NOISE-CON 2022 June 13-15 Lexington, KY

111Equation Chapter 1 Section 1

Feasibility study for remote psychoacoustic testing of human response to urban air mobility vehicle noise

Siddhartha Krishnamurthy[1]
Structural Acoustics Branch, NASA Langley Research Center
2 North Dryden Street, Mail Stop 463
Hampton, VA 23681

Stephen A. Rizzi[2]
Aeroacoustics Branch, NASA Langley Research Center
2 North Dryden Street, Mail Stop 461
Hampton, VA 23681

ABSTRACT

NASA will remotely administer a psychoacoustic test in late summer of 2022 as the first of two phases of a cooperative Urban Air Mobility (UAM) vehicle noise human response study. This study relies on the cooperation of multiple government agencies, academia, and industry to assemble a wide range of UAM vehicle sounds. This database of sounds will be used to create a rich database of human response to UAM noise that would be challenging for a single organization to acquire. The development of the remote test method to study human response to aviation noise was prompted by the novel coronavirus pandemic. The feasibility portion of the study described in this work will demonstrate and refine the remote test method for use in the implementation phase.

This paper details the method for remotely administering the psychoacoustic test and the sound stimuli to be used in the Feasibility Test. Comparisons of annoyance response data from previous in-person tests will be used to demonstrate the viability of the remote test method. The paper also describes an effort to determine if providing a contextual cue to test subjects influences the annoyance response.

1. INTRODUCTION

1.1. UAM Vehicle Noise Human Response Study

NASA seeks to remove barriers to the operation of Urban Air Mobility (UAM) vehicles [1]. UAM vehicles are a part of NASA's vision for Advanced Air Mobility, which seeks to develop new air transportation systems that move people and cargo between places previously not served or underserved by aviation [2]. Representative UAM vehicle concepts involve the use of electrically driven rotors, and the noise from these air vehicles in communities may restrict their operation.

To address this noise concern, NASA has been pursuing research to better understand the human response to UAM vehicle noise. It led the formation of the UAM Noise Working Group (UNWG), which consists of researchers from academia, industry, and government, to identify and address

1 siddhartha.krishnamurthy@nasa.gov
2 s.a.rizzi@nasa.gov

UAM noise issues. The UNWG published a white paper containing high-level goals intended to address noise barriers that may hamper UAM vehicle entry into service [3]. The white paper identified the need to perform laboratory studies to understand the perception of UAM vehicle noise and to also collect data on variations in perception between geographically different communities.

Based on these recommendations, members of the UNWG proposed that a psychoacoustic test be conducted using test facilities spanning multiple geographic locations to obtain data on community variation in response. Such a study would require the cooperation of multiple partners, but this cooperation could enable a more diverse collection of noise stimuli and human response. Overarching goals of the cooperative study, which is called the UAM vehicle noise human response study, are:

- Assemble a wide range of UAM vehicle sounds through cooperation between multiple agencies and organizations for use in human response studies.
- Conduct psychoacoustic tests using the database of UAM vehicle sounds to provide insights into human response to UAM vehicle noise that would be challenging, in terms of access to stimuli and a wide geographic demographic, for any single agency or organization to acquire.
- Assemble the stimuli and annoyance responses into a database that can be used by members of the UAM community for subsequent analyses.

The novel coronavirus outbreak made it challenging to conduct in-person psychoacoustic tests. For robustness against a novel coronavirus outbreak or similar future events, and to increase the amount of human response data, it was decided that a remote psychoacoustic test platform would be used for the UAM vehicle noise human response study.

1.2. Phases of the UAM Vehicle Noise Human Response Study

The UAM vehicle noise human response study is divided into a feasibility phase and an implementation phase. This paper details planning for the feasibility phase, which will be referred to as the Feasibility Test, that is scheduled to take place in late summer of 2022 with test subjects drawn from different geographic regions of the United States of America. Challenges and possibilities from the Feasibility Test will be used to plan the implementation phase, which will seek to answer research questions on human response to UAM vehicle noise.

The objectives of the Feasibility Test are:

1. Identify potential administrative and technical challenges in using the remote test platform.
2. Compare annoyance responses to those obtained for the same stimuli from a previous psychoacoustic test conducted in a controlled in-person test facility.
3. Determine if providing a contextual cue to test subjects produces a significant change in the annoyance response compared to not providing the cue.
4. Demonstrate the ability to rank sounds by their annoyance response.
5. Demonstrate the ability to compare responses from test subjects grouped by geographic location.

The focus of Objective 1. is on the administration and technical challenges that may arise from using the remote test platform. The objective does not focus on test subject responses to sounds. Section 3. describes Objective 1 in more detail.

Objective 2. determines how well the newly developed remote test platform and test method replicates a controlled in-person test. Other remote testing efforts have found good agreement to in-person laboratory testing [4, 5, 6]. Stimuli in the Feasibility Test will come from a previous in-person test conducted at the NASA Langley Research Center that determined the annoyance response to small unmanned aerial system (sUAS) flyover sounds and ground vehicle sounds [7]. Hence, this initial phase will not include UAM vehicle noise. Section 5. describes how the test will address Objective 2..

For Objective 3., examples of providing contextual cues are telling a test subject to respond how annoying they found a sound after "imagining they are at home" or "thinking about the past week." Contextual cues have been provided for in situ community response testing, such as instructions provided to participants in the Federal Aviation Administration's Neighborhood Environmental Survey [8], but it has not always been applied for laboratory psychoacoustic testing. To help relate laboratory tests to community tests, we seek to understand how the annoyance response changes when a contextual cue is provided to test subjects in a laboratory test. The NASA in-person test from which Feasibility Test stimuli will be drawn did not provide contextual cues to test subjects. Section 6. describes how the Feasibility Test intends to address Objective 3..

For brevity, further details on Objectives 4. and 5. will not be discussed in this paper. However, it is hoped that a ranking as described in Objective 4. would incentivize UAM vehicle manufacturers to release sounds of their vehicles as stimuli for psychoacoustic tests. Once the complete test subject pool is formed, Objective 5. involves categorizing test subjects into distinct geographic regions that are yet to be defined.

2. TEST SUBJECTS

The Feasibility Test will involve a minimum of 80 test subjects. The test subject pool will be divided in half to support Objective 3, with one half being provided a contextual cue, which will be defined in Section 6.2., and the other half not being provided contextual cue. An even number of subjects will be recruited from each yet-to-be-defined geographical region so that there are an equal number of subjects in each region within each group. Test subjects will be recruited from organizations that participate in the UNWG and will not be monetarily compensated for their participation. Test subjects must meet the following requirements:

- Be at least 18 years of age.
- Attest that they have no known hearing impairments.
- Not be aviation noise subject matter experts in their professional duties.
- Be able to access the internet from their own or work computers through a web browser.
- Have their own headphones (no earbuds) for listening to the test sounds.
- One third to two-thirds of the test subject pool must be female.

3. REMOTE TEST PLATFORM

The remote psychoacoustic test platform will be hosted with a NASA-approved cloud service. The server will send subjects the test through a web browser, which the subjects will access from their own computers. For the initial step, test subjects will log in using two-factor authentication and agree to informed consent and privacy act notices. Test subjects will listen to stimuli using their own headphones.

When test subjects enter the test application on their web browser, they will be guided through the test as illustrated in the Figure 1 flowchart. The green arrows indicate the normal flow of the test process. Orange arrows indicate where steps may be repeated. Red arrows lead to exiting the test. For each step, test subjects will have the ability to contact test support. Brief descriptions of problems encountered during each step will be noted by test support staff to achieve Objective 1., but the descriptions will not contain test subject information.

Test subjects then watch an introductory video on the test. It will mention NASA's interest in understanding human response to aviation noise and that the current test is on the feasibility of using a remote testing method. The video will give test subjects an overview of the remaining steps in the test. Test subjects will be able to exit the test at any point once the introduction video starts playing.

A request is then made for test subjects to enter the manufacturer and model of their computer and headphones. This information can be used to understand the spectral characteristics of the test subject sound system and may help explain outlier test subject responses. Test subjects may choose not to enter computer or headphone information.

Test subjects then enter the calibration session. The test application will not have access to the computer sound card and speaker system, and a relatively simple calibration process will be employed. Tentatively, it will involve test subjects rubbing their hands together and then adjusting the volume on their computer to match a recorded sound of someone else rubbing their hands together. After calibration, test subjects will be requested not to adjust the volume on their computers or headphones in subsequent steps.

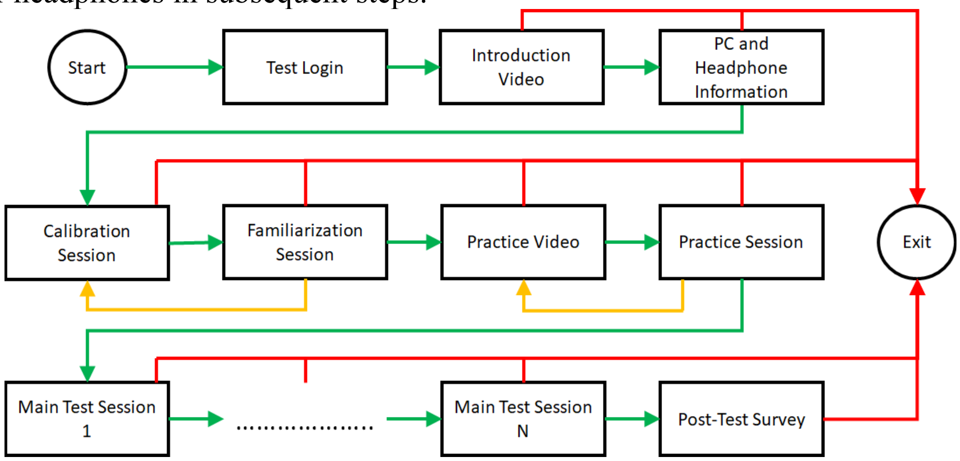

Figure 1: Remote Feasibility Test flowchart.

After calibration, test subjects will begin a familiarization session where they listen to a variety of sounds that will be played to them during the test. Test subjects will not answer any questions during the familiarization step. They will have the option to repeat the calibration step after listening to sounds in the familiarization step. If they choose not to repeat the calibration, test subjects will be reminded not to adjust the volume on their computers or headphones until the test is completed.

Next, the test application plays a video instructing test subjects on the mechanics of taking the test followed by a practice session. Subject responses during the practice session will not be included in the main test response database. Test subjects will have the option to repeat the practice video and session.

During a Feasibility Test session, sounds will be played to test subjects one at a time. After a sound is played, the question in Figure 2 is displayed to acquire the annoyance response to the sound. A test subject then moves the red slider to any point on the annoyance rating scale. Numerical annoyance ratings are real values ranging from 1.00 to 11.00. An annoyance rating of "2" corresponds with a rating of "Not at All Annoying," and a rating of "10" corresponds with a rating of "Extremely Annoying." After selecting an annoyance rating, a test subject will press another button, not shown in Figure 2, to play the next sound in the test.

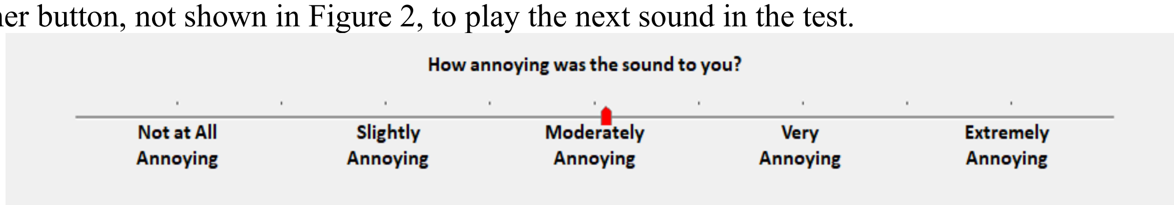

Figure 2: Test question prompt for each stimulus.

All the test stimuli will be stored on the cloud service. The cloud service will send each stimulus to each test subject when the subject is ready to respond to the stimulus. Stimuli presentation order will be stored on the cloud service, and it will be unique to each subject. The stimuli order for each test subject will be predetermined during test design to minimize presentation bias.

There will tentatively be four main test sessions, with breaks in between, to reduce test subject fatigue.

After completing the main test sessions, test subjects will be asked to answer two questions for a post-test survey. The first question is whether the test subject adjusted the volume on their headphones or computer after the familiarization step. If a test subject indicates that they adjusted their volume after the familiarization step, their test responses will not be analyzed. The second question allows the test subject to select multiple options from a list of 12 choices that indicate potential difficulties they may have encountered while taking the test. The default option of the 12 choices is that they had no difficulties. For brevity, all 12 options will not be listed for this paper, but examples of some choices include "Not able to properly hear test sounds," "Found calibration step difficult," and "I had distractions during the test." "Other issues" is also one of the options. Test subjects will be requested to elaborate on their selected options in a space that can accept typed text.

For each test subject, the cloud service will store an assigned test subject number. When test results are analyzed, only the test subject number and United States zip code will be associated with responses returned from the subject. Zip codes are needed for any geographic-based analysis (Objective 5), which is not discussed in this paper. Other test subject information, like name and email address, will not be used after the test is completed.

4. STIMULI

4.1. Previous In-Person Test

All stimuli for the Feasibility Test will be drawn from the NASA Design Environment for Novel Vertical Lift Vehicles psychoacoustic test, which is referred to as WGA-I [7]. This in-person test was conducted in February 2017 in the NASA Langley Research Center Exterior Effects Room (EER) [9]. The sounds were presented using the 3D sound reproduction capability of the EER. This test compared the annoyance response to sounds of sUAS and sounds of ground vehicles. It also found the annoyance response to sounds of aircraft with Distributed Electric Propulsion (DEP). A total of 103 nonunique sounds were played to WGA-I test subjects with a sampling rate of 44.1 kHz. Some of the stimuli were formed by simply adjusting the gain on other stimuli. Sixty-two of these sounds were of sUAS flyover recordings. Twelve sUAS flyover sounds were auralizations, which are sounds generated from numerical data [10]. Twenty sounds were of ground vehicle recordings. The remaining nine sounds were auralizations of DEP aircraft. Sound durations varied between 12 and 52 seconds, with a sample mean duration of approximately 24 seconds. Throughout the test, artificial ambient noise that recreated the ambient noise level observed at the loudest recording site was played as a three-minute loop.

WGA-I tested 38 test subjects in the EER. When test subjects were played a sound, they provided their response by answering the question "How annoying was the sound to you?" as displayed in Figure 2.

Among the metrics calculated for each stimulus was A-weighted Sound Exposure Level (SEL_A). One of the main results of WGA-I is regenerated in Figure 3. Here, augmented linear regression fit two lines to the data: one for responses to 46 of the 62 sUAS recorded flyover sounds and one for responses to the ground vehicle sounds, both as a function of the SEL_A of each sound. Responses to auralizations are not included in Figure 3. The 16 sUAS recorded flyovers that are excluded from Figure 3 are repetitions of included stimuli or are unique recordings that repeat flyover conditions using the same aircraft as included stimuli. Both regression lines in Figure 3 were assumed to have the same slope. Each marker in Figure 3 is the mean annoyance response, with a 95% confidence interval, to a sound from all 38 test subjects. An R^2 value of 0.82 indicates the regression line pair captures the annoyance variation relatively well. The SEL_A offset between the sUAS flyover sounds and ground vehicle lines is 5.64 dB. The offset indicates that an sUAS flyover will need an SEL_A that is approximately 5.64 dB lower to have a similar annoyance response as a ground vehicle.

To understand if the offset is significant, bootstrapped regressions, as described in Ref. [7], were run to regenerate 95% confidence intervals for regression model parameters and the coefficient of determination, R^2. As explained by Ref. [7], bootstrapping reduces R^2 by reintroducing variance into the response data through resampling. Reference [7] loosely associated having a lower R^2 confidence interval bound of 0.60 with a poor augmented linear regression result. Bootstrap regression results for the Figure 3 responses are shown in the first of row of Table 1. The lower percentile of the R^2 confidence interval remains above 0.60, further validating that the augmented linear regression forms a relatively good fit to the data. The lower percentile of the offset confidence interval is greater than 0 dB, indicating that there is a significant difference between responses to sUAS noise and ground vehicle noise.

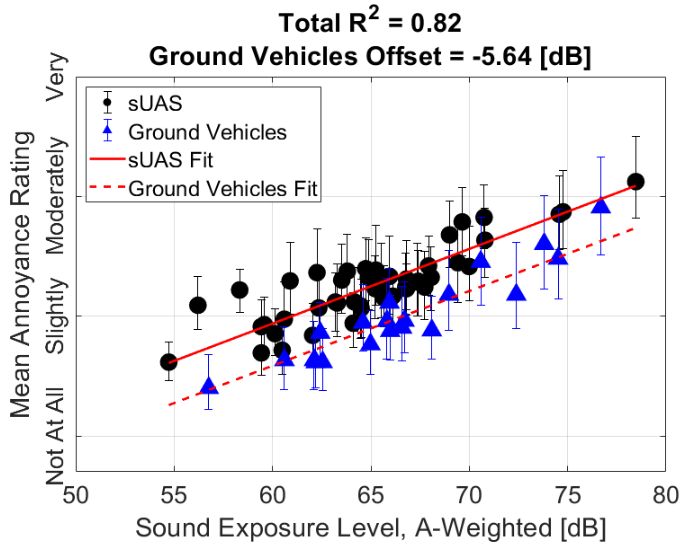

Figure 3: Augmented linear model regression results for SEL_A.

Table 1: Bootstrapped 95% confidence intervals (CI) for the augmented linear regression model.

Response Data	Median R^2	R^2 CI	Median Offset [dB]	Offset CI [dB]
WGA-I Response	0.71	[0.64, 0.78]	5.64	[4.3, 7.2]
WGA-I Response Without Sound 465	0.71	[0.63, 0.78]	5.84	[4.4, 7.4]

4.2. Feasibility Test Stimuli

A consideration to limit the remote main test duration to one hour will also limit the number of Feasibility Test stimuli. Long test durations may increase the likelihood of remote test subjects encountering distractions or exiting the test before completion. In the interest of test replication, it is preferred that all WGA-I stimuli be used in the Feasibility Test, but using all WGA-I stimuli will likely cause main test durations for most remote test subjects to exceed one hour. All the sounds needed to generate Figure 3 and the data in Table 1 will be used for the Feasibility Test in order to achieve Objective 2.. If only these sounds were used in the Feasibility Test, the main test duration is expected to be 45 minutes, which assumes a test subject response time of five seconds for each stimulus and five-minute breaks between main test sessions. The extent to which the 16 excluded sUAS recordings (see Section 4.1.) and the auralization are to be included will be influenced by the desire to keep the main test duration from exceeding one hour.

Test subject computers and headphones may not be able to reproduce sounds with absolute sample values greater than 1. Only one WGA-I sound, that of a Step Van ground vehicle drive-by, which Ref. [7] associates with sound identification number 465, has sample values greater than 1 Pa. All other WGA-I stimuli are available to be played to test subjects at their true absolute sound pressure levels of less than 1 Pa. While using a normalized version of Sound 465 in the test remains a possibility, the stimulus as currently given will be discarded. Bootstrapped regression parameters from responses excluding those of Sound 465 are given in the second row of Table 1. The R^2 confidence interval changed slightly. The offset between the response regression lines to sUAS sounds and ground vehicle sounds increased to 5.84 dB, and the confidence interval shifted slightly toward higher offset values. Removing Sound 465 caused a greater response difference between sUAS flyover and ground vehicle sounds.

Artificial ambient noise was played throughout WGA-I, as described near the beginning of Section 4.1.. The remote test platform will not have the capability to mix sounds for the Feasibility Test, so artificial ambient noise will be premixed with the stimuli before the test. Unlike in WGA-I, the artificial ambient noise will not continue to play between sound stimuli.

To closely match the sound presentation mode of WGA-I, the Feasibility Test stimuli will be presented binaurally to the test subjects. Binaural simulation will be used to render the monaural recordings to binaural stimuli using tracking data. A check will be made to ensure that all stimuli have absolute sound pressure values less than 1 Pa in both channels and that the relative levels between stimuli are maintained.

5. REPLICATING WGA-I RESULTS

5.1. Feasibility Test Outcomes

Results from the Feasibility Test will be compared with WGA-I results shown in Figure 3 and Table 1 in order to address Objective 2.. This comparison will initially be performed for the subject group that is not provided a contextual cue to match the WGA-I test conditions. Comparing the mean response and confidence intervals of each stimulus between the two tests is a possible analysis method, but there likely will be deviations, which may not necessarily indicate different overall results. Therefore, a more robust comparison will be based on aggregate results, such as those shown in Figure 3 and Table 1.

Analysis for Objective 2. mainly compares responses to sUAS and ground vehicles for just the Feasibility Test as was done for WGA-I responses in Figure 3.

gives the potential bootstrap regression outcomes, which are labeled O1-O3, of this analysis in terms of the R^2 and offset confidence intervals. The R^2 confidence interval threshold of 0.60 for outcomes O1 and O2 comes from the threshold set by Ref. [7], which was mentioned near the end of Section 4.1..

Table 2: Potential 95% bootstrap regression confidence interval (CI) outcomes for comparing Feasibility Test sUAS and ground vehicle responses.

Analysis Outcome Reference	R^2 CI	Offset CI [dB]
O1	¿0.60	¿0
O2	≤0.60	¿0
O3	-	≤ 0

Outcome O1 would indicate that the annoyance response difference between sUAS and ground vehicle sounds was the same as in Figure 3. For this outcome, there will need to be an additional pair of augmented linear regression analyses that compare Feasibility Test and WGA-I responses. One analysis will perform the regression between responses to sUAS sounds from the two tests, and the other analysis will compare their responses to ground vehicles. If offset confidence intervals from bootstrap regression from both analyses are less than or equal to zero dB, it would indicate that the remote test did not introduce an annoyance bias relative to in-person testing. Other bootstrap regression offset confidence interval results would indicate the introduction of an annoyance bias. Even with an annoyance bias, Outcome O1 would still show the Feasibility Test as capable of discerning the annoyance response difference between sUAS and ground vehicle noise.

Outcome O2 will warrant further analyses. One way it could occur is if the regression line slopes for the Feasibility Test responses to sUAS and ground vehicle sounds cannot be assumed as identical. There could still be a difference in the response to the two vehicle classes. The individual R^2 of the two vehicle classes will need to be checked. Higher response variance can also produce Outcome O2 as Section 5.2. will illustrate.

Outcome O3 would indicate that the Feasibility Test did not find a difference in the annoyance response between sUAS and ground vehicle sounds. For this result, the R^2 confidence interval can approach values that will occur if there was instead a single regression line, which may or may not overlap with 0.60. The single R^2 value of the augmented linear regression or the median R^2 value of the bootstrapped regression can help interpret Outcome O3. If either of these values are relatively high, one possibility is that the Feasibility Test subjects are inherently responding differently than WGA-I test subjects. Higher response variances in the Feasibility Test compared to those in WGA-I can cause these R^2 values to be relatively low, which may indicate problems with the testing platform and test method. Understanding Outcome O3 may involve analyzing the spectra of the computer and headphone sound systems of test subjects who provide this information.

5.2. Potential Effects of Increased Variance

Since there is less control over the Feasibility Test environment compared to WGA-I, there may be more variance in the Feasibility Test annoyance responses. To understand how increased variance in responses may affect test outcomes, the variance of the original WGA-I response data was artificially increased, then bootstrapped regressions were performed based on the artificial data.

Figure 4 illustrates how variance was increased with two examples. The black circles are the actual responses to two stimuli from WGA-I. The vertical green lines represent boundaries that are a certain fraction of the annoyance rating range between the mean response and extreme annoyance rating values of 1 and 11. This certain fraction, which defines the green line positions, is chosen to produce a particular variance increase percentage over all stimuli responses. In the Figure 4 example, a 10% variance increase was sought over all stimuli. WGA-I responses under the green arrows are shifted in the direction of the arrows by a linear ramp to reach the annoyance rating extremes of 1 or 11 to generate the blue circles, which Figure 4 offsets above the black circles. With the linear ramp, the lowest and highest annoyance ratings are shifted to 1 and 11, respectively, as shown for Stimulus Example 1. Responses under the green arrows that are closer to the green lines are shifted by a lesser amount. If a WGA-I annoyance rating is already at 1 or 11, no shift in the annoyance response occurs, as shown in Stimulus Example 2. This method attempts to minimize change in the mean response to each stimulus, which Figure 4 represents by how much the dashed red line (mean of artificial response) is shifted in annoyance ratings from the solid red line (mean of WGA-I response). Minimization performance is roughly measured by the root-mean-square error (RMSE) over all stimuli and is given by:

$$\overline{A}_{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\overline{A}_i - \overline{A}_{WGA-I,i}\right)^2}{N}} .$$ 22* MERGEFORMAT ()

In Equation 2, N is the number of test stimuli, $\overline{A}_i$ is the mean response to artificial stimulus i, and $\overline{A}_{WGA-I,i}$ is the mean response to the ith original WGA-I stimulus.

Artificially increasing the WGA-I response data variance as illustrated by Figure 4 produces the bootstrapped regression results in Table 3 for different variance increases over all responses (responses to Sound 465 were excluded). The R^2 and offset confidence intervals are for the augmented linear regression to sUAS and ground vehicle sound responses. The second column gives the RMSE, $\overline{A}_{RMSE}$, from Equation 2, for every variance increase condition, and they are relatively small. Table 3 shows the offsets between the sUAS and ground vehicle responses remain significant. The results suggest Outcome O2, where lower R^2 values warrant further analyses, can occur if the overall response variance in the Feasibility Test is increased by more than 20% compared to the WGA-I overall response variance.

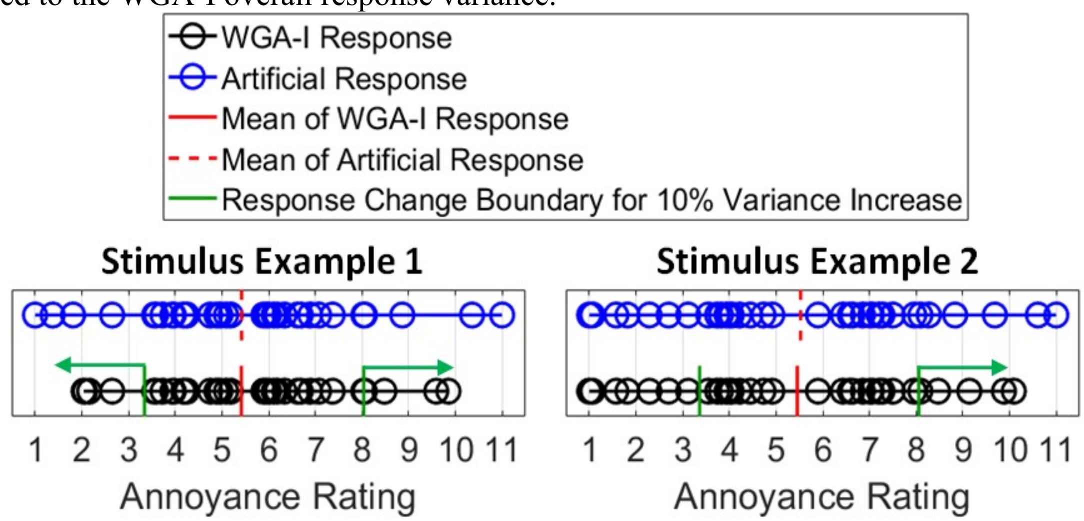

Figure 4: Artificially increasing annoyance response variance.

Table 3: Bootstrapped 95% confidence intervals (CI) for artificial response data. $\overline{A}_{RMSE}$ is in units of annoyance ratings.

Artificial Variance Increase Percentage	$\overline{A}_{RMSE}$	Median R^2	R^2 CI	Median Offset [dB]	Offset CI [dB]
5%	0.02	0.71	[0.62, 0.78]	5.81	[4.4, 7.4]
10%	0.04	0.70	[0.62, 0.77]	5.72	[4.2, 7.4]
20%	0.05	0.69	[0.60, 0.76]	5.69	[4.2, 7.3]
40%	0.10	0.66	[0.56, 0.75]	5.95	[4.3, 7.8]

6. EFFECT OF PROVIDING CONTEXTUAL CUE

6.1. Background

Examples of contextual cues were provided in Section 1., and previous laboratory psychoacoustic tests conducted at NASA provide some indication when contextual cues may affect annoyance responses. A contextual cue may mitigate absolute response differences when conducting a test in different environments. Reference [11] compared subjective responses from outdoor and indoor acoustic simulations without providing contextual cues. While there was high correlation between responses from the outdoor and indoor tests, the absolute response levels were different. Reference [12] instructed test subjects to imagine they are at home in three different test settings: in an anechoic chamber, in a semireverberant room, and while wearing earphones. Results from the three acoustic environments were equivalent. Similar findings appeared in Ref. [13], which provided identical test instructions to test subjects wearing headphones and to subjects in outdoor and indoor acoustic simulation rooms. In a preliminary test, Ref. [13] provided separate instructions to two different test subject groups and found that, for the same test location, the presence of a contextual cue does not appear to affect the annoyance response compared to not providing the cue. Reference [14] asked test subjects to respond to noise while not providing a time context and again while thinking about different times of a day. Having test subjects focus on nighttime appears to noticeably affect responses relative to when test subjects are asked to focus on noise during daytime. From these limited previous studies, a hypothesis for the Feasibility Test is that providing a contextual cue will not significantly affect responses if test subjects are not asked to think about nighttime noise.

Contextual cues have been developed in the interest of producing comparable survey results for community testing. The Community Response to Noise Team of The International Commission on the Biological Effects of Noise recommended that, when performing community noise testing, test subjects should be asked the following question: "Thinking about the last (… 12 months or so ...), when you are here at home, how much does noise from (… noise source ...) bother, disturb, or annoy you; Extremely, Very, Moderately, Slightly or Not at all [15]?" The items in parentheses may be replaced by more appropriate phrases. This recommendation provides a contextual cue through which test subjects respond to sounds. No such contextual cue was provided in the WGA-I laboratory test. Test subjects were simply asked, "How annoying was the sound to you?" after hearing the sound.

Objective 3. determines if providing a contextual cue significantly affects laboratory test responses. This question may be important for relating laboratory results to community testing where multiple sounds from an aircraft fleet may be heard. As discussed earlier, the pool of test subjects will be divided into two groups, where one group will be provided a contextual cue for the test and the other group won't be provided a contextual cue.

6.2. Specification of the Contextual Cue

Specification of the contextual cue began with the recommendation from Ref. [15]: "Thinking about the last (… 12 months or so ...), when you are here at home, how much does noise from (… noise source ...) bother, disturb, or annoy you?" The recommendation was modified by considering the following attributes of the Feasibility Test:

- Test subjects cannot be assumed to take the test in their home.
- The test will refer to stimuli as "sounds" and not "noise."
- The test is not a survey, and test subjects may have never experienced hearing the sounds.
- We are interested in relating responses to individual sounds and responses when hearing the sounds multiple times during a day.

After considering these attributes, the following question was generated to provide a contextual cue: "Imagine hearing this sound several times each day while outdoors and near your home, how annoying would this sound be to you?" It will be asked after each stimulus is played to the subjects receiving the contextual cue. Additionally, instructions provided during the Practice session will emphasize the context question to these subjects. The test subject group without a contextual cue will be asked "How annoying was the sound to you?" as shown in Figure 2 after each stimulus is played.

6.3. Analyzing Outcomes

Objective 3., determining if providing contextual cues affects responses to sounds, can be analyzed similarly to the results in Figure 3 and Table 1. Instead of analyzing responses between sUAS and ground vehicle sounds, responses will be analyzed between the test subject groups with and without a contextual cue. Figure 5 shows augmented linear regression to artificially created responses to sUAS sounds from the two groups of test subjects. The artificial responses in Figure 5 were created by randomly and uniformly perturbing each sUAS WGA-I response within 0.8 annoyance ratings from the original response. For the right plot, an extra step of shifting down one annoyance rating was taken for the responses hypothetically coming from the group provided a contextual cue. No annoyance rating was permitted to shift below 1 or above 11. The same analysis will be repeated for ground vehicle sound responses. If the results are similar to the left graph for both vehicle classes, it will support the hypothesis that providing a contextual cue does not provide a significant change in response.

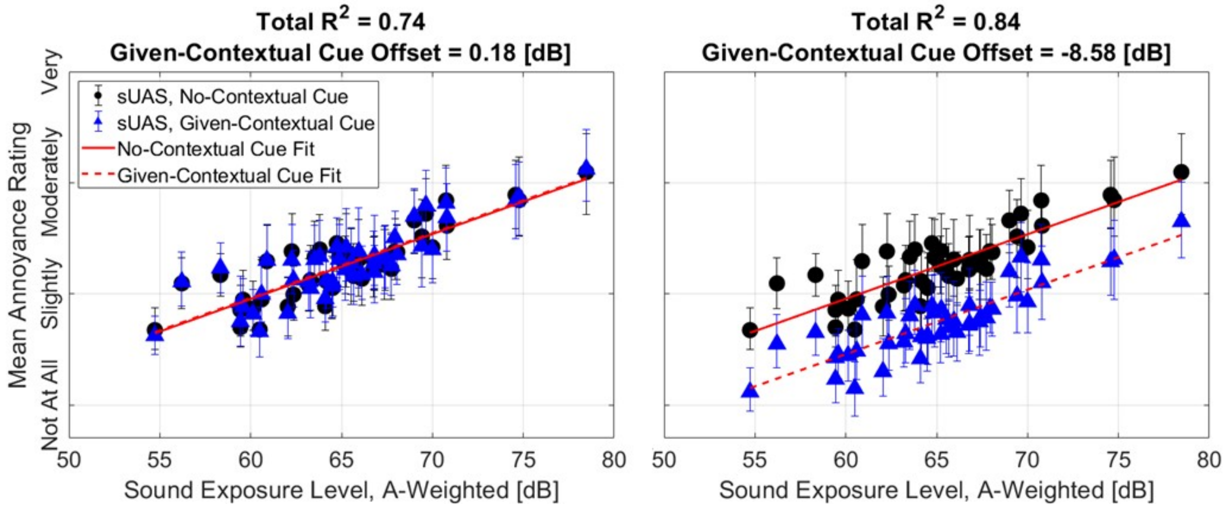

Figure 5: Two potential Feasibility Test response outcomes.

7. CONCLUSIONS AND CURRENT PLATFOM STATUS

An overview was described of the development and operation of a remote psychoacoustic test platform for gathering responses from test subjects during the feasibility portion of a UAM vehicle noise human response study. Stimuli will be presented that will help quantify where the remote test method replicates the results of an in-person test. The stimuli to be used for the Feasibility Test were described, although the final set of stimuli is still being determined. Analysis approaches were discussed that will be used to compare the remote test and in-person test responses and to determine if providing a contextual cue to test subjects significantly affects the annoyance responses.

The platform is currently being tested, and different components, such as videos and the calibration method, are being finalized. Execution of the Feasibility Test in late summer 2022 will guide any adjustments to the platform and test method before they are used for an actual annoyance study to UAM noise.

8. ACKNOWLEDGEMENTS

Development of the remote psychoacoustic test platform and execution of the feasibility portion of the UAM vehicle noise human response study is supported by the Revolutionary Vertical Lift Technology Project of the NASA Advanced Air Vehicles Program and the Structural Acoustics Branch at the NASA Langley Research Center. Analytical Mechanics Associates, Inc, is helping NASA develop and utilize the remote test platform for the Feasibility Test.

9. REFERENCES

1: , [1],
2: , [2],
3: , [3],
4, 5, 6: , [4, 5, 6],
7: , [7],
8: , [8],
9: , [9],
10: , [10],
11: , [11],
12: , [12],
13: , [13],
14: , [14],
15: , [15],