EASA

European Union Aviation Safety Agency

EASA FC06 SPECIFIC CONTRACT 02

Determination of a human dose-response with respect to single events of Urban Air Mobility-type vehicles

An Agency of the European Union

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Union Aviation Safety Agency (EASA). Neither the European Union nor EASA can be held responsible for them.

This deliverable has been carried out for EASA by an external organisation and expresses the opinion of the organisation undertaking this deliverable. It is provided for information purposes. Consequently it should not be relied upon as a statement, as any form of warranty, representation, undertaking, contractual, or other commitment binding in law upon the EASA.

Ownership of all copyright and other intellectual property rights in this material including any documentation, data and technical information, remains vested to the European Union Aviation Safety Agency. All logo, copyrights, trademarks, and registered trademarks that may be contained within are the property of their respective owners. For any use or reproduction of photos or other material that is not under the copyright of EASA, permission must be sought directly from the copyright holders.

Reproduction of this deliverable, in whole or in part, is permitted under the condition that the full body of this Disclaimer remains clearly and visibly affixed at all times with such reproduced part.

REPORT NUMBER:	**EASA.FC06.SC02.D1**
REPORT CLASSIFICATION:	UNCLASSIFIED
DATE:	01 December 2021
KNOWLEDGE AREA(S):	
DESCRIPTOR(S):	Noise, UAM, drones, noise annoyance, human perception
CUSTOMER:	EASA
CONTRACT NUMBER:	EASA.2020.FC06.SC02
OWNER:	European Union Aviation Safety Agency
DISTRIBUTION:	Limited
CLASSIFICATION OF TITLE:	**UNCLASSIFIED**

Author(s): R. Aalmoes, H.A. Lania, J. Choi

APPROVED BY:	AUTHOR	REVIEWER	MANAGING DEPARTMENT
M. Nagelsmit	R. Aalmoes	S.J. Hebly	NLR-AOSE

DATE: 01 December 2021

EASA FC06 SPECIFIC CONTRACT 02

Determination of a human dose-response with respect to single events of Urban Air Mobility-type vehicles

Royal Netherlands Aerospace Centre

NLR operates as an objective and independent research centre, working with its partners towards a better world tomorrow. As part of that, NLR offers innovative solutions and technical expertise, creating a strong competitive position for the commercial sector.

Anthony Fokkerweg 2 | 1059 CM Amsterdam | +31 88 511 3113 | info@nlr.nl | www.nlr.org

SUMMARY

This report describes a laboratory psycho-acoustic study on the annoyance of helicopter and rotor-based drone vehicles. Based on reported annoyance rates from single events, dose-response curves were derived for short-term annoyance of these vehicles. Measurements of UAM-type aircraft were used as input for the study, as well as recordings from helicopters. The participants were placed in front of a computer playing the sound events, and were subjected to the noise events via headphones. Each participant was presented with a number of events, and after each event, was asked to report their (subjective) annoyance on a non-discrete scale by moving an indicator on a slider between a value representing 'not annoyed at all' and 'extremely annoyed'. The presented stimuli corresponded to four multicopter drones, an air taxi, a small fixed wing model aircraft, and two helicopter sounds. These sounds were presented with sound exposure levels between 50 and 90 dB(A) SEL-A.

A total number of 40 participants from the area of Amsterdam participated in the test. The participants had a demographic diversity in age, gender, education, and residential area. Results from the measurements show expected reported higher annoyance rates for increased SEL values for each vehicle. Differences can be seen between the highly annoyance scores of the helicopters on one hand (lower), and the other vehicles on the other hand (higher), especially in SEL values above 70 dB(A). A curve-fitting algorithm has been applied to the discrete values using a logistic fitting function, in order to find the appropriate dose-response curve.

The number of air taxis in this study is very limited (1), and the current sample may not be representative for other new air taxis that are now being developed. As the annoyance is rated higher for drones and air taxis than with helicopters, it will be important to examine SEL values that will typically be expected when they operate in their expected setting.

CONTENTS

ABBREVIATIONS

ACRONYM	DESCRIPTION
dB	Decibel
M	Mean value
NLR	Royal NLR - Netherlands Aerospace Centre
PAV	Personal Air Vehicle
PSST	Psycho-acoustic Sound Scape Testing
SEL	Sound Exposure Level
UAM	Urban Air Mobility
VTOL	Vertical Take-off and Landing

1. Introduction

1.1 Purpose of the study

The purpose of this study was to conduct a psycho-acoustic study on the annoyance of helicopter and rotor-based drone vehicles. Based on reported annoyance rates from single events, dose-response curves were derived for short-term annoyance of these vehicles.

1.2 Scope of the study

This study had a limited scope to perform a laboratory experiment to measure single-event dose-responses from participants listening to UAM vehicles. In this study, both (unmanned) drone vehicles and Vertical Take-off and Landing (VTOL) vehicles that can be used to move people around, also known as air taxis, were considered. The study was limited in the sense that research on which metrics to use was done in earlier work [1][2]. It was also limited in the number of recordings that were evaluated and the number of participants in the study. No literature study was foreseen, but a short evaluation of previous noise-response studies was used as reference for this study.

1.3 Literature background

Typically, dose-responses are established for long-term average noise exposure using field studies, by mapping self-reported annoyance of the respondents living near airports to a calculated (outdoor) noise exposure [3][5]. In these studies, the noise annoyance depends not only on the amount of exposure, but also on the source of the noise, where rail and road traffic are considered less annoying than aircraft noise [4]. These field studies are not possible for noise events that are scarce or not common, such as new vehicles or operations of UAM. Laboratory studies are used instead, where single events are presented to test subjects, and their immediate annoyance is questioned. This short-term noise annoyance may give indications for long-term noise annoyance should those UAM operations become more common in the future.

Measurements of UAM-type aircraft were used as input for the study, as well as recordings from helicopters. The latter category was added for two reasons: first, helicopters are similar, rotor-equipped vehicles as drones, and some future UAM vehicles may also be configured similar to a helicopter. Second, the helicopter sound may act as a reference sound and it could be used to compare this study with other dose-response studies done for helicopters.

Typical UAM operations may involve take-off, landing, fly-over or hovering events, but in this study, only recorded fly-over sounds were considered. A recent study [6] indicates that hovering sounds can be considered more annoying than fly-overs at the same sound exposure level. Therefore, to make an objective comparison, only similar operations should be compared. Additionally, other effects that may be experienced and may influence annoyance, such as visual perception or ambient background sounds, were not part of this study. For the background sounds, it can be said that the recordings were made or selected with as little background noises as possible, but there may still have been remnants of limited background noise, such as bird sounds or wind/moving tree noises.

2. Study set-up

2.1 Introduction

This chapter describes the study set-up for this project. First, the methodology is described. Then, selection and demography of participants is described. After that, the hardware set-up is provided, including how the calibration of the experiment set-up was performed. Then, the procedure carried out over the participants is described.

2.2 Methodology

A single-event dose response laboratory study was performed with participants who evaluate UAM sounds by self-reporting an annoyance score per event. The participants were placed in front of a computer that played the sound events, and the participants heard the events through headphones. Each participant was presented with a number of events, and after each event, was asked to report their (subjective) annoyance on a non-discrete scale by moving an indicator on a slider between a value representing ‘not annoyed at all’ and ‘extremely annoyed’, See Figure 2. Each event was a predetermined sound sample from a UAM-vehicle or a helicopter and was played at different sound levels. The pre-determined sound levels were spaced in 10 dB steps of a SEL-A weighting, and only the sound level that the participants were subjected to was adjusted (not the event duration). This implies that no corrections were made to model atmospheric absorption, which would also have changed the (frequency) characteristics of the sound.
For each of the vehicles that were presented, the total number of highly-annoyed scores was counted per SEL_A sound exposure level. This lead to an annoyance curve with the SEL-A on the x-axis and the percentage of highly-annoyed on the y-axis. The criterium chosen for “highly annoyed” was a score over 72% of the presented annoyance-scale, in accordance to earlier studies [5][7]. As an alternative (lower highly-annoyed score), an annoyance score of 60% highly annoyance is presented as well (compatible with a 5-point annoyance scale).

2.3 Sound events (stimuli)

Table 1. Sound events used in study.

No	Id	Type of sound
1	PAV1	Air taxi
2	Heli1	Helicopter
3	Drone1	Multicopter drone
4	Drone2	Multicopter drone
5	Heli2	Helicopter
6	Drone3	Multicopter drone
7	Drone4	Multicopter drone
8	Fixedwing	Fixed wing aircraft model

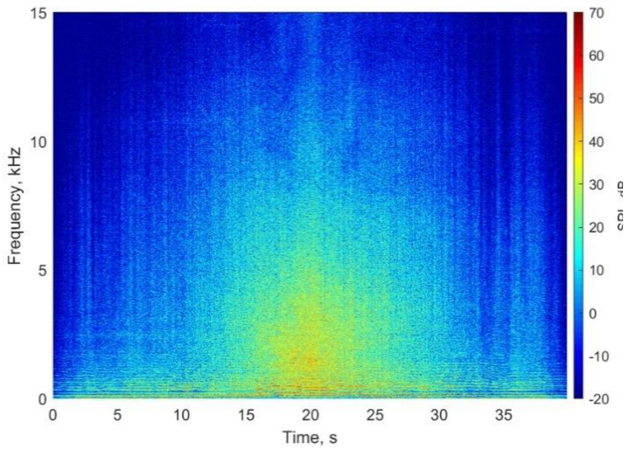
Table 1 shows the list of presented stimuli. They were played to the participants at the following noise levels: 50, 60, 70, 80, and 90 dB(A) SEL-A. The spectrograms of the sound events of the vehicles are displayed in Figure 1. It should be noted that only the measurements of the helicopters contained ground-reflections, recognizable by the fern leaf-shaped, ground interference pattern in the spectrograms. The

other recordings were made using microphones positioned over perfectly reflective ground-plates, which does not contain these ground-reflection patterns.
The length of each sound stimulus was approximately 30 seconds. When the noise of the vehicle was still audible at the beginning or end of the flyover, a short fade-in/out was applied. The total length of the sound evaluation session with headphones on was one hour, including three 3-minute breaks after each 15 minutes.

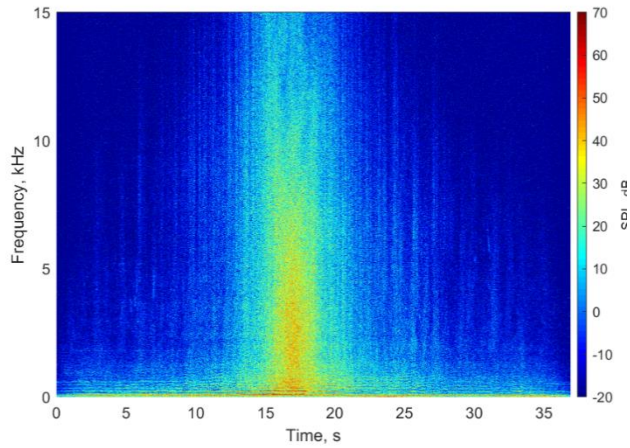
A total of 40 (8 samples x 5 levels) measured stimuli were foreseen. The sound samples were mono-recordings and were presented to the participants as mono sounds to both ears via the headphones.

The sound samples were randomly presented to the participants. Stimuli from the same vehicles (but with different sound levels) were not played subsequently, and each participant was presented with a different random order of events, to prevent any bias from the events sequence.

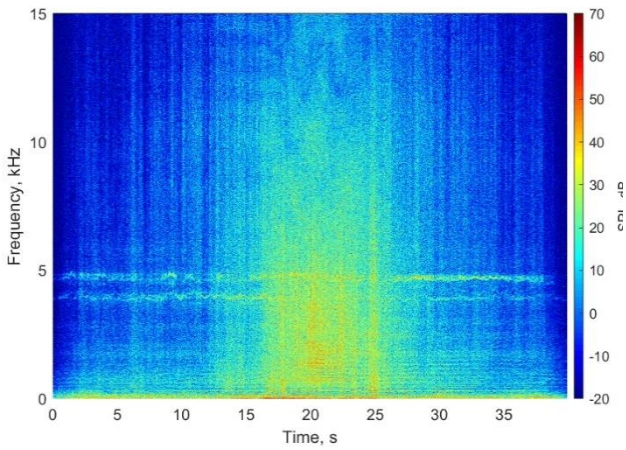
To get the participant acquainted with the test, two stimuli that were not used for the reporting were presented before the measured stimuli. These two practice stimuli were first the *fixedwing* sound played at 50 SEL-A, and second, the *heli1* sound played at 90 SEL-A. Counting both the practice stimuli and the measured stimuli, the total number of stimuli was 42 per participant.



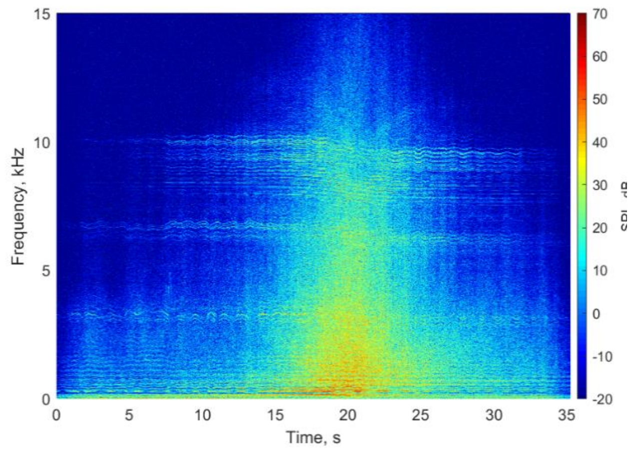
(a) Drone 1



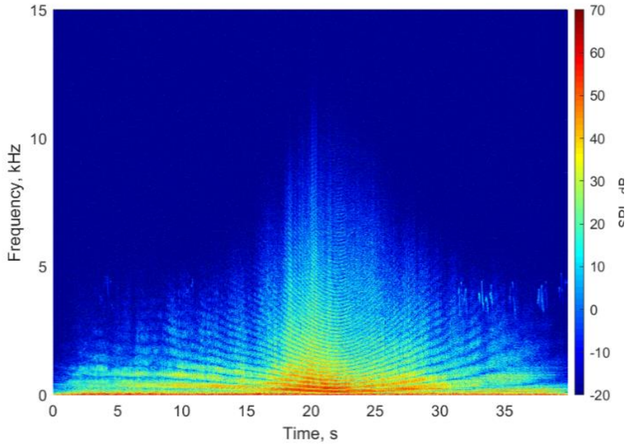
(b) Drone 2



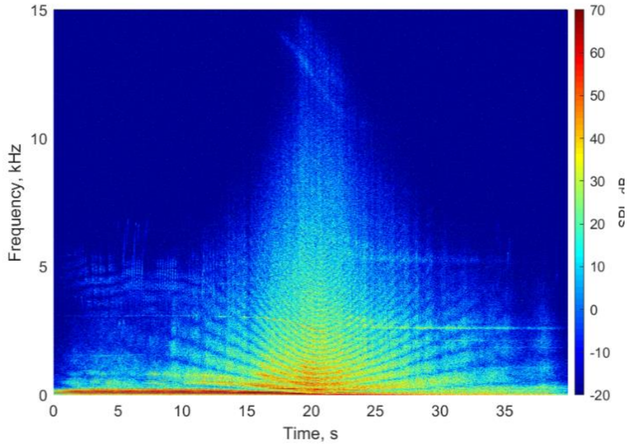
(c) Drone 3



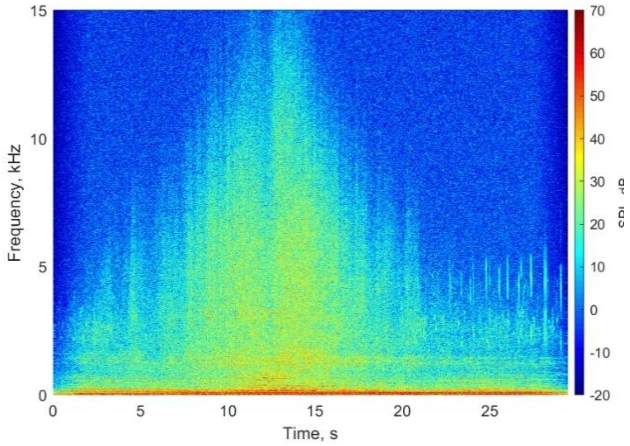
(d) Drone 4



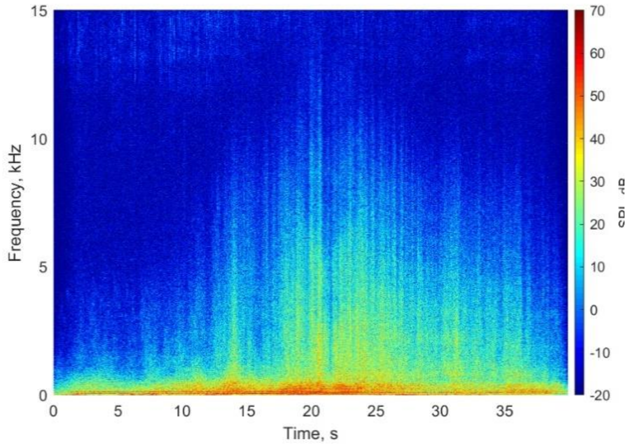
(e) Heli 1



(f) Heli 2



(g) fixedwing



(h) pav1

Figure 1. Spectrograms of the vehicles in the experiment.

2.4 Participants

Participants were gathered around the neighbourhood of Amsterdam, using a recruitment agency to obtain the candidates. Considering the limited number of participants in this set-up (approx. 40), a fully even representation of gender, education, age, etc. for a target population was difficult to obtain, but a certain representation of all considered age categories between 18-65, and a balanced representation of gender was aimed at. Additionally, because UAM operations are generally expected to take place in cities, the fact that the majority of participants came from the vicinity of Amsterdam was not deemed a problem. There were also participants recruited from more rural areas (outskirts of the city or actual rural areas) to get some representation from these areas.
Participants were asked their permission to conduct the test according to national and European GDPR (privacy) regulations by signing a consent form, and their identity is not made known in the results.

Participants were asked to have a self-reported good hearing. A hearing test for the participants was not part of this study: previous experiments using the same recruitment agency where hearing tests were performed showed that almost all candidates indeed did have a good hearing when they self-reported so. Participants received a small contribution for their participation (35 Euro each).

2.5 Experiment hardware, software and calibration

The hardware to perform the dose-response experiment consisted of an Alienware laptop running the Psycho-acoustic Sound Scape Testing (PSST) application that performed the dose-response study. PSST is a custom application developed at NLR for the purpose of sound evaluation by test subjects, See Figure 2. The laptop was equipped with an Audioquest Dragon Cobalt audio device and Beyerdynamic DT 770 M headphones. Two systems were prepared for the execution of these tests.

Figure 2. The evaluation screen for each sound event for the participant (in Dutch).

Sound levels of the simulation were calibrated using a HATS (Head and Torso Simulator), with installed B&K high fidelity microphones connected to National Instruments USB-4431 measurement equipment.

2.6 Procedure

Participants came to the office of NLR in Amsterdam for the experiment. They sat down at a desk that also contained the laptop for the sound evaluation. They received an instruction about the experiment and were asked to fill in a consent form and another form with personal detail such as gender, age, highest education, etc. Then, the participants put on the headphones and started the sound test.
The following cycle was performed for each of the presented stimuli:

1. A short, 4 seconds, 'white noise' was presented to the participant, to indicate the start of a new sample and to act as a neutralisation sound or "reset".
2. The sound event (stimulus) was presented. The length of the event was between 30-40 seconds, depending on the original recording. During the playback, the participant would adjust the slider to indicate their personal annoyance level (See Figure 2).
3. 5 seconds before the end of the sample, the OK-button would become enabled and the participant was allowed to continue, but only if the slider has been moved already.
4. After pressing the OK button, either this procedure was repeated, a scheduled break would take place, or in case of the last stimulus, the test was finished.

The sound test was interrupted for three 3-minute breaks where participants could take off their headphones and have a drink if they wanted. The total time of the sound evaluation was one hour as a maximum.

Before the start of these measured sound events, two (practice) sound events were presented using the same procedure as indicated above. This was intended to get the participants acquainted to the test procedure and to the sound levels of the test. These practice sounds were played at the lowest sound exposure level (50 SEL-A) and the highest sound exposure level (90 SEL-A), respectively.

At the end of the sound evaluation, the test subjects were asked to fill in a post-test questionnaire related to whether they recognized the sounds, their sensitivity to noise according to a Weinstein Noise Sensitivity Scale [9][10], and their attitudes towards drones. The results of these questionnaires were used for the final analysis.

3. Execution of experiment

3.1 General comments

The experiment with participants took place at NLR premises in Amsterdam, The Netherlands, between September 24 and October 15, 2021. A quiet room was selected with limited outside noise for the experiment, even though the Beyer dynamic headset provided good isolation from outside noise. Precautions were taken to manage COVID-19 pandemic issues although the number of COVID-19 infections in the Netherlands was low at the time and it did not have much impact on the participation. The anticipated number of participants of 40 persons was reached. Except for small events, which are mentioned in the next section, all participants completed the test without issues.

3.2 Anomalies, deviations

Due to a malfunction, the first two participants did get a pause screen. However, the first pause was only 1 minute, and the 2nd and 3rd pause allowed them to continue without a break, making these moments an optional break. After these two participants, this malfunction was corrected, so the other participants had three mandatory breaks of with each of at least 3 minutes. The impact of this anomaly is considered low as the total duration of the experiment was eventually shorter (40 minutes) than originally planned (1 hour), and the possibility still existed to have a break if desired.

Halfway in the completion of tests, we noticed that the drone3 sound sample was clipping (oversteered) and created a slight distortion when playing the sound on 90 SEL-A loudness at the peak level to the participants. Despite the fact that this sound was calibrated at SEL-A 90 dB, this unforeseen modified sound signature may have an influence on the reported annoyance.

A limited number of participants did not show up ("no show"), or could not identify themselves (a requirement to enter NLR premises for security reasons). This required the rescheduling of participants and making use of the reserve candidates that were scheduled at the end of the test period.

4. Results

4.1 Participation

A total number of 40 people were tested, and all participants managed to complete the full test. A total of 48% of the participants were female. The different age groups were all represented, but the age group of 56-65 years was represented by only one participant. Note that the participants were asked before the test whether they had good hearing, as it is generally known that hearing loss is more frequent with a higher age. Half of the participants self-reported to be living in a city, 27% reported to live in suburbs, and 23% reported to live in a village. The highest completed education of the participants was also mixed with the largest representation by people with an MBO education (post-secondary vocational education)[1], in line with what is expected from the Dutch population [11].

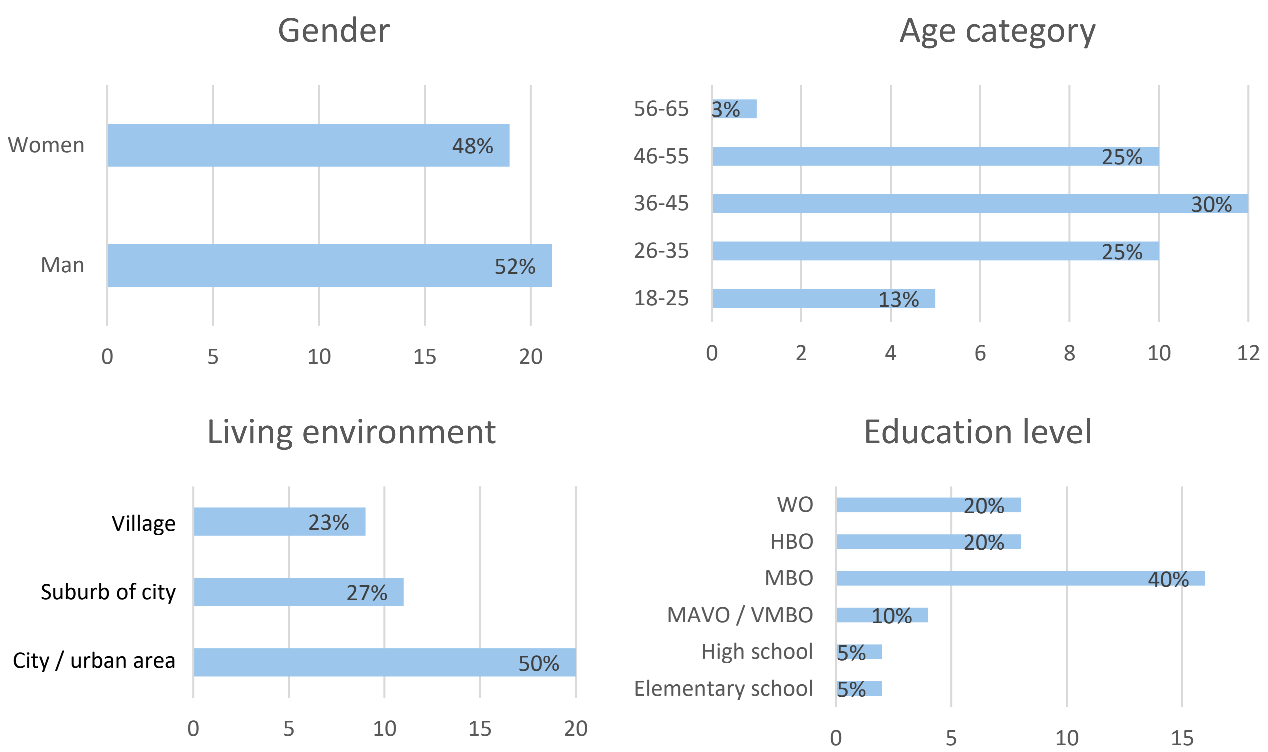


Figure 3. Demographic characteristics of participants.

4.2 Reported annoyance

The participants rated each sound by shifting a slider between 'not annoyed at all' and 'extremely annoying'. The score for each of the stimuli is between 0% and 100%, respectively, is presented here with mean value and confidence interval, See Figure 4. A confidence interval of 90% is also chosen instead of a more common 95% interval due to the limited number of participants. This confidence interval represents 90% of the number of 40 participants and equals the middle 36 values.

---

[1] MAVO is secondary school education. VMBO is pre-vocational education. HBO is higher professional education. MBO is post-secondary vocational education. WO is academic education.

As the helicopters seemed to produce a lower annoyance than other vehicles, this has been examined using a Paired sample t-test for the highest SEL-A value. With this test, two populations can be compared to see if there is a significant difference between the means of these populations. An outcome of $p< 0.05$ indicates that there is less than 5% chance of obtaining those results without there being a real effect. It showed that helicopters were considered least annoying for both *heli2* (M = 0.64, SD = 0.26) and *heli1* (M = 0.64, SD = 0.23) at SEL-A = 90, with p-values all < .001. Additional paired sample t test shows no difference in annoyance scores between *heli1* and *heli2*, $t(39) = 0.092$, $p = .927$.

4.3 Highly annoyance curves

The highly annoyed scores are the number of participants that rate a sound event over a certain annoyance threshold. As discussed in section 2.2, a threshold of over 72% is in current literature considered "highly annoyed". In this study, also the 60% threshold is presented as an alternative highly annoyed score. This is 60% conforms to 5-scale noise annoyance questionnaires where the upper two values are rated as highly annoyed.

A curve-fitting algorithm has been applied to found values using a logistic fitting function, in order to find the appropriate dose-response curve, See Equation 1. A logistic function is used here as it represents the expected "S-curve", where very low SEL-A value would result in a 0% highly annoyed score, and a very high SEL-A value would result in a score approaching 100%. With the fitting curves, a L=100 value is assumed, and the python/SciPy optimize.curve_fit method[2] is applied to find the matching x_0 and k values for each vehicle.

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \tag{1}$$

The results are provided in Figure 5. A combined curve for all UAM vehicles can be found in Figure 6. The parameters for the 72% highly-annoyed curve can be found in Table 2. As reported before, the drone3 sound on SEL-A 90 had clipping sound artefacts. We would therefore expect that this sound might be considered more annoying, but if we look at the curve, it is lower than expected. We therefore conclude that this clipping did not influence the measurement.

[2] See https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html

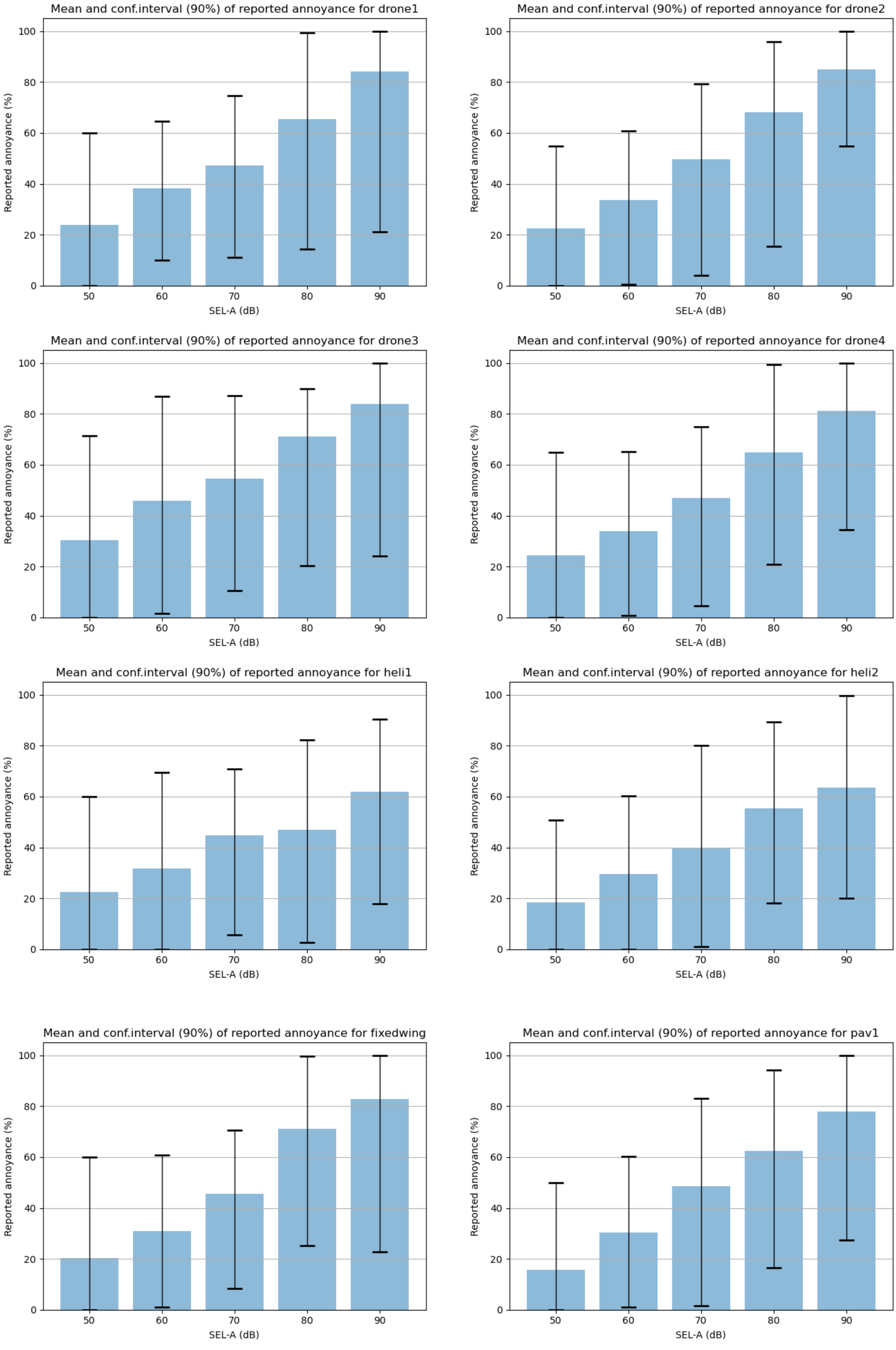


Figure 4. Reported annoyance per SEL-A value and event.

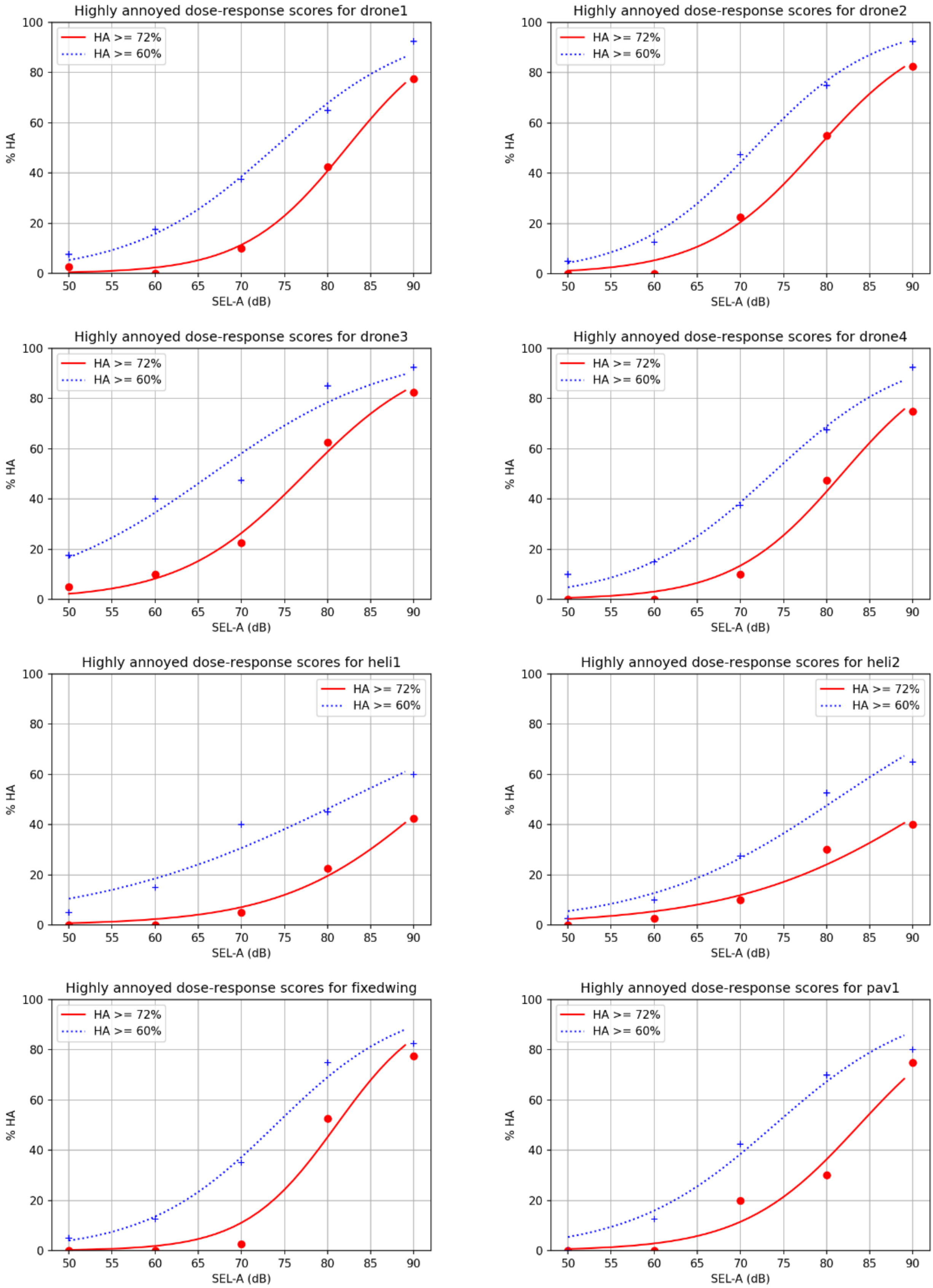


Figure 5. Highly-annoyed dose-response results and matching fitting curves for 40 participants in this study.

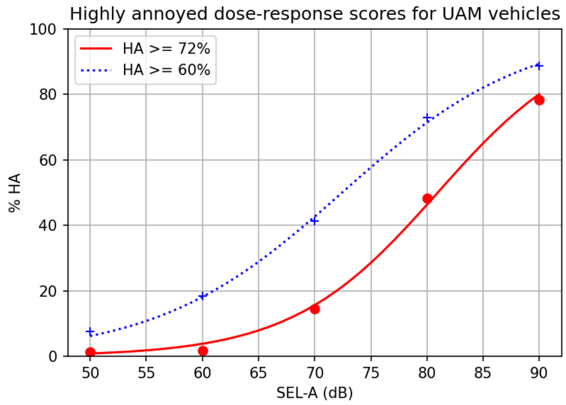


Figure 6. Dose-response for combined UAM vehicles (drone1, drone2, drone3, drone4, fixedwing, and pav1).

Table 2. Found logistic parameters for 72% Highly annoyed curves.

Vehicle	*k*	x_0
drone1	0.168	82.2
drone2	0.152	78.9
drone3	0.138	77.4
drone4	0.158	81.8
fixedwing	0.189	81.0
heli1	0.116	92.2
heli2	0.085	93.4
pav1	0.148	83.8
Combined UAM	0.154	81.0

4.4 Post-test questionnaire

Noise sensitivity was examined using the Weinstein Noise Sensitivity Scale filled in as part of the post-test questionnaire. The score, between 1 and 6 where a higher value indicates a higher sensitivity, was compared with the total number of (72%) highly annoyed scores for all vehicles that were given per participant. A Pearson correlation showed no significant relation between highly annoyed participants ($M = 10.38$, $SD = 5.96$) and their sound sensitivity score ($M = 3.59$, $SD = 0.91$), $r(39) = .21$, $p = .204$.
Just over half (22 of 40) of the participants recognized drone sounds in the test. Due to this low number, no correlation was calculated between participant's attitude towards drones and their annoyance scores.

5. Conclusion & discussion

This study presents dose-responses relationships for a new class of aircraft vehicles, namely four drones and one air taxi. Also, a small fixed wing aircraft has been included, and two helicopters that can be considered as reference measurements. A-weighted Sound Exposure Level (SEL) has been used to determine the input sound levels ('dose') for the stimuli. Stimuli have been presented using a headphone as a mono sound on both ears, so no binaural representation was used. Results from the measurements show expected reported higher annoyance rates for increased SEL values for each vehicle. Differences can be seen between the highly annoyance scores of the helicopters on one hand (lower), and the other vehicles on the other hand (higher), especially in SEL values above 70 dB(A). The fixed wing is also similarly rated as the drones, but it was also equipped with similar electric engines as the drones that created a similar noise sound. The air taxi (pav1) annoyance rates were similar as the drones, and was rated as more annoying than helicopters at the higher SEL values. The higher annoyance might be related to the unfamiliarity of drone and air taxi sounds in comparison to helicopter sounds, but this has not been examined. Another explanation of difference may be related to the inclusion of ground-reflection in the helicopter measurement, but this can easily be examined in a comparison study for this specific recording feature.

This was a study with a limited number of participants (40), but the number was sufficient to obtain statistical significance. The participants had a demographic diversity in age, gender, education, and residential area. But as all participants were tested in Amsterdam, there could be some difference with populations further away from a large city or from another European area. Also, attitude towards drones was asked at the end of the test, but was not further examined as just over half of all participants recognized the sounds as drone noise. Nevertheless, as all tests were conducted at NLR premises in Amsterdam, an aerospace institute, participants who know the institute could expect that the sound they heard were related some sort of aircraft sound, so a possible (negative) bias could have some influence here.

The number of air taxis in this study is very limited (1), and the current sample may not be representative for other new air taxis that are now being developed. This is noteworthy, as the configuration of this new class of vehicles can be much different from the one tested here, and therefore also their sound characteristics. Also, little is known about other air taxi's sound exposure levels and operations. As the annoyance is rated higher for drones and air taxis than with helicopters, it will be important to examine SEL values that will typically be expected when they operate in their expected setting.

References

[1] N. Böhnsich, C. Braun, Deliverable D3: Evaluation of Psychoacoustic Test, Specific Contract N° SC06 of Framework Contract N° EASA.2016.FC.21, 2020.

[2] N. Böhnsich, C. Braun, SC08 Evaluation of UAS/eVTOL noise impact - D1, Specific Contract N° SC08 of Framework Contract N° EASA.2016.FC.21, 2021.

[3] T.J. Schultz, Synthesis of social surveys on noise annoyance, *Journal of Acoustical Society of America*, Vol. 64, No. 2, August 1978.

[4] H.M.E. Miedema, C.G.M. Oudshoorn, Annoyance from Transportation Noise: Relationships with Exposure Metrics DNL and DENL and Their Confidence Intervals, *Environmental Health Perspectives*, Vol. 109, No 4, April 2001.

[5] H.M.E. Miedema, H. Vos, Exposure-response relationships for transportation noise, *Journal of Acoustical Society of America*, Vol 104, No. 6, December 1998.

[6] R. Aalmoes, N. Sieben, Visual and audio perception study on drone aircraft and similar sounds in an Urban Air Mobility setting, *International Congress and Exposition on Noise Control Engineering INTER-NOISE*, Washington DC, USA, 1-5 August, 2021.

[7] M. Brink, L. Giorgis-Allemand, D. Schreckenberg, A. Evrard, Pooling and Comparing Noise Annoyance Scores and “High Annoyance” (HA) Responses on the 5-Point and 11-Point Scales: Principles and Practical Advice, *International Journal Environmental Research and Public Health*, 2021, 18, 7339, https://doi.org/10.3390/ijerph18147339.

[8] FAA, Technical Support for Day/Night Average Sound Level (DNL) Replacement Metric Research Report Part 3, 14 June 2011, Retrieved from: https://www.faa.gov/about/office_org/headquarters_offices/apl/research/science_integrated_modeling/noise_impacts/media/6-14-2011_FinalReport_MetricsMestre_etal_061411_part3.pdf

[9] Weinstein, N. D. (1978). Individual differences in reactions to noise: A longitudinal study in a college dormitory. *Journal of Applied Psychology*, 63(4), 458–466. https://doi.org/10.1037/0021-9010.63.4.458

[10] Worthington, D. L., & Bodie, G. D. (2017). Weinstein Noise Sensitivity Scale (WNSS): (Weinstein, 1978). In *The Sourcebook of Listening Research: Methodology and Measures* (1st ed., pp. 475–481). Wiley-Blackwell. https://doi.org/10.1002/9781119102991

[11] CBS StatLine. Bevolking; onderwijsniveau; geslacht, leeftijd en migratieachtergrond. Website. Retrieved at 17 November 2021 from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82275NED/table?fromstatweb

European Union Aviation Safety Agency

Konrad-Adenauer-Ufer 3
50668 Cologne
Germany

Tel. +49 221 89990- 000
Mail research@easa.europa.eu
Web www.easa.europa.eu

An Agency of the European Union