i.e., higher during clear daylight and lower during foggy).



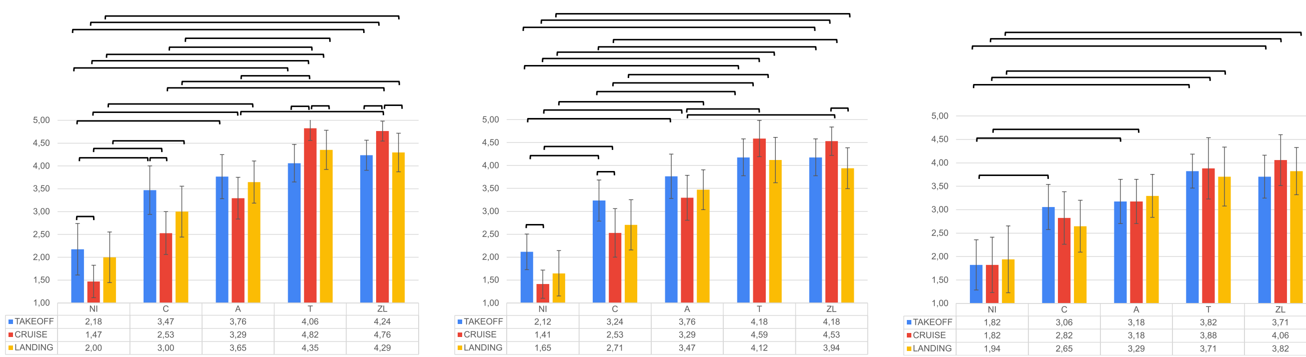
(a) Clear Daylight (b) Night (c) Foggy

Figure 7: Anxiety.



(a) Takeoff (b) Cruise (c) Landing

Figure 8: Trust.



(a) Clear Daylight (b) Night (c) Foggy

Figure 9: Predictability.

Preference Type	***p*-value**						
	FP	*VC*	*I*	*FP x VC*	*FP x I*	*VC x I*	*FP x VC x I*
AFCQ	.475	.498	**<.001**	.499	.129	**<.001**	.497
After FP	.070	//	**<.001**	//	.195	//	//
Overall	//	//	**<.001**	//	//	//	//

Table 1: ANOVA *p*-values and interactions of Flight Phases (FP), Visibility Conditions (VC), and Interfaces (I) for the three preference ranks.

Cognitive Workload Regarding cognitive workload, the results were influenced by the interfaces. For all the different flight phases, the interface that increased passenger cognitive workload the most was C. For NI, A, T, and ZL, cognitive workload scores rose when visibility became lower (i.e., lower during clear daylight and higher during foggy).

Perceived Safety Regarding safety, the results were influenced by the interfaces and by the interaction effect for visibility conditions on the interfaces. For all the different flight phases, the interface that increased passengers' sense of safety the most was T, followed steadily by ZL. Safety scores for NI, A, T, and ZL, were lowered

Flight Phase	Visibility Condition	Rank					*p*-value									
		NI	*C*	*A*	*T*	*ZL*	*NI-C*	*NI-A*	*NI-T*	*NI-ZL*	*C-A*	*C-T*	*C-ZL*	*A-T*	*A-ZL*	*T-ZL*
Takeoff	Clear Daylight	5	4	3	1	1	.493	.118	**.004**	**.001**	.252	.017	.006	.330	.215	.497
	Night	5	4	3	1	2	.385	.085	**<.001**	**.001**	.418	**.001**	.030	.022	.225	.385
	Foggy	5	2	2	1	2	.065	.171	**<.001**	.011	.493	.105	.475	.036	.370	.303
Cruise	Clear Daylight	5	4	3	1	2	.461	.016	**<.001**	**<.001**	.107	**<.001**	**<.001**	**<.001**	**<.001**	.499
	Night	5	4	3	1	2	**.005**	**<.001**	**<.001**	**<.001**	.273	**<.001**	**<.001**	**<.001**	**<.001**	.497
	Foggy	5	3	3	1	1	.018	.062	**.001**	**.003**	.493	.355	.197	.197	.081	.493
Landing	Clear Daylight	5	4	3	1	1	.056	**<.001**	**<.001**	**<.001**	.234	.009	.009	.272	.272	.500
	Night	5	4	3	1	1	**.005**	**<.001**	**<.001**	**<.001**	.271	**<.001**	.029	.020	.371	.229
	Foggy	5	4	3	1	1	.016	.008	**<.001**	**<.001**	.499	.123	.419	.188	.468	.419

Table 2: Ranking of preference associated with a specific phase and a specific visibility condition (AFCQ).

Preference Type	Rank					*p*-value									
	NI	*C*	*A*	*T*	*ZL*	*NI-C*	*NI-A*	*NI-T*	*NI-ZL*	*C-A*	*C-T*	*C-ZL*	*A-T*	*A-ZL*	*T-ZL*
After FP	5	4	3	1	2	.017	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.050
Overall	5	4	3	1	2	**<.001**	**<.001**	**<.001**	**<.001**	.389	**<.001**	**<.001**	**<.001**	.007	.212

Table 3: Ranking of preference for after FP (reported as grouped since no sig. difference was found for FP factor) and for overall.

by reducing visibility (i.e., higher during clear daylight and lower during foggy).

Regarding anxiety, the results were influenced by the flight phases and by the interfaces. For all the different flight phases, NI increased passengers' sense of anxiety the most, followed steadily by C. Anxiety scores for NI were higher by reducing visibility (i.e., higher during clear daylight and lower during foggy) and were higher in the vertical flight phases (i.e., takeoff and landing).

Trust Regarding trust, the results were influenced by the visibility conditions and by the interfaces. For all the different flight phases, the interface that increased the trust the most in passengers was T, followed consistently by ZL. Trust scores for A, T, and ZL, were lowered by reducing visibility (i.e., higher during clear daylight and lower during foggy).

Predictability Regarding predictability, the results were influenced by the visibility conditions, by the interfaces, and by the interaction effect for flight phases on the interfaces. For all the different flight phases, the interface that increased the path's predictability the most was T, followed steadily by ZL. Predictability scores for T and ZL decreased in the vertical flight phases (i.e., takeoff and landing) compared to cruise.

Preference (AFCQ) Regarding preference associated with a specific phase and a specific visibility condition, the results were influenced by the interfaces and by the interaction effect for visibility conditions on the interfaces. As flight phases and visibility conditions varied, the trends remained similar: T and ZL were found to be the most appreciated interfaces while NI was found to be the least valued.

Ranking (After flight phases (FP) and overall) As flight phases varied, T was found to be the most appreciated interface, followed by ZL, A, C, and NI. Similar trends were confirmed when expressing the overall preference, albeit in this regard T and ZL were found as comparable as well as w.r.t C vs. A.

5 Conclusions

In this paper, we aimed to investigate through the creation of a not interactive flight simulation the impact of using AR interfaces for passengers of AAVs as flight phases and visibility conditions change. To this aim, four different AR interfaces were compared in an online video-based study with 17 participants. A general trend that has been observed is that there has been an impact of the flight phases and visibility conditions which affected the feeling of safety and trust in passengers. Specifically, during vertical flight phases (i.e., takeoff and landing) was observed a decrease in understanding and increased anxiety whereas diminished visibility conditions deteriorated understanding, safety, and predictability while at the same time increasing cognitive workload and anxiety. The Tunnel and the Zip Line interfaces were rated as the best overall and were found in the majority of the cases as able to increase understanding of the flight situation, trust, and perception of safety in contrast to the experience with no interface. It should be noted that it was not always possible to identify one particular interface as significantly better than the others as flight phases or visibility conditions changed. This situation may lead to the study of mixed or adaptive interfaces, capable of changing or being replaced based on current flight and visibility conditions.

In general, the devised study allowed further investigation of the domain considered by Colley et al. [4], through additional analyses for flight phases (vertical takeoff and vertical landing) and visibility conditions (night and foggy). Moreover, the study allowed further investigation of the scenario addressed by Meinhardt et al. [19], through experimental analysis of visibility conditions changes, and more broadly, to expand the body of literature on the topic of UAM and AAVs.

The online video-based study carried out is limited in terms of external and ecological validity. In the future, a more realistic simulation through an interactive immersive VR experience will be implemented, which would allow us to evaluate the most promising HMIs mock-ups with higher validity. Moreover, another envisaged future work pertains to the analysis of different AR visualizations of air traffic and its impact on the user experience aspects of the passengers.

Acknowledgments

This study was carried out within the project MOST – Sustainable Mobility National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This paper reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

References

[1] E. T. Chancey and M. S. Politowicz. Public trust and acceptance for concepts of remotely operated urban air mobility transportation. In

Proc. of the Human Factors and Ergonomics Society Annual Meeting*, vol. 64, pp. 1044–1048. SAGE Publications Sage CA, 2020. doi: 10.1177/1071181320641251
[2] A. P. Cohen, S. A. Shaheen, and E. M. Farrar. Urban air mobility: History, ecosystem, market potential, and challenges. *IEEE Trans. on Intelligent Transportation Systems*, 22(9):6074–6087, 2021. doi: 10.1109/TITS.2021.3082767
[3] M. Colley, S. Krauss, M. Lanzer, and E. Rukzio. How should automated vehicles communicate critical situations? a comparative analysis of visualization concepts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 5(3), 2021. doi: 10.1145/3478111
[4] M. Colley, L.-M. Meinhardt, A. Fassbender, M. Rietzler, and E. Rukzio. Come fly with me: Investigating the effects of path visualizations in automated urban air mobility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 7(2), jun 2023. doi: 10.1145/3596249
[5] C. Dwork and M. Minow. Distrust of artificial intelligence: Sources & responses from computer science & law. *Daedalus, MIT Press*, 151(2):309–321, 2022. doi: 10.1162/daed_a_01918
[6] Eataly. Fico eataly world. `https://www.fico.it/it`. Accessed: 2023-12-26.
[7] T. Edwards and G. Price. eVTOL passenger acceptance. Technical report, NASA, ARC-E-DAA-TN76992, 2020.
[8] J. M. Ernst, H.-U. Doehler, and S. Schmerwitz. A concept for a virtual flight deck shown on an HMD. In J. Sanders-Reed and J. J. A. III, eds., *Degraded Visual Environments: Enhanced, Synthetic, and External Vision Solutions 2016*, vol. 9839, p. 983909. Int. Society for Optics and Photonics, SPIE, 2016. doi: 10.1117/12.2224933
[9] D. S. Ferrel, T. Carney, and S. R. Winter. Risk perception analysis of a small aircraft transportation system. *Jrnl. of Aviation Technology and Engineering*, 1(1):12, 2011. doi: 10.5703/1288284314634
[10] B. Foundation. Blender. `https://www.blender.org/`. Accessed: 2023-12-26.
[11] S. G. Hart. 18 - helicopter human factors. In E. L. Wiener and D. C. Nagel, eds., *Human Factors in Aviation*, Cognition and Perception, pp. 591–638. Academic Press, San Diego, 1988. doi: 10.1016/B978-0-08-057090-7.50024-2
[12] R. Häuslschmid, M. von Bülow, B. Pfleging, and A. Butz. Supporting trust in autonomous driving. In *Proc. of the 22nd Int. Conf. on Intelligent User Interfaces*, IUI, p. 319–329. ACM, 2017. doi: 10.1145/3025171.3025198
[13] C. Katins, S. S. Feger, and T. Kosch. Exploring mixed reality in general aviation to support pilot workload. In *Extended Abstracts of the 2023 CHI Conf. on Human Factors in Computing Systems*, CHI EA. ACM, 2023. doi: 10.1145/3544549.3585742
[14] I. C. Kleinbekman, M. A. Mitici, and P. Wei. eVTOL arrival sequencing and scheduling for on-demand urban air mobility. In *2018 IEEE/AIAA 37th Digital Avionics Systems Conf. (DASC)*, pp. 1–7, 2018. doi: 10.1109/DASC.2018.8569645
[15] M. Körber. Theoretical considerations and development of a questionnaire to measure trust in automation. In *Proc. of the 20th Congress of the Int. Ergonomics Association (IEA 2018) Volume VI: Transport Ergonomics and Human Factors (TEHF), Aerospace Human Factors and Ergonomics 20*, pp. 13–30. Springer, 2019. doi: 10.1007/978-3-319-96074-6_2
[16] D. Lenton. The measure of volocopter flying taxi. *Engineering & Technology*, 13(7/8):10–11, 2018. doi: 10.1049/et.2018.0712
[17] J. Leppink, F. Paas, C. P. Van der Vleuten, T. Van Gog, and J. J. Van Merriënboer. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45:1058–1072, 2013. doi: 10.3758/s13428-013-0334-1
[18] H. Li, X. Zhang, G. Shi, H. Qu, Y. Wu, and J. Zhang. Review and analysis of avionic helmet-mounted displays. *Optical Engineering*, 52(11):110901, 2013. doi: 10.1117/1.OE.52.11.110901
[19] L.-M. Meinhardt, M. Colley, A. Faßbender, and E. Rukzio. Stairway to heaven: A demonstration of different trajectories and weather conditions in automated urban air mobility. In *Adjunct Proc. of the 15th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '23, p. 311–313. ACM, 2023. doi: 10.1145/3581961.3610372
[20] Microsoft. Microsoft Hololens 2. `https://www.microsoft.com/it-it/hololens/hardware`. Accessed: 2024-01-06.
[21] L. Morra, F. Lamberti, F. G. Pratticò, S. La Rosa, and P. Montuschi. Building trust in autonomous vehicles: Role of virtual reality driving simulators in HMI design. *IEEE Trans. on Vehicular Technology*, 68(10):9438–9450, 2019. doi: 10.1109/TVT.2019.2933601
[22] T. Otte, N. Metzner, J. Lipp, M. S. Schwienhorst, A. F. Solvay, and T. Meisen. User-centered integration of automated air mobility into urban transportation networks. In *2018 IEEE/AIAA 37th Digital Avionics Systems Conf. (DASC)*, pp. 1–10, 2018. doi: 10.1109/DASC.2018.8569820
[23] M. N. Postorino and G. M. L. Sarné. Reinventing mobility paradigms: Flying car scenarios and challenges for urban mobility. *Sustainability*, 12(9), 2020. doi: 10.3390/su12093581
[24] Prochitecture. Blosm for blender. `https://prochitecture.gumroad.com/l/blosm?layout=profile`. Accessed: 2023-12-26.
[25] K. Raats, V. Fors, and S. Pink. Trusting autonomous vehicles: An interdisciplinary approach. *Transportation Research Interdisciplinary Perspectives*, 7:100201, 2020. doi: 10.1016/j.trip.2020.100201
[26] C. Reiche, R. Goyal, A. Cohen, J. Serrao, S. Kimmel, C. Fernando, and S. Shaheen. Urban air mobility market study. Technical report, NASA, HQ-E-DAA-TN70296, 2018.
[27] A. Riegler, A. Riener, and C. Holzmann. A research agenda for mixed reality in automated vehicles. In *Proc. of the 19th Int. Conf. on Mobile and Ubiquitous Multimedia*, MUM, p. 119–131. ACM, 2020. doi: 10.1145/3428361.3428390
[28] P. A. M. Ruijten, J. M. B. Terken, and S. N. Chandramouli. Enhancing trust in autonomous vehicles through intelligent user interfaces that mimic human behavior. *Multimodal Technologies and Interaction*, 2(4), 2018. doi: 10.3390/mti2040062
[29] K. Samson and P. Kostyszyn. Effects of cognitive load on trusting behavior – An experiment using the trust game. *PLOS ONE*, 10(5):1–10, 05 2015. doi: 10.1371/journal.pone.0127680
[30] A. Straubinger, R. Rothfeld, M. Shamiyeh, K.-D. Büchter, J. Kaiser, and K. O. Plötner. An overview of current research and developments in urban air mobility – Setting the scene for UAM introduction. *Jrnl. of Air Transport Management*, 87:101852, 2020. doi: 10.1016/j.jairtraman.2020.101852
[31] D. Szafir, B. Mutlu, and T. Fong. Communicating directionality in flying robots. In *Proc. of the Tenth Annual ACM/IEEE Int. Conf. on Human-Robot Interaction*, HRI, p. 19–26. ACM, 2015. doi: 10.1145/2696454.2696475
[32] R. M. Taylor. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness*, pp. 111–128. Routledge, 2017. doi: 10.4324/9781315087924-8
[33] T. H. Tran, F. Behrend, N. Fünning, and A. Arango. Single pilot operations with AR-glasses using microsoft hololens. In *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, pp. 1–7. IEEE, 2018. doi: 10.1109/DASC.2018.8569261
[34] Volocopter. Volocity specs. `https://assets.ctfassets.net/vnrac6vfvrab/73kYdf0o0kR7Y8XqAz9rEl/40bcf5c38552f6d1fcca71f7fe9736f3/20220607_VoloCity_Specs.pdf`. Accessed: 2023-12-26.
[35] T. von Sawitzky, P. Wintersberger, A. Riener, and J. L. Gabbard. Increasing trust in fully automated driving: Route indication on an augmented reality head-up display. In *Proc. of the 8th ACM Int. Symposium on Pervasive Displays*, PerDis '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3321335.3324947
[36] T. von Sawitzky, P. Wintersberger, A. Riener, and J. L. Gabbard. Increasing trust in fully automated driving: Route indication on an augmented reality head-up display. In *Proc. of the 8th ACM Int. Symp. on Pervasive Displays*, PerDis. ACM, 2019. doi: 10.1145/3321335.3324947
[37] A. Werner. New colours for autonomous driving: An evaluation of chromaticities for the external lighting equipment of autonomous vehicles. *Colour Turn*, (1), 2018.