Wind of Change: Investigating Information Visualizations for Passengers and Residents' Perception of Automated Urban Air Mobility

LUCA-MAXIM MEINHARDT*, Institute of Media Informatics, Ulm University, Germany
MARK COLLEY*, Institute of Media Informatics, Ulm University, Germany and Cornell Tech, U.S.
MARJANA TAHMID, Institute of Media Informatics, Ulm University, Germany
MAX RÄDLER, Institute of Media Informatics, Ulm University, Germany
ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

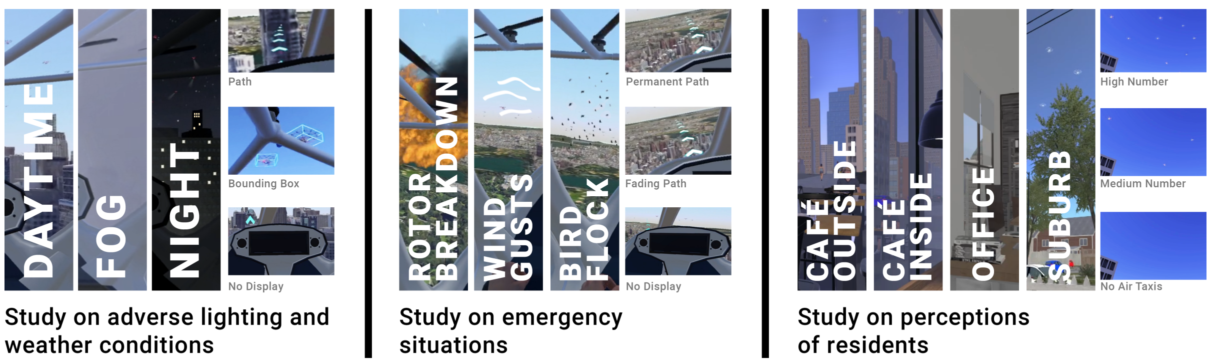


Fig. 1. Overview of the three different scenarios for UAM divided into three online studies. The first study explores passengers' perceptions of using UAM in adverse lighting and weather conditions. The second study investigates passengers' perceptions during emergency situations. The last study investigates residents' perceptions of UAM in the city and suburbs.

Automated Urban Air Mobility (UAM) is expected to improve passenger transportation but raises concerns about passenger trust and its impact on residents. We address these issues through three online studies using 360-degree videos. The study on adverse lighting and weather conditions (N=31) shows that bounding box visualizations of other air taxis positively impact passengers' perceived safety. Especially during fog, passengers' perceived safety is reduced when no additional information besides the path is shown. The second study (N=29) extends these findings to emergency scenarios (wind gusts, bird flocks, and rotor breakdown) and shows that visualizations of the alternate path increase perceived safety and trust. Finally, studying the impact of UAM on residents (N=29), we found significant concerns about visual pollution and privacy, especially in the

*Both authors contributed equally to this research.

Authors' Contact Information: Luca-Maxim Meinhardt, luca.meinhardt@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Mark Colley, mark.colley@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany and Cornell Tech, New York, U.S.; Marjana Tahmid, marjana.tahmid@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Max Rädler, max.raedler@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany; Enrico Rukzio, enrico.rukzio@uni-ulm.de, Institute of Media Informatics, Ulm University, Ulm, Germany.

ACM 2474-9567/2024/12-ART171
https://doi.org/10.1145/3699753

suburbs but also in offices. Our findings provide actionable insights into UAM perception and societal acceptability for future aerial transportation design and infrastructure.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; *Empirical studies in visualization.*

Additional Key Words and Phrases: urban air mobility, online study, mobility

ACM Reference Format:
Luca-Maxim Meinhardt, Mark Colley, Marjana Tahmid, Max Rädler, and Enrico Rukzio. 2024. Wind of Change: Investigating Information Visualizations for Passengers and Residents' Perception of Automated Urban Air Mobility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 4, Article 171 (December 2024), 31 pages. https://doi.org/10.1145/3699753

1 INTRODUCTION

Automated Urban Air Mobility (UAM) emerges as one possibility to reduce high congestion levels. For example, individuals in London, Paris, and Brussels experience over 130 hours of delays annually [51]. Considering the potential of UAM, the European Union Aviation Safety Agency (EASA) projects that this novel mode of transportation could be available within this decade [2]. In line with these predictions, pioneering startups like Lilium, Volocopter, and Ehang are ambitiously targeting service launches in selected cities by 2030 [23, 24, 41]. Some predictions go as far as to claim that by 2050, about 100,000 air taxis will be in service worldwide [26].

At the same time, advances in battery technology have pioneered electrically powered air taxis, reducing local emissions to zero. Specifically, Volocopter's electrical air taxi VoloCity has a range of 35 km, which is insufficient to transport passengers between city centers and airports in the majority of the world's megacities [69]. These vehicles – commonly referred to as electrical vertical take-off and landing (eVTOL) aircraft – are particularly advantageous because they require minimal urban space for launching and landing. In the initial phase of UAM, it is anticipated that onboard pilots will operate eVTOLs manually. This is primarily due to passengers' higher level of trust in onboard pilots compared to automated operations or intermediate levels, such as remote pilots [8, 15, 43, 69]. Nevertheless, the ultimate goal remains to establish a fully automated service in the near future without human pilots involved [15, 69], distinguishing automated UAM from current urban aerial vehicles such as helicopters. Additionally, with the advent of low-latency network standards such as 5G, air taxis are expected to be interconnected, enhancing safety through communication with other vehicles and infrastructure, similar to developments in the automotive industry [33].

While UAM could be beneficial to solving the congestion problem and improving mobility, (market) studies show that passengers' trust and perceived safety highly influence whether potential passengers are willing to use fully automated air taxis [2, 3, 14]. To address this, Colley and Meinhardt et al. [12] evaluated the effects of visualizing path information for automated air taxis. They found that "the path line visualization via chevrons performed highest for most dependent variables [trust, perceived safety, ...]" [12, p. 1]. However, their study was limited to ordinary travel scenarios, neglecting adverse lighting and weather conditions as well as emergency situations. While the impact of extraordinary conditions on passenger trust has been well-studied in the field of automated vehicles [11], comparable research in the UAM sector is missing, leaving a gap in strategies to maintain high levels of passengers' trust during extraordinary situations. In line with this, Bauranov and Rakas [5] recommended broadening the scope of research to include extraordinary situations for UAM such as "*accident scenario planning, bird strike risk, loss of control, and risk due to wind gusts*" [5, p. 24]. Further, we argue that extraordinary situations could lead to negative feelings such as decreased perceived safety. Thus, visualizations such as path lines of the trajectory [12] might counteract them. Moreover, the influence of the introduction of UAM extends beyond passengers to include urban and suburban residents. These individuals may experience auditory and visual disturbances due to the operation of air taxis [5]. Moreover, privacy issues, previously stated for unmanned drones [68], could similarly occur due to UAM [77]. Therefore, we adopted a dual focus: we examine the impact of UAM on passengers during extraordinary scenarios such as adverse lighting and weather conditions and emergency situations but also evaluate the impact on residents, providing a more holistic view of the societal impact of UAM. Thus, this work is guided by the three research questions (RQs):

RQ1 *What are the effects of different visualizations on UAM users in different adverse lighting and weather conditions (i.e., daytime, fog, and night)?*
RQ2 *What are the effects of different visualizations on UAM users in emergency situations (i.e., rotor breakdown, a flock of birds approaching, and wind gusts?*)
RQ3 *What effects does the availability of different numbers of air taxis have on residents in the city (in a café, outside a café, in an office) and in the suburbs?*

To address these RQs, we created specific scenarios for each, incorporating three conditions for comparison that focused on safety (RQ1, RQ2) and social aspects (RQ3) [5]. These simulations were divided into three online studies that incorporated 360deg videos and gaze-tracking via webcam. The first study focused on the effects of adverse lighting and weather conditions on UAM users, the second focused on emergency scenarios, and the third assessed the residential perception concerning varying numbers of air taxis. Our findings reveal that the acceptance of UAM both as passengers and residents could be high. We found that fog was perceived as very critical, as well as the rotor breakdown scenario. Having a display with additional information led to the highest positive impact, e.g., on perceived safety during adverse lighting conditions and on SA for emergency situations. Hence, we imply that showing the future trajectory can be considered an "add-on" feature.

Contribution Statement:

- We created 10 scenarios relevant to the UAM context, including 3 adverse lighting and weather scenarios, 3 emergency scenarios, and 4 urban and suburban scenarios from a residents' perspective
- The results of three 360deg video-based online within-subject studies (N_1=31, N_2=29, and N_3=29) suggest that acceptance of UAM could be high among both passengers and affected residents.

2 RELATED WORK

This work extends previous studies on path visualizations for airborne robots, such as helicopters, remotely controlled aircraft, and drones, and takes the context of UAM research within the field of HCI into account.

2.1 Visualization Support for Pilots

While we focus on visualizations' effect on novice users, visualizations for pilots of helicopters help to conceptualize and integrate prior work. Waanders et al. [70] discovered that pilots' situation awareness (SA) could be improved by a 3D representation of the surrounding environment in situations of poor visibility. A comparison of Head-Down-Displays (HDD) and Head-Up-Displays (HUD) revealed that pilots favored the 3D environmental display on the HDD when flying at high altitudes. Conversely, HUD was the preferred choice during landing or when flying at low altitudes [47]. They posited that this preference stemmed from the pilots' reliance on HDD for navigational assistance at high altitudes, while HUD improved visibility, especially in adverse weather conditions. This is consistent with a qualitative study by Minotra and Feigh [45] on the cognitive demands of ship-based landings. They found that pilots initially depend on the HDD inside the helicopter while approaching the ship and then switch to visual cues for landing. More experienced pilots tend to switch to visual flight earlier.

Concepts also exist for remote pilots as an intermediate level of automation for UAM. They operate multiple flying taxis simultaneously, rather than just a single one [8]. Therefore, Calhoun et al. [7] explored visualizations for these predicted paths. They concluded that their path visualization design could help potential remote pilots save time when switching attention between different remotely controlled aircraft. Unlike remote pilots, Szafir et al. [64] studied various methods to indicate drones' future flight direction by attaching a light band to the drone. This research was extended by Walker et al. [72], who examined different path visualizations to communicate the motion intent of automated drones. By leveraging Augmented Reality (AR), they showed users four distinct visualizations, determining that their designs significantly improved the understanding and predictability of drone behavior.

2.2 Human-Computer Interaction for Urban Air Mobility

Kim et al. [35] conducted a conference workshop focused on User Experience in UAM, involving experts in the automotive field due to the novelty of UAM. This workshop's findings were further explored in a subsequent publication by Lim et al. [38]. Their analysis indicated that the workshop participants perceived a transition from an initial emphasis on *safety* and *acceptance* to a more mature phase where *comfort* became a significant concern. However, they did not propose explicit solutions to these aspects. A workshop with a similar focus by Edwards and Price [14], who also explored the passengers' needs in the early stages of UAM. Their workshop, conducted with aviation professionals, resulted in six primary categories of concerns: perceived safety, noise and vibration, passenger well-being, and environmental concerns. Furthermore, the authors suggested measures to mitigate these issues. One of their key recommendations is to create a high-fidelity simulator for a better understanding of passenger needs and to study the effects of rotor noise and vibration inside the cabin. Meinhardt and Colley et al. [43] conducted a workshop with professional helicopter pilots evaluating automation and visualization possibilities for helicopter piloting for UAM in the future. The workshop mainly focused on visualizing air traffic (e.g., showing flight paths to increase SA), avoiding obstacles, and map visualization. The closest work to ours is the paper by Colley and Meinhardt et al. [12]. They, based on the visualizations in the automotive domain, evaluated the effects of different trajectory visualizations in an online video-based study. Finding that the chevron line was the most appropriate, they then evaluated in a VR study the effect of this chevron line and the presence of other air traffic. Therefore, we also use the same chevron line in our study. This was looked at more frequently when there was other air traffic, and they found that the chevron line enhanced trust and provided a more predictable outlook for the future trajectory of the air taxi.

3 USER STUDIES' GENERAL APPROACH

We conducted three distinct online studies to investigate passengers' and resident's perceptions of UAM during different scenarios. This section outlines the general approach we used for all studies. Detailed explanations for each specific study can be found later in the paper (refer to section 4 for the study on adverse lighting and weather conditions, section 5 for the study on emergencies and section 6 for the study on perceptions of residents).

3.1 Scenario Creation

Scenarios for each study were created using the Unity Game Engine version 2022.3.4f1 [66]. For the first two studies, the Google Map Tiles API for Unity was employed to create a realistic model of New York City. These two studies were designed from the perspective of a passenger (see subsection 4.1 and subsection 5.1). Therefore, a virtual camera was positioned above the shoulders of an avatar inside a 3D model of a Volocopter 2X.

The third study required more ground-level details and indoor scenarios across different metropolitan areas. Therefore, we used additional assets for Unity [1, 20, 34] to create scenarios inside and outside a café, in an office, and the suburbs. Additionally, we positioned the 3D model of the Volocopter 2X in the sky to simulate the presence of air taxis. All scenarios were recorded with the VR Panorama 360 PRO Renderer [49], producing equirectangular 360deg videos that were hosted on Vimeo. Using 360deg videos, we ensured a high immersion [16], suggesting that "360-degree videos offer a unique sense of presence and immersion not possible to achieve using traditional videos" [57, p. 1450]. Using this method, we were able to address a high diversity of participants [63], creating accessibility to people who do not own VR glasses.

3.2 Gaze Tracking

We used `WebGazer.JS` [50] to track the participants' gaze as they watched the 360deg videos. This allowed us to analyze the areas of interest (AOI) the participants looked at during the study. Our approach accounted for the movement of the field of view inherent in 360deg videos to identify the correct AOIs accurately. We predefined these AOIs, including the air taxis' interior, display, and ego path. A detailed description of the technical setup can be found in Appendix A.

While the gaze data analysis for the three studies yielded significant interaction effects between the scenarios $\times$ Visualizations $\times$ AOI, the results did not provide any insightful implications. On the one hand, this can be explained by the inconsistent behavior of the participants, inducing a high alteration of their gaze. On the other hand, the gaze-tracker in our online study environment was not accurate enough to provide detailed insights. Therefore, we decided not to include them in the paper but to include them in the appendix (see subsection B.2, subsection C.2 and subsection D.2)

3.3 Study Procedure

Each session began with an initial introduction, followed by a consent form. The eye tracker was then calibrated. The nine (study 1 and 2) or 12 (study 3) conditions were then presented in random order. After all conditions, participants were given a demographic questionnaire to complete. A background script ensured that the videos were viewed in full screen, without skipping or replaying, and that the exposure duration was consistent. It also required a FullHD monitor that logged the participant's viewport, eye position, and screen resolution to an external database at 10Hz. To determine whether participants paid attention during the survey, we implemented attention check questions that appeared randomly throughout the survey. Thus, we discarded participants if they failed at least two attention check questions or when they had not submitted any eye gaze data. We did not control the distance to the screen for eye tracking.

3.4 Statistical Data Analysis

Before every statistical test, we checked the required assumptions (e.g., normality distribution via Shapiro-Wilk test [61]). For non-parametric data, we used the aligned rank transformation (ART) for repeated measures by Wobbrock et al. [75] as the typical ANOVA is inappropriate with non-normally distributed data. For the posthoc test, we used the Dunn test with Holm correction. For the analysis, R in version 4.4.1 and RStudio in version 2023.04.2 were employed. All packages were up to date in July 2024.

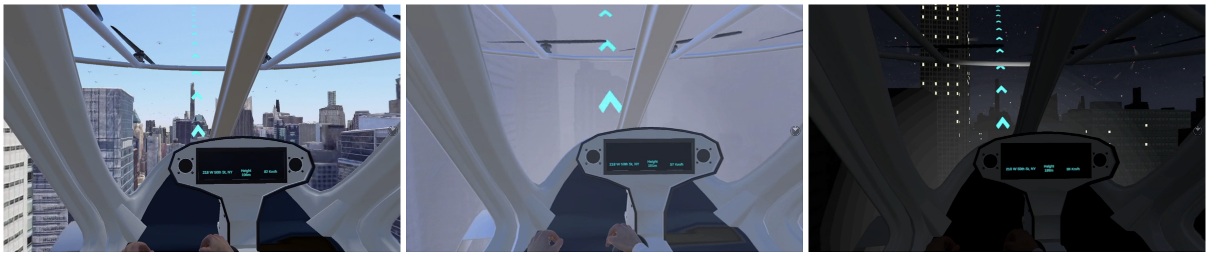
We did not formally analyze the participants' open feedback. We report anecdotal quotes discussed by two authors, summarizing the main points reported by the participants. The response rate for the study on adverse lighting and weather conditions and the study on emergency situations was 100%. While also all participants replied regarding open feedback in the study on the perceptions of residents, 23/32 answers indicated that they had no additional feedback (e.g., "no", "none", "thanks"). Thus, the response rate was 28.13%.

3.5 Participant Recruiting

We recruited participants via prolific.co and calculated the required sample size before each experiment by an a-priori power analysis using G*Power [19]. To achieve a statistical power of 80% with an alpha level of .05, 30 participants should result in an expected medium effect size [21] for a two-way ANOVA. This number is based on the reported results by Colley and Meinhardt et al. [12], which correspond to an effect size of 0.25 for Predictability. For the resident perception study, the G*Power analysis resulted in 28 required participants.

The experimental procedure followed the guidelines of our university's ethics committee and adhered to regulations regarding the handling of sensitive and private data, anonymization, compensation, and risk aversion. Compliant with our university's local regulations, no additional formal ethics approval was required.

4 STUDY ON ADVERSE LIGHTING AND WEATHER CONDITIONS



(a) Daytime (b) Fog (c) Night

Fig. 2. Views on daytime, fog, and night conditions.

To answer RQ1, we designed three scenarios: daytime, foggy weather, and at night *Scenario* (see Figure 2).

4.1 Materials

We created nine simulation videos showcasing a journey in a Volocopter 2X over New York City, facilitated by the Google Map Tiles API for Unity [66] (see Figure 2).
In these simulations, the Volocopter followed a predetermined route, involving various maneuvers such as turns in both directions and changes in altitude. A windshield display (WSD) provides users with a view limited to a range of 200 meters. To maintain a consistent and neutral appearance, we used a turquoise color for all augmented objects displayed on the WSD, aligning with established design principles [73].
Additionally, we included other air taxis in the simulations based on assumptions from related work. Colley and Meinhardt et al. [12] already showed that the presence of other air taxis influences passengers' trust and perceived safety. However, predictions for the number of air taxis operating in New York vary [25, 53–55]. Rajendran and Zack [55] predicts 150 air taxi operations per hour in New York. With 603 predicted vertiports [25] in New York, it is estimated that about 500 air taxis might be in operation at any time. Hence, for our simulation, we included 500 Volocopter 2X air taxis operating on randomly timed pre-defined trajectories. For this user study, the nine videos consist of three scenarios, each paired with three different visualizations:

Adverse lighting and weather conditions.

- **Daytime.** This is the baseline scenario, also used by Colley and Meinhardt et al. [12] (see Figure 2a).
- **Fog.** In AV research, fog is used to simulate scenarios that seem risky but are safe due to the enhanced perception capabilities of the vehicles [29, 74, 76] (see Figure 2b).
- **Night.** We included night scenarios, which also reduce sight but not as badly as in fog conditions (see Figure 2c).

Visualizations.

- **Path without display.** Based on Colley and Meinhardt et al. [12], we always showed the future path of the air taxi. However, for this condition, no further information was displayed on the screen.
- **Path with display.** In this visualization, we added the air taxi's current height, location, and speed in a display in the center of the air taxi.

- **Bounding Box.** For this visualization, in addition to the path and the display, we highlighted the recognition of other air taxis by visualizing turquoise bounding boxes around them (see Figure 1). This is based on prior work in the automotive domain where recognized vehicles are highlighted [10, 13, 76].

4.2 Procedure

For a detailed procedure of this study, refer to subsection 3.3. Before the conditions, the participants were provided with a preamble to set the scenario as follows:

You'll be watching several 360-degree videos, imagining yourself as a passenger in a self-flying taxi flying over a city. The taxi sticks to a certain route and handles all aspects of flying. While you can't control the air taxi, you can use your mouse to look around in the 360-degree video. Each video lasts about 2 minutes and shows the air taxi in different adverse lighting and weather conditions; further, there will be other air taxis visualized differently. After each video, you are expected to share your thoughts about your feelings in such a potential Scenario.

After each video, participants were asked to respond to the questionnaires as detailed in subsection 4.3. The length of one video was approx. 120s. Previous work in UAM used videos with 50s [12], which had also been used in automated vehicle research [10] (or even less [46]). In general, related work showed that duration is not the determining factor for immersion. For example, Zhang et al. [78] found that when comparing 3min, 7min, and 11min videos, either 3min or 7min created the highest immersion. Therefore, we argue that our 2min videos are sufficient to induce immersion to be able to answer our RQs. On average, each session spanned 43 minutes. Participants received a compensation of 6£.

4.3 Measurements

As the primary passenger concerns related to air taxis are trust and perceived safety in automation [2, 3, 14], our first measurement was perceived safety using four 7-point semantic differentials ranging from -3 (anxious/agitated/unsafe/timid) to +3 (relaxed/calm/safe/confident) [18]. Next, trust was measured using the *Predictability/Understandability* and *Trust* subscales of the *Trust in Automation* questionnaire [36].

As outlined by Hoff and Bashir [30], trust is a multifaceted construct with three dimensions: dispositional trust, situational trust, and learned trust. Particularly, dispositional trust can differ based on both internal and external variables. The latter is influenced by factors such as system type, system complexity, task difficulty, workload, perceived risks and benefits, organizational environment, and task framing, as highlighted by Holthausen et al. [31] and Müller et al. [46]. Further, we incorporated the NASA TLX [27] to measure cognitive load as Samson and Kostyszyn [60] showed that cognitive load correlates with trust. Additionally, the situation awareness rating technique (SART) questionnaire [65] was utilized to measure the perceived quality of SA [17], which could forecast "how a person will choose to act on that situation awareness" [17, p. 86]. With high-quality SA, passengers are more prone to react calmly towards air taxis and their post-automation effects [6, 44], allowing automation to undertake the piloting task.

After all conditions, the Immersion subscale of the Technology Usage Inventory (TUI) [37] was utilized for participants to confirm adequate immersion. Participants were also asked for their agreement whether they were comfortable ("I would feel comfortable living close to a take-off station for air taxis."), concerned about infrastructure ("I am concerned about the impact of air taxi infrastructure (landing pads, charging stations) on the cityscape.") and whether they thought UAM would change behavior ("I believe air taxis will change human behavior significantly.") on 7-point Likert scales (1=Totally Disagree to 7=Totally Agree).

Finally, participants could provide open feedback ("Do you have any other feedback?") regarding the visualizations and scenarios ("What did you find particularly good/particularly bad about the visualizations? What would you change?") and how they think that UAM is changing human behavior ("How do you believe will air taxis alter human behavior?"). This information was given via text fields.

4.4 Participants

31 participants (mean age = 45.4, SD = 13.0, range: [25, 75]; Gender: 16 female, 14 male, 1 non-binary) took part in the study. 10 live in a rural, 12 in a suburban and 9 in an urban setting. Regarding their employment, participants stated that they were employees (21), jobseekers (3), self-employed (2), or college students (1). Four indicated "Other". 18 participants reported having a college, 10 having a high school, and one reported having a middle school degree. On a 5-point Likert scale (*1 = Strongly Disagree — 5 = Strongly Agree*), participants showed medium interest in UAM (*M*=3.45, *SD*=1.31), were skeptical whether UAM would ease their lives (*M*=3.32, *SD*=1.33), and were skeptical about whether they become reality by 2033 (*M*=2.77, *SD*=1.28).

The average score for Immersion, used to assess the results' reliability, was moderate to high (*M*=19.77, *SD*=5.25). This represents an even higher Immersion than a previously reported VR UAM study (*M*=18.88, *SD*=3.95; [12]). After all conditions, participants stated they were rather comfortable (*M*=4.13, *SD*=2.00) but were also concerned about infrastructure (*M*=4.39, *SD*=1.87) and believed UAM to alter humans' behavior (*M*=4.71, *SD*=1.62).

4.5 Results

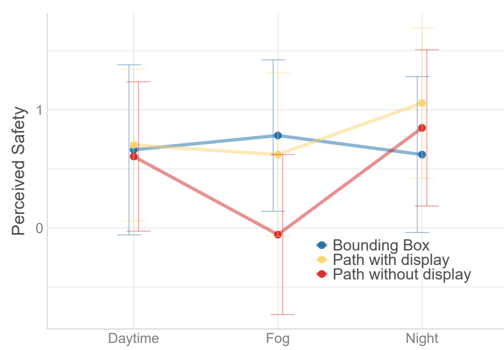


Fig. 3. Interaction effect *Visualization* × *Scenario* on perceived safety.

4.5.1 Perceived Safety. We found a significant main effect of *Visualization* on perceived safety ($F(2, 60) = 4.20$, p=0.020). We also found a significant interaction effect of *Visualization* × *Scenario* on perceived safety ($F(4, 120) = 3.41$, p=0.011; see Figure 3). During Daytime and Night, perceived safety was almost equal for all *Visualization*. However, for Fog, it was lower with Path without display compared to the other *Visualization*.

4.5.2 Cognitive Load. The ART found a significant main effect of *Visualization* on cognitive load ($F(2, 60) = 4.81$, p=0.012). However, a post-hoc test found no significant differences.

4.5.3 Trust and Understanding. The ART found no significant effects on Trust. However, there is a significant main effect of *Visualization* on Understanding ($F(2, 60) = 10.12$, p<0.001).

A post-hoc test found that Bounding Box (*M*=3.67, *SD*=0.89; p_{adj}=0.024) and Path plus display (*M*=3.71, *SD*=0.89; p_{adj}=0.027) were significantly higher in terms of understanding compared to Path without display (*M*=3.38, *SD*=0.89; p_{adj}=0.024).

4.5.4 Situation Awareness. No significant effects were found for SA.

4.5.5 Final Assessment and Ranking. In general, participants rated the visualizations as necessary (*M*=4.07, *SD*=1.29; scale 1 to 5) and useful (*M*=4.35, *SD*=.95).

In particular, Friedman's ANOVA ($\mathcal{X}^2(2) = 21.74$, $p < 0.001$) found differences for the Bounding Box and Path with display, as both were rated significantly higher than the Path without display.

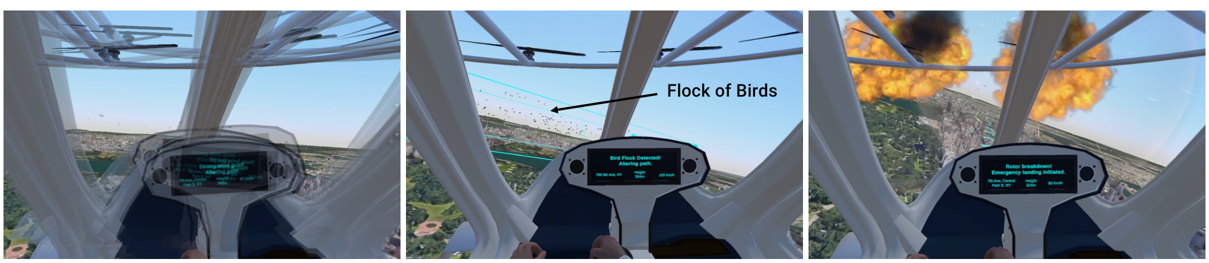
However, Friedman's ANOVA ($\mathcal{X}^2(2) = 4.32$, $p = 0.12$) found no significant differences in which scenario more visualization was preferred.

4.5.6 Open Feedback.

Visualization and Scenarios. The open feedback on the visualizations and scenarios shows that most participants were satisfied with the existing visualizations and mentioned that they felt more comfortable and in control. For example, one participant stated, "*I loved that the air taxi had a display of where I was going and the altitude. This made me feel more comfortable.*" However, 12 individuals also asked for more detailed and customizable information, such as routes of other air taxis, to enhance their comfort and trust. In particular, they desired "*more info on the screen about what [the air taxi] is seeing and about anything that comes near it.*" Also, while the daytime and night conditions were not of a concern, the foggy condition was for at least two participants, with comments such as "*When we had fog, it was terrible and kind of scary not to be able to see anything around me.*".

Human Behavior Change. The open feedback on the potential impact of air taxis on human behavior reveals a mix of optimism and skepticism. On the positive side, many participants anticipate that air taxis will ease traffic congestion, reduce commute times, and increase overall mobility. One participant noted, "*[...] Humans will have much more time to spare for themselves rather than sitting in traffic.*" Another stated, "*With the introduction of this futuristic mode of transportation, people will have more flexibility and convenience in their daily lives.*" However, affordability was frequently mentioned as a concern, with one participant saying, "*I think they will make it easier for the rich to travel, but harder for the poor until air taxis become affordable for all.*" Safety and noise pollution were other areas of concern. One respondent mentioned, *"I believe air taxis could become a flight hazard, potentially leading to accidents both in the air and on the ground."*

5 STUDY ON EMERGENCY SITUATIONS



(a) Wind gusts (b) Bird flock (c) Rotor breakdown

Fig. 4. Overview of the conditions in the emergency situation study.

To answer RQ2, we designed three scenarios consisting of different emergency situations (see Figure 4).

5.1 Materials

Bauranov and Rakas [5] suggested "*including accident scenario planning, bird strike risk, loss of control, and risk due to wind gusts*"[5, p. 24]. Hence, we designed three emergency situations: Rotor breakdown, a sudden wind gust, and a sudden occurrence of a bird-flock. Coping with various dangerous situations, research on unmanned aircraft already employs a *sense-and-avoid* mechanism for altering flight paths, which enables the aircraft to detect potentially dangerous objects and adjust its trajectory accordingly [39, 42, 56]. We adapted this approach for emergency scenarios in air taxis, enabling them to alter their path to avoid obstacles such as flocks of birds and wind gusts or to signal a path change for emergency landings. Based on Colley and Meinhardt et al. [12] and similar to the aforementioned study, we always showed the future path of the air taxi via chevrons.

Emergency Situations.

- **Wind Gust.** Due to high buildings in urban metropolitan areas, turbulences can cause sudden wind gusts that affect the trajectory of automated aircrafts [22, 48, 52]. Our scenario includes a sudden gust of wind that causes the air taxi to shake and eventually change its trajectory (see Figure 4a).
- **Flock of Birds.** Sudden arrival of a bird flock: A bird flock requires the air taxi to adjust its path (see Figure 4b).
- **Rotor breakdown.** A rotor breakdown requires the air taxi to initiate an emergency landing (see Figure 4c).

Visualizations.

- **Permanent Path.** When altering the path due to an emergency, the original flight path remains visible but transitions to a grey color for the entire duration of the flight. Additionally, surrounding air taxis are highlighted with bounding boxes to signify their detection by the air taxi. In addition, a notification appears on the screen to inform passengers about the emergency situation and to indicate that the flight path will be altered.
- **Fading Path.** When altering the path, the original trajectory turns grey, remains visible for 5 seconds, and then gradually fades away. Additionally, nearby air taxis are outlined with bounding boxes to indicate that they have been detected by the air taxi. In addition, a notification appears on the screen to inform passengers about the emergency situation and to indicate that the flight path will be altered.
- **Path without display.** Only the current path is shown in the WSD [12], while the original path is not shown anymore. Further, no information is displayed on the screen.

The scenario was based on the Unity environment already used for the previous study on adverse lighting and weather conditions. Additional assets used for the emergency scenario include the exposition for the rotor breakdown [32] and the flock of birds [67].

5.2 Procedure and Measurements

This study used the same procedure (see Section 4.2) and measurements (see Section 4.3) as in the study on adverse lighting and weather conditions. We explicitly did **not** mention the possibility of emergency situations to not bias participants. On average, each session spanned 49 minutes. Participants received a compensation of 6£.

5.3 Participants

29 participants (mean age = 34.6, SD = 11.2, range: [18, 64]; Gender: 12 female, 15 male, 2 non-binary) took part in the study. 4 live in a rural, 13 in a suburban, and 13 in an urban setting. Regarding their employment, participants stated that they were employees (15), jobseekers (2), self-employed (5), or college students (4). Three indicated "Other". 22 participants reported having a college, 7 having a high school, and two reported having a vocational degree. On a 5-point Likert scale (*1 = Strongly Disagree — 5 = Strongly Agree*), participants showed medium interest in UAM (M=3.13, SD=1.64), were skeptical whether UAM would ease their lives (M=3.10, SD=1.59), and were skeptical about whether they become reality by 2033 (M=2.97, SD=1.52). The average Immersion score [37], used to assess the results' reliability, was moderate to high (M=19.48, SD=5.94). After all conditions, participants stated they were relatively comfortable (M=4.21, SD=2.24) but were also concerned about infrastructure (M=4.45, SD=2.15) and believed UAM to alter humans' behavior (M=5.00, SD=1.81).

The participants **did not** overlap with the study participants on adverse lighting and weather conditions.

5.4 Results

5.4.1 Perceived Safety. The ART found a significant main effect of *Visualization* ($F(2, 56) = 12.54$, p<0.001) and of the emergency situation on perceived safety ($F(2, 56) = 14.80$, p<0.001). A post-hoc test found that Fade Path (*M*=0.90, *SD*=1.76; p_{adj}=0.025) and Permanent Path (*M*=0.78, *SD*=1.79; p_{adj}=0.047) were significantly higher in terms of perceived safety compared to the path without display (*M*=0.11, *SD*=1.96). A post-hoc test found that the flock of birds (*M*=1.13, *SD*=1.69; p_{adj} <0.001) and the Wind Gust scenario (*M*=0.82, *SD*=1.72; p_{adj}=0.002) yielded a higher perceived safety compared to Rotor Breakdown (*M*=-0.17, *SD*=1.94).

5.4.2 Cognitive Load. We found a significant main effect of emergency situation on cognitive load ($F(2, 56) = 8.09$, p<0.001). A post-hoc test found that Rotor Breakdown was significantly higher (*M*=10.45, *SD*=6.42) in terms of cognitive load compared to Bird Flock (*M*=7.46, *SD*=5.98; p_{adj}=0.005) and Wind Gust (*M*=7.83, *SD*=5.92; p_{adj}=0.015).

5.4.3 Trust and Understanding. The ART found a significant main effect of *Visualization* ($F(2, 56) = 17.28$, p<0.001) and of emergency situations on trust ($F(2, 56) = 5.63$, p=0.006). In detail, a post-hoc test found that the fade path (*M*=3.53, *SD*=1.13; p_{adj} <0.001) and the Permanent Path (*M*=3.62, *SD*=1.17; p_{adj} <0.001) yielded significantly higher trust compared to the path without display (*M*=2.81, *SD*=1.21). The post-hoc test found no significant differences for the emergency situations.

There was also a significant main effect of *Visualization* on understanding ($F(2, 56) = 45.04$, p<0.001). A post-hoc test found that the fade path (*M*=3.83, *SD*=0.81; p_{adj} <0.001) and Permanent Path (*M*=3.87, *SD*=0.75; p_{adj} <0.001) were significantly higher in terms of understanding than the path without display (*M*=2.73, *SD*=1.00).

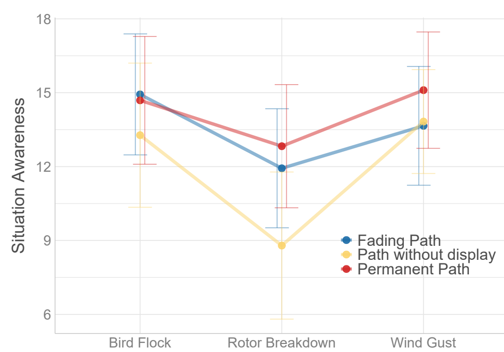


Fig. 5. Interaction effect *Visualization* × *Scenario* on SA.

5.4.4 Situation Awareness. The ART found a significant main effect of *Visualization* ($F(2, 56) = 6.44$, p=0.003) and of emergency scenario on SA ($F(2, 56) = 9.41$, p<0.001). The ART also found a significant interaction effect of *Visualization* × *Scenario* on SA ($F(4, 112) = 2.52$, p=0.045; see Figure 5). In detail, the gained SA consistently ranked lowest for the condition path without display. However, in the scenario, Bird Flock, the SA was nearly equal for the fading and Permanent Path visualizations. Interestingly, in the scenario, Wind Gust, the fade path resulted in lower perceived SA compared to other visualizations. These findings underscore the intricate interplay between visualization type and scenario in affecting SA levels.

5.4.5 Final Assessment and Ranking. Participants rated the visualizations as necessary (*M*=4.48, *SD*=.69; scale 1 to 5) and useful (*M*=4.72, *SD*=.45). A Friedman's ANOVA ($\mathcal{X}^2(2) = 46.67$, $p < 0.001$) found significant differences for the visualizations. Hence, the Permanent Path was clearly preferred (*Mdn*=1.00), followed by the Fading Path (*Mdn*=2.00) and the path without display (*Mdn*=3.00). All post-hoc comparisons were significant.

5.4.6 Open Feedback.

Visualization and Scenarios. The open feedback implies that the rotor breakdown scenario especially threatened participants. On the one hand, participants were reassured by the information provided, as one mentioned: "*I was glad when the screen told me it was in landing mode when a rotor broke. It put my mind at ease instantly.*". Further, another participant appreciated the Fading Path visualization, as she *liked the [WSD] that had the Fading Paths so you knew when it was changing paths.*. On the other hand, participants mentioned a desire for more detailed information during emergency situations. Specific suggestions for improvement included better visualization of the flight path, recommending to "*[...] provide more information about why the path is changing [...]*" Others suggested adding flashing lights or audio announcements to draw attention to important information.

Human behaviour change. The open feedback regarding the potential impact of air taxis on human behavior is similar to the one for the first study, including the potential for increased efficiency in mobility. One participant noted, "*Humans can get to where they want faster without the traffic [congestion]*", while another mentioned that air taxis could "*free up space and traffic on the ground.*" Similar to the first study, affordability was another issue, with one participant stating, "*I think they'll likely be too expensive for most people.*" Others speculated on the societal changes that could occur, such as one who said air taxis "*will greatly reduce the need for private cars,*" potentially leading to more pedestrian-friendly cities.

6 STUDY ON PERCEPTIONS OF RESIDENTS

(a) Café inside (b) Café outside (c) Suburb (d) Office

Fig. 6. Overview of the four scenarios in the study on resident perceptions.

To address RQ3, we created four scenarios that encompass metropolitan areas, each focusing on residents' perceptions (see Figure 6).

6.1 Materials

The four scenarios to explore the perception of residents on UAM are described below:

- **Café inside.** One sits inside at a café window. Other Nearby people engage in conversations (see Figure 6a).
- **Café outside.** One sits outside the same café. People talk nearby and vehicles drive by (see Figure 6b).
- **Suburb.** One sits in the front yard. Few vehicles and people are present (see Figure 6c).
- **Office.** One sits in an office on a higher floor (i.e., the 50th floor, ≈160 m above the ground; see Figure 6d).

The acoustic damping of the windows inside the café and the office was modeled based on the distance and realistic values for safety glass. We selected these four scenarios to cover diverse environments: from busy urban areas with more visual and auditory noise to quieter suburban areas. Additionally, proximity to the noise source, as in the elevated office scenario, might reduce acceptance levels. We altered the *number of air taxis* from None to Medium and High, with High representing double the Medium amount. While existing studies offer different predictions on New York's air traffic [25, 53–55]. Assuming 603 vertiports in New York [25], we estimated an average of 500 air taxis (Volocopter 2X model) in flight at any moment. Based on internal tests, we added 500 (Medium) or 1000 (High) air taxis to our simulation, following predetermined trajectories at random intervals.

6.2 Procedure

For this study, we altered the introductory text to:

You'll be watching several 360-degree videos, imagining yourself as a person on the ground or in an office in various scenarios. Depending on the condition, there will be air taxis flying in the air. You can use your mouse to look around in the 360-degree video. Each video lasts about 2 minutes. After each video, you are expected to share your thoughts about your feelings in such a potential scenario. You will also see a newspaper in front of you. This will contain an article, which you will be questioned upon later.

In contrast to the previous studies, we introduced a newspaper as an additional element to enhance external validity. Engaging in a secondary task, like reading a newspaper, simulates real-world scenarios such as visiting a café, working at an office, or enjoying leisure time at home. Each session lasted, on average, 46 minutes, and participants received compensation of 7.5£ for their participation. For this study, the compensation was higher as we had anticipated a longer duration due to three more conditions compared to the other two studies.

6.3 Measurements

We modified the measurements for this study to focus on residents rather than air taxi passengers. However, we kept the measurement of perceived safety [18]. Additionally, we used four single items by Yedavalli and Mooberry [77] on 5-p Likert scales (1=Strongly disagree to 5=Strongly agree). These items addressed concerns about noise, altitude, visual pollution, safety, and personal privacy induced by UAM.

Finally, we measured how comfortable participants felt in the situation ("I feel comfortable in this situation", same 5-p Likert scale). This was used in previous research (e.g., [59]) to assess the general comfort level. After all conditions, participants could give open feedback if they thought UAM would change human behavior.

6.4 Participants

29 participants (mean age = 42.7, SD = 15.2, range: [24, 77]; Gender: 14 female, 15 male, 0 non-binary) took part in the study. 4 live in a rural, 12 in a suburban and 13 in an urban setting. Participants stated that they were employees (19), jobseekers (2), self-employed (2), or college students (3). Three indicated "Other". 20 participants reported having a college, 8 having a high school, and one reported having a vocational degree.

On 5-point Likert scales, participants showed medium interest in UAM (M=3.24, SD=1.24), were skeptical whether UAM would ease their lives (M=3.24, SD=1.18), and were skeptical about whether they become reality by 2033 (M=2.97, SD=1.35). The average score for Immersion [37], which was used to assess results' reliability, was moderate to high (M=20.14, SD=6.64). After all conditions, participants stated they were rather comfortable (M=3.90, SD=2.11) but were also concerned about infrastructure (M=3.76, SD=2.17) and believed UAM to alter humans' behavior (M=4.86, SD=1.68). The participants **did not** overlap with the participants of the other studies.

6.5 Scenario Comparison

We divided the results into three comparisons to explain the impact of different environmental and local factors on residents' perceptions of UAM. These comparisons are Café Inside vs. Café Outside, Café Outside vs. Suburb, and Café Inside vs. Office.

- **Café Inside vs. Outside:** This comparison allows us to isolate the effects of visual and auditory pollution by keeping the environment constant while varying the level of exposure. Specifically, being inside the café dampens both the visual and auditory pollution from the air taxis.

- **Café Outside vs. Suburb:** This comparison highlights the role of location in influencing residents' perceptions. It provides insights into how the suburban environment contrasts with the urban environment outside a café.
- **Café Inside vs. Office:** This comparison focuses on the effects of elevation, as the office environment is presumably at a higher level than the café. The change in height affects visual and noise pollution and provides insight into how height affects residents' perception.

6.6 Results Comparison Café Inside vs Outside

6.6.1 Perceived Safety and Comfort. The ART found a significant main effect of *Scenario* ($F(1, 28) = 4.49$, p=0.043) and of *number of air taxis* on perceived safety ($F(2, 56) = 6.56$, p=0.003). It was rated higher inside the Café (M=1.44, SD=2.01) than outside (M=1.37, SD=2.03). However, a post-hoc test found no significant differences between the *number of air taxis*. Further, the ART found a significant main effect of comfort for *Scenario* ($F(1, 28) = 14.52$, p<0.001) and *number of air taxis* ($F(2, 56) = 9.98$, p<0.001). Comfort was higher inside (M=4.25, SD=0.99) than outside (M=4.07, SD=1.11) the Café. Further, a post-hoc test found that the rated comfort of no air taxis was significantly higher (M=4.47, SD=0.98) to a high (M=3.97, SD=1.09; p_{adj}=0.006) and medium *number of air taxis* (M=4.05, SD=1.03; p_{adj}=0.016).

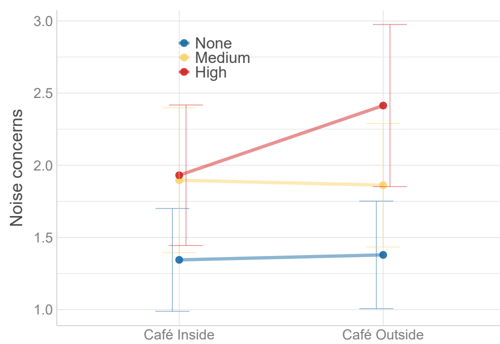


Fig. 7. Interaction effect on noise concern.

6.6.2 Visual and Noise Pollution. We found a significant main effect in terms of the *Scenario* ($F(1, 28) = 12.06$, p=0.002) and of *number of air taxis* on noise concerns ($F(2, 56) = 14.36$, p<0.001). Additionally, the ART found a significant interaction effect of *Scenario* × *number of air taxis* on noise concerns ($F(2, 56) = 3.81$, p=0.028; see Figure 7). With no exposure to UAM, the concern was little but strongly increased with exposure, especially for the Office and the Suburb scenario, for which concern was almost equal. The concerns consistently increased with higher exposure to UAM.

Besides noise concerns, we also found a significant main effect of *number of air taxis* on visual pollution concern ($F(2, 56) = 14.52$, p<0.001). A post-hoc test found that high (M=2.74, SD=1.62; p_{adj} <0.001) and medium *number of air taxis* (M=2.38, SD=1.47; p_{adj} <0.001) were significantly higher in terms of visual pollution concerns compared to no air taxis (M=1.38, SD=0.93).

6.6.3 Altitude and Privacy Concerns. We found a significant main effect of *Scenario* ($F(1, 28) = 9.58$, p=0.004) and of *number of air taxis* on altitude concerns ($F(2, 56) = 13.76$, p<0.001). Altitude concerns were higher outside (M=1.82, SD=1.24) the Café than inside (M=1.66, SD=1.10). A post-hoc test found that high (M=2.00, SD=1.28; p_{adj}=0.001) and medium *number of air taxis* (M=1.90, SD=1.22; p_{adj}=0.003) were significantly higher in terms of concerns about the altitude of UAM compared to no air taxis (M=1.31, SD=0.86).

Additionally, the ART found a significant main effect on privacy concerns of *Scenario* ($F(1, 28) = 10.31$, p=0.003) and of *number of air taxis* ($F(2, 56) = 4.94$, p=0.011). In particular, privacy concerns were higher outside (M=1.72, SD=1.16) than inside the café (M=1.61, SD=1.09). A post-hoc test found that high (M=1.86, SD=1.30, ; p_{adj}=0.050) and medium number of air taxis ((M=1.74, SD=1.05; p_{adj}=0.037) were significantly higher in terms of privacy concerns compared to no UAM (M=1.40, SD=0.95).

6.6.4 Correctness of Questions. We fitted a logistic mixed model (estimated using ML and Nelder-Mead optimizer) to predict correctness with scenario and number of air taxis (formula: correct ~scenario * number of air taxis). The model included the participants as random effect (formula: ~1 | participants). The model's total explanatory power is moderate (conditional R2 = 0.18), and the part related to the fixed effects alone (marginal R2) is 0.10. The model's intercept, corresponding to scenario = Café Inside and the number of air taxis = None, is at -0.08 (95% CI [-0.86, 0.71], p = 0.850). Within this model:

- The effect of the number of air taxis [Medium] is statistically significant and positive (beta = 2.02, 95% CI [0.68, 3.35], p = 0.003; Std. beta = 2.02, 95% CI [0.68, 3.35])
- The effect of the number of air taxis [High] is statistically significant and positive (beta = 1.74, 95% CI [0.48, 3.00], p = 0.007; Std. beta = 1.74, 95% CI [0.48, 3.00])

6.7 Results Comparison Café Outside vs Suburb

6.7.1 Perceived Safety and Comfort. The ART found a significant main effect of *Scenario* ($F(1, 28) = 6.45$, p=0.017) and of *number of air taxis* on perceived safety ($F(2, 56) = 11.93$, p<0.001). The ART also found a significant interaction effect of *Scenario* × *number of air taxis* on perceived safety ($F(2, 56) = 15.40$, p<0.001; see Figure 8a). Perceived safety was the same for the medium *number of air taxis*. It decreased between the café outside to suburbs with a high *number of air taxis* and increased with no UAM.

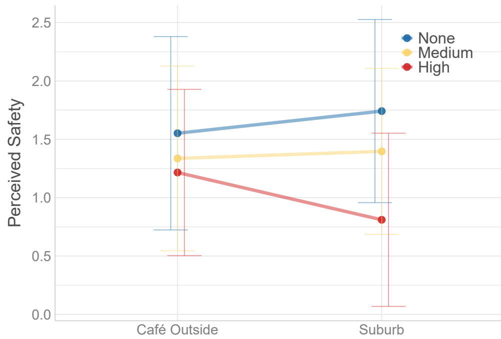
The ART found a significant main effect of *number of air taxis* ($F(2, 56) = 17.07$, p<0.001) and a significant interaction effect of *Scenario* × *number of air taxis* on comfort ($F(2, 56) = 4.85$, p=0.011; Figure 8b). While comfort was lower in the suburb with any number of air taxis compared to Café outside, it increased without UAM.

6.7.2 Visual and Noise Pollution. We found a significant main effect of *Scenario* ($F(1, 28) = 20.85$, p<0.001) and of *number of air taxis* on noise concerns ($F(2, 56) = 26.08$, p<0.001). Moreover, the ART found a significant interaction effect of *Scenario* × *number of air taxis* on noise concerns ($F(2, 56) = 5.66$, p=0.006; see Figure 8c). While the noise concern was approx. equal for the scenarios, it increased strongly in the suburbs with the increasing number of air taxis.

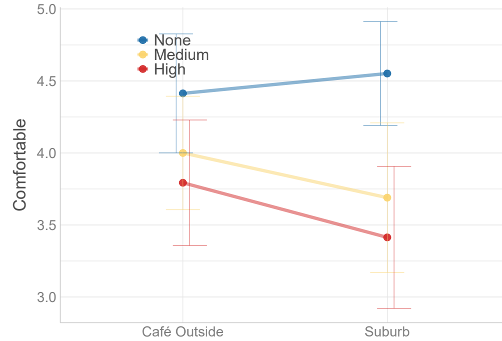
The ART found a significant main effect of *Scenario* ($F(1, 28) = 19.61$, p<0.001) and of *number of air taxis* on visual pollution concern ($F(2, 56) = 23.43$, p<0.001). The ART found a significant interaction effect of *Scenario* × *number of air taxis* on visual pollution concern ($F(2, 56) = 7.80$, p=0.001; see Figure 8d). While visual pollution concern was about the same for the scenarios, it increased in the suburbs with the increasing number of air taxis.

6.7.3 Altitude and Privacy Concerns. The ART found a significant main effect of *number of air taxis* on altitude concerns ($F(2, 56) = 14.95$, p<0.001). A post-hoc test found that High (M=2.28, SD=1.45; p_{adj} <0.001) and Medium (M=2.09, SD=1.27; p_{adj}=0.002) were significantly higher in terms of concerns about the altitude of UAM compared to no air taxis (M=1.40, SD=0.92).

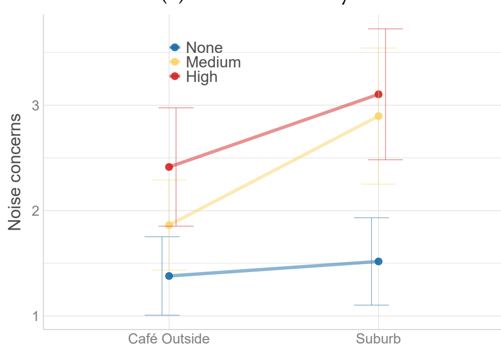
The ART found a significant main effect of *Scenario* ($F(1, 28) = 16.49$, p<0.001) and of *number of air taxis* on privacy concerns ($F(2, 56) = 19.68$, p<0.001). Further, there is an interaction effect of *Scenario* × *number of air taxis* on privacy concerns ($F(2, 56) = 8.95$, p<0.001; see Figure 9). While privacy concerns were approx. equal for the scenarios, it increased in the suburbs with the increased number of air taxis.



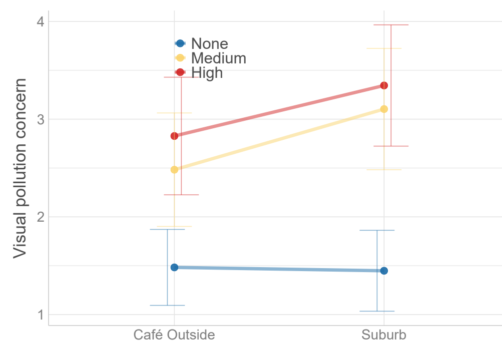
(a) Perceived safety



(b) Comfort



(c) Noise concern



(d) Visual pollution concern

Fig. 8. Interaction effects between Café Outside and Suburb on perceived safety, comfort, noise, and visual pollution concerns

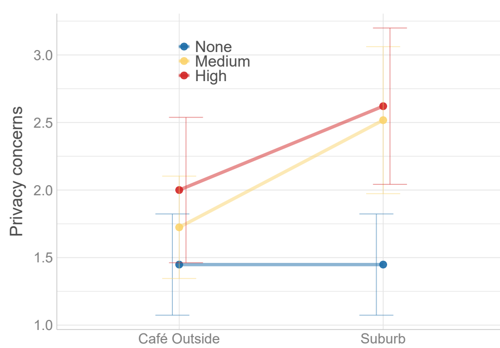


Fig. 9. Interaction effect on privacy concern.

6.7.4 Correctness of Questions. We fitted a logistic mixed model (estimated using ML and Nelder-Mead optimizer) to predict correctness with scenario and number of air taxis (formula: correct ~scenario * number of air taxis).

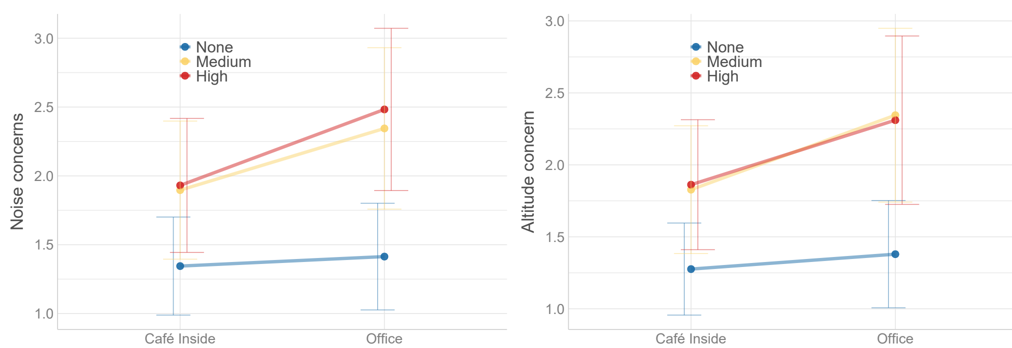
The model included participants as random effect (formula: ~1 | participants). The model's total explanatory power is weak (conditional R2 = 0.03) and the part related to the fixed effects alone (marginal R2) is 0.02. The model's intercept, corresponding to scenario = Café Inside and number of air taxis = None, is at 0.98 (95% CI [0.14, 1.81], p = 0.022). Within this model, we found no significant effects.

6.8 Results Comparison Café Inside vs Office

6.8.1 Perceived Safety and Comfort. The ART found a significant main effect of *number of air taxis* on perceived safety ($F(2, 56) = 4.56$, p=0.015). It was rated higher inside the Café (M=1.44, SD=2.01) than in the Office (M=1.43, SD=2.02). However, a post-hoc test found no significant differences regarding *number of air taxis*.

For the *Scenario* ($F(1, 28) = 6.71$, p=0.015) and *number of air taxis* ($F(2, 56) = 14.46$, p<0.001), there was a main effect on comfort, rated higher inside the Café (M=4.25, SD=0.99) than the Office (M=4.14, SD=1.08). A post-hoc test found that the condition with no air taxis was significantly higher (M=4.52, SD=0.94) in terms of comfort compared to a high (M=3.97, SD=1.06; p_{adj}=0.002) and medium (M=4.10, SD=1.04; p_{adj}=0.020) number of air taxis.

6.8.2 Visual and Noise Pollution. The ART found a significant main effect of *number of air taxis* on noise concerns ($F(2, 56) = 17.48$, p<0.001). Further, there was a significant interaction effect of *Scenario* × *number of air taxis* on noise concerns ($F(2, 56) = 6.11$, p=0.004; see Figure 10a). While the noise concern was approx. equal for the scenarios, it increased strongly in the suburbs with the addition of air taxis to the scene. Further, the ART found a significant main effect of *number of air taxis* on visual pollution concern ($F(2, 56) = 25.32$, p<0.001). A post-hoc test found that a high (M=2.64, SD=1.61; p_{adj} <0.001) and medium (M=2.47, SD=1.50; p_{adj} <0.001) number of air taxis were significantly higher in terms of visual pollution concerns compared to no UAM (M=1.38, SD=0.95; p_{adj} <0.001).



(a) Noise Concerns (b) Altitude Concerns

Fig. 10. Interaction effect of altitude and noise concerns between the Café Inside and the Office scenario

6.8.3 Altitude and Privacy Concerns. The ART found a significant main effect of *number of air taxis* ($F(2, 56) = 15.37$, p<0.001) and interaction effect of *Scenario* × *number of air taxis* on altitude concerns ($F(2, 56) = 5.01$, p=0.010; see Figure 10b). While the altitude concern was approx. equal for the scenarios, it increased in the suburbs with the addition of air taxis. Further, there was a significant main effect of *number of air taxis* on privacy concerns ($F(2, 56) = 11.87$, p<0.001). Post-hoc tests found that a high number of air taxis was significantly more privacy concerning (M=1.83, SD=1.20) than no air taxis (M=1.38, SD=0.95; p_{adj}=0.020) and that a medium number of air taxis was significantly higher (M=1.91, SD=1.23) compared to no air taxis (M=1.38, SD=0.95; p_{adj}=0.011).

6.8.4 Correctness of Questions. We fitted a logistic mixed model (estimated using ML and Nelder-Mead optimizer) to predict correctness with scenario and number of air taxis (formula: correct ~scenario * number of air taxis). The model included participants as random effect (formula: ~1 | participants). The model's total explanatory power is moderate (conditional R2 = 0.25) and the part related to the fixed effects alone (marginal R2) is of 0.16. The model's intercept, corresponding to scenario = Café Inside and number of air taxis = None, is at -0.08 (95% CI [-0.88, 0.72], p = 0.847). Within this model:

- The effect of scenario [Office] is statistically significant and positive (beta = 2.05, 95% CI [0.69, 3.42], p = 0.003; Std. beta = 2.05, 95% CI [0.69, 3.42])
- The effect of the number of air taxis [Medium] is statistically significant and positive (beta = 2.05, 95% CI [0.69, 3.42], p = 0.003; Std. beta = 2.05, 95% CI [0.69, 3.42])
- The effect of the number of air taxis [High] is statistically significant and positive (beta = 1.77, 95% CI [0.49, 3.06], p = 0.007; Std. beta = 1.77, 95% CI [0.49, 3.06])

6.9 Open Feedback

The open feedback from participants, who empathize as urban and suburban residents, provides a multifaceted view of how air taxis might affect human behavior. On the positive side, UAM was described as having the potential to "*revolutionize human behavior by transforming urban mobility.*". In particular, participants frequently mentioned the potential for increased efficiency in daily life, with comments like "*It will make it easier for [people] to get around to where they need to go in less time*". One suggested that people would be able to move to more suburban areas due to the reduced commute time. Further, "*air taxis will lead to less traffic on the street*".

However, concerns were also highlighted. Noise pollution was a significant issue, with participants stating, "*I did not like the noise component of the air taxis [...] it would make being outside less pleasant*". Safety and privacy were also points of interest, with comments like "*I believe we'll be more anxious about our safety and privacy*" and "*people might feel watched and opt to stay inside more*". One participant even mentioned that the "*reduction in commute times will further reduce our capacity for waiting, so patience will further become less prevalent.*" Issues of affordability and social inequality were also raised, encapsulated by the comment, "*The rich will use air taxis while the poor live close to work*," highlighting potential class divisions that the technology could exacerbate.

7 DISCUSSION

Colley and Meinhardt et al. [12] explored the information needs of UAM passengers during ordinary flights without sudden changes and investigated the effects of different visualizations. They stated that "*[...] future work should evaluate the effect of these visualizations in a plethora of situations, such as emergencies*" [12, p. 16]. They assumed that information visualizations during extraordinary situations might be even more relevant, which aligns with Bauranov and Rakas [5] stating that further research is needed for emergency situations in the context of UAM. Therefore, we created six extraordinary flight scenarios. We further extended the research of Yedavalli and Mooberry [77] by creating four scenarios from residents' perspectives in metropolitan urban and suburban areas. The scenarios were evaluated in three online studies using 360deg videos. We had a dual focus on (1) the passenger perspective during adverse lighting and weather conditions (N=31) and emergency situations (N=29) and (2) the urban and suburban resident perspective (N=29).

7.1 Information Needs in Various Scenarios

Our findings indicate a relationship between passengers' information needs and their perceived safety across different scenarios. Addressing RQ1, during low visibility conditions such as fog, passengers reported lower levels of perceived safety when the WSD showed the path without additional information on the screen. This contrasts with the daytime and night conditions, where perceived safety was rated higher (see Figure 3). Notably, the introduction of additional on-screen information, such as speed and altitude, significantly improved perceived safety in foggy conditions, bringing it almost in line with daytime and night scenarios. While this is not a strong indicator, this insight suggests that, contrary to Colley and Meinhardt et al. [12], the augmented air taxi path in the WSD might not be needed in adverse lighting and weather conditions. Also, during emergency situations, not showing additional information on the display yielded a significantly lower SA rating compared to the conditions, including the display inside the air taxi. Nonetheless, the display inside the UAM is important.

Previous research by Woide et al. [76] suggests that augmenting the WSD with additional data can enhance trust for low visibility scenarios in AVs. However, our study did not reveal statistically significant differences in trust levels based on the information displayed, again suggesting that the WSD might not be necessary. Regarding RQ2, we found that the rotor breakdown scenario was associated with the lowest levels of SA when the air taxi's path was displayed without additional information on the screen. Similar to perceived safety in the fog scenario, the SA increased with the introduction of additional information on the display. This finding is consistent with qualitative feedback from participants who expressed a need to provide as much information as possible, even when passengers are not in control of the air taxi. This statement also aligns with Appel et al. [4], who emphasized the importance of transparency throughout the entire customer journey in fostering acceptance and trust for UAM. Our study validates and extends these findings by demonstrating that the need for transparency and SA is not limited to standard flight conditions; it is even more critical in extraordinary circumstances such as emergencies or adverse lighting and weather conditions. However, we conclude that basic information on the screen is most relevant and that higher transparency is not always needed. This could be for various reasons, such as information saturation, or that additional information is not necessary in unfamiliar scenarios without the ability to intervene during the automated flight. Regarding WSD, future work should be conducted to confirm its reduced relevance.

7.2 Relevance of Participants' Perspective

RQ3 addresses the contrasting viewpoints of residents and passengers towards UAM, highlighting the complexity of its societal impact. While passengers see UAM as a convenient and efficient solution to urban congestion, residents express broader concerns related to social well-being and equity. Our findings regarding the residents' opinions partly contrast with Yedavalli and Mooberry [77], who found that residents' safety concerns are less for rural areas than for urban areas. However, we found that perceived visual and noise pollution increases in suburban areas compared to urban areas. Expanding on the work of Yedavalli and Mooberry [77], our study explored the nuances of residents' privacy concerns based on their location (inside vs. outside) and height (high vs. low). Our data suggest that residents experience diminished privacy concerns when they are inside a building, which is consistent with intuitive expectations. Interestingly, elevation also significantly influences privacy perceptions. We found that concerns increased when residents were at higher elevations, likely due to the closer proximity to air taxis at these heights. At lower elevations, the presence of tall buildings may act as natural obstructions, reducing the visibility of air taxis and thereby alleviating privacy concerns. Yet the divergence between passengers' and residents' perspectives necessitates a multi-faceted approach in the development and implementation of UAM, considering the diverse interests of all stakeholders. Furthermore, our research reveals differences in risk perception between the two groups. Specifically, we found that the participants in the study on residents' perceptions felt less comfortable using UAM and **less** concerned about the infrastructure needed for UAM. This is interesting as non-users are disproportionately affected by these infrastructure changes. It seems that not having the experience of using the UAM in the study led to reduced risk perception for the infrastructural changes. Therefore, we argue that to study UAM holistically, both perspectives have to be considered simultaneously, especially in the design of infrastructure. This holistic approach is crucial for balancing the varied and sometimes conflicting interests involved in UAM's societal integration.

7.3 Acceptance of a New Form of Mobility

To our knowledge, this work is the first to consider the perspectives of UAM passengers and residents affected by UAM alike. Therefore, we could address the question of whether UAM could be accepted as a new form of mobility. Our data suggests that if key factors like cost [4] are addressed, UAM will likely gain acceptance. We found high levels of perceived safety, even in scenarios considered dangerous. Furthermore, high acceptance persisted even when simulating extraordinary UAM conditions. Our results also show initial acceptance in various urban and suburban settings by residents (see RQ3). As these findings are based on a rather short simulation, they need to be verified in a longitudinal and more externally valid setting; however, the initial exposure did not lead to neophobia or fear of the new mobility form. Whether these insights hold true in various cultural and socio-demographic contexts and over longer durations should be evaluated in future work.

7.4 Methodology Considerations

Studying the impact of unreleased devices or, in our case, mobility concepts is difficult and potentially dangerous. Usually, researchers use envisioned scenarios (thus avoiding the need to visualize a future), sketches, or higher fidelity approaches such as simulations. Also, the visualization method (e.g., VR or monitor-based) can be varied [28]. In our study, we opted for a Unity-based visualization of New York City to explore specific scenarios of interest. While we acknowledge the limitations of reduced external validity inherent to online video-based studies (as detailed in subsection 7.6), we chose this approach over VR to maximize participant diversity, which is evident, for example, by the wide variety of participants' ages [40]. Focusing on U.S.-based participants was a strategic decision, informed by the country's advanced planning in UAM airspace concepts and its role in technological leadership. This focus also made the densely populated urban landscape of New York a fitting context for our study [5]. Although the findings of our gaze data were inconclusive, we contend that its integration encouraged participants to pay closer attention to the videos. However, our remote online study did not permit strict control over the proper use of the gaze tracker. To improve our approach, future studies could include a validation process for the calibration phase to ensure that calibration is completed successfully. In addition, incorporating real-time feedback during the calibration phase could alert participants to problems with tracking quality, allowing for immediate correction. Further, alternative methods for remote gaze-data collection, such as those proposed by Wagner and Colley et al. [71], might have yielded more precise results. However, this tool requires external libraries and, therefore, lacks accessibility. To further the field, we plan to open-source our gaze-tracking solution, offering a resource for future researchers in this area.

7.5 Practical Implications

In line with previous work [12], we found visualizations to improve acceptance and trust for passengers. This also holds true for non-standard and potentially dangerous scenarios. Participants' feedback indicates that the information on the display inside the air taxi is most important as the absence of the display reduced the perceived safety during adverse lighting and weather conditions and the SA during emergency situations. This indicates that the display is more important than the projected future path on the WSD; therefore, the projected trajectory via WSD proposed by Colley and Meinhardt et al. [12] could be considered an "add-on" feature rather than essential. While WSDs are expected to improve the user experience in the automotive domain [58], simple displays inside the air taxi can already increase trust. Therefore, WSDs might not be necessary at all.

Based on our study, we also expect UAM to be acceptable to residents. Especially in urban environments, the additional noise and visual pollution were not perceived as too negative. However, UAM still raised concerns about noise, privacy, and visual pollution. It will be crucial to provide sufficient levels of noise reduction and to ensure the public's right to privacy. One solution is to use predefined aerial corridors [5] redirecting UAM. This would allow air taxis to avoid flying directly over residential areas (especially suburban areas) or operate at higher altitudes. However, as operating at higher altitudes would reduce energy/fuel efficiency, an appropriate trade-off must be found, also considering sustainability aspects. Finally, to protect residents' privacy, computation should happen mainly (or completely) on the air taxi.

7.6 Limitations and Future Work

Our video-based online study's lack of real-world risks reduces its external validity and generalizability. Participants experienced the scenarios in a controlled, consequence-free environment without actual dangers. As a result, they may not have had the same emotional and psychological responses that would occur in real-life situations involving potential emergencies. This absence of genuine danger and physical immersion likely influenced how participants perceived safety and trust. Therefore, our findings may not fully apply to real-world contexts. Additionally, the immersion in our study might be reduced compared to setups such as virtual reality or motion chairs [62]. Hence, future research should include user studies with motion simulations to accurately replicate air taxi flights, including vibrations [14] and wind gust movements. However, our approach of using 360deg videos in an online study ensured a high diversity of participants, including those who do not own VR glasses, bridging the gap between VR studies and ordinary video studies. While including participants with some experience in air taxi flights, such as helicopter pilots, could have increased ecological validity [43], UAM is expected to be used for public transportation. Therefore, having a diverse participant pool is advantageous. Despite these limitations, the TUI score for immersion was relatively high in all three studies (approx. 20 on a scale of 4-28). While the number of participants was sufficient to avoid Type II errors, as calculated by a priori analysis (see subsection 3.5), and included people of different ages, an even wider range of people with different backgrounds needs to be considered to achieve greater generalizability of the results.

8 CONCLUSION

This research provides insights into passengers' and residents' perspectives on automated Urban Air Mobility. Conducted through three online studies incorporating 360deg video (N_1=31, N_2=29, and N_3=29), our work addresses key gaps in existing literature for extraordinary flight scenarios and the perspectives of urban and suburban residents. We focused on adverse lighting and weather conditions, emergency situations, and residents' perceptions in the city and the suburbs. Our findings reveal that the perceived safety and trust of UAM is high among both passengers and residents. However, passengers perceived specific scenarios like foggy conditions and rotor breakdowns as particularly critical. Introducing a display with additional information, such as speed and altitude, significantly improved perceived safety and SA, especially during these critical scenarios. Interestingly, while the future trajectory visualization was noticed, we consider it more as an "add-on" rather than a core feature for enhancing SA and perceived safety. From the residents' perspective, our study shows diversity towards UAM. While there is general acceptance, privacy and noise pollution concerns were evident, particularly among suburban residents, those situated outdoors, and individuals at higher elevations. This suggests that while UAM is primarily viewed as a positive development, there are specific conditions under which public opinion may shift. This emphasizes the need for a holistic approach that considers the perspectives of passengers and residents, particularly in the planning and design of UAM infrastructure.

OPEN SCIENCE

The theme for LimeSurvey using the `WebGazer.JS` [50] plugin together with all necessary instructions is available via https://github.com/Max-Raed/LimeGazer. We will provide the scenarios to interested researchers.

REFERENCES

[1] 3D Paw. 2022. *Coffee Shop interior*. 3D Paw. https://assetstore.unity.com/packages/3d/props/interior/coffee-shop-interior-209622

[2] European Union Aviation Safety Agency. 2021. *Study on the societal acceptance of Urban Air Mobility in Europe*. European Union Aviation Safety Agency. Retrieved March 11, 2022 from https://www.easa.europa.eu/sites/default/files/dfu/uam-full-report.pdf

[3] Christelle Al Haddad, Emmanouil Chaniotakis, Anna Straubinger, Kay Plötner, and Constantinos Antoniou. 2020. Factors affecting the adoption and use of urban air mobility. *Transportation Research Part A: Policy and Practice* 132 (2020), 696–712. https://doi.org/10.1016/j.tra.2019.12.020

[4] Patricia B Appel, Henrike Böck, and Andreas Riener. 2023. What would you expect? Insights from a customer journey analysis of passenger drone flights. , 10 pages.

[5] Aleksandar Bauranov and Jasenka Rakas. 2021. Designing airspace for urban air mobility: A review of concepts and approaches. *Progress in Aerospace Sciences* 125 (2021), 100726. https://doi.org/10.1016/j.paerosci.2021.100726

[6] S. Brandenburg and E. M. Skottke. 2014. Switching from manual to automated driving and reverse: Are drivers behaving more risky after highly automated driving?. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, New York, NY, USA, 2978–2983. https://doi.org/10.1109/ITSC.2014.6958168

[7] Gloria Calhoun, Heath Ruff, Chad Breeden, Joshua Hamell, Mark Draper, and Christopher Miller. 2013. Multiple Remotely Piloted Aircraft Control: Visualization and Control of Future Path. In *Virtual, Augmented and Mixed Reality. Systems and Applications*, Randall Shumaker (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 231–240.

[8] Eric T. Chancey and Michael S. Politowicz. 2020. Public Trust and Acceptance for Concepts of Remotely Operated Urban Air Mobility Transportation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1 (2020), 1044–1048. https://doi.org/10.1177/1071181320641251

[9] Junming Chen, Haruya Ishikawa, and Olli Etuaho. 2023. *Equilib*. https://github.com/haruishi43/equilib#

[10] Mark Colley, Benjamin Eder, Jan Ole Rixen, and Enrico Rukzio. 2021. Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 155, 11 pages. https://doi.org/10.1145/3411764.3445351

[11] Mark Colley, Svenja Krauss, Mirjam Lanzer, and Enrico Rukzio. 2021. How Should Automated Vehicles Communicate Critical Situations? A Comparative Analysis of Visualization Concepts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 94 (sep 2021), 23 pages. https://doi.org/10.1145/3478111

[12] Mark Colley, Luca-Maxim Meinhardt, Alexander Fassbender, Michael Rietzler, and Enrico Rukzio. 2023. Come Fly With Me - Investigating the Effects of Path Visualizations in Automated Urban Air Mobility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., IMWUT* 7, 2 (2023), 52. https://doi.org/10.1145/3596249

[13] Mark Colley, Max Rädler, Jonas Glimmann, and Enrico Rukzio. 2022. Effects of Scene Detection, Scene Prediction, and Maneuver Planning Visualizations on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 49 (jul 2022), 21 pages. https://doi.org/10.1145/3534609

[14] Thomas Edwards and George Price. 2020. eVTOL Passenger Acceptance. , 27 pages. https://ntrs.nasa.gov/citations/20200000532 Contractor Report.

[15] EHang. 2020. *The future of Transportation: White Paper on Urban Air Mobility Systems*. EHang. https://www.ehang.com/app/en/EHang%20White%20Paper%20on%20Urban%20Air%20Mobility%20Systems.pdf

[16] Ahmed Elmezeny, Nina Edenhofer, and Jeffrey Wimmer. 2018. Immersive Storytelling in 360-Degree Videos: An Analysis of Interplay Between Narrative and Technical Immersion. *Journal For Virtual Worlds Research* 11, 1 (April 2018). https://doi.org/10.4101/jvwr.v11i1.7298

[17] Mica R Endsley, Stephen J Selcon, Thomas D Hardiman, and Darryl G Croft. 1998. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 42. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications, Los Angeles, CA, USA, 82–86.

[18] Stefanie M. Faas, Andrea C. Kao, and Martin Baumann. 2020. A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle – Pedestrian Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376484

[19] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[20] Finward Studios. 2021. *Suburb Neighborhood House Pack (Modular)*. Finward Studios. https://assetstore.unity.com/packages/3d/environments/urban/suburb-neighborhood-house-pack-modular-72712

[21] David C. Funder and Daniel J. Ozer. 2019. Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science* 2, 2 (2019), 156–168. https://doi.org/10.1177/2515245919847202 arXiv:https://doi.org/10.1177/2515245919847202

[22] David Galway, Jason Etele, and Giovanni Fusina. 2008. Modeling of the Urban Gust Environment with Application to Autonomous Flight. In *AIAA Atmospheric Flight Mechanics Conference and Exhibit*. American Institute of Aeronautics and Astronautics, Reston, Virigina. https://doi.org/10.2514/6.2008-6565

[23] Lilium GmbH. 2020. *Lilium agrees partnership with Dusseldorf and Cologne/Bonn airports*. Lilium GmbH. Retrieved May 20, 2023 from https://lilium.com/newsroom-detail/lilium-partnership-dusseldorf-cologne

[24] Volocopter GmbH. 2022. *Volocopter Conducts First Crewed eVTOL Flight in France*. Volocopter GmbH. Retrieved May 20, 2022 from https://www.volocopter.com/newsroom/first-crewed-evtol-flight/

[25] Julien Haan, Laurie A. Garrow, Aude Marzuoli, Satadru Roy, and Michel Bierlaire. 2021. Are commuter air taxis coming to your city? A ranking of 40 cities in the United States. *Transportation Research Part C: Emerging Technologies* 132 (2021), 103392. https://doi.org/10.1016/j.trc.2021.103392

[26] Manfred Hader. 2018. *Urban air mobility poised to become a fast-growing new market*. Roland Berger. Retrieved June 3, 2022 from https://www.rolandberger.com/en/Insights/Publications/Passenger-drones-ready-for-take-off.html

[27] Sandra G. Hart. 1988. Helicopter Human Factors. In *Human Factors in Aviation*. Elsevier, Amsterdam, The Netherlands, 591–638. https://doi.org/10.1016/B978-0-08-057090-7.50024-2

[28] Philipp Hock, Mark Colley, Ali Askari, Tobias Wagner, Martin Baumann, and Enrico Rukzio. 2022. Introducing VAMPIRE – Using Kinaesthetic Feedback in Virtual Reality for Automated Driving Experiments. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Seoul, Republic of Korea) *(AutomotiveUI '22)*. Association for Computing Machinery, New York, NY, USA, 204–214. https://doi.org/10.1145/3543174.3545252

[29] Philipp Hock, Johannes Kraus, Marcel Walch, Nina Lang, and Martin Baumann. 2016. Elaborating Feedback Strategies for Maintaining Automation in Highly Automated Driving. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) *(Automotive'UI 16)*. Association for Computing Machinery, New York, NY, USA, 105–112. https://doi.org/10.1145/3003715.3005414

[30] Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors* 57, 3 (2015), 407–434.

[31] Brittany E. Holthausen, Philipp Wintersberger, Bruce N. Walker, and Andreas Riener. 2020. Situational Trust Scale for Automated Driving (STS-AD): Development and Initial Validation. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) *(AutomotiveUI '20)*. Association for Computing Machinery, New York, NY, USA, 40–47. https://doi.org/10.1145/3409120.3410637

[32] Hovl Studio. 2022. *3D Fire and Explosions*. Hovl Studio. https://assetstore.unity.com/packages/vfx/particles/fire-explosions/3d-fire-and-explosions-176981

[33] Antonio Imbruglia, Daniela Cancila, and Marina Settembre. 2023. 5G Communication and Security in Connected Vehicles. *Ada Lett.* 42, 2 (apr 2023), 109–113. https://doi.org/10.1145/3591335.3591351

[34] INIRROR. 2020. *Office Interior Archviz*. INIRROR. https://assetstore.unity.com/packages/3d/environments/urban/office-interior-archviz-155701

[35] Young Woo Kim, Cherin Lim, Seul Chan Lee, Sol Hee Yoon, and Yong Gu Ji. 2021. The 1st Workshop on User Experience in Urban Air Mobility: Design Considerations and Issues. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 175–177. https://doi.org/10.1145/3473682.3477440

[36] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, Sebastiano Bagnara, Riccardo Tartaglia, Sara Albolino, Thomas Alexander, and Yushi Fujita (Eds.). Springer International Publishing, Cham, 13–30.

[37] Oswald Kothgassner, A Felnhofer, N Hauk, E Kastenhofer, J Gomm, and I Krysprin-Exner. 2013. Technology Usage Inventory. https://www.ffg.at/sites/default/files/allgemeine_downloads/thematische%20programme/programmdokumente/tui_manual.pdf. *Manual. Wien: ICARUS* 17, 04 (2013), 90. [Online; accessed: 05-JULY-2020].

[38] Cherin Lim, Young Woo Kim, Yong Gu Ji, Solhee Yoon, and Seul Chan Lee. 2022. Is This Flight Headed Downtown? : User Experience Considerations for Urban Air Mobility. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, Simone Barbosa, Cliff Lampe, Caroline Appert, and David A. Shamma (Eds.). ACM, New York, NY, USA, 1–7. https://doi.org/10.1145/3491101.3519852

[39] Yucong Lin and Srikanth Saripalli. 2017. Sampling-Based Path Planning for UAV Collision Avoidance. *IEEE Transactions on Intelligent Transportation Systems* 18, 11 (2017), 3179–3192. https://doi.org/10.1109/TITS.2017.2673778

[40] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 143, 14 pages. https://doi.org/10.1145/3411764.3445488

[41] Guangzhou EHang Intelligent Technology Co. Ltd. 2021. *EHang Joins European Union's AMU-LED Project to Demonstrate Urban Air Mobility*. EHang. Retrieved May 20, 2022 from https://www.ehang.com/news/728.html

[42] Walter Lucia, Giuseppe Franzè, and Mario Sznaier. 2020. A Hybrid Command Governor Scheme for Rotary Wings Unmanned Aerial Vehicles. *IEEE Transactions on Control Systems Technology* 28, 2 (2020), 361–375. https://doi.org/10.1109/TCST.2018.2880936

[43] Luca-Maxim Meinhardt, Mark Colley, Alexander Faßbender, Michael Rietzler, and Enrico Rukzio. 2023. Up, Up and Away - Investigating Information Needs for Helicopter Pilots in Future Urban Air Mobility. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 304, 6 pages. https://doi.org/10.1145/3544549.3585643

[44] Natasha Merat, A. Hamish Jamson, Frank C.H. Lai, Michael Daly, and Oliver M.J. Carsten. 2014. Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour* 27 (2014), 274 – 282. https://doi.org/10.1016/j.trf.2014.09.005

[45] Dev Minotra and Karen M. Feigh. 2020. An Analysis of Cognitive Demands in Ship-Based Helicopter-Landing Maneuvers. *Journal of the American Helicopter Society* 65, 4 (2020), 1–11. https://doi.org/10.4050/JAHS.65.042009

[46] Tobias Müller, Mark Colley, Gülsemin Dogru, and Enrico Rukzio. 2022. AR4CAD: Creation and Exploration of a Taxonomy of Augmented Reality Visualization for Connected Automated Driving. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 177 (sep 2022), 27 pages. https://doi.org/10.1145/3546712

[47] Thomas Münsterer, Tobias Schafhitzel, Michael Strobel, Philipp Völschow, Stephanus Klasen, and Ferdinand Eisenkeil. 2014. Sensor-enhanced 3D conformal cueing for safe and reliable HC operation in DVE in all flight phases. In *Degraded Visual Environments: Enhanced, Synthetic, and External Vision Solutions 2014 (SPIE Proceedings)*, Jeff J. Güell and Jack Sanders-Reed (Eds.). SPIE, Baltimore, Maryland, USA, 90870I. https://doi.org/10.1117/12.2050377

[48] Craig Murray, Murray Ireland, and David Anderson. 2014. On the response of an autonomous quadrotor operating in a turbulent urban environment. *AUVSI Unmanned Systems 2014* 2.

[49] OliVR. 2021. *VR Panorama 360 PRO Renderer*. OliVR. https://assetstore.unity.com/packages/tools/video/vr-panorama-360-pro-renderer-35102

[50] Alexandra Papoutsaki, Aaron Gokaslan, James Tompkin, Yuze He, and Jeff Huang. 2018. The eye of the typer: a benchmark and analysis of gaze behavior during typing.. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications (ETRA)*. ACM, New York, NY, USA, 16–1.

[51] Bob Pishue. 2021. 2021 INRIX global traffic scorecard.

[52] Erich J. Plate and Harald Kiefer. 2001. Wind loads in urban areas. *Journal of Wind Engineering and Industrial Aerodynamics* 89, 14-15 (2001), 1233–1256. https://doi.org/10.1016/S0167-6105(01)00159-3

[53] A. Pukhova, C. Llorca, A. Moreno, C. Staves, Q. Zhang, and R. Moeckel. 2021. Flying taxis revived: Can Urban air mobility reduce road congestion? *Journal of Urban Mobility* 1 (2021), 100002. https://doi.org/10.1016/j.urbmob.2021.100002

[54] Suchithra Rajendran and Jake Shulman. 2020. Study of emerging air taxi network operation using discrete-event systems simulation approach. *Journal of Air Transport Management* 87 (2020), 101857. https://doi.org/10.1016/j.jairtraman.2020.101857

[55] Suchithra Rajendran and Joshua Zack. 2019. Insights on strategic air taxi network infrastructure locations using an iterative constrained clustering approach. *Transportation Research Part E: Logistics and Transportation Review* 128 (2019), 470–505. https://doi.org/10.1016/j.tre.2019.06.003

[56] Subramanian Ramasamy, Roberto Sabatini, and Alessandro Gardi. 2014. Avionics sensor fusion for small size unmanned aircraft Sense-and-Avoid. In *2014 IEEE Metrology for Aerospace (MetroAeroSpace)*. IEEE, New York, NY, USA, 271–276. https://doi.org/10.1109/MetroAeroSpace.2014.6865933

[57] Jorge Reyna. 2018. The Potential of 360-degree Videos for Teaching, Learning and Research. Valencia, Spain, 1448–1454. https://doi.org/10.21125/inted.2018.0247

[58] Andreas Riegler, Philipp Wintersberger, Andreas Riener, and Clemens Holzmann. 2019. Augmented Reality Windshield Displays and Their Potential to Enhance User Experience in Automated Driving. *i-com* 18, 2 (2019), 127–149. https://doi.org/10.1515/icom-2018-0033

[59] Jan Ole Rixen, Teresa Hirzle, Mark Colley, Yannick Etzel, Enrico Rukzio, and Jan Gugenheimer. 2021. Exploring Augmented Visual Alterations in Interpersonal Communication. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 730, 11 pages. https://doi.org/10.1145/3411764.3445597

[60] Katarzyna Samson and Patrycjusz Kostyszyn. 2015. Effects of cognitive load on trusting behavior–an experiment using the trust game. *PloS one* 10, 5 (2015), e0127680.

[61] S. S. Shapiro and M. B. Wilk. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52, 3/4 (1965), 591–611. http://www.jstor.org/stable/2333709

[62] YAW Motion Simulator. 2023. *Virtual Reality Motion Simulator - Yaw VR*. YAW Motion Simulator. Retrieved Jul 19, 2024 from https://www.yawvr.com/

[63] Katta Spiel, Oliver L. Haimson, and Danielle Lottridge. 2019. How to do better with gender on surveys: a guide for HCI researchers. *Interactions* 26, 4, 62–65. https://doi.org/10.1145/3338283

[64] Daniel Szafir, Bilge Mutlu, and Terry Fong. 2015. Communicating Directionality in Flying Robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (Portland, Oregon, USA) *(HRI '15)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/2696454.2696475

[65] Richard M Taylor. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness*. Routledge, Abingdon, UK, 111–128.

[66] Unity Technologies. 2019. *Unity*. Unity Technologies. https://unity.com/

[67] Unluck Software. 2021. *Bird Flock Bundle*. Unluck Software. https://assetstore.unity.com/packages/3d/characters/animals/birds/bird-flock-bundle-25576

[68] John Villasenor. 2013. Observations from above: unmanned aircraft systems and privacy. *Harv. JL & Pub. Pol'y* 36 (2013), 457.

[69] Volocopter GmbH. 2021. *The roadmap to scalable urban air mobility*. Volocopter GmbH. Retrieved June 3, 2023 from https://static1.squarespace.com/static/5d27bb3e330ac30001dc14fd/t/6165c265c89c61365aa517f5/1634058869091/Volocopter+v2+Roadmap+to+Scalable+UAM.pdf

[70] Tim Waanders, T. Münsterer, and M. Kress. 2013. Sensor supported pilot assistance for helicopter flight in DVE. In *Degraded Visual Environments: Enhanced, Synthetic, and External Vision Solutions 2013 (SPIE Proceedings)*, Kenneth L. Bernier and Jeff J. Güell (Eds.). SPIE, Baltimore, Maryland, USA, 873704. https://doi.org/10.1117/12.2015783

[71] Tobias Wagner, Mark Colley, Daniel Breckel, Michael Kösel, and Enrico Rukzio. 2024. UnitEye: Introducing a User-Friendly Plugin to Democratize Eye Tracking Technology in Unity Environments. In *Proceedings of Mensch Und Computer 2024 (MuC '24)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3670653.3670655

[72] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating Robot Motion Intent with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI '18)*. Association for Computing Machinery, New York, NY, USA, 316–324. https://doi.org/10.1145/3171221.3171253

[73] Annette Werner. 2018. New Colours for Autonomous Driving: An Evaluation of Chromaticities for the External Lighting Equipment of Autonomous Vehicles. https://doi.org/10.25538/tct.v0i1.692 Number: 1.

[74] Philipp Wintersberger, Tamara von Sawitzky, Anna-Katharina Frison, and Andreas Riener. 2017. Traffic Augmentation as a Means to Increase Trust in Automated Driving Systems. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter* (Cagliari, Italy) *(CHItaly '17)*. Association for Computing Machinery, New York, NY, USA, Article 17, 7 pages. https://doi.org/10.1145/3125571.3125600

[75] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

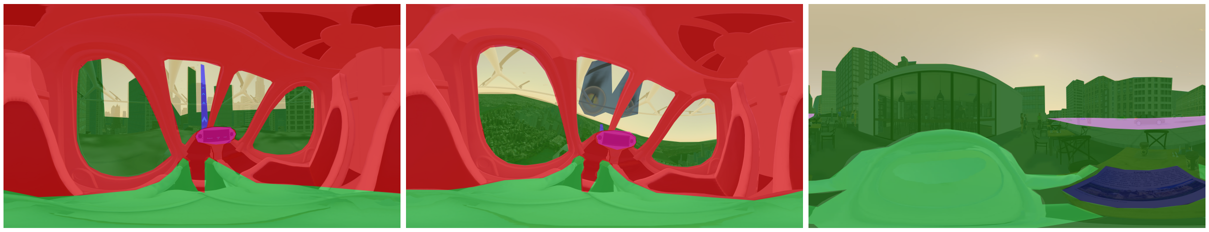
[76] Marcel Woide, Mark Colley, Nicole Damm, and Martin Baumann. 2022. Effect of System Capability Verification on Conflict, Trust, and Behavior in Automated Vehicles. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Seoul, Republic of Korea) *(AutomotiveUI '22)*. Association for Computing Machinery, New York, NY, USA, 119–130. https://doi.org/10.1145/3543174.3545253

[77] Pavan Yedavalli and Jessie Mooberry. 2019. An assessment of public perception of urban air mobility (UAM). , 2046738072–1580045281 pages.

[78] Chenyan Zhang, Aud Sissel Hoel, Andrew Perkis, and Saman Zadtootaghaj. 2018. How Long is Long Enough to Induce Immersion?. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. 1–6. https://doi.org/10.1109/QoMEX.2018.8463397

APPENDIX

A GAZE TRACKING EVALUATION



(a) Study on adverse lighting and weather conditions.

(b) Study on different emergency situations.

(c) Study on resident perceptions - café scene.

Fig. 11. Examples of segmentation masks per study. The figure shows the equirectangular videos.

To enable remote gaze-tracking for the 360deg videos, yaw and pitch values were recorded to capture the participants' field-of-view (FOV) movements. Utilizing this data, each participant's FOV was reconstructed from the equirectangular video using the Python library Equilib [9].

Simultaneously, we integrated `WebGazer.JS` [50] into LimeSurvey and embedded each 360deg video into the survey via *iFrames*. The gaze data was then matched with the corresponding segmentation mask of each video on a frame-by-frame basis, identifying the relevant object class under focus. By considering the yaw and pitch values for each frame, we accounted for the movement of the field of view inherent in 360deg videos, ensuring accurate identification of the correct AOIs.

The segmentation masks were manually annotated using Adobe After Effects (version 23.3.0). A color-coding scheme was established to represent the relevant object classes (air taxi interior, air taxi path, display, sky, surroundings, explosion, flock of birds, newspaper, road). These colors were subsequently aligned with the gaze points. It should be noted that not all object classes occurred in every video. The segmentation masks for one exemplary video per study can be seen in Figure 11.

Finally, the *percentage fixation* was calculated by dividing the number of fixations per object class by the total number of eye gaze logs per participant.

B STUDY ON ADVERSE LIGHTING AND WEATHER CONDITIONS

B.1 Descriptive Data

Variable	Levels	Min	q_1	$\widetilde{x}$	$\bar{x}$	q_3	Max	s	IQR
perceived safety	Daytime	-3.00	-0.75	1.00	0.66	2.00	3	1.80	2.75
	Fog	-3.00	-1.25	0.75	0.45	2.00	3	1.84	3.25
	Night	-3.00	-0.50	1.00	0.84	2.50	3	1.77	3.00
	all	-3.00	-0.75	1.00	0.65	2.00	3	1.80	2.75
cognitive load	Daytime	1.00	3.00	10.00	9.55	15.00	20	5.66	12.00
	Fog	1.00	4.00	10.00	9.59	14.00	20	5.98	10.00
	Night	1.00	3.00	10.00	9.61	15.00	20	5.84	12.00
	all	1.00	3.00	10.00	9.58	15.00	20	5.81	12.00
Trust	Daytime	1.00	3.00	3.50	3.49	4.00	5	1.01	1.00
	Fog	1.00	3.00	4.00	3.49	4.00	5	1.06	1.00
	Night	1.00	3.00	4.00	3.60	4.00	5	1.09	1.00
	all	1.00	3.00	4.00	3.53	4.00	5	1.05	1.00
understanding	Daytime	1.00	3.00	3.75	3.61	4.50	5	0.96	1.50
	Fog	1.50	3.00	3.50	3.48	4.00	5	0.85	1.00
	Night	1.75	3.00	3.50	3.66	4.25	5	0.87	1.25
	all	1.00	3.00	3.50	3.58	4.25	5	0.90	1.25
Situation awareness	Daytime	4.00	12.00	15.00	15.85	19.00	36	6.63	7.00
	Fog	-4.00	11.00	14.00	15.23	19.00	39	7.47	8.00
	Night	3.00	12.00	15.00	16.14	19.00	36	6.37	7.00
	all	-4.00	11.50	15.00	15.74	19.00	39	6.83	7.50

Table 1. table of scores for the v on resident perceptions grouped by scenario.

B.2 Eye Gaze Data

While multiple main effects were observed, a significant interaction effect subsumed these, suggesting that the combined influence of the variables is more impactful than their individual contributions, i.e., we found a significant interaction effect of *Scenario* × *Visualization* × area of interest (AOI) fixation ($F(16, 480) = 8.84$, $p<0.001$; see Figure 12).

In general, participants rarely looked at the interior of the air taxi. However, in conditions with additional information displayed on the screen (path with display and bounding box), the sky garnered more attention. Conversely, in non-daytime scenarios, participants focused less on the environment. Further, the path, which was consistently visible in all conditions, received relatively little visual attention.

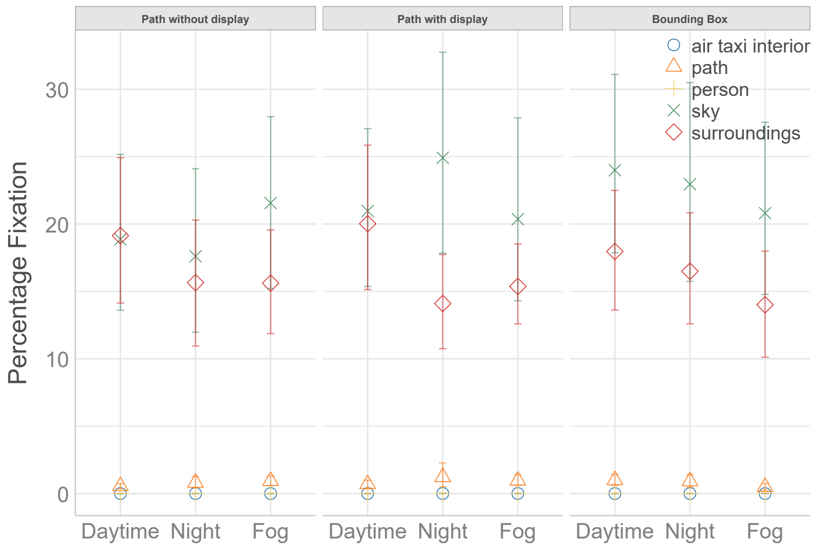


Fig. 12. **Study on Adverse Lighting and Weather Conditions:** Interaction effect *Scenario* × *Visualization* × *AOI* on AOI fixation

C STUDY ON EMERGENCY SITUATIONS

C.1 Descriptive Data

Variable	Levels	Min	q_1	$\widetilde{x}$	$\bar{x}$	q_3	Max	s	IQR
perceived safety	Bird Flock	-3.00	0.00	1.50	1.13	2.62	3	1.69	2.62
	Rotor Breakdown	-3.00	-1.75	0.00	-0.17	1.25	3	1.94	3.00
	Wind Gust	-2.75	-0.25	1.00	0.82	2.25	3	1.72	2.50
	all	-3.00	-0.75	0.75	0.60	2.00	3	1.86	2.75
cognitive load	Bird Flock	1.00	3.00	5.00	7.46	12.50	20	5.98	9.50
	Rotor Breakdown	1.00	5.00	10.00	10.45	16.00	20	6.42	11.00
	Wind Gust	1.00	3.00	6.00	7.83	12.00	20	5.92	9.00
	all	1.00	3.00	7.00	8.58	14.00	20	6.23	11.00
Trust	Bird Flock	1.00	3.00	4.00	3.51	4.00	5	1.16	1.00
	Rotor Breakdown	1.00	2.00	3.50	3.09	4.00	5	1.26	2.00
	Wind Gust	1.00	2.00	4.00	3.36	4.00	5	1.21	2.00
	all	1.00	2.00	4.00	3.32	4.00	5	1.22	2.00
understanding	Bird Flock	1.00	3.00	3.75	3.56	4.25	5	0.99	1.25
	Rotor Breakdown	1.00	2.75	3.50	3.35	4.25	5	1.06	1.50
	Wind Gust	1.50	2.75	3.75	3.52	4.38	5	0.97	1.62
	all	1.00	3.00	3.50	3.48	4.25	5	1.01	1.25
Situation awareness	Bird Flock	-6.00	9.50	14.00	14.30	19.00	30	6.97	9.50
	Rotor Breakdown	-9.00	7.00	11.00	11.18	16.00	29	7.09	9.00
	Wind Gust	-2.00	10.00	15.00	14.20	18.00	26	6.01	8.00
	all	-9.00	9.00	13.00	13.23	18.00	30	6.83	9.00

Table 2. table of scores for the study on resident perceptions grouped by scenario.

C.2 Eye Gaze Data

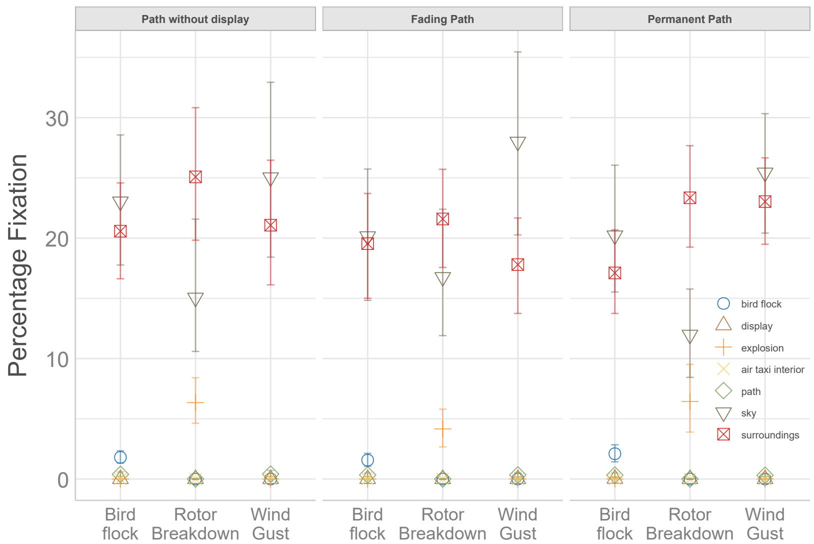


Fig. 13. **Study on Emergency Situations:** Interaction effect *Scenario* × *Visualization* × *AOI* on AOI fixation

While multiple main effects were observed, a significant interaction effect subsumed these, suggesting that the combined influence of the variables is more impactful than their individual contributions, i.e., the ART found a significant interaction effect of *Scenario* × *Visualization* × *AOI* on AOI fixation $F(24, 696) = 14.10$, p<0.001; see Figure 13).

D STUDY ON RESIDENT PERCEPTIONS

D.1 Descriptive Data

Variable	Levels	Min	q_1	$\widetilde{x}$	$\bar{x}$	q_3	Max	s	IQR
perceived safety	Café Inside	-3	0.12	2.25	1.44	3.00	3	2.01	2.88
	Café Outside	-3	0.00	2.00	1.37	3.00	3	2.03	3.00
	Office	-3	0.00	2.50	1.43	3.00	3	2.02	3.00
	Suburb	-3	0.00	2.00	1.32	3.00	3	1.98	3.00
	all	-3	0.00	2.12	1.39	3.00	3	2.00	3.00
Noise concerns	Café Inside	1	1.00	1.00	1.72	2.00	5	1.21	1.00
	Café Outside	1	1.00	1.00	1.89	2.50	5	1.27	1.50
	Office	1	1.00	1.00	2.08	3.00	5	1.46	2.00
	Suburb	1	1.00	2.00	2.51	4.00	5	1.64	3.00
	all	1	1.00	1.00	2.05	3.00	5	1.43	2.00
concerns about the altitude of UAM	Café Inside	1	1.00	1.00	1.66	2.00	5	1.10	1.00
	Café Outside	1	1.00	1.00	1.82	2.00	5	1.24	1.00
	Office	1	1.00	1.00	2.01	3.00	5	1.45	2.00
	Suburb	1	1.00	1.00	2.02	3.00	5	1.32	2.00
	all	1	1.00	1.00	1.88	3.00	5	1.29	2.00
privacy concerns	Café Inside	1	1.00	1.00	1.61	2.00	5	1.09	1.00
	Café Outside	1	1.00	1.00	1.72	2.00	5	1.16	1.00
	Office	1	1.00	1.00	1.80	2.00	5	1.21	1.00
	Suburb	1	1.00	2.00	2.20	3.00	5	1.42	2.00
	all	1	1.00	1.00	1.83	2.00	5	1.24	1.00
visual pollution concerns	Café Inside	1	1.00	1.00	2.07	3.00	5	1.47	2.00
	Café Outside	1	1.00	2.00	2.26	4.00	5	1.50	3.00
	Office	1	1.00	2.00	2.25	4.00	5	1.50	3.00
	Suburb	1	1.00	2.00	2.63	4.00	5	1.69	3.00
	all	1	1.00	2.00	2.30	4.00	5	1.55	3.00
comfort	Café Inside	1	4.00	5.00	4.25	5.00	5	0.99	1.00
	Café Outside	1	4.00	4.00	4.07	5.00	5	1.11	1.00
	Office	1	3.00	5.00	4.14	5.00	5	1.08	2.00
	Suburb	1	3.00	4.00	3.89	5.00	5	1.30	2.00
	all	1	3.00	4.50	4.09	5.00	5	1.13	2.00

Table 3. Table of scores for the study on resident perceptions grouped by scenario.

D.2 Eye Gaze Data

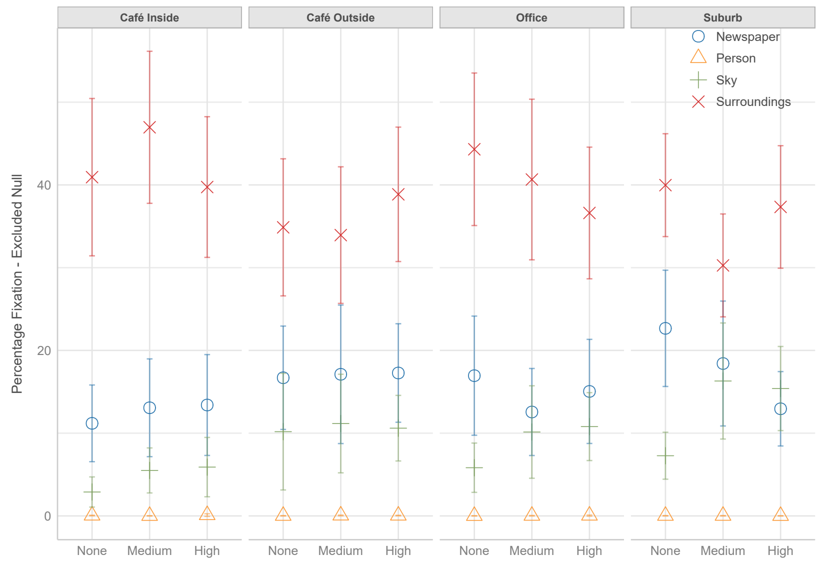


Fig. 14. **Study on Resident Perceptions:** Interaction effect *Scenario* × *Visualization* × *AOI* on AOI fixation

The ART found a significant interaction effect of *Scenario* × *number of air taxis* × *AOI* on AOI fixation ($F(18, 501) = 3.12$, $p<0.001$; see Figure 14). Participants did not show consistent behavior over the different conditions besides not looking at their own character. The surroundings were the part that was most looked at, followed consistently by the newspaper. However, the relative frequencies alter strongly.