

Passenger Experience of Simulated Urban Air Mobility Ride Quality: Responses to Large-Scale Motion

Bernard D. Adelstein, William B. Toscano, Fernando A. Espinosa, Patricia S. Cowings

Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA 94035 USA

Abstract: Twenty-three participants took 10-min solo Urban Air Mobility quadrotor flights as passengers on two separate days in a six-dof large-motion simulator. One flight was in a rotor speed (i.e., RPM) controlled model; the other was under rotor blade pitch (i.e., collective) control. Both were flown in the same modeled turbulence. When ranked across test conditions, the severity of participants’ self-reported simulator sickness symptoms paralleled acceleration-derived predictions of motion sickness likelihood in the following worst-to-best order: 1) RPM control; 2) collective control; and 3) preflight while still on the vertiport pad. Various objective measures revealed potential impacts of flight roughness on the learning of a visuo-manual reaction-time task and on heart and breathing rate indicators of preflight/inflight passenger stress.

Keywords: urban air mobility, ride quality, passenger acceptance, motion sickness, simulator sickness.

1. INTRODUCTION

As envisioned, electrically powered Vertical Take-Off and Landing (eVTOL) aircraft for Urban Air Mobility (UAM) will transport passengers across urban areas on trips ranging between 10 and 70mi (Crown Consulting, 2018, pp. 14 and 18). With departures and landings at vertiports typically located in urban canyons and level flight at relatively low altitudes (1500 to 4000ft) (Deloitte, 2020, p. 26) and relatively slow speeds (40 to 150mph, depending on vehicle category) (Booze-Allen and Hamilton, 2018, p. 88; Crown Consulting, 2018, p.17), UAM passengers may be exposed to substantial turbulence and wind buffeting. While the first generation of these vehicles to go into service is expected to be flown by onboard pilots, the long-term goal is for these aircraft to fly autonomously under the supervision of ground-based monitoring systems and personnel (Deloitte, 2020).

The conceptual block diagram in Figure 1 shows that vehicle ride quality (RQ) is experienced by passenger and, if present, pilot. However, while the passenger and pilot have different performance and comfort needs, the ultimate indicator of UAM industry success will be return paying customers. Beyond simply avoiding motion sickness, a broad demographic of busy passengers will demand an unimpeded ability to read, use personal electronic devices, and engage in conversation. Thus, passengers’ comfort and productivity will significantly influence their acceptance of UAM RQ.

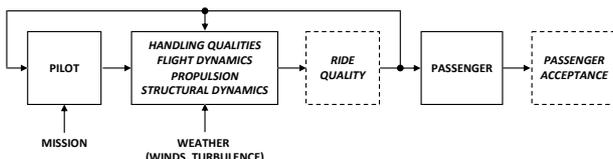


Figure 1. Passenger and pilot ride quality experience.

Contributors to perceived aircraft RQ listed by Griffin (1990, pp. 43-44) can be divided into induced environment (e.g.,

large-scale motion, vibration, noise, temperature, air quality) and habitability (e.g., seating, windows, lighting) factors. NASA simulator studies for commercial jet transport, summarized by Leatherwood et al. (1980), made seminal contributions to our understanding of human tolerance to multi-axis whole-body vibration that underpin current standards (International Standards Organization, 1997). However, none of these earlier simulator studies examined concurrent large-scale vehicle motion, a recognized nauseogenic factor (Golding, 2006a). Further, the applicability of this prior RQ knowledge obtained from large aircraft for vibration (as well as interactions between combined vibration and sound) to the UAM-eVTOL domain has not yet been determined.

Recent analytic and empirical research at NASA has examined handling qualities (HQ) and RQ for a set of in-house reference designs representative of a range UAM industry vehicle configurations (Silva et al., 2018). Malpica and Withrow-Maser (2020) analyzed the relative merits of rotor speed (i.e., RPM) and blade pitch (i.e., collective) control architectures for quadrotor-class concept vehicles variously sized to carry from one to six passengers. Specifically, Withrow-Maser et al. (2021) related electric motor sizing for eVTOL versions of some of the NASA reference designs to HQ. Withrow-Maser et al. (2022) and Aires et al. (2022) reported on a pilot-in-the-loop, moving-base simulator study conducted to empirically validate theoretically-predicted HQ (Schuet et al., 2020) for six-passenger versions of the NASA quadrotor reference vehicle. In the study, test pilots flew quadrotor models embodying collective control and a number of instantiations of RPM control tuned for different levels of heave (i.e., vertical motion) disturbance rejection. Each pilot provided formal HQ ratings as well as assessments of vehicle RQ after flying brief mission task elements and a representative urban landing approach.

Here, we describe a separate follow-on study in which passenger participants experienced whole-mission (i.e., takeoff, level cruising, and landing) “air taxi” flights for Withrow-Maser et

al.’s (2022) quadrotor models operated under RPM and collective control. Our goals in this new study were 1) to assess differences between the two control implementations in terms of RQ from the passenger’s perspective and 2) to examine the experimental utility of a selection of subjective (i.e., rating) and objective (i.e., performance and physiological monitoring) metrics in making these assessments.

2. MATERIALS AND METHODS

2.1 Flight Simulation

Study participants took two simulated eVTOL quadrotor flights at the NASA Ames Research Center’s (ARC’s) Vertical Motion Simulator (VMS) facility. The VMS is capable of delivering six-degree-of-freedom (dof) cab motion—three in rotation plus three with large-scale translation—as well as computer-generated through-the-window views consistent with the cab and modeled vehicle motion. For this study, motion was governed by the modeled vehicle response to pilot inputs and turbulence as shown in Figure 1. In addition, audio cueing representing UAM rotor noise was presented via cabin-mounted speakers to both enhance the sense of immersion and mask external sounds from the VMS motion drives. Sound levels from both sources were further attenuated by the circumaural headphones worn by all simulator riders as part of VMS standard operation procedures for communication with facility staff and the research team at the control room console. No other environment factors, such as seat vibration, were introduced. The cab thermostat was set to $\sim 72^\circ\text{F}$ for the study runs.

Because the study’s two flights were pre-recorded by the ARC test pilot, all participants experienced the exact same flight conditions. One flight was under collective control (henceforth termed COL) and the other under RPM control, which were identical to the COL1 and RPM1 models described in detail by Withrow-Maser et al. (2022). The same Dryden turbulence model (McFarland & Duisenberg, 1995) tuned for minimum airspeed (Aires et al., 2022; Withrow-Maser et al., 2022) was invoked for both flights. The COL controlled model with heave disturbance rejection bandwidth (DRB) greater than 2rad/s and theoretical Level 1 HQ presented a much smoother ride than the RPM version with a 1rad/s heave DRB and borderline Level 1 to 2 HQ (Withrow-Maser et al., 2022).

On its flights, the quadrotor model, depicted in Figure 2, departed southeast from a virtual vertiport on the rooftop (155 ft above mean sea level) of the Fifth and Mission Parking Garage in San Francisco, looped to the south and then returned to land at the same vertiport from the southwest. The COL flight path was 9.38mi long and lasted 10min 12s, while the RPM flight covered 10.52mi and lasted 10min 38s. Both flights cruised at a nominal 60kn and 500ft altitude. Additionally, the Dryden turbulence was further scaled by a multiplying factor that linearly increased from zero at 155ft to 2.5 times higher than Aires et al.’s (2022) levels for altitudes at and above 480ft. While such rescaling had little impact under COL control, as described next, it delivered a sufficiently pronounced vertical motion response to turbulence for the RPM flight condition.

In order to select experiment flight profiles that were not overly provocative, we computed the likelihood of study participants being brought to the point of vomiting in accordance

with guidance in ISO 2631-1, Annex D, “Guide to the effects of vibration on the incidence of motion sickness” (International Standards Organization, 1997). This computation addresses vertical motion between 0.1 and 0.5Hz, the direction and frequency range known to trigger motion sickness (Griffin, 1990, pp. 297-310). The resultant measured acceleration (as opposed to the combined visual-plus-motion simulation commands) in the vertical direction and weighted by the ISO motion-sickness factor, W_i , is plotted in Figure 3 for both flights. The dashed blue line indicates the average weighted acceleration for the whole flight while the red portions denote the intervals above the 0.5m/s^2 level for which the ISO motion sickness guidance is validated. The COL flight’s 0.11m/s^2 root-mean-square (rms) weighted z-axis acceleration corresponds to an estimate that 0.9% (1 in 111) of unadapted men and women will vomit; the RPM profile’s 0.46m/s^2 weighted rms level represents a 3.9% (1 in 25) likelihood of vomiting. In order to mitigate the risk of vomiting, no more than one flight was taken per day with at least two days off between participant’s two flights. Moreover, flight durations were kept to half the 20-min period for which the ISO prediction is validated. Finally, participants were repeatedly and firmly instructed to stop doing any tasks and, if necessary, end their flight should they feel overly nauseated.

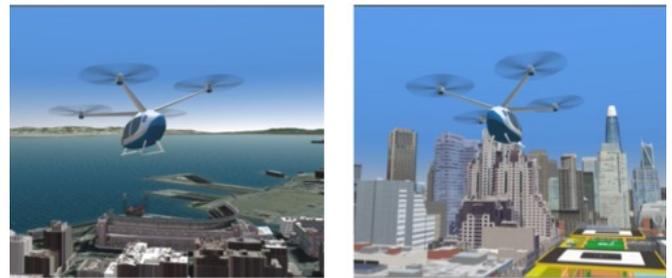


Figure 2. Quadrotor in flight over San Francisco viewed from a virtual chase plane: (left) departing (right) returning to the vertiport.

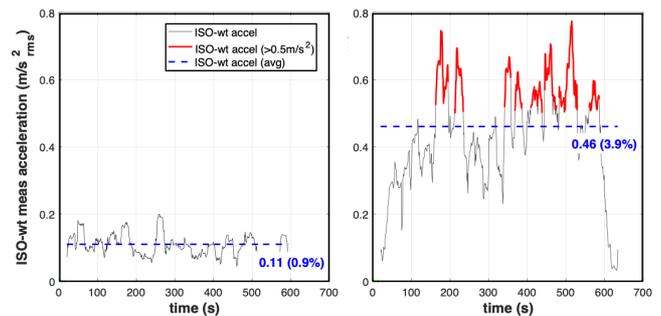


Figure 3. Vertical acceleration profiles for COL (left) and RPM (right) flight after applying ISO motion-sickness weighting factor.

Because the flights were pre-recorded, participants were seated alone in the VMS two-seat T-Cab. This afforded privacy to the individual participants, thereby reducing the potential for distraction from assigned tasks. More importantly, single occupancy obviated requirements for COVID-19 masking and any consequent breathing or visual impediments.

2.2 Tasks

Participants performed a five-minute visually-mediated manual task called the Psychomotor Vigilance Test, or PVT (Dinges & Powell, 1985). The PVT was done on each flight day, once preflight while at the vertiport and again while at

cruise altitude. Following each PVT, participants provided responses to the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993) and assessments of their workload using the NASA Task Load Index, or NASA-TLX (Hart & Staveland, 1988). Participants were reminded each time to restrict their SSQ and TLX responses solely to their perceptions and performance during the immediately preceding PVT. The SSQ and TLX responses for the inflight PVT were requested only after the vehicle had landed at the vertiport. The custom “NASA Study” version of the NASA PVT+ app (Arsintescu et al., 2019), installed on an Apple iPad Air, was enhanced to include the SSQ and a link to a separate NASA-TLX app. The study monitor prompted the participant on the VMS via their headset when to unstow, run a task, and restow the iPad. An electronic version of Golding’s (2006b) short-form Motion Sickness Susceptibility Questionnaire (MSSQ-Short) was also added to the enhanced PVT+ app. The MSSQ, however, was administered just once at enrollment prior to any PVT, SSQ, and TLX training or testing.

The primary purpose of the iPad PVT task was to elicit visual attention to, and manual interaction with a handheld device—an activity intended as a surrogate for onboard information processing or entertainment activities that future UAM commuters are expected to engage in. Because the PVT requires a manual input as soon as a response prompt (i.e., latency counter) appears on the screen, the app provides a clear indication of when the task is not being attended to. Participants were free to observe the computer-generated scenery outside through the VMS cab windows when iPad use was not required. Participants were informed that they should set aside the iPad for as long as needed if they found its use too provocative in terms of motion sickness. None, however, did so.

2.3 Physiological monitoring

Objective physiological data were collected preflight and inflight by portable chest- and wrist-worn instrumentation. The Equivital EQ02 chest system recorded two-lead electrocardiography (ECG), heart and respiration rates, skin temperature, and triaxial accelerometry. The Empatica E4 wrist unit measured blood volume pulse (BVP), triaxial accelerometry, electrodermal activity (EDA), and skin temperature. These physiological signals can all potentially be correlated to motion sickness signs (Kennedy & Frank, 1985).

2.4 Procedure.

The study protocol was reviewed and received approval by the NASA Institutional Review Board (IRB). Nonpilot volunteers with no prior VMS motion experience were recruited from the ARC employee community via an online centerwide announcement. Candidates also needed to have normal or corrected-to-normal vision; had to be able to use or learn to use an iPad; and, due to the confined cabin space, not have claustrophobia. We did not discourage any individual from participating based on prior motion sickness history (or MSSQ score). The absence of VMS experience was intended to enhance the sense of novelty for a first flight encounter with new type of vehicle. Participants were not offered any remuneration for their involvement in this study.

After providing their informed consent and receiving medical clearance, participants were invited to their first session. At the

start of the first session, participants used the PVT+ app to enter their age, gender, and handedness information and completed their MSSQ. After training the PVT+ task and SSQ and TLX data entry, the Equivital and Empatica devices were fitted to the participant and proper functioning of the sensors was verified. After a briefing by staff on VMS safety and evacuation procedures, participants were seat-belted in the cab’s left seat and donned the communication headset. The staff member then turned off the interior cab lights, exited, and closed the door. Following participant confirmation that they were ready, the cab was moved to its starting location at the center of the VMS’s travel range at which point the simulation commenced with the UAM vehicle on the vertiport pad awaiting takeoff.

The study monitor then prompted the participant to unstow the iPad and complete the preflight PVT followed immediately by the SSQ and TLX and restow the iPad when done. After the participant confirmed their readiness, the simulated flight commenced. At 180s after takeoff, the study monitor called for the participant to unstow the iPad, complete the PVT, and then restow the iPad. The participant was prompted to again unstow the iPad after landing at the vertiport and complete the SSQ and TLX for the just completed inflight PVT. After restowing the iPad, the VMS cab was brought back to the dock. The participant exited the cab and privately doffed and returned the chest and wrist devices. Throughout the flights, an over-the-shoulder cabin camera enabled the VMS operators and study staff to monitor participant wellbeing and progress through the assigned iPad tasks. The VMS safety briefing was not offered and all participants declined PVT/SSQ/TLX refresher training at their second session. Otherwise, procedures followed those of the first session.

2.5 Study design

The study followed a within-subjects repeated-measures design, with each participant performing tasks preflight and inflight and providing responses for two levels of the primary independent variable: vehicle flight control condition (i.e., COL or RPM) and the consequent roughness of the ride. First and second session assignment of COL and RPM flight condition was balanced between two groups (denoted as Cfirst and Rfirst) to counter potential order effects. Based on a nonparametric “sign test” approach, a posited large directional effect of 75% of participants providing a higher SSQ (the dependent variable of main interest) for RPM than for COL flights predicated a sample size of $n = 23$ for $\alpha < 0.05$ significance at 0.8 statistical power. To enable an exact order balance, 24 participants were requested in the IRB protocol.

3. RESULTS

3.1 Demographics and susceptibility

In all, 23 (19M/4F) participants completed all study conditions; a 24th could not start due to scheduling conflicts. Hence, one more participant was assigned to the Cfirst ($n_c = 12$) than the Rfirst ($n_r = 11$) group. Participant ages ranged between 20 and 67yr (median 38; mean 42.8). Participants’ pre-study MSSQ responses revealed them to be less predisposed to motion sickness (median total score 4.0) than other general populations, e.g., Golding’s 257 individuals (2006b) whose data (median total score 11.3) were used to develop the short-form MSSQ. The significant difference between these two MSSQ

distributions ($D_{max} = 0.354$, $p < 0.05$; Kolmogorov-Smirnov $D_{crit(23,0.05)} = 0.275$) was due to over half of our self-selected volunteers scoring within the lowest, i.e., [0-5] susceptibility bin, a rate more than double the proportion of Golding’s larger population, as illustrated by the histograms in Figure 4. This observation may in part reflect the gender imbalance in our study pool, as women are generally more prone to motion sickness (Griffin, 1990, p.287). Finally, it is important to note that in this case between-groups (i.e., Cfirst/Rfirst) MSSQ scores did not differ (Mann-Whitney $U = 53$, $z = 0.802$, $p_{2-tail} > 0.4$).

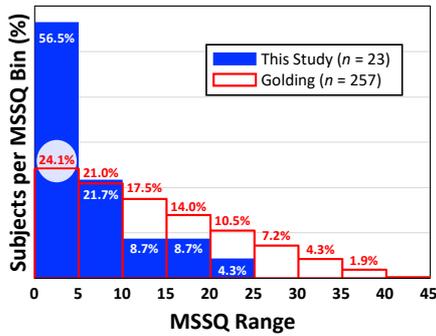


Figure 4. MSSQ distribution for this study vs. Golding (2006b).

3.2 SSQ, PVT, and TLX

Significant within-participant pre-to-inflight increases in “Total Severity” of simulator sickness symptoms (SSQ-TS) (Kennedy et al., 1993) were noted by one-sided Wilcoxon matched-pair sign-rank tests (Bonferroni corrected for 6 possible contrasts, $p_{crit} = 0.05/6 = 0.0083$) for both flight conditions. Inflight RPM (median SSQ-TS: 14.96) was higher than the same-day preflight baseline (median SSQ-TS: 0) ($T^+ = 168.5$, $n = 18$, $p < 0.0005$). Similarly, inflight COL (median SSQ-TS: 7.48) was higher than its same-day baseline (median SSQ-TS: 0) ($T^+ = 103.5$, $n = 15$, $p < 0.006$). A one-sided Wilcoxon matched-pair sign-rank test also indicated that participants’ inflight RPM SSQ-TS scores were significantly higher than those from the COL condition ($T^+ = 172$, $n = 20$, $p < 0.0053$). Further, no evidence of a Cfirst/Rfirst order effect was seen for SSQ-TS (COL: Mann-Whitney $U = 54$, $z = 0.752$, $p_{2-tail} > 0.4$; RPM: $U = 59.5$, $z = 0.403$, $p_{2-tail} > 0.6$). In summary, simulator sickness severity reports increased in flight and this effect was more pronounced for the RPM than the COL control condition.

Comparisons were made of the cumulative distribution of SSQ-TS scores from the four (preflight/inflight, RPM/COL) conditions and the calibration data employed by Kennedy et al. (1993) to develop the SSQ. Figure 5 shows the distributions of the SSQ-TS scores generated by our 23 participants overlaid on Kennedy et al.’s canonical data. The SSQ-TS distribution from the inflight COL condition did not differ significantly ($\chi^2 = 1.167$, $df = 2$, $p = 0.56$) from the Kennedy et al. data compiled from the experiences of 1100+ subjects in nine different flight simulators. On the other hand, inflight RPM SSQ-TS distribution was very significantly different ($\chi^2 = 21.78$, $df = 2$, $p < 0.00002$) and, as inferred from Figure 5 and the preceding Wilcoxon matched pair sign-rank tests, more severe than Kennedy et al.’s distribution. Similarly, the SSQ-TS score distributions from the preflight RPM ($\chi^2 = 8.377$, $df = 1$, $p < 0.005$) and COL ($\chi^2 = 4.174$, $df = 1$, $p < 0.05$) indicate that the non-motion conditions were more benign.

Spearman rank correlation analyses indicated that participants’ childhood and past-decade motion sickness history as quantified by MSSQ was related to their reported inflight SSQ-TS sickness symptom severity for both RPM ($r_s = 0.439$, $n = 23$ observations, $p < 0.05$) and COL ($r_s = 0.416$, $n = 23$, $p < 0.05$) conditions. Despite being less susceptible than the broader general population, the participants predisposition to motion sickness can still be inferred to be predictive of elevated SSQ reports.

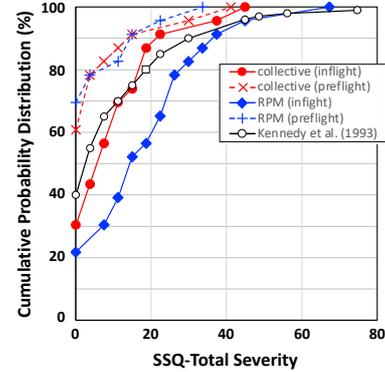


Figure 5. Cumulative SSQ-TS distributions for this study vs. Kennedy et al. (1993).

Arsintescu et al. (2017) list five overlapping metrics employed to analyze PVT data, all of which are based on observer Reaction Time (RT). In this case, since our initial interest is simply whether our participants were attending to the assigned PVT task, we first checked lapse count (the cumulative number of RTs exceeding 500ms) and observed that two participants’ average number of lapses (respectively, 11 and 12.5) were notably higher than the 1.14 average for the remaining 21. Of the two, one also had a high count preflight. During post-session debriefing, both participants indicated being distracted by out-the-window views.

Next, we examined participants’ median RT after correcting for the iPad’s internal processing latency (Arsintescu et al., 2017). Each median was computed from the 48 (on average) responses provided by each individual for each preflight or inflight condition. Because normality was not violated, a two-way repeated-measures Analysis of Variance (ANOVA) was conducted directly on median RT. After applying the Greenhouse-Geisser (G-G) sphericity correction for repeated-measures analyses (in this case, G-G epsilon = 0.726), the ANOVA demonstrated a significant main effect for preflight/inflight, COL/RPM condition levels ($F_{2,2,45.8} = 6.639$, $p < 0.003$) and the interaction between conditions and Cfirst/Rfirst group order ($F_{2,2,45.8} = 6.743$, $p < 0.003$). Bonferroni-corrected post hoc t-tests ($p_{crit} = 0.05/28 = 0.00179$ with $k(k-1)/2 = 28$ possible contrasts for $k = 2 \times 4$ order-by-flight levels) for the repeated-measures design revealed a significant 21.7ms (8.5%) preflight-to-inflight increase in median RT for the RPM condition on Day 2 (i.e., for the Cfirst group) ($t = 4.678$, $df = 11$, $p_{2-tail} < 0.0007$). Further, there was a significant 21.1ms (7.9%) improvement preflight in median RT for the combination of the Cfirst and Rfirst groups ($t = 3.737$, $df = 22$, $p_{2-tail} < 0.0011$). As plotted in Figure 6, these findings together are taken to indicate that participants’ performance and, by inference, their continued learning of the PVT task was to some

degree hampered by receiving the rougher RPM ride on their second day.

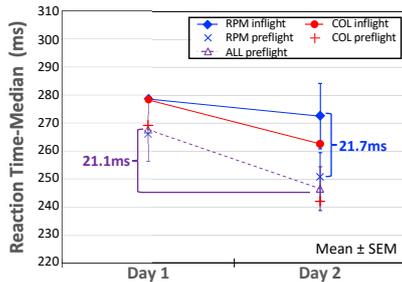


Figure 6. PVT Median RT as a function of test day and flight condition.

Participants’ NASA-TLX ratings did not provide any insight into passenger workload for the following reasons. First, workload ratings exhibited little *within*-rater range between preflight/inflight, RPM/COL conditions; on average, the maximum-to-minimum range was 10.3 of the TLX scale’s full 100-point span. This suggests that the PVT for which they were asked to provide workload ratings did not constitute a particularly rich, impactful, or therefore appropriate task to evaluate. Second, with standard deviations between 21.1 and 22.7 points, *between*-rater variability for each of the four conditions was double the average rater’s range, which indicates participants’ brief training on how to standardize and use the NASA-TLX rating scale was likely insufficient. This is not surprising, given that NASA-TLX, like other workload and HQ rating techniques, is typically employed by experienced test pilots and other highly trained professionals.

3.3 Physiological monitoring

A variety of sophisticated heart rate variability (HRV) metrics reflecting the state of the sympathetic and parasympathetic nervous systems (SNS and PNS) were extracted from the Equivital EQ02’s ECG recordings using Kubios HRV Standard (www.kubios.com) freeware analysis tools. SNS and PNS govern the body’s response to and recovery from a variety of physical and psychological stressors and together regulate the body’s organ systems to maintain homeostatic balance. HRV was examined over three 5-minute intervals—preflight during the PVT, first half of flight from takeoff (with the first part of inflight PVT), the second half of flight until landing (and the last part of the PVT).

Of the Kubios metrics, the only statistically significant finding was for the root-mean-square of successive time differences (RMSSD) from one heartbeat interval to the next, which is a direct expression of variability. RMSSD data were log-transformed to correct for non-normality, as determined by the Lilliefors/Kolmogorov-Smirnov test, and then subjected to a two-way repeated-measures ANOVA, which indicated a significant interaction ($F_{2,1,43.7} = 4.629, p < 0.014$; G-G epsilon = 0.416) between RPM/COL condition and Rfirst/Cfirst presentation order. The interaction is explained in terms of an order/day effect. Participants in the Rfirst group exhibited a statistically significant drop ($t = 3.370, df = 32, p_{2\text{-tail}} = 0.002$; $p_{crit} = 0.0083$, Bonferroni corrected) in terms of $\text{Log}_{10}(\text{RMSSD})$ during their second session when they experienced the more benign COL flight while the Cfirst group did not change significantly between days as depicted in Figure 7.

RMSSD is an indicator of PNS tone—the degree of recovery from stress or, effectively, “relaxation.” Thus, the observed RMSSD drop suggests that a rougher initial flight primes the Rfirst group to be less relaxed when returning on a subsequent day for their second, albeit gentler, COL flight. This observation points to the potential importance of the first flight’s RQ on passenger acceptance. Further testing is needed to confirm this observation and to also determine whether such an effect would be attenuated after more than two flights.

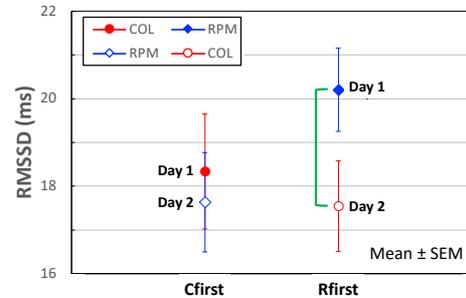


Figure 7. RMSSD by group and flight.

A two-factor repeated-measures ANOVA on respiration rate (RR), time-averaged over each PVT interval, only demonstrated a significant main effect of preflight/inflight COL/RPM condition ($F_{2,1,44.4} = 3.676, p < 0.03$; G-G epsilon = 0.704). Post hoc t-tests show that RR rose significantly from 16.93 ± 0.55 (mean \pm SEM) breaths per minute (bpm) preflight to 18.13 ± 0.61 bpm inflight for the COL condition ($t = 3.446, df = 23, p_{2\text{-tail}} = 0.0023$; $p_{crit} = 0.0083$ for 6 possible contrasts). The comparable contrast for the RPM condition, from 16.91 ± 0.57 bpm preflight to 17.19 ± 0.67 bpm inflight, however, was not significant ($t = 0.537, df = 23, p_{2\text{-tail}} = 0.6$). Depressed RR is indicative of higher stress. Thus, one can posit that participants’ anticipatory preflight stress dissipated once the smooth COL flight is experienced while the RPM condition does not offer comparable relief.

A two-factor repeated measures ANOVA on chest skin temperature only revealed a significant main preflight-inflight/COL-RPM effect ($F_{1,4,30.4} = 14.332, p < 0.0002$; G-G epsilon = 0.482). Post hoc t-tests indicate the effect was driven by a preflight-to-inflight rise in temperature regardless of control condition (COL: $\Delta = 1.96 \pm 0.10^\circ\text{F}, t = 18.77, df = 23, p_{2\text{-tail}} < 10^{-14}$; RPM: $\Delta = 1.75 \pm 0.30^\circ\text{F}, t = 5.81, df = 23, p_{2\text{-tail}} < 10^{-5}$; $p_{crit} = 0.0083$). It is possible that participants may simply have gotten warmer the longer they sat in the VMS cab.

To date, measurements from the Empatica E4 device have not yielded any significant results. At least for EDA, this may be attributable to dry sensors that contact the wrist rather than more sensitive surfaces such as the palm or finger pads and a slow, 4-sample/s data rate (cf. Benedek & Kaernbach, 2010).

4. DISCUSSION

Our finding that empirically obtained participant SSQ-TS scores (the study’s main dependent variable) aligned with the “RPM > COL > preflight” motion sickness severity ranking order predicted by the International Standards Organization’s (1997) guidance helps satisfy two of this study’s goals. First, the finding demonstrates the utility of SSQ ratings in the subjective assessment of large-scale vertical motion and as a tool

for examining passenger acceptance of RQ even when our participants' MSSQ results indicate them to be less predisposed to motion sickness. Second, we were able to corroborate from the passenger perspective that RQ suffers under RPM control as was seen in previous HQ analyses and pilot testing. While our observations for SSQ are easily understood and align with expectations, the secondary effects we detected (e.g., rougher flight hampering learning for at least one type of visual-manual task and a rougher first flight priming participants to be less relaxed even when their second flight is benign), while more nuanced, could still impact UAM passenger acceptance.

For this, a first UAM passenger study specifically targeting passenger response to motion, we were cautious and incremental in choosing our stimulus levels. In future studies, we will of course seek replication of the effects reported here, especially for the secondary findings. Observed effects could be strengthened not only through more challenging, e.g., longer, simulated flights with the opportunity for more SSQ repetitions, but also by recruiting an even gender balance and MSSQ screening to better match participant motion sickness susceptibility to the broader population. Finally, subsequent studies should allow for different inflight passenger tasks, better physiological instrumentation, and the incorporation of vibration and enhanced out-the-window viewing to augment simulation realism and test provocativeness.

ACKNOWLEDGEMENTS

This work was supported by NASA's RVL T Project. The authors thank the VMS staff, especially the engineers at the console, M. Blanken and D. Benn; ARC test pilot, G. Hardy; medical monitor, Dr. R. Pelligra; K. Kato for PVT+ app development; C. Malpica for UAM domain insight; and, most of all, the study's 23 participants.

REFERENCES

- Aires, J, Withrow-Maser, S, Ruan, A, Malpica, C, & Schuet, S. (2022). Analysis of handling qualities and power consumption for Urban Air Mobility (UAM) eVTOL quadrotors with degraded heave disturbance rejection and control response. *VFS 78th Annual Forum and Technology Display*, Fort Worth TX.
- Arsintescu, L, Mulligan, JB, & Flynn-Evans, EE (2017). Evaluation of a psychomotor vigilance task for touch screen devices. *Hum Factors*, 59(4), 661-670.
- Arsintescu, L, Kato, KH, Hilditch, CJ, Gregory, KB, & Flynn-Evans, E (2019). Collecting sleep, circadian, fatigue, and performance data in complex operational environments. *JoVE*, (150). DOI: 10.3791/59851
- Benedek, M, & Kaernbach, C (2010). A continuous measure of phasic electrodermal activity. *J Neurosci Methods*, 190, 80-91.
- Booz-Allen and Hamilton (2018). *Urban Air Mobility (UAM) Market Study*. Final Report, NASA contract numbers 80HQTR17F0127 and NNH13CH54Z.
- Crown Consulting Inc. (2018). *Urban Air Mobility (UAM) Market Study*. Final Report (presentation), NASA contract number 80HQTR17F.
- Deloitte (2020). *UAM Vision Concept of Operations (ConOps) UAM Maturity Level (UML) 4 Version 1.0*. Contract Report, NASA contract number 80HQTR19F0107.
- Dinges, DI, & Powell, JW (1985). Microcomputer analysis of performance on a portable, simple visual RT task sustained operations. *Behav res meth instrum comput*, 17, 652-655.
- Golding JF (2006a). Motion sickness susceptibility. *Auton Neurosci*, 129, 67-76.
- Golding JF (2006b). Predicting individual differences in motion sickness susceptibility by questionnaire. *Pers Individ Dif*, 41, 237-48.
- Griffin, MJ (1990). *Handbook of Human Vibration*. London, UK: Elsevier Academic Press.
- Hart, SG, & Staveland, LE (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Hancock, PA, Meshkati, N (eds.). *Human Mental Workload. Advances in Psychology*, 52. Amsterdam: North Holland. pp. 139-183.
- International Standards Organization (1997). *Mechanical vibration and shock - Evaluation of human exposure to whole-body vibration - Part 1: General requirements*, ISO 2631-1:1997(E).
- Kennedy, RS, & Frank, LH (1985). *A Review of motion sickness with special reference to simulator sickness* (report NAVTRAEQUIPCEN 81-C-0105-16). Orlando, FL: Naval Training Equipment Center.
- Kennedy, RS, Lane NE, Berbaum, KS, & Lilienthal, MG (1993). Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol*, 3, 203-20.
- Leatherwood, JD, Dempsey, TK, & Clevenson, SA (1980). A design tool for estimating passenger ride discomfort within complex ride environments. *Hum Factors*, 22(3), 291-312.
- Malpica, C, & Withrow-Maser, S (2020). Handling qualities analysis of blade pitch and rotor speed controlled eVTOL quadrotor concepts for urban air mobility. *VFS International Powered Lift Conference*, San Jose, CA.
- McFarland, RE, & Duisenberg, K (1995). *Simulation of Rotor Blade Element Turbulence*. NASA TM-108862.
- Schuet, S, Malpica, C, Lombaerts, T, Kaneshige, J, Withrow-Maser, S, Hardy, G, and Aires, J (2020). A modeling approach for handling qualities and controls safety analysis of electric air taxi vehicles. *AIAA Aviation Forum*, virtual, Paper AIAA 2020-3188.
- Silva, C, Johnson, W, Antcliff, KR, & Patterson, MD (2018). VTOL Urban Air Mobility Concept Vehicles for Technology Development," *Aviation Technology, Integration, and Operations Conference, AIAA Aviation Forum*, Atlanta, GA, Paper AIAA 2018-3847.
- Withrow-Maser, S, Malpica, C, & Nagami, K (2021). Impact of handling qualities on motor sizing for multicopter aircraft with urban air mobility missions. *VFS 77th Annual Forum and Technology Display*.
- Withrow-Maser, S, Aires, J, Ruan, A, Malpica, C, & Schuet, S (2022). Evaluation of heave disturbance rejection and control response criteria on the handling qualities evaluation of urban air mobility (UAM) eVTOL quadrotors using the Vertical Motion Simulator. *VFS Aeromechanics for Advanced Vertical Flight Technical Meeting*, San Jose, CA.