NASA/CR–20205005471

Independent Analyses of Galveston QSF18 Social Survey

Sanford Fidell
Fidell Associates, Inc., Woodland Hills, California

Richard Horonjeff
RDH Acoustics, Peterborough, New Hampshire

Barbara Tabachnick
Independent Consultant, Canoga Park, California

Steven Clark
Scientific Telephone Samples, Rancho Santa Margarita, California

August 2020

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/CR–20205005471

Independent Analyses of Galveston QSF18 Social Survey

Sanford Fidell
Fidell Associates, Inc., Woodland Hills, California

Richard Horonjeff
RDH Acoustics, Peterborough, New Hampshire

Barbara Tabachnick
Independent Consultant, Canoga Park, California

Steven Clark
Scientific Telephone Samples, Rancho Santa Margarita, California

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract 80LARC17C0003

August 2020

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

INDEPENDENT ANALYSES OF GALVESTON QSF18 SOCIAL SURVEY

Sanford Fidell, Richard Horonjeff, Barbara Tabachnick and Steven Clark

PREPARED FOR:

AMA Purchase Order T3-0407-FY19
NASA Langley Research Center
Structural Acoustics Branch
Hampton, VA 23681

PREPARED BY:

Fidell Associates, Inc.
23139 Erwin Street
Woodland Hills, CA 91367

16 June 2020

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Tables

List of Figures

1 EXECUTIVE SUMMARY

This document is the final report of independent analyses of social survey and other information collected during NASA's QSF18 field study in Galveston, TX (Page *et al.*, 2020). The report presents the findings of descriptive and inferential analyses conducted *after* completion of the field study and of a series of data quality control screenings. These analyses were focused on test participants' annoyance judgments and dose-response relationships. The report also discusses implications of these findings for future X-59 community response testing.

Major findings, some of which arose from more than one analysis, are noted immediately below, and in greater detail, in subsequent sections of the report.

Data Set Findings

- A pre-screening of case records in both the prompt response and delayed response data sets was necessary to avoid various ambiguities and uncertainties in data archiving.
- Of 7,068 prompt response database records, 4,169 (59%) merited analysis for present purposes. Of the 2,855 delayed response records, 1,952 (68%) merited analysis for present purposes.
- The screened data produced analyzable results for three degrees of annoyance judgments of low-amplitude sonic booms: slightly and greater, moderately and greater, and highly (very or extremely) degrees of annoyance.
- Only 33 (0.8%) of the prompt and 15 (0.8%) of the delayed annoyance judgments fell in the high annoyance category.
- Of the 500 panelists responding to participation confirmation requests, 354 (71%) who provided one or more prompt responses remained after screening, and 386 (77%) who provided one or more delayed responses remained after screening.
- After screening, 728 case records in which a panelist was determined to be at home, and 295 case records where a panelist was determined to be at work, were available for location-based analyses of annoyance judgments.
- Half of the booms (27) occurred within 30 minutes of the time of occurrence of a prior boom, complicating efforts to establish a reliable association between dose and annoyance judgments.

Dose-response findings

- High annoyance was only rarely self-reported. Only about 2% of panelists reported high prompt annoyance judgments at 75 PLdB, while between 3% and 4% did so at 85 PLdB;

- Reminder messaging mode and panelist location both seemed to systematically affect the prevalence of reported annoyance, but multi-level logistic modeling failed to confirm statistically significant effects other than steeper high-annoyance dosage-response slope for e-mail than for text reminders;

- Panelists showed signs of sensitization to low-amplitude sonic booms after the first three days of participation in the QSF18 data collection exercise, with the prevalence of annoyance (to any degree) increasing two-fold thereafter;

- Judgments of the annoyance of exposure to low-amplitude sonic booms increased when rattle and vibration were noticed (typically at lower exposure levels);

- The notice of rattle and vibration increased as expected with boom sound level. At 65 PLdB, approximately 10% of panelists reported noticing rattle or vibration; at 75 PLdB about 30% did; and at 85 PLdB, 50% did;

- Evidence of a non-zero asymptote in dose-response relationships was found for slightly or greater degrees of annoyance judgments;

- Delayed (end-of-day) judgments of annoyance with cumulative exposures to low-amplitude sonic booms were increasingly likely to increase in degree ("slightly", "moderately", and "highly" annoying) as the duration of participation in the data collection exercise increased. (This finding might be an artifact of a partial confounding in the study design of boom presentation level with study duration);

- Both prompt and delayed judgments of the annoyance of exposure to low-amplitude sonic booms increased when rattle and vibration were noticed (typically at lower exposure levels);

- The prevalence of judged annoyance of booms increased with elapsed time for both delayed and prompt annoyance judgments, over all three degrees of annoyance judgments (slightly or more greatly annoyed, moderately or more greatly annoyed, and highly annoyed). The QSF18 study design did not permit tests of whether this greater annoyance over time was due to cumulative noise exposure or annoyance with the multiple responses required; and

- Effects of panelist location on dose-response relationships were consistently absent, but lacking in generalizability due to the small subsample of events for which the panelists' locations could be ascertained. The large "Undetermined" location category does in fact contain geocoded positions from which it may be possible to resolve an additional number of at-home and at-work panelists for the statistical analyses.

- No statistically significant differences were observed in annoyance judgments at any exposure level associated with the sending of reminder messages. The dose-response curve was steeper for e-mail than for text reminders, however. As a consequence, text reminders were associated with greater frequency of reports of high annoyance at noise levels below 75 PLdB, while e-mail messages were associated with greater frequency of high annoyance reports at higher noise levels. However, no statistically significant differences between text and e-mail messages were observed for responses of at least moderate or at least slight annoyance.

- The prevalence of annoyance in all degrees was lower for text reminder messages than for no and e-mail reminder messages. The lower prevalence rate may have been due to the addition of considerable numbers of "not "noticed"" responses to the database by text reminder messages (and to a more limited degree, by e-mail reminder messages as well). Overall, the frequency of reports of high annoyance was greater for text than for e-mail reminder messages, but this finding may have simply been an artifact of the greater number of text messages and noise events at exposure levels lower than 75 PLdB. The steeper dosage-response slope for e-mail than for text messages shows greater annoyance when text messages were associated with lower noise doses and greater annoyance when e-mail messages were associated with higher noise doses.

- Additional analyses in which differences in delayed responses were examined between those who received only morning and evening reminder messages and those who received messages after each noise event reinforced the finding of steeper slopes for e-mail than text messages. Again, greater prevalence of high annoyance (and also at least moderate annoyance) was found with text messages at lower noise levels, and greater prevalence of high annoyance (and at least moderate annoyance) was found with e-mail messages at higher sound levels (above a PLDNL of 20 PLdB). Further, a stronger Dose-response relationship was found for high annoyance judgments for panelists issued multiple messages than for panelists issued only two reminders messages. Greater incidence of high annoyance was consistently associated with multiple reminder messages than with only two messages.

Descriptive Findings

- Panelists rendered the preponderance of the prompt annoyance judgments over a Perceived Level range of 65 to 85 PLdB. They rendered the preponderance of the delayed annoyance judgments over a Day-Night Average Perceived Level range between 20 and 40 PLdB;

- The average annoyance judgment rate for prompt responses differed by reminder message group: 10% for no reminder messages, 21% for e-mail reminders, and 32% for text reminders;

- The proportions of male and female panelists recruited for the panel sample were representative of the population of Galveston, but QSF18 panelists were older and more highly educated than Galveston's population. It is not known whether the non-stratified, address-based sample was representative of Galveston's population with respect to other demographic variables (*e.g.*, ethnicity, home ownership, household size, composition, annual income, *etc*.);

- The average rate of end-of-day annoyance judgments was 56%;

Inferential Findings

- Increased latency of annoyance judgments was associated with e-mail reminder messages when compared with either text reminders or no reminders. The percentage of panelists providing prompt responses within 30 minutes was 70% for the no-reminder group, 44% for the e-mail group and 77% for the text reminder group; and

- Increased latency of annoyance judgments was also associated with panelist location at the time of the boom. Of the At-Home panelists, 52% provided prompt responses within 30 minutes, while 45% of the At-Work did so.

Table 1 summarizes the findings of all of the inferential analyses. The rows of Table 1 describe the findings of analyses of the effects shown in the leftmost column. The second column entry indicates the questionnaire item that produced the analyzed responses. The third column describes the analysis type, while the fourth describes the degrees of annoyance judgments ("not at all annoyed" through "extremely annoyed") to which the findings pertain. The fifth column categorizes whether the analyzed annoyance judgments were associated with single boom events or to cumulative (end-of-day) exposure. The sixth column summarizes the finding of the analysis, while the seventh column shows the *p* value for the analysis. The only *p* values that may be unambiguously interpreted as unlikely to have arisen by chance alone are highlighted in **blue** in Column 7 of Table 1. The final column provides a reference to the table reporting detailed findings of the analysis.

Table 1: Summary of findings of inferential analyses

EFFECT OF INTEREST	QUESTIONNAIRE ITEM	ANALYSIS TYPE	ANNOYANCE CATEGORY	PROMPT/DELAYED RESPONSE	FINDING	*P*-VALUE[1]	TABLE
Reminders on annoyance	Multiple *vs.* two reminder messages	MLM logistic	Highly Annoyed	Delayed	No message more annoying[2]	**<.001**	Table 66
↓	Reminder messages vs. none	MLM logistic	Highly Annoyed	Prompt	No difference	0.55	Table 19
↓	Multiple *vs.* two reminder messages	MLM logistic	Moderately Annoyed	Delayed	No difference	0.062	Table 67
↓	Reminder messages *vs.* none	MLM logistic	Moderately Annoyed	Prompt	No difference	0.438	Table 20
↓	Multiple *vs.* two reminder messages	MLM logistic	Slightly Annoyed	Delayed	No difference	0.857	Table 68
↓	Reminder messages *vs.* none	MLM logistic	Slightly Annoyed	Prompt	No difference	0.353	Table 21
↓	E-mail *vs.*Text reminder message	MLM logistic	Highly Annoyed	Delayed	Text more annoying[2]	**<.001**	Table 66
↓	↓	MLM logistic	Highly Annoyed	Prompt	Text more annoying[2]	**<.001**	Table 19
↓	↓	MLM logistic	Moderately Annoyed	Delayed	Text more annoying[2]	**<.001**	Table 67
↓	↓	MLM logistic	Moderately Annoyed	Prompt	No difference	0.756	Table 20
↓	↓	MLM logistic	Slightly Annoyed	Delayed	Text more annoying	0.009	Table 68
↓	↓	MLM logistic	Slightly Annoyed	Prompt	No difference	0.205	Table 21
↓	Multiple *vs.* two reminder messages by noise interaction	MLM logistic	Highly Annoyed	Delayed	Steeper slope of Dose-response for e-mail reminders	**<.001**	Table 66

EFFECT OF INTEREST	QUESTIONNAIRE ITEM	ANALYSIS TYPE	ANNOYANCE CATEGORY	PROMPT/DELAYED RESPONSE	FINDING	*P*-VALUE[1]	TABLE
↓	Reminder message/no-reminder by noise interaction	MLM logistic	Highly Annoyed	Prompt	No difference	0.478	Table 22
↓	Multiple *vs.* two reminder messages by noise interaction	MLM logistic	Moderately Annoyed+	Delayed	No difference	0.015	Table 67
↓	Reminder message/no-reminder by noise interaction	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.2	Table 20
↓	Multiple *vs.* two reminder messages by noise interaction	MLM logistic	Slightly Annoyed+	Delayed	No difference	0.545	Table 68
↓	Reminder message/no-reminder by noise interaction	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.419	Table 21
↓	E-mail/Text reminder by noise interaction	MLM logistic	Highly Annoyed	Delayed	Steeper slope of dose-response for e-mail reminders	**<.001**	Table 66
↓	↓	MLM logistic	Highly Annoyed	Prompt	Steeper slope of dose-response for e-mail reminders	.001	Table 19
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	Steeper slope of dose-response for e-mail reminders	**<.001**	Table 67
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.741	Table 20
↓	↓	MLM logistic	Slightly Annoyed+	Delayed	Steeper slope of dose-response for e-mail reminders	0.006	Table 68
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.314	Table 21

EFFECT OF INTEREST	QUESTIONNAIRE ITEM	ANALYSIS TYPE	ANNOYANCE CATEGORY	PROMPT/DELAYED RESPONSE	FINDING	*P*-VALUE[1]	TABLE
Sequential Effects on annoyance	Duration of participation in study	MLM logistic	Highly Annoyed	Delayed	Later exposure more annoying	**<.001**	Table 44
↓	↓	MLM logistic	Highly Annoyed	Prompt	Later exposure more annoying	0.023	Table 41
↓	↓	MLM logistic	Moderately Annoyed+	Delayed	Later exposure more annoying	**<.001**	Table 45
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	Later exposure more annoying	**<.001**	Table 42
↓	↓	MLM logistic	Slightly Annoyed+	Delayed	Later exposure more annoying	**<.001**	Table 46
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	Later exposure more annoying	0.006	Table 43
Effect of location on annoyance judgment	Panelist at-home *vs.* at work	MLM logistic	Highly Annoyed	Prompt	No difference	0.992	Table 29
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.074	Table 30
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.943	Table 31
↓	Location known *vs.* Undetermined	MLM logistic	Highly Annoyed	Prompt	No difference	0.572	Table 32
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.734	Table 33
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.383	Table 34
↓	Home/Work by noise interaction	MLM logistic	Highly Annoyed	Prompt	No difference	0.954	Table 29
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.151	Table 30

EFFECT OF INTEREST	QUESTIONNAIRE ITEM	ANALYSIS TYPE	ANNOYANCE CATEGORY	PROMPT/DELAYED RESPONSE	FINDING	*P*-VALUE[1]	TABLE
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.75	Table 31
↓	Location determined by noise interaction	MLM logistic	Highly Annoyed	Prompt	No difference	0.72	Table 32
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	No difference	0.72	Table 33
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	No difference	0.449	Table 34
Rattle or vibration	Boom Noticed *vs.* not noticed	MLM logistic	Highly Annoyed	Prompt	Greater annoyance when noticed	0.005	Table 47
↓	↓	MLM logistic	Moderately Annoyed+	Prompt	Greater annoyance when noticed	**<.001**	Table 48
↓	↓	MLM logistic	Slightly Annoyed+	Prompt	Greater annoyed when noticed	**<.001**	Table 49
Effect of reminders on attrition	Reminder message *vs.* no-reminder	ANOVA planned comparison	Number of days in study	Delayed	No effect	0.619	Table 56
↓	E-mail *vs.* Text reminder message	ANOVA planned comparison	E-mail *vs.* Text reminder message	Delayed	No effect	0.66	Table 57
Effect of reminder message on latency of annoyance judgment	Reminder message *vs.* no-reminder message	MLM regression	Time between exposure and response	Prompt	Greater latency with reminder	0.002	Table 53
↓	E-mail *vs.* Text reminder message	MLM regression	Time between exposure and response	Prompt	Greater latency with e-mail	**<.001**	Table 53

EFFECT OF INTEREST	QUESTIONNAIRE ITEM	ANALYSIS TYPE	ANNOYANCE CATEGORY	PROMPT/DELAYED RESPONSE	FINDING	*P*-VALUE[1]	TABLE
Missing data interpretation	Boom not noticed *vs.* imputed	NA	NA	NA	(Too few missing data)	NA	NA

[1] Interpretations of *p* values numerically equal to or greater than .001 are dubious because of the large number of inferential tests performed, and the consequent enhanced risk of false positive errors in the present set of analyses. Statistically meaningful *p* values are highlighted in **blue** in this table.

[2] Interpretations of main effects can be misleading when they are part of statistically significant interactions.

THIS PAGE INTENTIONALLY LEFT BLANK

2 INTRODUCTION

2.1 Purpose of Report

This document is the final report of independent analyses of social survey and other information collected during NASA's QSF18 field study in Galveston, TX (Page *et al.*, 2020). In addition to inferring dose-response functions, analyses were conducted of the conditional effects of:

- e-mail *vs.* telephone text message reminders to panelists on dose-response functions and attrition rates;
- the frequency (schedule) of test flights on the reliability of association of responses with exposures and response latencies;
- sequential biases on the degree of annoyance judgments;
- panelist recruitment methods on the demographic representativeness of the panel sample of the Galveston population;
- the locations of panelists at the time of exposure;
- missing data;
- binning of noise exposure into intervals; and
- exposure on reports of rattle and vibration.

The results of the present analyses may differ somewhat from those of similar analyses reported by Page *et al.* (2020) due to differences in numbers and types of case records considered eligible for analysis. Unless otherwise noted (as in Appendix A), the dose quantity is expressed as either Perceived Level (for prompt annoyance judgments), or as Day-Night Average Perceived Sound Level (for delayed annoyance judgments), both in units of PLdB.

2.2 Context of QSF18 Data Collection Exercise

2.2.1 *Purpose of X-59 low boom flight demonstrator ("LBFD") program*

The purpose of NASA's X-59 low boom flight demonstration program is to collect information for consideration by U.S. (FAA) and international (ICAO) civil aviation regulatory agencies about community response to exposure to low-amplitude sonic booms. Such information may persuade these agencies to reverse their long-standing prohibitions on overland supersonic flight. The most credible and compelling evidence that NASA can develop for this purpose is that which is consistent and readily comparable with the agencies' longstanding regulatory rationales.

For practical regulatory purposes, the term "community response" implies little more than the prevalence of a consequential degree of neighborhood-wide annoyance with long-term, cumulative, outdoor residential aircraft noise exposure.[1] Regulatory agencies do not consider individuals' personal annoyance with exposures to single aircraft noise events in non-residential settings as useful indications of "community response" to cumulative aircraft noise exposure.[2]

2.2.2 *Relationship of QSF18 data collection to prior field studies*

NASA's QSF18 data collection effort was the most recent in a series of risk reduction field studies sponsored by the Langley Research Center in which F-18 aircraft performed inverted supersonic dive maneuvers to produce low-amplitude sonic booms. NASA had previously sponsored two similar pilot studies (Fidell *et al.*, 2012 and Page *et al.*, 2014) among residents of base housing at Edwards Air Force Base, a military community familiar with and economically linked to exposure to expected sonic booms.

The broad goal of the prior studies was to explore methodological risks of field investigations of community response to the low-amplitude carpet booms to be produced by future deployments of NASA's X-59 low boom flight demonstrator ("LBFD") aircraft. The QSF18 study was intended in large part to investigate the risks of address-based and panel (rather than independent respondent) sampling in a community unfamiliar with exposure to sonic booms. Panelists were not formally interviewed at specific times, but rather instructed to self-report their annoyance with single event and cumulative exposure to low-amplitude sonic booms.

2.2.3 *Exposure-related differences between F-18 inverted dive and X-59 carpet booms*

The shock waves produced by the F-18 supersonic diving maneuver employed in the QSF18 data collection differ from the carpet booms that will be produced in straight and level flight by NASA's X-59 experimental aircraft. The shockwave shapes produced by the inverted dive maneuver are strongly location-dependent, double booms whose waveforms do not closely resemble those of classic N-wave carpet booms. Further, the inverted dive booms are highly susceptible to perturbation by meteorological conditions within small noise exposure (footprint) areas. In contrast, the carpet booms to be produced by high altitude X-59 flights will insonify far larger areas, and their cross-track characteristics are more homogeneous. In particular, geographic variability in the level and waveshape of carpet booms across the 30+ mile width of the boom corridor will be reduced with respect to the localized impulsive exposure produced by the inverted dive maneuver (Morgenstern *et al.*, 2012; Maglieri, 2019).

---

[1]*Per* "Federal Agency Review of Selected Airport Noise Analysis Issues" (FICON, 1992)

[2] Individual reactions to sonic booms are nonetheless of intrinsic interest, and are arguably useful as predictor and potentially explanatory variables in statistical (as opposed to regulatory) analyses. No use is made of information about personal reactions to individual sonic booms in exposure-based regulatory decision making, however, since no reasonable regulatory policy can plausibly address heterogeneity in individuals' reactions to either single event or cumulative aircraft noise exposure.

2.3 Overview of QSF18 Data Collection and Current Analyses

NASA selected Galveston, TX as a boom-naïve, coastal site suitable for exposure to low-amplitude sonic booms produced by an off-shore, inverted supersonic dive maneuver. The present set of inferential analyses was conducted primarily to investigate the effects identified in §2.1 on judgments on both prompt (single event) and delayed (cumulative, or end-of-day) reporting of boom-induced annoyance.[3] Descriptive analyses were also conducted to investigate other methodological issues.

2.4 Organization of Report

Chapter 1 is an Executive Summary. Chapter 3 describes quality control examinations of data sets prepared by Page *et al.* (2020). Chapters 4 and 5 derive various dose-response relationships from the QSF18 data by regression techniques. Chapter 6 analyzes QSF18 data for relationships other than dose-response functions. Chapter 7 derives dose-response relationships by means other than regression analysis, while Chapter 8 examines the demographic representativeness of the panel sample of the Galveston population. Chapter 9 addresses a variety of reminder message and response latency issues.

Chapter 10 discusses implications of the findings of the present analyses for future X-59 community response testing with carpet booms. Chapters 11 and 12 provide a glossary and a list of references. Chapters 13 through 16 are informative appendices.

[3] The terms "prompt" and "single event", as well as the terms "delayed" and "end-of-day", are used interchangeably in this report to distinguish the two types of annoyance judgments of low-amplitude sonic booms that panelists were asked to make. The terms "prompt" and "delayed" emphasize the time frames of annoyance judgments. The terms "single event" and "end-of-day" emphasize the distinction between the annoyance of individual and cumulative forms of impulsive noise exposure.

THIS PAGE INTENTIONALLY LEFT BLANK

3 PRE-ANALYSIS DATA QUALITY CHECKS

Page *et al.* (2020, p. 17) describe the QSF18 data collection exercise as exposing 496 Galveston-area panelists to 51 low-amplitude sonic booms over the course of nine days.[4] Separate, NASA-provided Excel workbooks containing case records of prompt (single event) and delayed (end-of-day) questionnaires served as the starting point for the current analyses. Panelists within both data sets were identified only by anonymized case identifiers.

Throughout this chapter references are made to variables (data fields) in the case record workbooks. Descriptions of these variables are provided in Table 76 (page 155) of this report.

The single event data set was sorted by Participant_ID and Boom_ID (and the end-of-day data set by Participant_ID and Test_day) to obtain an initial familiarity with the two data sets. Informal preliminary inspections of the sorted data raised a number of questions. In consultation with NASA, a number of quality control checks were performed on each data set prior to analysis. The purpose was to produce a groomed data set from which an unambiguous set of analyses could be performed. The various steps taken to create a groomed data set are described in detail in the following subsections, and also in Appendix D.

The preliminary sorting revealed multiple records with identical combinations of Participant_ID and Boom_ID (prompt responses) or Participant_ID and Test_day (delayed responses).The first stage of data screening therefore consisted of eliminating all but one occurrence of such records, so that each case record eligible for analysis was unique. Since multiple records for a given case (identical participant and boom number) did not always contain identical values for all response variables, reconciling such occurrences became the second stage of the screening. The third screening stage related to the coding of panelists' locations at the time of exposure to a boom. Elimination of some records due to location uncertainties occurred at this stage. At the fourth and final stage, a number of additional variables (data field codings) were also checked for interpretability.

3.1 Replicated Data Records

Both the prompt and delayed response data sets contained replicated records. The prompt judgment (single event) data set contained multiple records for the same panelist/sonic boom combinations, while the delayed judgment (end-of-day) data set contained multiple records for the same panelist/test day pairs. This subsection describes the findings regarding such replicated case records, as well as the criteria used for excluding the entirety of some replicate sets from analysis.

---

[4] Page *et al.* (p.17) cite 51 low-amplitude sonic boom exposures. Although the single event data base provided by NASA numbers the Boom_ID entries as 1 to 52, the first of the 52 dive maneuvers produced no boom in the target neighborhood.

3.1.1 *Prompt response judgments*

<u>Replicated Records</u>

Some case records in the prompt response data set provided by NASA contained identical Panelist_ID and Boom_ID fields. Table 2 shows counts of numbers of replicated records in the data set. The first column classifies replicate case record sets by the number of times a Participant_ID/Boom_ID pair was found. Redundant case records varied from as few as two to as many as seven replicated records. The second column restates the first column as the replicate group size. (A group size of 1 indicates no replicates.) The third column tabulates the number of cases in which a replicate set of that group size was found. The fourth column tabulates the number of records involved (group size multiplied by number of cases). The last column shows the number of redundant records in each group (number of cases multiplied by the group size minus one).

For example, the top row of Table 2 shows tallies 5,621 single (non-replicated) case records. The second row counts 641 instances of duplicate records, for 641 redundant cases. The third row shows 42 instances of triplicate case records, for 84 redundant cases.

The 7,068 records in the original data set contained a total of 6,312 unique cases. The original set, therefore, included 756 replicated records which were considered ineligible for the current analyses. The inconsistent replicate cases shown in the next to last row of the table are discussed below. Once these were excluded from analysis, 6,041 cases remained.

Table 2: Instances of replicated prompt response case records in the single event database

Record Clusters	Group Size	Cases	Records Involved	Extra Records
Unique records	1	5,621	5,621	0
Duplicate records	2	641	1,282	641
Triplicate records	3	42	126	84
Quadruplicate records	4	4	16	12
Quintuplicate records	5	2	10	8
Sextuplicate records	6	1	6	5
Septuplicate records	7	1	7	6
Subtotals		**691**	**1447**	**756**
Totals		6,312	7,068	756
Inconsistent replicate cases		271		
Remaining cases		**6,041**		

Inconsistent responses within replicated records

Response fields[5] "E6," "annoy," "vibration," and "startle" did not always contain identical values across all records within the replicate sets cited in Table 2. Inconsistent replicate records were found for 144 panelists, or 35.3% of the panelists. Of the 691 cases with two or more records 420 cases contained consistent codings within one or more response fields. Such cases were retained for analyses. The remaining 271 case records, which exhibited one or more fields with inconsistent variable codings, were excluded from the current analyses.

As a result of this screening process, two of the original 408 panelists were eliminated from the pool. (Each of the two panelists had contributed only one response record to the data set over the course of the study.)

3.1.2 *Delayed annoyance judgments*

Replicated Database Records

Some records in the delayed annoyance judgment set provided by NASA had identical Participant_ID and Test_Day fields. Table 3 shows counts of numbers of replicated records in the data set. The first column classifies replicate record sets by the number of times a Participant_ID /Test_Day pair was found. These varied from as few as two to as many as four replicated records. The second column restates the first column as the replicate group size (a group size of 1 indicates no replicates). The third column tabulates the number of cases where a replicate set of that group size was found. The fourth column tabulates the number of records involved (group size multiplied by number of cases). The last column shows the number of redundant records in each group (number of cases multiplied by the group size minus one).

Table 3: Instances of replicated delayed response case records

Record Clusters	Group Size	Cases	Records Involved	Extra Records
Unique records	1	2,548	2,548	0
Duplicate records	2	137	274	137
Triplicate records	3	9	27	18
Quadruplicate records	4	2	8	6
Subtotals		**148**	**309**	**161**
Totals		2,696	2,857	161
Inconsistent replicate cases		14		
Remaining cases		**2,682**		

[5] Please see Table 76 in the Glossary for a list of data set variables names and definitions.

For example, the top row of Table 3 tallies 2,548 unique (unreplicated) records. The second row counts 137 instances of duplicate records. The third row shows nine instances of triplicate records, of which 18 case records were redundant.

A total of 2,696 unique cases were found out of the 2,857 records in the data set. Hence, 161 redundant records were excluded from analysis. The 14 inconsistent replicate cases shown in the next to last row of the table are discussed below. After replicated inconsistent cases were excluded, 2,682 cases remained.

Inconsistent responses within replicated records

Response fields "D1," "HA," and "D3_1" did not always contain identical values across all records within the replicate sets cited in Table 3. Inconsistent replicate records were found for 14 panelists (2.8% of the panelist pool). Of the 148 cases with two or more records, 134 contained consistent codings in each response field. These cases were retained for analyses. The remaining 14 contained one or more fields with inconsistent codings and were eliminated from analysis. None of the remaining 408 panelists within the replicate sets cited in Table 3 were excluded from analysis upon completion of this step.

Figure 6-28 of Page *et al.* (2020) shows that more than 300 replicated delayed (end-of-day) annoyance judgments were recorded for the Galveston panelists. Page *et al.* also note in §6.2.1.2 that these responses were archived in a manner that did not permit determination of which of the multiple redundant submissions were valid.[6] (Some of them may have been false alarms – that is, annoyance judgments mis-attributed to booms.) Page *et al.* nonetheless appear to have included at least some replicated responses in their dose-response analyses. The present analyses excluded all such case records from consideration.

3.2 Additional Screening Procedures

3.2.1 *Prompt response judgments*

Further inspection of the prompt response data set resulted in the application of eight additional screening criteria to the 6,041 records passing the replicate tests described in §3.1.1. These are described in the following subsections. The abbreviations in parentheses at the end of each subheading identify each of these criteria later in Figure 2.

[6] As noted on page 101 of Page *et al.* (2019), "Unfortunately, with the information gathered, it was impossible to accurately label some of the duplicate submissions as good/accurate and others as inaccurate."

Match Type and Generic Location Determinations (MT)

The Match_Type field in the prompt response database provided the basis for associating each case record with one of four generic panelist location categories. Records that could not be assigned to one of these categories were deemed ineligible for analysis.

The nine Match_Type codings are shown in Table 4. Their dispositions to a generic location category shown in the rightmost column. The four disposition categories were:

(1) Undetermined - the panelist geographic location (latitude and longitude) was successfully determined at the time of the survey but the location had not been further ascribed to one of the following three generic location categories (Match_Type = 1),

(2) At Home – the panelist geographic location was successfully matched to their home address (Match_Type = 2) or convincing evidence was obtained from the end-of-day survey to place them at home at the time of the boom (Match_Type = 6),

(3) At Work – the panelist geographic location was successfully matched to their employment address (Match_Type = 3) or convincing evidence was obtained from the end-of-day survey to place them at work at the time of the boom (Match_Type = 7), and

(4) Elsewhere - the panelist geographic location was obtained but did not match that of their home or employment address (Match_Type = 4).

The remaining Match_Types of 5, 8 and 9 with pink shading in Table 4 provided no information from which a generic location (At Home, At Work or Elsewhere) could be determined.

Table 4: Match type codings

Type	Coding Key	Disposition
1	Matched SE*, latitude/longitude (some of which are clearly erroneous)	Undetermined
2	Matched SE, at home	At-home
3	Matched SE, at work	At-Work
4	Matched SE, "somewhere else" locations manually geocoded	Elsewhere
5	Matched SE, undetermined location (no latitude/longitude information)	Not used
6	Have SEs for day but no time match, DS** indicates home at time of boom	At Home
7	Have SEs* for day but no time match, DS indicates work at boom time	At Work
8	Have SEs for day but no time match, DS indicates not at home or at work at time of exposure (no latitude/longitude information)	Elsewhere
9	Have SEs for day but no time match, and no DS fall back (Assume not heard - no latitude/longitude information)	Elsewhere

*SE: "Single Event" data set of prompt judgments of annoyance of individual booms

**DS: "Daily Survey" data set of cumulative exposure annoyance judgments

Numbers of case records in each category following removal of the redundant case records are shown in Figure 1. Blue bars denote match types considered eligible for analysis – a total of

5,836 case records. Pink bars (match types 5, 8 and 9) were excluded from the set of analyzable cases. The bar heights total to the 6,041 case records shown at the bottom of Table 2.

The majority of case records (70%) fell in the "Undetermined" category. Another 22% of the records could be reliably assigned to panelists' locations as either At Home or At Work. Another 5% could be reliably assigned to somewhere other than home or work (Elsewhere). The location of the remaining 3% could not be determined. At the conclusion of match type screenings, 5,836 case records remained eligible for analysis, a reduction of an additional 205 case records after replicate screening.

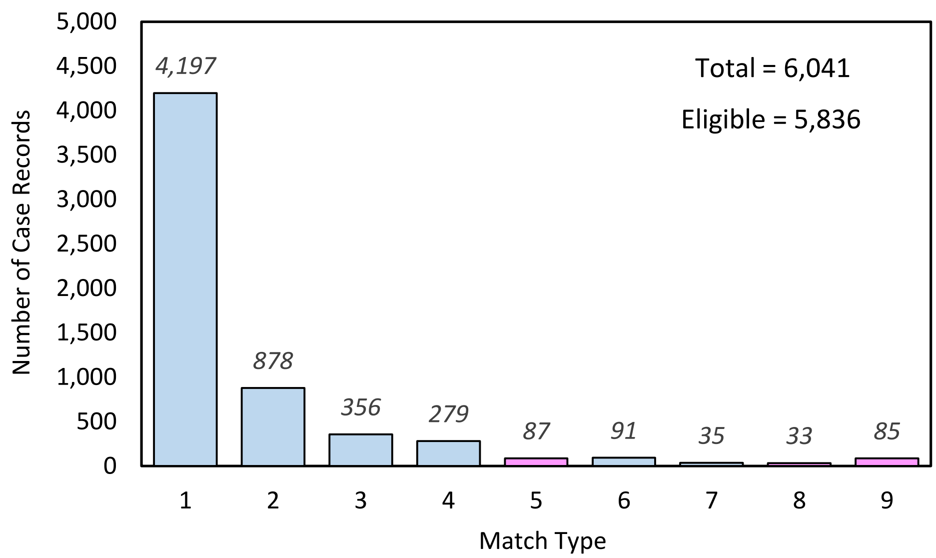

Figure 1. Distribution of match type categorizations

Table 5 shows the resulting distribution of cases records by generic location category.

Table 5: Generic location determinations

Generic Location	Match Type	Count	Percent
Undetermined	1	4,197	69.5%
Home	2, 6	969	16.0%
Work	3, 7	391	6.5%
Elsewhere	4	279	4.6%
Unknown	5,8,9	205	3.4%
Total		6,041	100.0%
Eligible for analysis		5,836	92.0%

Reminder Messages Sent Prior to Occurrence of Boom (Early)

Of the 6,041 unique, analyzable case records, 494 were found for which reminder messages had been sent *prior* to the time of occurrence of an associated boom. These mis-timed reminders were associated with four booms identified in the "Sent Early" columns of Table 9 on page 26. A TRUE entry in this table indicates the reminder message was sent prematurely.

There were no available means for determining whether these cases should be treated as false alarms or whether they were attributable to a particular boom event. While responses received between the time of the reminder message and the time of the boom could probably be considered as false alarms, those received afterward might have been due either to false alarms or to actual boom events. All records associated with such boom events were excluded from the set of case records eligible for analysis because of this uncertainty.

Latitude and Longitude (Lat/Lon)

When the latitude and longitude of the participant's location could not be determined, these two parameters had been coded as zeros. All such records were considered ineligible for analyses.

Annoyance Judgments for Unnoticed Booms (NH/AN)

Of the 6,041 unique, analyzable case records, seven cases remained in which a conflict existed between the record coding the panelist as not hearing the boom, but the annoyance judgment was coded as slightly or more greatly annoyed. No means were available for correcting such records, so those case records were excluded from the data set eligible for analysis.

Conflicting "noticed" / Not "noticed" Data Fields (H<>E6)

Two fields in the data set – E6 and Heard – purport to report the same condition (did or did not hear the boom). In most cases, the values in these two fields agreed. In some cases, however, they did not. Since it could not be determined which of these parameter fields was the more reliable, records with conflicting entries in the two fields were considered ineligible for analysis.

Latency of Prompt Annoyance Judgments (Latency)

Lengthy delays were observed in a number of cases between the occurrence of a boom and the time a response began. These delays ranged from a few minutes to as long as four days. It was assumed for current analysis purposes that panelists reached all single event judgments promptly and that delays occurred only in submitting and/or archiving the annoyance judgments. The prevalence of such lengthy delays and their implications are discussed in greater detail in §6.10, beginning on page 93.

In addition to the delays, some case records were found in which the participant responded as much as 9.4 hours *before* the occurrence of a boom. All 162 case records in which the response preceded the boom were excluded from the analyzable data set.

Exclude a Case Too Far in Future (EF)

Page *et al.* (2020) created the screening parameter Exc_Future to mark records in which the annoyance judgment was submitted more than 900 seconds (15 minutes) *before* the boom event. The reasoning for such a generous acceptance criterion was unclear, as it would seem any pre-boom annoyance judgments would be suspect. The previously described latency criterion trapped all cases coded as too far in the future with the exception of one. This case indicated the response was provided 12 minutes *after* the boom. Since it could not be determined why this record was coded as occurring prior to the boom a decision was made to retain this record for analysis.

Not Used in PL Analysis (UIPL)

Some records were coded in the USED_IN_ANALYSIS_PL field as not having been used in the Page *et al.* (2020) analyses. Many of these records contained a value of zero in the PL field, while others with non-zero values must not have been used for other reasons. An additional 1,240 case records were considered ineligible because they were considered ineligible for analysis in Page *et al.* (2020).

Summary of additional criteria

Figure 2 summarizes the number of case records affected by the screening criteria in this subsection. Note that the screening criteria are not mutually exclusive: in some instances, a case record failed to meet more than one acceptance criterion. A total of 1,872 additional case records were excluded from analysis after the exclusion criteria shown in Figure 2 were applied.

Remaining analyzable records

Table 6 summarizes the effects of the various screening criteria on the numbers of analyzable records and panelists from the data set. The table is a summary of Table 108 beginning on page 181 of Appendix D that details the reduction in analyzable records by individual test participant.

The first row in Table 6 shows the number of records (7,068) and panelists (408) in the original data set. In the second row, after excluding 756 replicated records with identical Participant_ID and Boom_ID combinations, 6,312 unique cases remained. No panelists were eliminated in this step. In the third row an additional 276 case records where one or more response variable fields did not match across the replicates were excluded. As a result of these exclusions 2 panelists were eliminated from the pool. In the last row another 1,872 case were eliminated based on the additional screenings leaving a total of 4,169[7] case records for analysis. In this step 52 panelists were removed from the pool leaving a total of 354 providing one or more prompt annoyance judgements. The blue-shaded cells in this table carry forward to those in blue in Table 7.

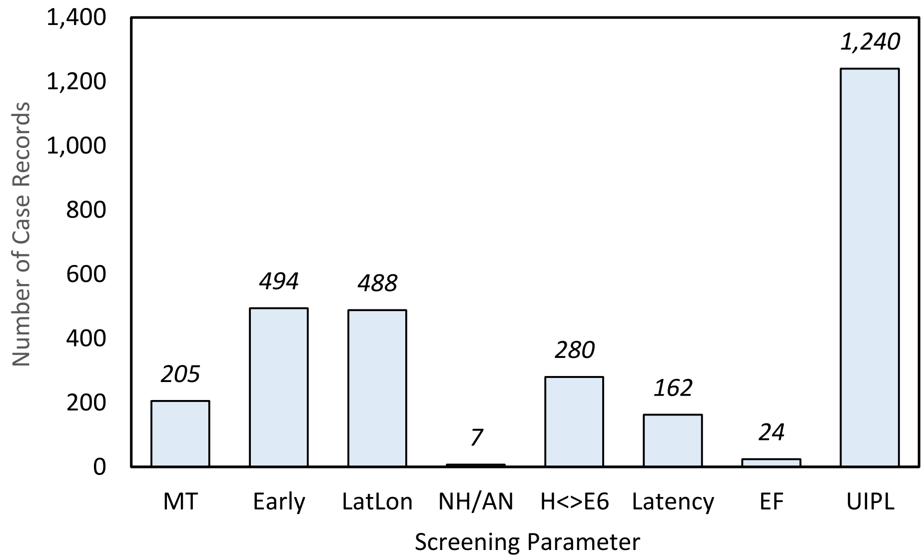

Figure 2. Numbers of single event records not meeting additional screening criteria

Table 6: Reductions in prompt response case records panelists eligible for analysis

Stage	Records	Records		Panelists	
		Reduction	Remaining	Reduction	Remaining
---	Total Records	---	7,068	---	408
1	Unique Case Records	756	6,312	0	408
2	Non-identical Replicate Cases	271	6,041	2	406
3	Additional Screenings	1,872	4,169	52	354

Further insight into reductions in case records may be gained by examining the reductions with prompt response reminder groups. Table 8 shows the breakdown. The two lefthand columns of Table 8 replicate those in Table 6. The numbers in the records reduction column of Table 6 may be found in the righthand column of Table 8 (please see corresponding blue-highlighted cells).

[7] Page *et al.* (2020, p. 17) state that 5,796 of 7,068 prompt case records in the original database could be associated with a measurable noise level. This figure is 1,627 greater than the number found eligible for the present analyses. No explanation is apparent for the discrepancy.

The entries in the three reminder group columns provide the added detail. The top row of the table shows the numbers of original data set records found for each of the reminder groups.[8]

For each of the three screening stages (rows) the number of data set records eliminated for each reminder group (columns) is shown. Immediately below each of these numbers is the percentage of total reminder group records represented by these numbers. In the last row the total number of records eliminated for each reminder group is shown. The total loss for the no-reminder group was 32%, while the loss for the average of the e-mail and text reminder groups was 43% (a factor of 1.36 over the no-reminder group).

Table 7: Ineligible prompt response case records by reminder group

Stage	Screening Stage	Reminder Group			All Groups
		No Rem	E-mail	Text	
---	Total Records	1,352	2,134	3,582	7,068
1	Unique Case Records	80	270	406	756
		5.92%	12.65%	11.33%	10.70%
2	Non-Identical Replicate Cases	44	99	128	271
		3.25%	4.64%	3.57%	3.83%
3	Additional Screenings	305	547	1,020	1,872
		22.56%	25.63%	28.48%	26.49%
	Total Records	429	916	1,554	2,899
		31.73%	42.92%	43.38%	41.02%

Figure 3 plots the percentages shown in Table 8. The graphic shows three clusters of colored bars. The color coding denotes the reminder group. The clusters represent the three screening stages shown in the table. Unique case record losses were those with the same Participant_ID / Boom_ID entry beyond the first occurrence. Non-identical replicate losses were those of the first occurrence of a Participant_ID / Boom_ID pair where one or more fields of the response fields did not agree. The additional screenings are those described above. Noteworthy are the following observations:

- Approximately twice the number of non-unique case records occurred in the e-mail and text reminder groups as were found in the no-reminder group.
- Non-identical replicate response codings were found with about the same frequency across all reminder groups.

[8] The No Rem (no reminder) group consists of reminder group numbers 1 and 3. The E-mail reminder group corresponds to reminder group 3. The Text reminder group corresponds to reminder group 4. Please see §3.3.1 on page 24 for additional details on reminder group reassignment during the data grooming phase.

- The additional screening tests results in increasing losses from the no-reminder group to the text reminder group. NASA may be best equipped to determine any explanations for these findings.

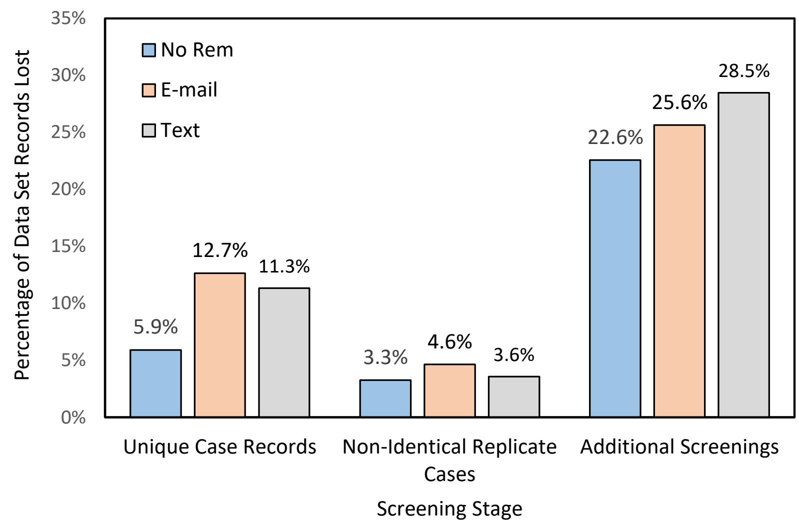

Figure 3. Ineligible prompt response case records by reminder group

3.2.2 *Delayed response judgments*

<u>Not Used in PLDN Analysis</u>[9]

Some records were coded as not having been used in the Page *et al.* (2020) data analysis (all of these records contained a zero value in the PLDN database field). An additional 729 so-coded cases not previously excluded by the replicate record screening process were considered ineligible for analysis.

<u>Remaining analyzable records</u>

Table 8 summarizes the detail of Table 109 on page 193 by showing the manner in which the various screening criteria reduced both the number of analyzable records and the number of panelists from the full data set. A detailed breakdown of the reduction in analyzable records by test participant is provided in Appendix D.

---

[9] For the sake of consistency with ANSI/ASA standard acoustical terminology, the abbreviation for Perceived Day-Night Average Sound Level is represented in this report as PLDNL. In this instance, however, the term refers only to the name of a data set field field.

Table 8: Reduction in delayed response case records and panelists eligible for analysis

Stage	Records	Records		Panelists	
		Reduction	Remaining	Reduction	Remaining
---	Total Records	---	2,855	---	427
1	Unique Case Records	160	2,695	1	426
2	Non-Identical Replicate Cases	14	2,681	0	426
3	Additional Screenings	729	1,952	41	385

The first row of Table 8 shows the number of records (2,855) and panelists (427) in the original data set. After eliminating 160 replicated records with identical Panelist_ID and Boom_ID combinations, 2,695 unique cases remained and one panelist was eliminated. After screening the replicate case record sets for inconsistent response variables, an additional 14 records were excluded from analysis, leaving 2,681 records. No panelists were excluded at this stage, leaving 426. Following the application of one additional screening criterion (the "not used in PL field"), 729 case records were excluded, leaving 1,952 analyzable records. This additional screening also eliminated 41 further panelists, leaving 385 in the data set.

3.3 Anomalies Observed in the Prompt Response Data Set

3.3.1 *No reminder messages sent*

Table 9 lists the reminder messaging status for each of the 52 booms. A "TRUE" entry in the "Not Sent" column indicates the reminder message was not sent following the boom; a FALSE entry means it was. Reminder messages were not sent for five of the booms. For analysis purposes, all responses from the e-mail or text message groups associated with these booms were included with those from the no-reminder message group (that is, e-mail reminder group 3 became group 1 and text reminder group 4 became group 2).

Table 9: Reminder message sending and timing issues

Boom ID	Not Sent	Early	Boom ID	Not Sent	Early	Boom ID	Not Sent	Early
1	FALSE	TRUE	19	FALSE	FALSE	37	FALSE	FALSE
2	FALSE	FALSE	20	FALSE	FALSE	38	FALSE	FALSE
3	FALSE	FALSE	21	FALSE	FALSE	39	FALSE	FALSE
4	TRUE	FALSE	22	FALSE	FALSE	40	FALSE	FALSE
5	TRUE	FALSE	23	FALSE	FALSE	41	FALSE	FALSE
6	FALSE	TRUE	24	FALSE	FALSE	42	FALSE	FALSE
7	FALSE	FALSE	25	FALSE	FALSE	43	FALSE	FALSE
8	FALSE	FALSE	26	FALSE	FALSE	44	FALSE	FALSE
9	FALSE	FALSE	27	FALSE	FALSE	45	FALSE	FALSE
10	TRUE	FALSE	28	FALSE	FALSE	46	FALSE	FALSE

Boom ID	Not Sent	Early	Boom ID	Not Sent	Early	Boom ID	Not Sent	Early
11	TRUE	FALSE	29	FALSE	FALSE	47	FALSE	FALSE
12	FALSE	FALSE	30	FALSE	TRUE	48	FALSE	FALSE
13	FALSE	FALSE	31	FALSE	FALSE	49	FALSE	FALSE
14	FALSE	TRUE	32	FALSE	FALSE	50	FALSE	FALSE
15	FALSE	FALSE	33	FALSE	FALSE	51	FALSE	FALSE
16	FALSE	FALSE	34	FALSE	FALSE	52	FALSE	FALSE
17	FALSE	FALSE	35	TRUE	FALSE			
18	FALSE	FALSE	36	FALSE	FALSE			

3.3.2 *Uncertain panelist locations at times of impulsive exposure*

Table 10 summarizes location estimates for panelists derived from the Federal Information Processing Standards (FIPS) codes corresponding to the latitude/longitude coordinates assigned to them in the spreadsheets provided for the current analyses. Of the 5,941 estimates of panelist locations at times of exposure in the database supplied by NASA, 5,304 (89%) were within Galveston County, TX. No location estimates were available for 8% of the panelists, while another 3% of the panelist locations were in counties elsewhere in Texas (some in adjacent Harris County, some in counties hundreds of miles distant from Galveston); in states other than Texas; and even in South America.

An indeterminate percentage of the panelists may have nonetheless been located in areas outside of the target area for exposure to low-amplitude sonic booms at the time of the annoyance report. Lacking any independent information, it was not possible to verify panelist location estimates. Out-of-study-area location estimates were *not* excluded from analysis, on the assumption that it was more likely that location estimates were erroneous than that panelists were actually located outside of the study area.[10]

3.4 Additional Preparations for Analyses

Sequence numbers were added to both the prompt response (single event) and cumulative (end of day) data sets. Responses for each panelist were numbered sequentially across the entire test period in the single event data set, while days on which each panelist participated were numbered sequentially in the delayed response data set. This sequential numbering created a variable that allowed panelists' persistence in the data collection process to be quantified as number of boom events (or days) for which panelists provided annoyance judgments.[11]

---

[10] Locations outside of Galveston county may have arisen erroneously from the geolocation application software that panelists approved as their actual location, and would have been categorized as Match_Type = 1.)

[11] The sequence numbering took place after responses ineligible for analysis had been removed from the data set.

Table 10: Estimated locations of panelists based on FIPS code latitude/longitude coordinates in prompt response database

FIPS Code	County Name	State Name	Frequency	Percentage
48167	Galveston County	Texas	5,304	89.278%
Unknown	Unknown	Unknown	476	8.012%
Colombia, South America	NA	Colombia	4	0.067%
06073	San Diego County	California	4	0.067%
12011	Broward County	Florida	2	0.034%
22019	Calcasieu Parish	Louisiana	2	0.034%
48021	Bastrop County	Texas	1	0.017%
48039	Brazoria County	Texas	10	0.168%
48055	Caldwell County	Texas	1	0.017%
48091	Comal County	Texas	1	0.017%
48121	Denton County	Texas	4	0.067%
48143	Erath County	Texas	2	0.034%
48149	Fayette County	Texas	1	0.017%
48157	Fort Bend County	Texas	7	0.118%
48175	Goliad County	Texas	2	0.034%
48185	Grimes County	Texas	1	0.017%
48187	Guadalupe County	Texas	4	0.067%
48201	Harris County	Texas	88	1.481%
48209	Hays County	Texas	1	0.017%
48239	Jackson County	Texas	6	0.101%
48255	Karnes County	Texas	1	0.017%
48291	Liberty County	Texas	2	0.034%
48309	McLennan County	Texas	1	0.017%
48321	Matagorda County	Texas	2	0.034%
48339	Montgomery County	Texas	2	0.034%
48395	Robertson County	Texas	1	0.017%
48423	Smith County	Texas	5	0.084%
48439	Tarrant County	Texas	2	0.034%
48469	Victoria County	Texas	2	0.034%
48491	Williamson County	Texas	2	0.034%
		Total	**5,941**	**100.000%**

The four missing responses to the “Did you hear a sonic thump?” questionnaire item in the single event data set were treated as the panelist not noticing the boom, and the corresponding annoyance judgment was “not at all annoying.” Binary variables were created (“dummy coded”) to capture effects of such comparisons as “ReminderNone” and “HomeWork” from more complex variables reflecting estimates of reminder message status and panelist location.

3.5 Selection of Exposure Bin Widths for Graphic Illustration

Since logistic regression uses individual case records to derive dose-response relationships, the binning of response variables within sound level intervals is not relevant for these analyses. If the data had been binned prior to a logistic analysis, the actual noise doses associated with each case would have been associated with the mid-point of the respective bin, in effect, reducing the precision of the noise dose variable. A two-decibel interval was chosen for the graphics in this report. This bin width was a compromise between excessively-fine gradation (creating excessive visual clutter on the ordinate) and an excessively-coarse gradation (obscuring trends in the distribution function). Figure 4 illustrates the rationale for this decision by showing histogram plots of differing bin widths for the 4,169 records in the analyzable prompt response data set.

The bin widths shown in the figure range from 0.5 PLdB $\leq$ PL $\leq$ 8 PLdB in multiples of 2 PLdB. The general shape of the histogram is unaffected over this range, but the fine detail becomes lost when a bin width exceeding 2 PLdB is adopted. On the other hand, a bin width of 0.5 PLdB displays too much detail, and partially obscures the general shape of the histogram.

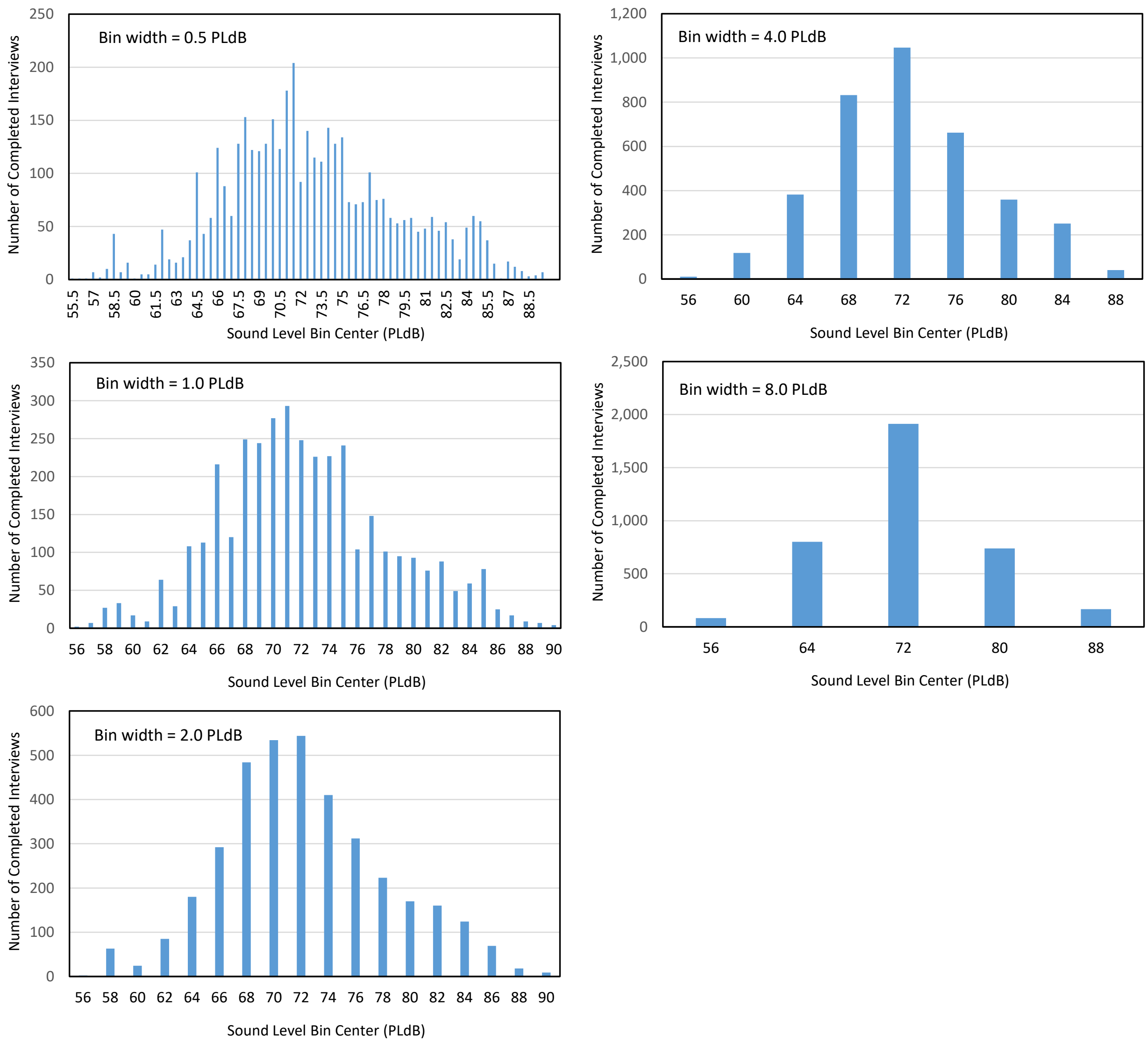

Figure 4. Examples of differing histogram bin widths for representing a range of noise doses

4 REGRESSION-BASED DOSE-RESPONSE ANALYSES

In a similar inferential strategy to that adopted by Lee, Rathsam, and Wilson (2019), dose-response relationships were constructed using Version 6.8 of the HLM (Hierarchical Linear Modeling) software (Raudenbush *et al.*, 2004). Two-level (measures within panelists) regression models were constructed to specify binary outcomes for the various degrees of predicted dependent variables (prevalence of annoyance).[12]

4.1 Logic and Limitations of the Current Analyses

Multi-level regression models take into account individual differences among panelists who made repeated annoyance judgments of low amplitude booms. The "levels" in multi-level models represent nesting or clustering. Observations at the first level of analysis are "nested" within the second level of analysis. In the present case, panelists are identified at the second level of analysis, while their individual annoyance judgments constitute the first level of analysis. Figure 5 provides an example of the analysis layout for a two-level model of delayed annoyance judgments with three panelists. The corresponding layout for prompt annoyance judgments would substitute boom events for days in the first level of analysis.

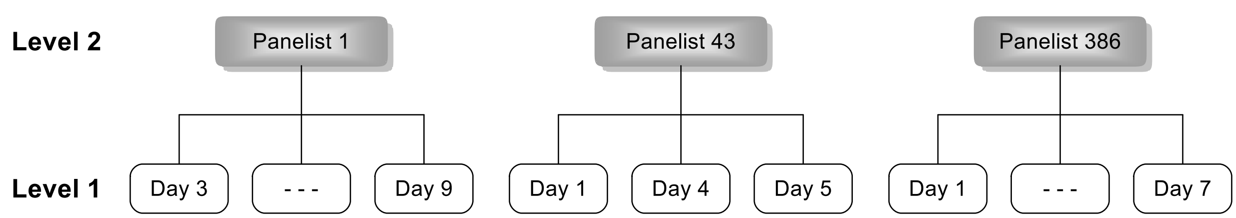

Figure 5. Example of multi-level analysis layout, showing days of exposure at the first level of analysis and panelists at the second analysis level

Predictor variables were added at the level at which they occurred. For example, noise exposure is a first-level predictor because it is measured for each day (or for each event, in the case of prompt annoyance judgments). A predictor such as e-mail *vs.* text reminder messages is assigned at the panelist level as a second-level predictor.

Interactions (for example, between noise exposure level and some other predictor) are created through coding as the product of exposure and the value of the predictor, and then added to the model as an additional predictor. Interaction terms can be formed between any variables, either within or across levels. For example, if an interaction between exposure and reminder messages is of interest, it is considered a cross-level interaction. Even though an interaction

[12] Chapter 15 further discusses the terminology and interpretation of the HLM multi-level regression modeling summary tables in this chapter.

contains components at both levels of the model, it is considered a predictor at the level of the lowest component, in this case level 1, for exposure.

Multi-level regression analyses address individual differences across panelists, while at the same time accounting for non-independence of annoyance judgments within panelists.[13] These differences may be thought of as created by panelists operating on different underlying levels of annoyance, but on the same function relating dose to probability of annoyance. Thus, the models specified random intercepts at level 1 and level 2, but fixed slopes for all predictors of annoyance.

The underlying correlations between predicted and predictor variables control the interpretability and practical utility of the findings of multi-level logistic regression analysis. If the correlations between second level (individual panelist) and first level variables, such as the prevalence of annoyance, are small, they limit the utility of the findings of multi-level regression analyses. Both the significance (*p* values) of relationships quantified by logistic regression, and the strength of association of variables, require attention in assessing the importance and utility of the findings of multi-level regression analyses.

The large number of inferential tests conducted in the course of the present analyses inflates the number of Type I (false positive) errors. In the context of the present analyses, findings with *p* values numerically greater than .001 should not be considered as unlikely to have occurred by chance alone. Further, very wide confidence intervals on odds ratios imply that multi-level logistic regression analysis is unable to usefully estimate the strength of association between predicted and predictor variables. This in turn implies that the strength of association in such cases is unknown for practical purposes, or in other words, that multi-level regression analysis may not reveal whether effects are trivial or meaningful.

Panelists were required to rate their annoyance judgments on a five-point category scale (not-at-all annoyed, slightly annoyed, moderately annoyed, very annoyed and extremely annoyed). Three differing degrees of annoyance (slightly or more annoyed, moderately or more annoyed, and highly [very or more] annoyed) were analyzed by three separate multi-level logistic regression models.[14] The three separate models may approximate the summary dose-response curves inferred from an ordinal regression model[15] in which the dependent variable is a function of the categorical

---

[13] Multi-level modeling is needlessly complicated in the absence of strong individual differences in, for example, the annoyance of exposure to sonic booms. Recall that individual differences in annoyance judgments play no formal role in aircraft noise regulatory policy decisions in any event. The technique is, however, more closely aligned with underlying statistical assumptions, such as independence of errors. It also provides additional statistical power in the current application for answering questions about methodological issues, such as the effects of reminder messages.

[14] Only 33 judgments were made at the extremely annoying degree of annoyance out of a total 4,169 responses. These judgments were too few in number to serve as a basis for inferential analyses.

[15] A multi-level ordinal regression model analyzes the multiple dose-response relationships in the context of annoyance judgments nested within panelists. An ordinal model is functionally similar to the three binary models of the present analyses, with the exception that the multiple binary modeling strategy does not evaluate “very or more greatly annoyed”. The last unanalyzed category threshold could not provide useful information in the present case

annoyance judgments. Results quantify the probability that an annoyance judgment is of equal or greater intensity than the category scale value. Accordingly, three dose-response relationships are defined for each ordinal regression analysis.

Multi-level ordinal analyses with two-level models using annoyance categories as dependent variables did not yield a solution ("failed to converge"). This was due to the very small numbers of annoyance judgments in the two highest annoyance categories, as noted by Lee *et al.* (2019). The maximum likelihood algorithm used by HLM was able to produce a stable solution only when the three highest degrees of annoyance were combined into a single category. Such combination defeats any advantage of ordinal regression – which simultaneously compares each annoyance category with all categories of higher annoyance– over the series of logistic multi-level analyses conducted. Non-multi-level ordinal regression analysis was rejected as an alternative, because it does not account for non-independence of panelists' annoyance judgments (*i.e.*, multiple annoyance judgments from the same panelist.) This concern is especially cogent in the present data, in which the numbers of annoyance judgments contributed to the data set by each panelist varied widely (per §6.2 on page 65 *et seq.*).

4.2 Types of Annoyance Judgments Solicited from Panelists

Two time frames were of interest in the QSF18 data collection for assembling information about community response to sonic booms for consideration by FAA and ICAO. Panelists' annoyance judgments to exposure to low-amplitude sonic booms were sought both for short- term (prompt, for individual boom events) and cumulative (delayed, at the end of the day) impulsive exposures. The annoyance of both forms of exposure was of interest due to concerns that annoyance judgments made in these two timeframes could be differentially susceptible to situational influences. Such situational factors include startle and distraction from ongoing tasks (in the case of exposure to individual booms), and short-term adaptation (in the case of delayed annoyance judgments of multiple exposures).

Judgments of the annoyance of individual sonic booms are referred to herein as "prompt" annoyance judgments. This convention emphasizes that the immediacy of the annoyance judgment is of concern, before the influence of startle, vibration, rattle or other potential short-term influences on annoyance judgments may begin to fade.

Likewise, judgments of the annoyance of cumulative daily exposure to sonic booms are referred to as "delayed" annoyance judgments. This terminology emphasizes concern for the temporal integration of annoyance due to multiple exposures to sonic booms over the course of a day, rather than for the annoyance of individual booms. Delayed judgments of the annoyance of cumulative exposure were also sought to provide information more readily comparable to that

---

because (1) far too few "extremely annoyed" judgments were made to support a meaningful analysis; and (2) the conventional "highly annoyed" designation combines the two highest annoyance categories.

which informs subsonic aircraft noise regulatory decisions based on FICON (1992) recommendations.

The prompt/delayed distinction is arguably relevant for present purposes because FAA certifies noise levels produced by aircraft offered for sale in the U.S. on the basis of individual overflight noise emissions (*per* Federal Aviation Regulations Part 36), but regulates aircraft noise for community-level environmental impact assessments and land use compatibility planning policy purposes on the basis of annual average daily residential exposure (*per* Federal Aviation Regulations Part 150). Thus, for regulatory analyses, the exposure of interest is that of the residence.

4.3 Descriptive Statistics of Panelists' Annoyance Judgments

4.3.1 *Prompt annoyance judgments*

Table 11 shows the distribution of annoyance judgments for prompt responses over the five annoyance categories available to panelists for the 4,169 cases found eligible for analysis in §3.2.1. The annoyance categories are shown in the first column. The number of annoyance judgments eligible for analysis is shown in the second column, and the third shows each category's percentage of the total (4,169). The fourth column shows the cumulative percentage of annoyance judgments obtained for the level of annoyance and all lower levels.

The paucity of judgments in degrees greater than moderate annoyance – less than 4% of the data – is readily apparent. This paucity complicates inference of any relationship between exposure level and annoyance for judgments of moderate or greater annoyance. However, at the slightly or more greatly annoyed degree of annoyance, there may be sufficient data to estimate how the prevalence of annoyance grows with level beyond just a few percent.

Table 11: Numbers of prompt annoyance judgments to the question "How much did the sonic thump bother, disturb, or annoy you?"

Category	Frequency	Percent of judgments	Cumulative Percent
Not at all annoyed	3,703	88.82%	88.82%
Slightly annoyed	324	7.77%	96.59%
Moderately annoyed	109	2.61%	99.21%
Very annoyed	23	0.55%	99.76%
Extremely annoyed	10	0.24%	100.00%
Total	4,169	100.00%	

4.3.2 *Delayed annoyance judgments*

Table 12 shows the QSF18 distribution of annoyance judgments for delayed responses over the 1,952 cases found eligible for analysis in §3.2.2. Like Table 11, Table 12 also shows very small numbers of annoyance judgments of moderate or greater degrees of annoyance (less than 4% of the data), and hence, the dearth of information required to support inference of relationships between long-term exposure level and consequential annoyance. At the slightly or more greatly annoyed level of judged annoyance, the data can support more substantive analyses.

Table 12: Numbers of delayed annoyance judgments to questionnaire item "Over the course of your day, how much did the sonic thumps bother, disturb, or annoy you?"

Category	Frequency	Percent	Cumulative Percent
Not at all annoyed	1,709	87.6%	87.6%
Slightly annoyed	170	8.7%	96.3%
Moderately annoyed	58	3.0%	99.2%
Very annoyed	11	0.6%	99.8%
Extremely annoyed	4	0.2%	100.0%
Total	1,952	100.00%	

4.4 Form of Relationships Developed by Multi-level Logistic Modeling

The general multi-level logistic regression equation is:

$$\boldsymbol{p(annoyed = yes) = \frac{e^y}{1+e^y}}$$ **Equation 1**

where:

$y = \beta_{00} + (\beta_{10})(NoiseDose)$

and

y: like each of the two terms on the right-hand side of the equation, y is an unspecified quantity that is sensitive to both a slope that is applied to the dose as well as to an intercept (where the dose would be equal to zero). It enters as an exponent in the probability of annoyance equation, so that an increasing dose increases the probability of annoyance to a particular or greater degree

β_{00}= intercept for noise dose (in units of dB or PLdB as the metric implies)

β_{10}= average dose-response relationship; *i.e.*, slope for *NoiseDose* (in PLdB): either PLDNL (delayed responses) or PL (prompt responses), averaged over all panelists.

NoiseDose = noise dose in metric-implied units

This equation characterizes a population-specific, rather than a unit-specific, model. In other words, the model describes the annoyance judgments of a group of individuals rather than the opinions of any given individual. While individual differences could be taken into account in deriving dose-response inferential statistics, they are both prospectively unknown and of no utility for regulatory applications. Since generalization intended for regulatory analyses is to populations, not individuals, parameter estimates are based on a population-average model rather than a unit-specific model, parameter estimates are based on a population-average model rather than a unit-specific model. HLM was applied to the totality of both the prompt response and the delayed response data sets after cleaning. The model assumed the relationship shown in Equation 1. A discussion of the findings is presented in the two following subsections.

4.4.1 *Multi-level model of prompt annoyance judgments*

Table 13, Table 14, and Table 15 show the derived β_{00} and β_{10} parameter estimates for multi-level logistic modeling of dose and annoyance judgments for the entire prompt response data set of 4,169 cases. The three tables correspond to the three annoyance levels: (a) slightly or more greatly annoyed, (b) moderately or more greatly annoyed, and (c) highly (very or extremely) annoyed. The parameter estimates are the values of intercepts and slopes in the multi-level equation of probability of annoyance (Equation 1). A fuller explanation of table column headings may be found in Appendix C.

The tables also show the statistical significance of the effect of boom level for each of three annoyance criteria. The odds ratio is the ratio of the odds of an outcome (*e.g.*, a high annoyance judgment) at one level of the noise (*e.g.*, 40 PLdB) to the odds of that same outcome after a 1 PLdB increase in noise (*e.g.*, 41 PLdB). Thus, the odds ratio quantifies the increased likelihood of a particular annoyance judgment increasing from 0 (in this case, not annoyed) to 1 (annoyed), given a one PLdB increase in noise exposure. This concept is described more fully in Appendix C.

Thus, if the probability of being annoyed (*vs.* not annoyed) were 0.6 at a Perceived Level of 40 PLdB, and the odds ratio were 1.126, then the probability of being annoyed at a Perceived Level of 41 PLdB would increase to 0.6 * 1.126 (or 0.68). An increase from 41 PLdB to 42 PLdB would thus increase the odds as 0.68 * 1.126, or by a factor of 0.76. Conversely, if the odds ratio were less than 1, say 0.9, then the probability of annoyance at a Perceived Level of 41 PLdB would *decrease* to 0.6 * 0.9 (or 0.54). An odds ratio of 1.00 means there would be no change in the probability of annoyance with increasing sound level, so that the resulting dose-response function would be flat.

Table 13 displays parameter values for the equation predicting the probability of high annoyance (HA) from the dose in PLdB. The predicted probability of high annoyance is calculated in Equation 1 using the values of -13.517 and 0.119 for the parameters β_{00} and β_{10}, respectively. Table 14 and Table 15 display values for predicting MA (at least moderate annoyance judgments) and SA (at least slight annoyance judgments).

Table 13: Logistic multi-level model of prompt Dose-Response relationship between PL (in PLdB) and high annoyance judgments

Parameter	Parameter estimate	Std Error	t-ratio	Approx. df	p-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.517	1.345	-10.08	353	<.001			
NoiseDose (PL) (β_{10})	0.119	0.018	6.77	4,167	<.001	1.126	1.088	1.166

Table 14: Logistic multi-level model of prompt Dose-Response relationship between PL and moderate or greater annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-11.970	0.813	-14.72	353	<.001			
NoiseDose (PL) (β_{10})	0.117	0.011	10.88	4,167	<.001	1.125	1.101	1.149

Table 15: Logistic multi-level model of prompt Dose-Response relationship between PL and slight or greater annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-8.942	0.504	-17.72	353	<.001			
NoiseDose (PL) (β_{10})	0.095	0.006	14.68	4,167	<.001	1.100	1.086	1.114

Figure 6 graphs the relationships between prevalence of annoyance and Perceived Level for the three degrees of annoyance judgments. The figure plots panelists' annoyance judgments for the three annoyance levels in the upper panel. Each data point shows the percentage of panelists exposed to a Perceived Level dose falling within a 2 PLdB wide bin *vs.* the sound level center point of the bin. The dashed lines plot the logistic regression curves using Equation 1 and the values of β_{00} and β_{10} in the tables. The lower panel of the figure shows the number of panelists providing annoyance judgments within each dose bin. The preponderance of the annoyance judgments lies in the dose range 65 PLdB ≤ PL ≤ 80 PLdB. The ~12 - 15 PLdB lateral separations among regression curves are a rough indication of the distance from one absolute annoyance judgment category label to the next.[16]

[16] Such differences are further quantified in the CTL analyses of Chapter 7.

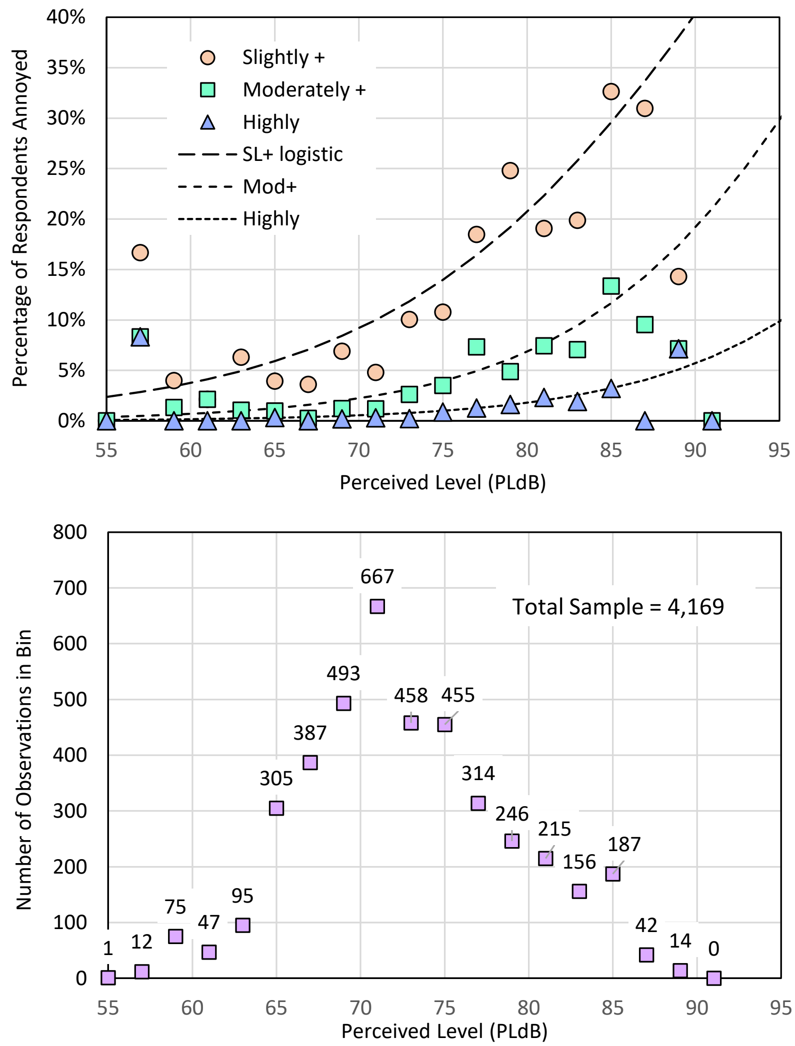

Figure 6. Logistic multi-level regression model fits to prompt annoyance judgments

The goodness of fit of the dose-response curves of Figure 6 is noteworthy. The curves for all degrees of annoyance fit the data well. However, a more nuanced examination of the slight or greater annoyance regression curve suggests an interesting phenomenon. A slight over-prediction may be seen over the range of 64 to 74 PLdB. However, at higher exposure levels the curve fits the data points quite well. This pattern of over prediction at low exposure levels may be an artifact of the assumption of a zero asymptote at low exposure levels.

A zero asymptote may not be the most appropriate assumption, as noted in recent dose-response studies of subsonic aircraft noise [Mestre *et al.*, 2017; Mestre *et al.*, 2016]. The logistic regression model assumes the percentage of annoyed individuals approaches zero at low doses. If provision were made for a non-zero lower asymptote of about 3 – 4 %, then the relationship for Slight or greater annoyance would more closely fit the entire range of data points. This observation is explored in greater detail in §7.2.1 on page 100.

4.4.2 *Multi-level model of delayed annoyance judgments*

Table 16, Table 17 and Table 18 show the derived β_{00} and β_{10} parameter values for multi-level logistic modeling of dose and annoyance judgment for the entire delayed response data set of 1,952 cases. Each table corresponds to one of the three annoyance levels: (a) slightly or more greatly, (b) moderately or more greatly, and (c) highly (very or more greatly) annoyed. The parameter estimates are the values of intercepts and slopes in the multi-level equation (Equation 1) of probability of annoyance.

Table 16 displays the Equation 1 parameter values for the equation predicting the probability of high annoyance (HA) from the noise dose in Perceived Level Day-Night Average Level (PLDNL) in PLdB. The predicted probability of high annoyance is calculated using the values of –7.248 and 0.085 for the parameters β_{00} and β_{10}, respectively. Table 17 and Table 18 display values for predicting moderate or greater annoyance judgments and slight or greater annoyance judgments, respectively.

Table 16: Logistic multi-level model of delayed Dose-Response relationship between PLDNL and high annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.248	0.336	-21.57	385	<.001			
NoiseDose (β_{10})	0.085	0.011	7.92	1,950	<.001	1.089	1.089	1.112

Table 17: Logistic multi-level model of delayed Dose-Response relationship between PLDNL and moderate or greater annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-6.719	0.394	-17.04	385	<.001			
NoiseDose (β_{10})	0.110	0.012	9.03	1,950	<.001	1.116	1.090	1.144

Table 18: Logistic multi-level model of delayed Dose-Response relationship between PLDNL and slight or greater annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-5.186	0.379	-13.67	385	<.001			
NoiseDose (β_{10})	0.102	0.011	9.09	1,950	<.001	1.107	1.107	1.1314

The upper panel of Figure 7 graphs the relationships between prevalence of annoyance and noise dose for the three degrees of annoyance. Each data point shows the percentage of panelists exposed to a Perceived Day-Night Average Level dose falling within a 2 PLdB wide bin *vs.* the sound level center point of the bin. The dashed lines plot the logistic regression curves using Equation 1 and the values of β_{00} and β_{10} in the tables. The lower panel of the figure shows the number of panelists providing annoyance judgments within each dose bin. The preponderance of the experimental data lies in the dose range 20 PLdB ≤ PLDNL ≤ 40 PLdB. As with the prompt annoyance judgments shown in Figure 6, the curves are displaced laterally between annoyance levels by about 12 PLdB.

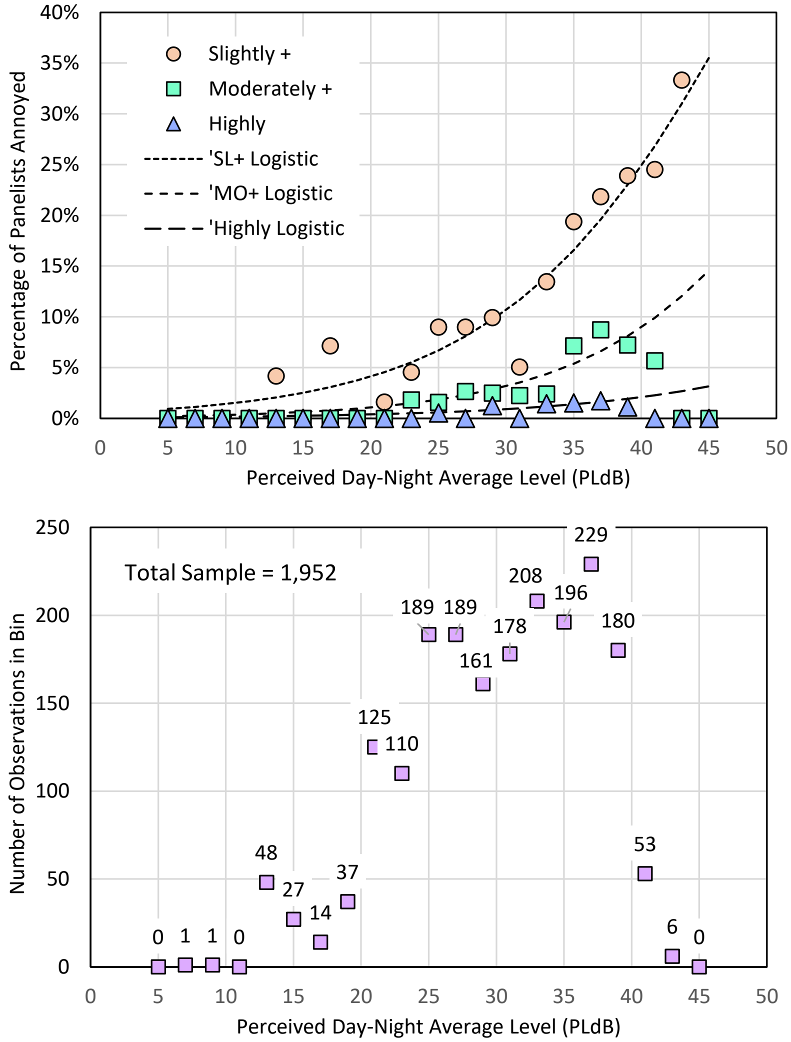

Figure 7. Logistic multi-level model fits to delayed response data points

5 CONDITIONAL DOSE-RESPONSE ANALYSES

The analyses described in this chapter focus on questions of whether reminder messages and panelist generic location at the time of boom exposure had material effects on observed dose-response relationships.

5.1 Reminder Messages

Panelists were randomly assigned at the time of recruitment to three groups receiving:

(1) no-reminder messages,
(2) reminder messages via e-mail, or
(3) reminder messages via text message on their mobile telephones.

Separate multi-level logistic regression (MLM) inferential analyses of the effects of providing reminder messages on observed dose-response relationships were conducted for prompt and delayed annoyance judgments.[17] Analyses were conducted separately for each degree of annoyance: (1) slight or greater annoyance, (2) moderate or greater annoyance, and (3) high (very or greater) annoyance. The percentage of panelists was the dependent variable, and Perceived Level was the independent variable. Separate analyses were conducted for delayed and prompt responses.

The results are presented in tabular form with separate tables for each degree of annoyance. The tables show parameter estimates for the intercept, the slope for the main effect of noise, and the slopes for each additional continuous or binary predictor variable. Parameter estimates were derived using a multi-level modeling equation of probability of annoyance, as described in Appendix C. Tables of results also show the statistical significance of the effect of each predictor, including the noise measure, after adjusting for all other predictors.

5.1.1 *Effects of reminder messages on prompt annoyance judgments*

Table 19 displays descriptive statistics for each of the above reminder message groups. The first row shows the number of active panelists[18] associated with each group. The second row tabulates the number of analyzable self-reports, while the third row calculates the average number of annoyance judgments *per* panelist. The fourth row calculates each group's response rate, the number of completed annoyance judgments divided by the theoretical maximum number (number of active panelists multiplied by 52 booms).

[17] All panelists received morning and evening reminder messages for delayed annoyance judgments. Differences were therefore examined between those who additionally received messages after each event and those who did not. These analyses are described in §9.2.

[18] An active panelist is one who contributed one or more annoyance judgments which passed the screening process developed to determine eligibility for analysis, as described in Chapter 3.

Figure 8 plots the response rates shown in Table 19. Note that numbers of annoyance judgments *per* panelist increases twofold from the no-reminder to the e-mail group. A threefold increase was observed from the no-reminder to the text reminder group. The 31.7% response rate for the text reminder group is a considerable improvement over the live agent, independent-sample, single-contact attempt telephone response rate of about 2% reported by Fidell *et al.* (2019).

Table 19: Distributions of numbers of panelists and numbers of annoyance judgments by reminder message category for prompt annoyance judgments

Quantity	Reminder Group			Totals
	None	**E-mail**	**Text**	
Number of Panelists	172	113	123	408
Number of Annoyance judgments	923	1,218	2,028	4,169
Annoyance judgments *per* panelist	5.37	10.78	16.49	
Response rate[1]	10.32%	20.73%	31.71%	

[1] (Number of annoyance judgments) / (number of panelists x 51 booms)

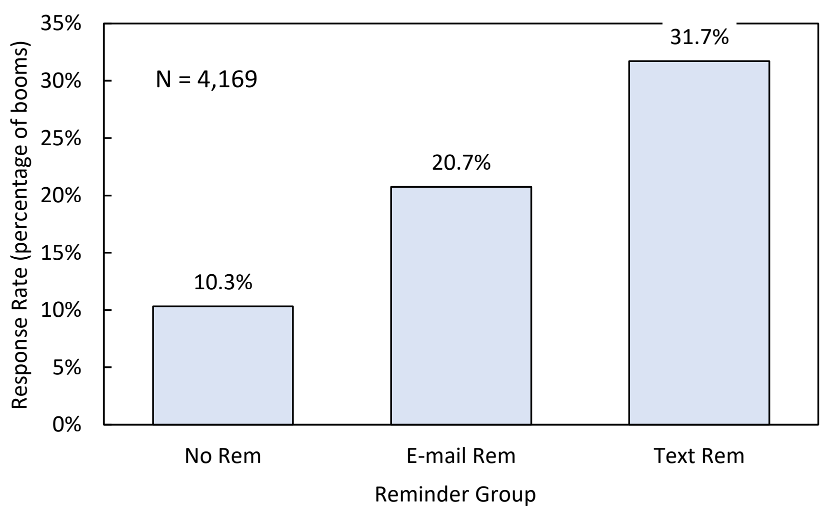

Figure 8. Prompt annoyance judgment rate by reminder message group

Figure 9 plots the annoyance judgments made by panelists in each of the three reminder message groups. Three panels plot these data separately for each degree of annoyance. Within each panel percentages of panelists annoyed are plotted separately for each reminder message group on the ordinate, while associated Perceived Levels are plotted on the abscissa.

In the top panel of the figure, at the slight or greater level of annoyance, little or no difference is apparent between the e-mail reminder message group (green, square plotting symbols) and no-reminder message group (orange circle plotting symbols). In contrast, a difference <u>is</u> apparent between the text reminder message group and the other two. Speculatively, the lower prevalence of annoyance in the text reminder message group might have been due to a significant number of panelists not noticing the boom who provided a not-al-all-annoyed judgment in response to the text message prompt. As such, text reminder messages would have had the effect of reducing the fraction of panelists describing themselves as slightly (or even more greatly) annoyed. Further analysis and discussion of this topic may be found in §9.1, beginning on page 119 of this report.

The distribution of numbers of completed annoyance judgments *per* dose exposure bin is shown in Figure 10 for each reminder message group. As indicated in the second row of Table 19, the greatest numbers of annoyance judgments were made by panelists in the text message group. The next greatest number of annoyance judgments were made by panelists in the e-mail reminder message group; and the least number of annoyance judgments were made by panelists in the no-reminder message group. The preponderance of judgments in each of the messaging groups lay between 65 PLdB ≤ PL ≤ 80 PLdB.

The possible effects of reminder message type on observed dose-response relationships are analyzed in further detail in the underlined subsections below.

<u>Multi-level Logistic Regression for Prompt Reminders</u>

The first question of interest was whether the issuance of reminders affected the dose-response relationship. To that end:

- the data set was dichotomized by reminder and no-reminder to form a single binary variable, labeled ReminderNone. The addition of this variable was performed in SPSS (IBM, 2019), prior to processing by HLM; and
- the two reminder categories were contrasted to form another variable, EmailText, in which no-reminder events were ignored. The addition of this variable was also performed in SPSS prior to processing by HLM.

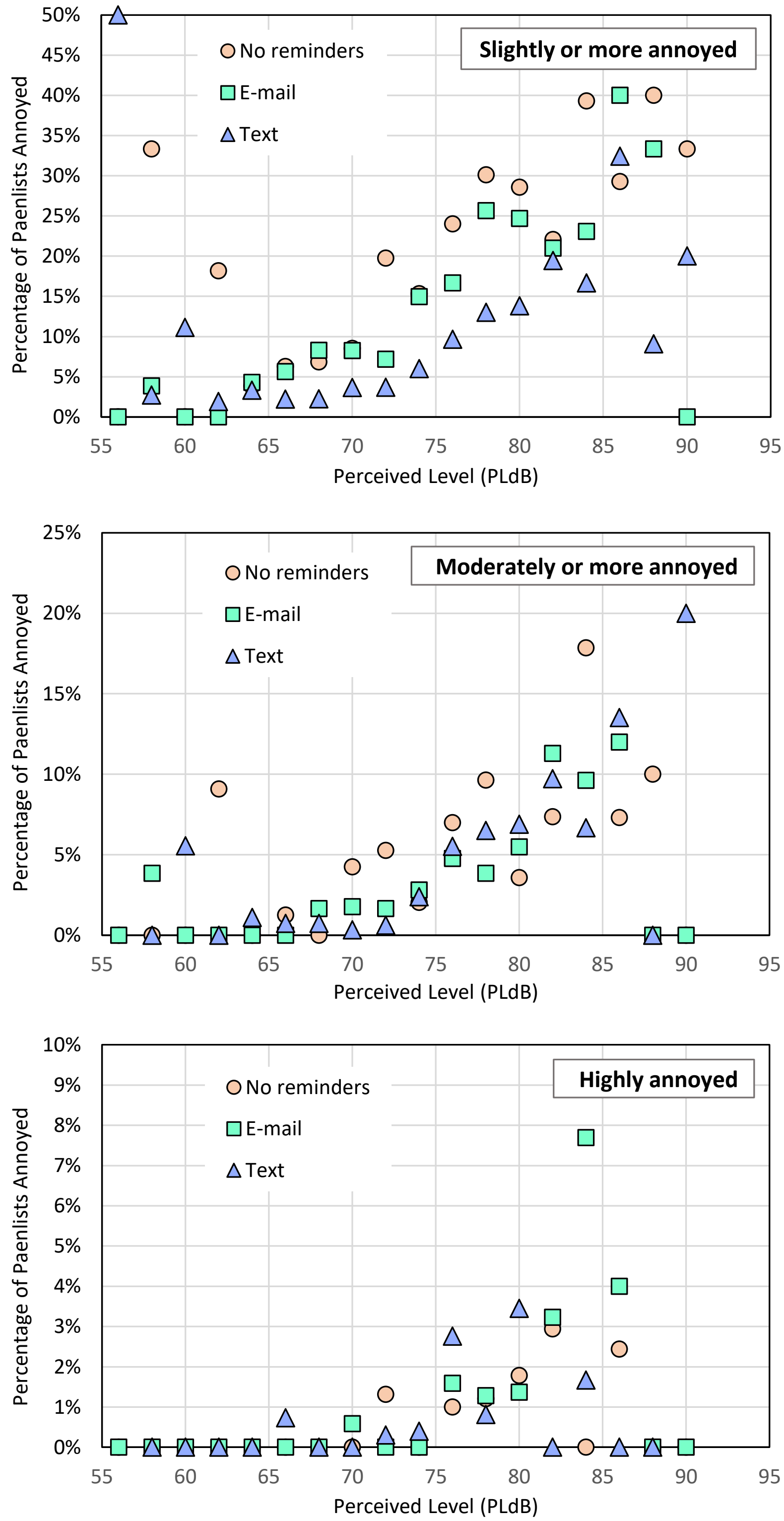

Figure 9. Prompt response annoyance judgments for three reminder messaging modes

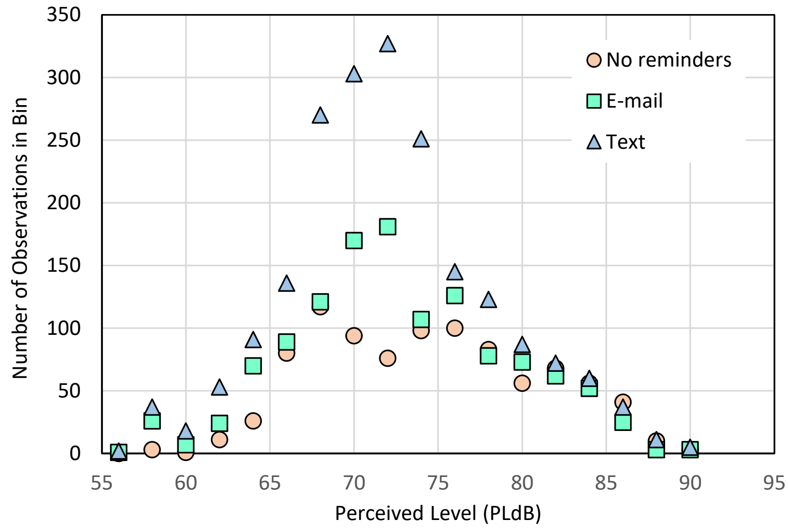

Figure 10. Numbers of prompt response annoyance judgments *per* exposure bin for three reminder messaging modes

The creation of these two variables enabled two comparisons: one in which issuance of reminders could be compared with a lack of reminders, and the other in which the two types of reminders, e-mail and text, could be compared. Equation 2 shows the multi-level logistic regression equation for predicting annoyance from five variables (the first three are main effect variables and the last two are interaction effect variables):

- noise exposure;
- whether a reminder message was sent;
- if sent, the type of reminder (e-mail or text);
- the interactions between whether a reminder message was sent and noise exposure; and
- the interactions between type of reminder message and noise exposure.

$$p(annoyed = yes) = \frac{e^y}{1+e^y} \qquad \textbf{Equation 2}$$

where:

$y=\beta_{00} + (\beta_{10})(PL) + (\beta_{20})(ReminderNone) + (\beta_{30})(EmailText) + (\beta_{40})(ReminderNone)(PL) + (\beta_{50})(Reminder\ Type)(PL)$

and

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion level of annoyance;

β_{00}= intercept for noise dose (in units of dB or PLdB as the metric implies)

PL is the noise dose in Perceived Level in PLdB

ReminderNone is the code for whether a reminder was sent: 0=no-reminder, 1=Reminder;

EmailText is the code for type of reminder (-0.5 for text reminders, 0 for no-reminder message, and +0.5 for an e—mail reminder message);

ReminderNone x PL is an interaction term (the product of PL times the *ReminderNone* reminder message code) which tests the interaction between sending a reminder message and not sending one and noise dose; that is, whether the dose-response relationship varies with the presence or absence of reminder messages; and

EmailText by PL is another interaction term (the product of PL times *EmailText* code) which tests the interaction between type of reminder and noise dose, that is, whether the dose-response relationship differs by type of reminder message.

Table 20, Table 21, and Table 22 show results of single-event models in which the noise dose in Perceived Level and reminders predict probabilities high, moderate or greater, and slight or greater annoyance, respectively. These tables establish which of the reminder message types and their interactions with exposure yielded significantly different dose-response functions. Any parameter in the table with a *p* value $\leq .001$ (highlighted in light green) was considered significant for current purposes.

Table 20 shows the parameter values for Equation 2 for the Highly Annoyed analysis. The green shaded cells indicate a statistically significant relationship. The example below illustrates how the parameter estimates from the table are placed in the equation:

$$y=\beta_{00} + (\beta_{10})(PL) + (\beta_{20})(ReminderNone) + (\beta_{30})(EmailText) + (\beta_{40})(ReminderNone)(PL) + (\beta_{50})(Reminder\ Type)(PL)$$

$$=-13.325 + (0.114)(PL) + (-1.155)(ReminderNone) + (-9.017)(EmailText) + (0.018)(Reminder)(PL) + (0.118)(Reminder\ Type)(PL)$$

Table 20 and Table 22 provide values for predicting MA (at least moderate annoyance) and SA (at least slight annoyance), respectively.

An odds ratio greater or smaller than 1.00 for any parameter in these tables suggests an association between that parameter and the proportion of annoyed individuals. A strong association for an interaction (*e.g.*, reminder message type by PL) means the dose-response slope for one condition (*e.g.*, e-mail) is steeper than for the other condition (*e.g.*, text). A strong association for a main effect (*e.g.*, e-mail *vs.* text) suggests that annoyance is greater for one condition than the other (*e.g.*, e-mail greater than text). Note, however, that main effect associations are ambiguous if a main effect is part of an interaction.

Table 20: Prompt response model of high annoyance due to Perceived Level of single event exposure, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept(β_{00})	-13.325	1.465	-9.10	353	<.001			
NoiseDose (β_{10})	0.114	0.019	6.00	4,163	<.001	1.121	1.080	1.163
ReminderNone (β_{20})	-1.155	1.934	-0.60	4,163	.550	0.315	0.007	13.953
EmailText(β_{30})	-9.017	2.470	-3.65	4,163	<.001	0.000	0.000	0.015
ReminderNone x PL (β_{40})	0.018	0.253	0.71	4,163	.478	1.018	0.969	1.070
EmailText x PL (β_{50})	0.118	0.033	3.58	4,163	.001	1.125	1.055	1.201

Table 21: Prompt response model of moderate or greater annoyance due to Perceived Level of exposure, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-11.052	1.497	-7.38	353	<.001			
NoiseDose (β_{10})	0.107	0.020	5.43	4,163	<.001	1.130	1.071	1.157
ReminderNone (β_{20})	-1.417	1.825	-0.78	4,163	.438	0.242	0.007	8.675
EmailText (β_{30})	-0.636	2.047	-0.31	4,163	.756	0.529	0.010	29.219
ReminderNone x PL (β_{40})	0.016	0.024	0.67	4,163	.200	1.016	0.010	1.065
EmailText x PL (β_{50})	0.008	0.027	0.33	4,163	.741	1.009	0.957	1.064

Table 22: Prompt response model of slight or greater annoyance due to Perceived Level of exposure, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-8.357	0.134	-13.03	353	<.001			
NoiseDose (β_{10})	0.088	0.012	7.15	4,163	<.001	1.072	1.045	1.099
ReminderNone (β_{20})	-1.023	1.101	-0.93	4,163	.353	0.360	0.042	3.108
EmailText (β_{30})	1.530	1.207	1.27	4,163	.205	4.619	0.434	49.156
ReminderNone x PL (β_{40})	0.117	0.014	0.81	4,163	.419	1.012	0.984	1.041
EmailText x PL (β_{50})	-0.0216	0.016	-1.01	4,163	.314	0.984	0.954	1.0158

The statistical significance of an effect is assessed in terms of the probability that the association occurred by chance alone, and expressed as a *p* value. A *p* value depends not only on strength of association, but also on the variability (standard error) and the degrees of freedom ("approx. df") of the variable. In the present context, a *p* value less than .001 was considered indicative of a significant effect. Parameter estimates are used to generate prediction equations, but cannot be directly interpreted as strength of association because they depend on the scale of measurement of the variable. For example, the noise dose is measured in decibels, whereas *ReminderNone* is coded as 0 or 1.

The three tables are consistent in the sense that the *Intercept* and *NoiseDose* parameters always exhibit strong associations. They are inconsistent in the sense that only the highly annoyed analysis (Table 20) showed a significant association with any parameter other than *Intercept* and *NoiseDose*.

The statistically significant *EmailText x PL* interaction for responses of high annoyance indicates that the dose-response relationship varied significantly with type of reminder message: e-mail reminder messages produced a stronger relationship between noise and annoyance than did text messages. Thus, the dose-response curve was steeper for e-mail than for text reminders.

A statistically significant difference in annoyance was observed for high annoyance judgments associated with e-mail *vs.* text reminder messages. However, it is difficult to interpret the main effect of reminder message type in light of the significant interaction between reminder message type and noise dose.

No statistically significant effects of reminder messages were observed on dose-response relationships for either the moderately or more greatly annoyed or the slightly or more greatly annoyed judgments, nor were reminder messages found to be related to annoyance judgments in these degrees. This lack of significance is difficult to reconcile with the relationship displayed in the upper panel of Figure 9, for which the trend of text messaging data points lies below the trends of no-reminders or e-mail messaging points. No immediately obvious explanation is available for the lack of statistical significance.

Form of Reminder Follow-up Analysis of Significant Interactions

The statistically significant interaction between *EmailText* reminder messages and noise exposure for Highly Annoyed judgments shown in Table 20 indicates that the two forms of reminder messages lead to different dose-response relationships, *i.e.*, different slopes for judgments of a high degree of annoyance. A follow-up analysis to explore this difference evaluated separate dose-response relationships for each form of reminder message. Table 23 and Table 24 show the results of separate dose-response investigations for e-mail and text reminder messages, respectively. Although each form of reminder messages is associated a with statistically significant dose-response relationships, the odds ratios indicated a stronger relationship (steeper slopes) for e-mail reminder messages than for text messages (the more the odds ratio exceeds 1.0 the steeper the positive slope of the function). Conversely, the more the odds ratio lies below 1.0, the greater the slope in the negative direction.

$$p(HA = yes) = \frac{e^y}{1+e^y}$$ **Equation 3**

where:

$y = \beta_{00} + \beta_{10}\ (PL)$

Table 23: Prompt response model of high annoyance due to Perceived Level of exposure for events with associated e-mail reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-18.267	1.663	-10.98	95	<.001			
NoiseDose (β_{10})	0.182	0.023	7.96	1,216	<.001	1.201	1.148	1.256

Table 24: Prompt response model of high annoyance due to Perceived Level of exposure for events with associated text reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.159	1.830	-5.55	113	<.001			
NoiseDose (β_{10})	0.074	0.024	3.16	2,026	.002	1.077	1.029	1.128

The upper panel of Figure 11 plots the annoyance prevalence rates and dose-response curves from Table 23 and Table 24. The lower panel displays numbers of annoyance judgments associated with each data point in the upper panel. The greater strength of association (slope) for the e-mail reminder message type (green triangles) suggested by the greater odds ratio for e-mail (1.201) in Table 23 compared with text messaging (1.077) in Table 24 is apparent in this plot. The red line in Figure 11 shows greater prevalence of annoyance at low noise doses (below 75 PLdB) for text messages, while the green line shows greater prevalence of annoyance at high noise doses (above 75 PLdB) for email messages.

The difference in intercepts between e-mail messages (-18.267) and text messages (-10.159) suggests greater prevalence of annoyance when events are accompanied by text messages. As evident in the figure, this observation applies only to noise doses above a Perceived Level of 75 PLdB. The numbers of data points in the two groups are similar above 75 PLdB where the curves begin to diverge.

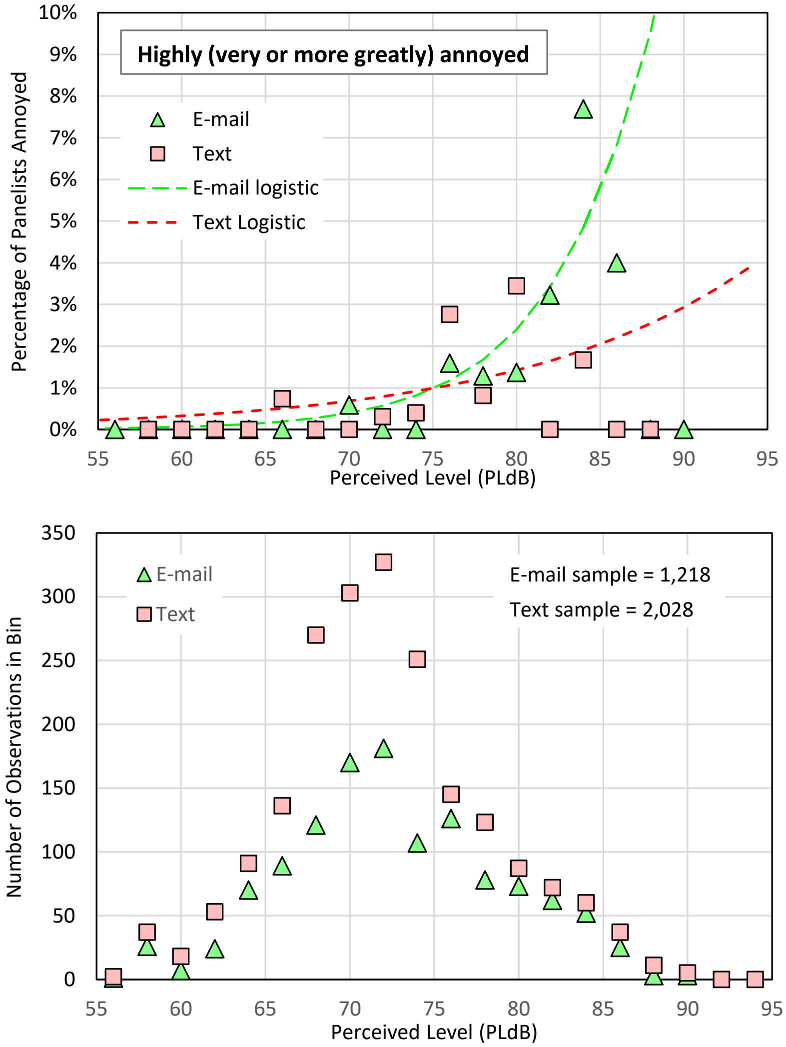

Figure 11. Dose-response relationship for high annoyance judgments for e-mail and text message reminder groups

Post hoc Analysis of Differences between Text Reminders and the Other Two Reminder Groups

Figure 11 suggests that text reminders for high annoyance judgments are associated with judgments of a lower prevalence rate of annoyance and a shallower dose-response curve than for e-mail reminders. This begs the question as to whether text reminder results might differ not only from e-mail reminders, but also from no reminders. A model was therefore constructed in which text reminder messages were coded as "1" and the other two reminder categories (email and no-reminder messages) were combined and coded as "0". Table 25 shows results of a single-event model in which PL and text *vs.* other reminder message strategies predict annoyance probabilities at "high" annoyance; that is, judgments of "very" or "extremely" annoyed.

Table 25: Prompt response model of high annoyance due to Perceived Level of single event exposure, text *vs.* other reminder, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-16.653	1.276	-13.05	353	<.001			
NoiseDose (β_{10})	0.159	0.017	9.26	4,165	<.001	1.172	1.133	1.212
TextOther (β_{20})	6.650	2.147	3.10	4,165	.002	772.936	11.507	51919.1
TextOther x PL (β_{30})	-0.085	1.028	-3.06	4,165	.003	0.918	0.869	0.980

Table 25 suggests a difference in dose-response curves between responses of high annoyance even though the *p* value of 0.002 fails to the meet the criterion of .001.[19] The main effect of text vs. other reminder message categories is not independently interpretable, however, because it is part of an interaction. The exceptionally wide confidence interval on the odds ratio for the main effect indicates a great deal of uncertainty in the magnitude of the effect.

Nonetheless, at the risk of over-interpreting such findings, a post hoc analysis was conducted separately to examine the high annoyance judgments made with text message reminders and those with judgments made in the other two reminder message conditions. Table 26 (text reminders) and Table 27 (e-mail and no-reminder) show the results of this analysis. The tables suggest a shallower slope for high annoyance for judgments associated with text messages than for the other two message categories (NoiseDose parameter values of 0.74 and 0.156 for text and e-mail / no-reminder groups, respectively). The top panel of Figure 12 plots dose-response data points and multi-level logistic regression curves for the two groups. Note that the curves cross at about 76 PLdB, implying that events with exposure levels lower than about 76 PLdB followed by text reminder messages have a greater prevalence of high annoyance than those followed by e-mail or no reminders. At greater exposure levels, boom events followed by text messages have a lower prevalence of high annoyance than those followed by e-mail or no reminder messages.

[19] The choice of a *p* value for the definition of a "significant" statistical effect depends on the costs and payoffs of false positive and false negative outcomes. The current selection ($p = .001$) is a conservative one, but relaxing it would not greatly change the number of inferential tests classified as yielding significant results. As with all exploratory *post hoc* analyses (*i.e.*, those conducted after an experimenter has had an opportunity to examine the findings), even more stringent definitions of "statistical significance" are appropriate than for pre-planned analyses.

Table 26: Prompt response model of high annoyance due to Perceived Level of exposure for events with associated text reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.159	1.8303	-5.55	113	<.001			
NoiseDose (β_{10})	0.074	0.024	3.16	2,026	.002	1.077	1.028	1.126

Table 27: Prompt response model of high annoyance due to Perceived Level of exposure for events with associated e-mail or no reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-16.410	1.250	-13.128	324	<.001			
NoiseDose (β_{10})	0.156	1.169	9.25	2,139	<.001	1.169	1.131	1.208

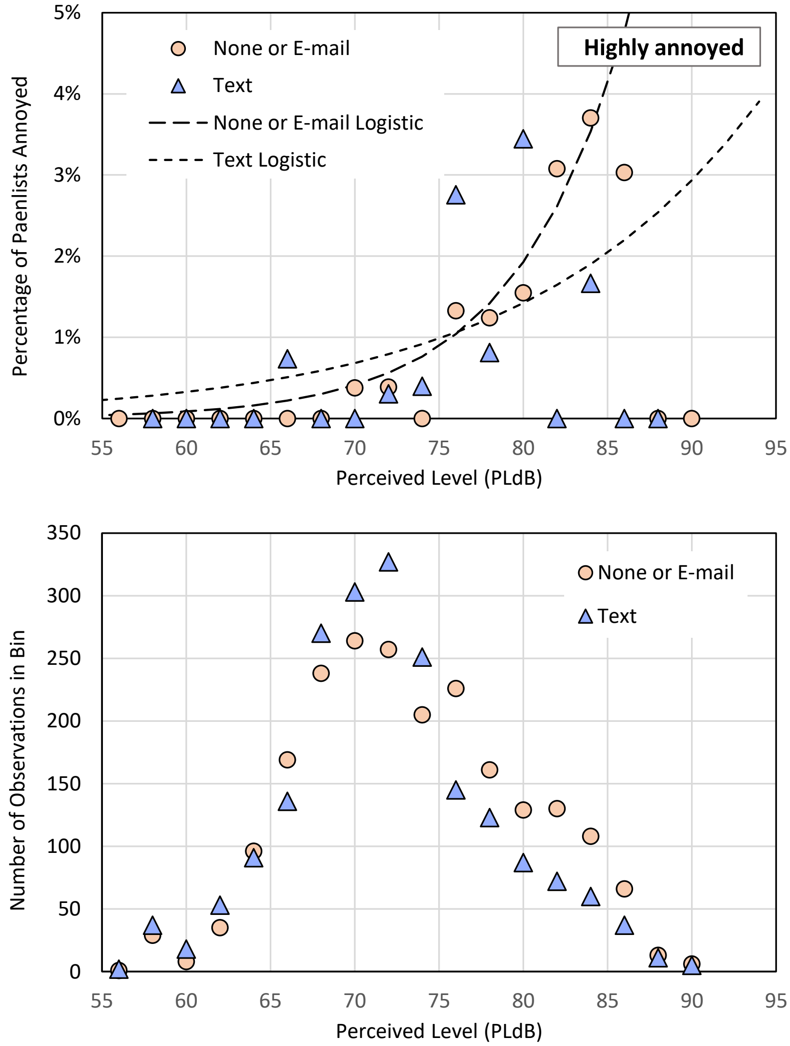

Figure 12. Dose-response relationship for high annoyance judgments for text message and e-mail and no-reminder message groups combined

The two panels of Figure 13 show data points for Slightly and greater and Moderately and greater degrees of annoyance for the sake of completeness. Although systematic differences between the none or e-mail and text reminder groups seem apparent in the two panels, they were not found to be statistically significant by multi-level logistic regression. The numbers of observations associated with these data points are the same as shown in the lower panel of Figure 12.

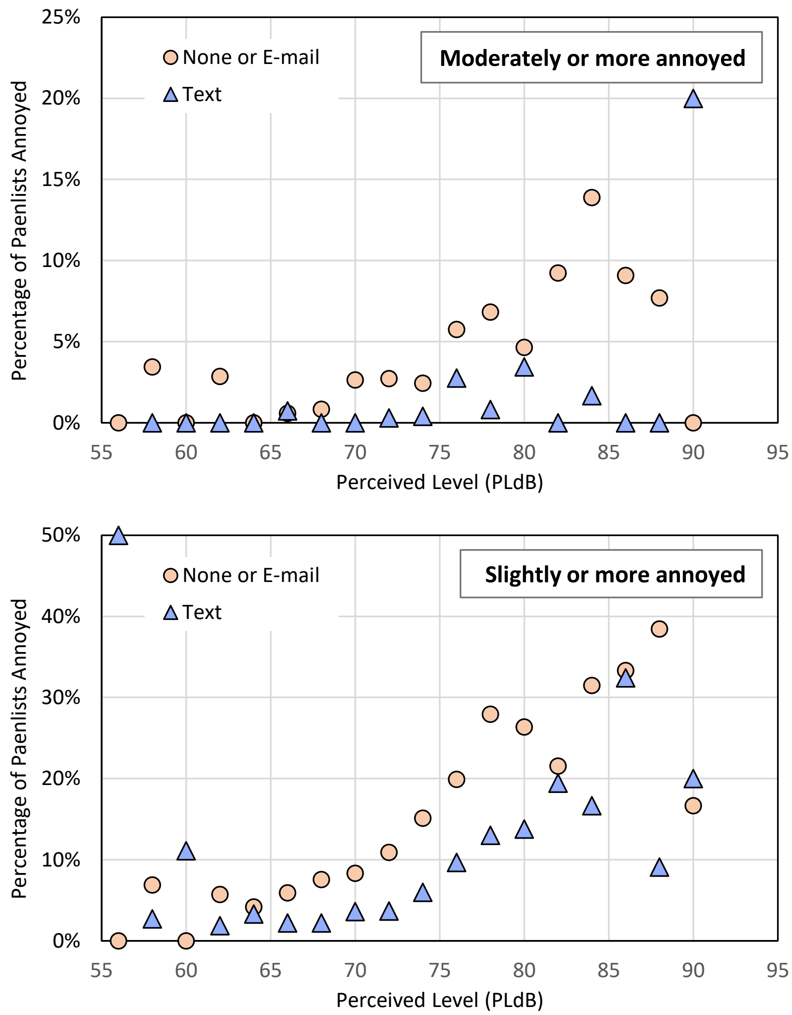

Figure 13. Slight and greater and moderate or greater annoyance judgments for text message and e-mail plus no-reminder groups combined

5.1.2 *Summary of effects of reminder messages*

At no level of annoyance were statistically significant differences observed between whether or not reminder messages were issued after each noise event. E-mail reminder messages were associated with a stronger relationship (*i.e.*, steeper slope) between exposure levels and annoyance judgments than text reminder messages for judgments of high annoyance. E-mail reminder messages were associated with a stronger relationship between exposure levels and annoyance judgments than text reminder messages for judgments of high annoyance. In general, text reminder messages were associated with greater annoyance judgments than e-mail reminder messages at Perceived Levels less than 75 PLdB, but e-mail reminder messages were associated with greater annoyance judgments at higher Perceived Levels.

5.2 Panelist Locations (Prompt Annoyance Judgments)

The spatially non-uniform noise dose created by the dive maneuver necessitated determination of panelists' geographic locations at the time of each boom to assign a dose to each annoyance judgment. Page *et al.* (2020) expended considerable effort in determining these locations. This effort provided the basis for determining whether panelists were in one of three generic locations, (1) At Home, (2) At Work or (3) Elsewhere at the time of exposure. However, for only about a quarter of the analyzable case records could geographical locations be translated to the two generic locations of at-home or at-work. These cases provided an initial basis for examining potential dose-response differences between these two environments. No significant difference was found in prevalence of annoyance between the two locations.

Panelists' generic locations were determined only for prompt annoyance judgments. Table 28 shows the distribution of panelist locations at the time of exposures, while Figure 14 plots the same information. At Home records shown in the table are those classified by Page *et al.* (2020) as Match Types 2 and 6. At Work records are those classified as Match Types 3 and 7. The Undetermined records are those categorized as Match Type 1, of which only a small percentage were believed to be neither at home nor at work. Records with all other Match Types were excluded for various reasons from the pool eligible for analysis.

Table 28: Distribution of panelist locations at times of exposures

Panelist Location	Frequency	Percent
At Home	728	17.5%
At Work	295	7.1%
Undetermined	3146	75.5%
Totals	**4,169**	**100.0%**

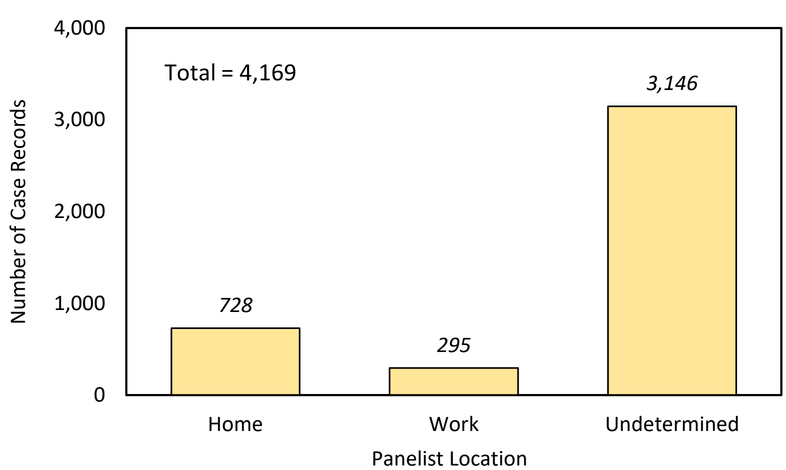

Figure 14. Prompt annoyance judgments by panelists in known locations at time of exposure

A variable was formed to represent planned comparisons between generic home and work locations, while ignoring the annoyance judgments in which generic location was Undetermined. Another binary variable was formed to indicate whether location was determined. HLM software was unable to determine standard errors for a model in which both these variables were analyzed together. Separate analyses are therefore reported for (1) home *vs.* work and (2) generic known *vs.* generic unknown panelist location.

Figure 15 shows annoyance judgments made by At Home and At Work panelists. The upper panel plots slightly or more greatly annoyed judgments. The middle panel plots moderately or more greatly annoyed judgments, while the bottom panel plots high annoyance judgments. For all three degrees of annoyance, At Home panelists show higher annoyance prevalence rates than did those At Work. Once again, the paucity of annoyance judgments at the high degree of annoyance precludes any strong conclusions. However, the trends of the lower degrees of annoyance suggest that a similar trend is likely to be observed at higher exposure levels.

Figure 16 shows the numbers of judgments used for the calculations in each sound level bin. Most of these annoyance judgments lie in the 75 PLdB $\leq$ PL $\leq$ 90 PLdB range. Fewer At Work than At Home panelists provided annoyance judgments.

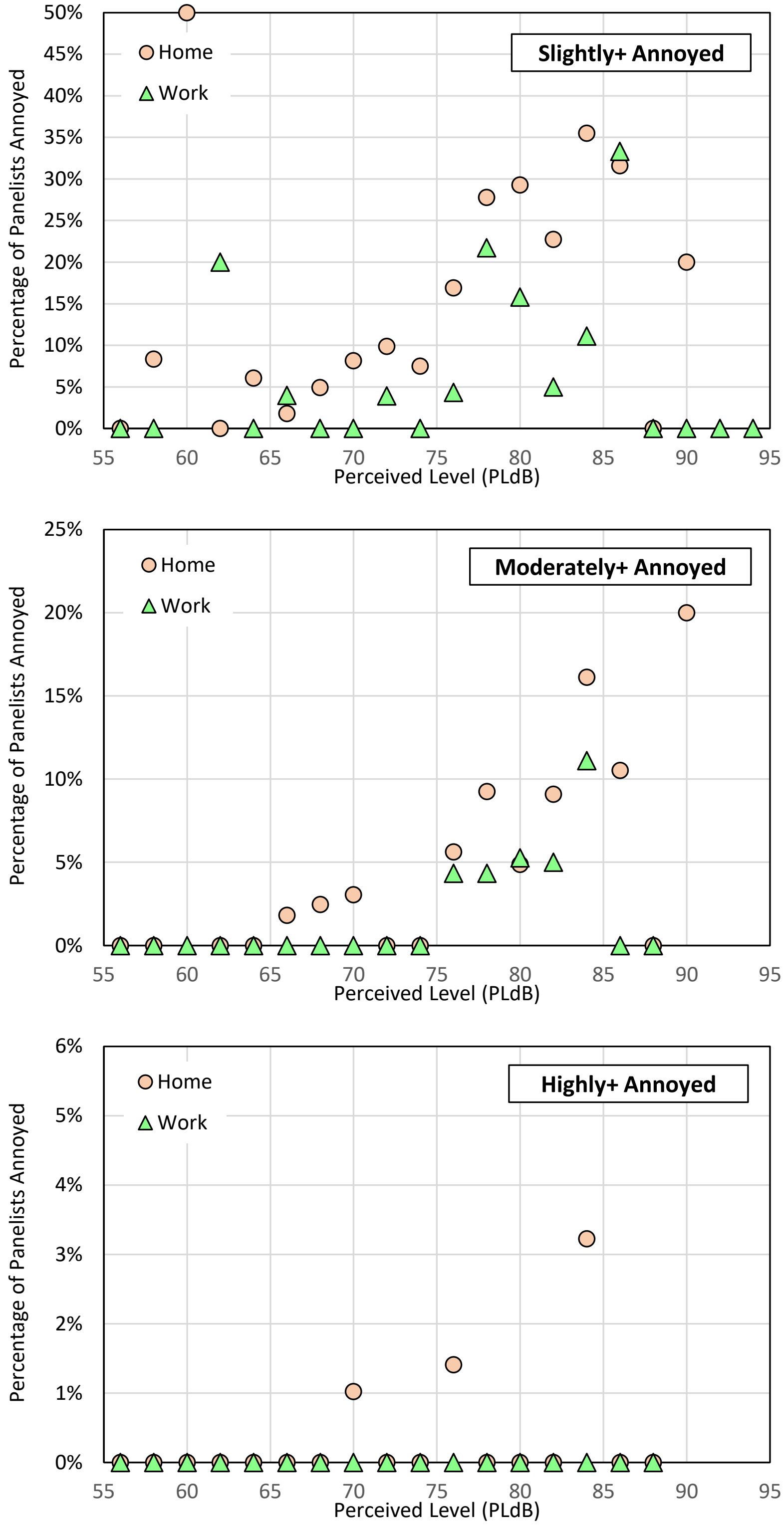

Figure 15. Prompt annoyance judgments for At Home and At Work Panelist Locations

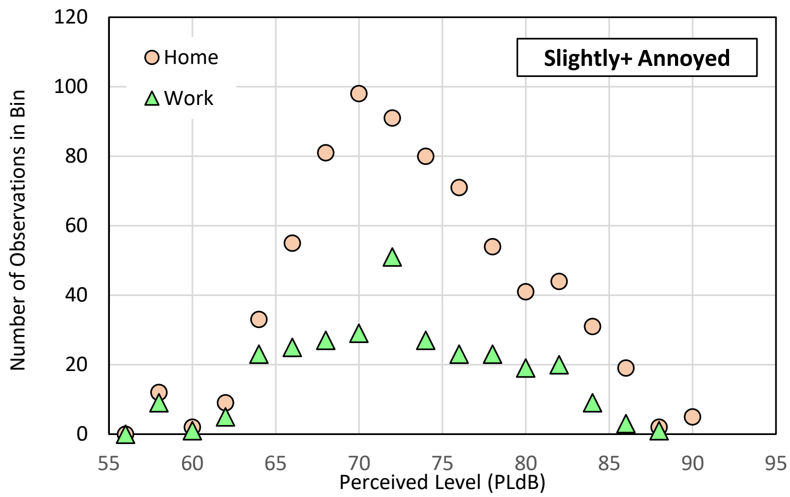

Figure 16. Numbers of prompt annoyance judgments *per* exposure bin for At Home and At Work annoyance judgments

The multi-level logistic regression equation for predicting annoyance from noise dose, panelist generic location (At Home or At Work), as well as the interaction between generic location and noise exposure is:

$$p(annoyed = yes) = \frac{e^y}{1+e^y} \qquad \textbf{Equation 4}$$

where:

$y=\beta_{00} + \beta_{10}\,(PL) + (\beta_{20})(HomeWork) + (\beta_{30})(HomeWork)(PL)$

and

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion level of annoyance;

PL is the Perceived Level noise dose in PLdB

HomeWork is the code for location (1 for Home, 0 for generic Undetermined [Match_Type = 1[20]], and -1 for Work);

Table 29,Table 30, and Table 31 show results of multi-level logistic single-event models in which Perceived Level and generic location (home *vs.* work) are utilized to predict annoyance defined by the three annoyance criteria: (1) HA (high annoyance) includes annoyance judgments indicating that panelists described themselves as very or extremely annoyed; (2) MA (panelists described themselves as at least moderately annoyed), and (3) SA (panelists described themselves

[20] Match_Type = 1 indicates the panelists' GPS location was autodetected correctly. However, it does not indicate whether the panelist was at home, at work, or somewhere else.

as at least slightly annoyed); all after adjustment for Perceived Level (PL). The interaction between location and noise dose is labeled HomeWork x PL in the tables.

For example, the probability of high annoyance is calculated from Equation 4 and Table 29 as:

$$y=\beta_{00} + \beta_{10} (PL) + (\beta_{20})(HomeWork) + (\beta_{30})(HomeWork)(PL)$$

$$-13.480+0.118(PL)+(-0.019)(HomeWork+(.001)(HomeWork)(PL)$$

Table 29: Prompt response model of high annoyance due to Perceived Level of exposure, panelist location, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.480	1.396	-9.66	353	<.001			
NoiseDose (β_{10})	0.118	0.018	6.47	4,165	<.001	1.126	1.086	1.167
HomeWork (β_{20})	-0.019	1.807	-0.10	4,165	.992	0.982	0.028	33.897
HomeWork x PL (β_{30})	.001	0.024	0.058	4,165	.954	1.001	0.955	1.050

Table 30: Prompt response model of moderate or greater annoyance due to Perceived Level of exposure, panelist location, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-12.300	0.902	-13.64	353	<.001			
NoiseDose (β_{10})	0.094	0.007	14.08	4,165	<.001	1.098	1.084	1.113
HomeWork (β_{20})	2.159	1.209	1.79	4,165	.074	8.665	0.811	92.636
HomeWork x PL (β_{30})	-.012	0.016	-1.44	4,165	.151	0.977	0.947	1.008

No significant differences were observed in any degree of annoyance judgments between home and work locations among panelists whose locations were determined. (Recall, however, that many panelists' locations were not determined.) Very large confidence interval widths for odds ratios in these tables indicate a great deal of uncertainty in resolving the strength of the effect of location on two classes of annoyance judgments (HA and MA) through multi-level logistic regression. Nevertheless, this small subsample of panelists whose location was determined reveals no statistically significant evidence of differences in annoyance prevalence rates between At Home *vs.* At Work panelists.

Table 31: Prompt response model of slight or greater annoyance due to Perceived Level of exposure, panelist location, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-8.907	0.522	-17.06	353	<.001			
NoiseDose (β_{10})	0.095	0.007	14.30	4,165	<.001	1.099	1.085	1.114
HomeWork (β_{20})	0.051	0.719	0.07	4,165	.943	1.053	0.257	4.305
HomeWork x PL (β_{30})	0.003	0.010	0.32	4,165	.750	1.003	0.983	1.024

Table 32, Table 33, and Table 34 show results of multi-level logistic single-event models in which Perceived Level and panelist locations predicted annoyance judgments in three degrees: (1) HA (high annoyance); (2) MA (at least moderate annoyance), and (3) SA (at least slight annoyance), respectively; all after adjustment for Perceived Sound Level (PL). "Determined x PL" in the tables labels the interaction between determined panelist locations and noise dose.

The multi-level logistic regression equation for predicting annoyance from noise exposure, whether location is determination, and the interaction between location determination and noise exposure is:

$$\boldsymbol{p(annoyed = yes) = \frac{e^y}{1+e^y}} \qquad \textbf{Equation 5}$$

where:

$y=\beta_{00} + (\beta_{10})(PL) + (\beta_{20})(Determined) + (\beta_{30})(Determined)(PL)$

and:

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion degree of annoyance;

PL is the noise dose in Perceived Level

Determined is whether panelist location could be designated as Home or Work with certainty (a location that could be determined as Home or Work was coded as "1", while a location that could not be determined as either was coded as "0")

Table 32 displays values for the equation for predicting high annoyance (HA):

$y = \beta_{00} + (\beta_{10})(PL) + (\beta_{20})(Determined) + (\beta_{30})(Determined)(PL)$

$y=-13.321+0.118(PL)+(-1.188)(\text{ Determined})+(0.015)(\text{ Determined by PL})$

Table 32: Prompt response model of high annoyance due to Perceived Level of exposure, whether panelist location could be determined, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.321	1.156	-8.53	353	<.001			
NoiseDose (β_{10})	0.118	0.020	5.77	4,165	<.001	1.125	1.081	1.171
Determined (β_{20})	-1.188	3.216	-0.57	4,165	.572	0.162	0.000	88.716
Determined x PL (β_{30})	0.015	0.041	.036	4,165	.720	1.015	0.965	1.102

Table 33: Prompt response model of at moderate or greater annoyance due to Perceived Level of exposure, whether panelist location could be determined, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-11.979	0.809	-14.80	353	<.001			
NoiseDose (β_{10})	0.117	0.012	10.89	4,165	<.001	1.125	1.101	1.149
Determined (β_{20})	0.052	0.130	0.33	4,165	.734	1.054	0.770	1.442
Determined x PL (β_{30})	0.015	0.042	0.36	4,165	.720	1.015	.935	1.102

Table 34: Prompt response model of slight or greater annoyance due to Perceived Level of exposure, whether panelist location could be determined, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-8.675	0.552	-17.61	353	<.001			
NoiseDose (β_{10})	0.092	.007	12.99	4,165	<.001	1.097	1.081	1.112
Determined (β_{20})	-1.038	1.190	-0.87	4,165	.383	0.354	0.034	3.649
Determined x PL (β_{30})	0.012	0.016	0.76	4,165	.449	1.012	0.981	1.044

These analyses show no evidence of any relationships between annoyance and whether location was determined, nor did determination of location affect dose-response relationships. However, interpretation must be tempered in the light of the mixed nature of location estimates for the responses from panelists whose location was undetermined. These annoyance judgments include responses from panelists At Home, At Work, and Elsewhere. That is, both the determined and undetermined location categories include panelists who were At Home or At Work. However, only the undetermined locations also include panelists who were elsewhere. Therefore, this is an extremely weak test of the difference between the two categories. The extremely wide confidence limits for the odds ratio for the high annoyance criterion speaks further to the uncertainty associated with this analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

6 ADDITIONAL ANALYSES

The analyses described in this Chapter focus initially on long-term participation of panelists in the nine-day QSF18 study, without regard for dose-response issues. Noise exposure level in these analyses is considered a covariate (for which a statistical adjustment is made) rather than as an effect of interest. Other analyses described in this Chapter include analyses of rattle and vibration onset levels, annoyance judgment latencies, and missing data.

6.1 Panelist Participation Rates in QSF18 Data Collection

This subsection discusses the numbers of participants remaining in the pool at various stages in the recruitment, test period and analysis. Prompt and delayed annoyance judgments are treated separately.

6.1.1 *Prompt annoyance judgments*

Rows 1 through 4 of Table 35 restate the contents of Table 5-4 of Page *et al.* (2020), and adds two additional rows. The table shows attrition in panelist participation at various stages from the initial recruitment of 544 panelists to the 354 panelists remaining in the single event database following quality control checks described below. Row 5 of Table 35 shows the number of unique Participant_ID codes found in the prompt response ("SE, or "single event") database supplied by NASA. The number of unique participant IDs found in the database (row 5) differs for unknown reasons by 68 from the number of panelists Page *et al.* (2020) believed were participating at the completion of the QSF18 data collection (row 4).

Table 35: Attrition in numbers of QSF18 prompt annoyance judgments (*per* Table 5-4, page 32 of *Page et al.*, 2020)

Row	Quantity	Count	Percent Remaining	Comment
1	Number of panelists initially recruited	544		
2	Number of panelists who responded to confirmation requests	500		
3	Panelists at start of testing	496	99.2%	Attrition due to unavailability
4	Panelists remaining at completion of testing	476	95.2%	Attrition due to failures of confirmed panelists to submit any responses
5	Number of Unique Panelist IDs in database	408	81.6%	Reduction from 476, probably due to inability to determine panelist locations
6	Number of Unique Panelist IDs after removal of unanalyzable records	354	70.8%	Total Attrition: (500-354)/500 = 29.2%

The data grooming process described in the subsections that follow eliminated all of some panelists' records. Row 6 of Table 35 shows the number of panelists remaining in the database following the grooming. The net effect of attrition and screening for eligibility for the current analyses is shown in the lower right-hand cell of the table. The 146 (500-354) panelists with no remaining analyzable records represent 29% of the original recruitment of 500 panelists.

6.1.2 *Delayed annoyance judgments*

In a manner similar to Table 35, Table 36 recreates Table 5-4 of Page *et al.* (2020) with two additional rows. The table shows attrition in panelist participation at various stages, from the initial recruitment of 544 panelists to the actual number of panelists contributing one or more annoyance judgments in the groomed delayed response database. Rows 1 through 4 restate the quantities found in Table 5-4 of Page *et al.* Row 5 shows the number of unique Participant_ID codes found in the DS database.

The data grooming process described in the subsections that follow eliminated all records for some participants. Row 6 shows the number of panelists remaining in the database who provided one or more delayed response annoyance judgments after screening. (Although all panelists were believed to be active, some panelists apparently made no end-of-day annoyance judgments.) The effect of attrition is shown in the lower right-hand cell of the table. The 114 (500-386) panelists with no remaining annoyance judgment records represent 23% of the original panel.

Table 36: Delayed judgment panelist attrition (*per* Table 5-4, page 32 of Page *et al.*, 2020)

Row	Quantity	Count	Percent Remaining	Comment
1	Number of panelists initially recruited	544		
2	Number of panelists who responded to confirmation requests	500		
3	Panelists at start of testing	496	99.2%	Attrition due to unavailability
4	Panelists remaining at completion of testing	476	95.2%	Attrition due to failures of confirmed panelists to submit any responses
5	Number of Unique Participant_IDs found in the end-of-day database	427	85.4%	Reduction from 476 for unknown reasons by Page *et al.*
6	Number of Unique Participant_IDs after removal of unanalyzable records	386	77.2%	Attrition: (500-386)/500 = 23%

6.1.3 *Summary of prompt and delayed response self-report rates*

Table 37 summarizes annoyance judgment opportunity completion rates in the NASA-provided data sets. The lefthand column of the table shows various stages in the data acquisition and analysis process. The remaining columns provided statistics on the two data sets, single event and end-of-day. The top row of the table shows the maximum possible number of annoyance judgments: shown (51 booms multiplied by 496 initial panelists for prompt responses, and 9 days multiplied by 496 initial panelists for delayed responses). The second row shows the number of judgments reported by Page *et al.*, 2020. These values are also shown as percentages of maximum possible values. The third row reports the numbers of records in the two data sets provided by NASA. The fourth row shows the numbers remaining following the data screening process described in §3. Of the maximum number of possible annoyance judgments, 16% of the prompt and 44% of the delayed responses were available for analysis.

Table 37: Annoyance judgment opportunity completion rates

Screening Stage	Single Event (Prompt Response)		Cumulative Exposure (Delayed, End of Day Response)	
	Number of Annoyance Judgments	Percent of Maximum	Number of Annoyance Judgments	Percent of Maximum
Maximum Possible	25,296	---	4,464	---
Number of judgments reported by Page *et al.*, 2020	11,869	47%	2,855	64%
Number in data sets provided by NASA	7,068	28%	2,855	64%
Screened records	4,169	16%	1,952	44%

6.2 Annoyance Judgment Rates in Groomed Data Set

This subsection discusses the numbers of both prompt and delayed annoyance judgments. The annoyance judgments discussed below are limited to the case records that satisfied the screening tests described in §3.1 and §3.2.

6.2.1 *Self-report rates for prompt annoyance judgments*

One or more analyzable annoyance judgments remained for 354 panelists following data record screening. This number represents 83% of the 427 panelists found in the database, and 71% of the original 500 panelists recruited. Table 38 shows the distribution of the 354 panelists making varying numbers of annoyance judgments. The greatest number of judgments by any panelist was 40. The first three columns of the table (under Density) group the numbers of judgments into range bins (1-5, 6-10, *etc*.) shown in column one Column two shows the number of panelists who provided a total number of annoyance judgments in that range, and column three converts this number into a percentage of the 354 panelist total. The last two columns (under Cumulative) show cumulative information which indicates the percentage of panelists providing N or more judgments.

Figure 17 plots the distribution of percentages shown in column three of Table 38. The abscissa plots the number of judgments shown in column one and the ordinate plots the percentages. For example, 39.0% of the panelists contributed between 1 and 5 annoyance judgments. Another 18.9% of panelists contributed between 6 and 10 judgments. The sum of all bar heights equals 100%.

Table 38: Distribution of panelists contributing one or more prompt annoyance judgments

Density			Cumulative	
Number of Judgments	**Number of Panelists**	**% of Panelists**	**Number of Judgments**	**% of Panelists**
1-5	138	39.0%	1 or more	100.0%
6-10	67	18.9%	6 or more	61.0%
11-15	35	9.9%	11 or more	42.1%
16-20	27	7.6%	16 or more	32.2%
21-25	35	9.9%	21 or more	24.6%
26-30	31	8.8%	26 or more	14.7%
31-35	15	4.2%	31 or more	5.9%
36-40	6	1.7%	36 or more	1.7%
41-45	0	0.0%	41 or more	0.0%
46-50	0	0.0%	46 or more	0.0%
51-55	0	0.0%	51 or more	0.0%
Totals	354	100.0%		

Figure 18 shows the cumulative percentages of panelists contributing N or more prompt annoyance judgments. Starting with the leftmost column, every active panelist, by definition, completed one or more judgments. However, only 61% made more than 6 judgments, and only 42% contributed more than 11 judgments. Only 14.7% of panelists contributed 26 or more annoyance judgments.

Each of the 354 panelists contributing annoyance judgments to the total of 4,169 useable case records had 52 annoyance judgment opportunities (impulsive exposure events.) The average number of judgments was therefore 4,169/354 = 11.8 *per* panelist. The overall annoyance judgment rate was thus 11.8/52 = 0.226.

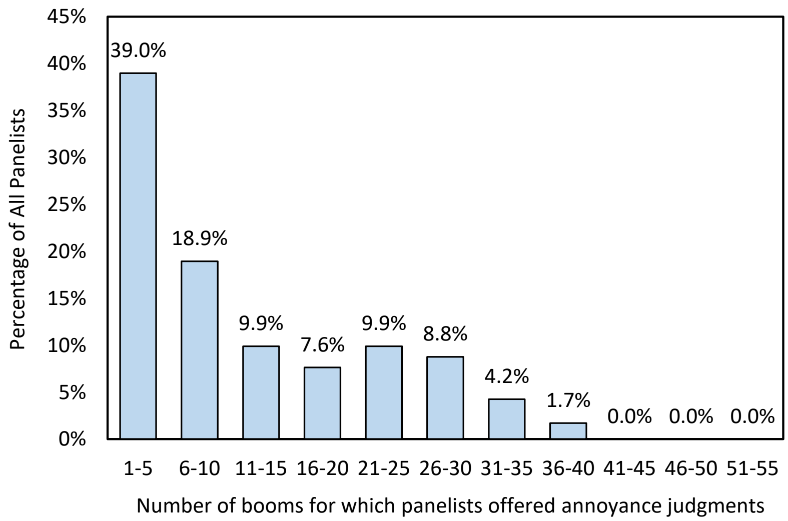

Figure 17. Histogram of percentages of panelists contributing prompt annoyance judgments by numbers of boom events

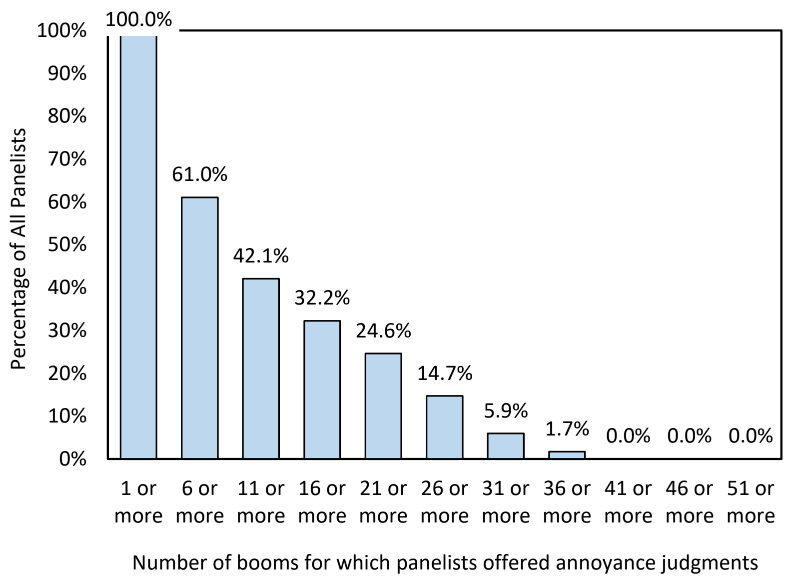

Figure 18. Histogram of cumulative percentages of panelists contributing prompt annoyance judgments to N or more boom events

6.2.2 *Self-report rates for delayed annoyance judgments*

Following case record screening, one or more analyzable annoyance judgments remained for 386 panelists. Table 39 shows the numbers of panelists contributing varying numbers of end-of-day annoyance judgments. The table divides the panelists into the nine groups (rows.) The table is divided into two sections. The left-hand side shows density results, while the right side shows cumulative information.

The leftmost column of the Density group in Table 39 shows numbers of annoyance judgments received from panelists (0 to 9). The second column shows the numbers of panelists completing each of these numbers of judgments. The third column shows the percentage of the original 427 panelists represented by these numbers.

For the Cumulative display, the first column shows the number of annoyance judgments made. The second column shows the numbers of panelists making these judgments, and the third column shows the second column's percentage of the original 427 panelists. The top row of the table (shaded in orange) shows that 41 (9.6%) of the original 427 panelists provided no end-of-day annoyance judgments. The 386 panelists who made at least one judgment represent 90.4% of the 427 panelists. These numbers are identified in boldface to indicate the starting point of the current analysis. The 386 panelists represent 77.2% of the 500 panelists originally recruited.

Table 39: Distribution of panelists contributing delayed response judgments

Density			Cumulative		
# of Self-Reports	**Number of Panelists**	**% of Panelists**	**# of Annoyance judgments**	**Number of Panelists**	**% of Panelists**
0	41	9.6%	None	427	100.0%
1	45	10.5%	1 or more	**386**	**90.4%**
2	45	10.5%	2 or more	341	79.9%
3	38	8.9%	3 or more	296	69.3%
4	38	8.9%	4 or more	258	60.4%
5	44	10.3%	5 or more	220	51.5%
6	42	9.8%	6 or more	176	41.2%
7	39	9.1%	7 or more	134	31.4%
8	49	11.5%	8 or more	95	22.2%
9	46	10.8%	9	46	10.8%

Figure 19 plots column 3 of the above table by showing the frequency with which panelists provided varying numbers of delayed (end-of-day) annoyance judgments. The orange-shaded leftmost column in the figure shows the percentage of panelists who made no annoyance judgments (top row of Table 39). Roughly 10% of the panelists fell into each of the ten daily self-report groups. The average number of annoyance judgments provided was 4.6 *per* panelist, while the average response rate across panelists was 51%.

Figure 20 shows the cumulative percentages of panelists contributing N or more prompt annoyance judgments. Starting with the next to leftmost column, 90% of the original panelists completed one or more annoyance judgments. However, only 41% provided more than six annoyance judgments, and only 22% contributed eight or more.

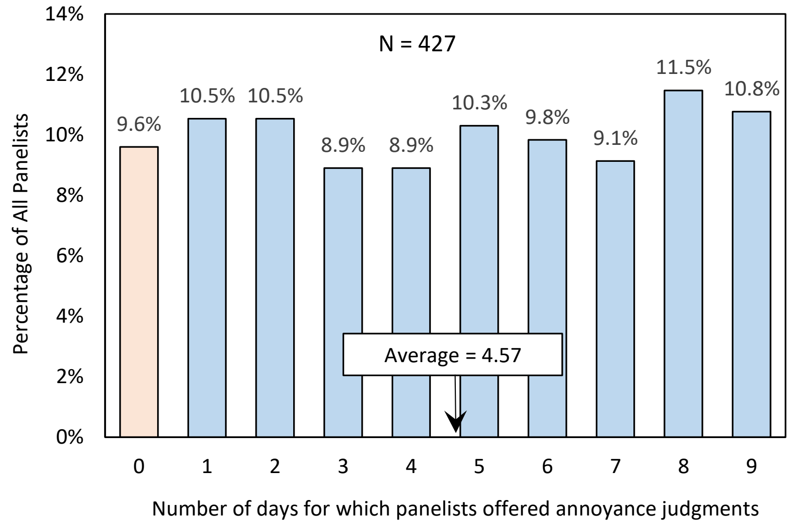

Figure 19. Frequency of participation in delayed annoyance judgments

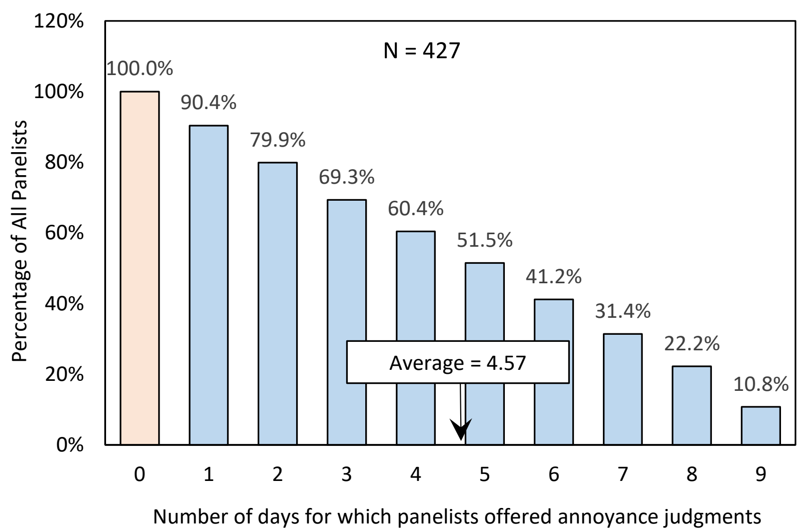

Figure 20. Cumulative frequency of participation in delayed annoyance judgments

6.3 Duration of Participation by Panelists

This subsection examines effects of the panel sample (repeated measures) design on annoyance judgments. In principle, panelists could have either sensitized, or habituated, or neither over repeated booms in the case of prompt responses, or repeated days of exposure in the case of delayed responses. In the case of prompt responses, duration of participation was measured as the sequence number of booms to which each panelist responded. In the case of delayed annoyance judgments, duration of participation was measured as the sequential day of exposure for each panelist.

6.3.1 *Sequence effects of duration of participation on prompt responses*

Figure 21 illustrates differences in annoyance judgments in varying degrees over the course of the QSF18 data collection exercise. The complete set of analyzable prompt annoyance judgments was separated into three sequential groups: (1) judgments made on days 1 through 3; (2) judgments made on days 4 through 6; and (3) judgments made on days 7 through 9. The top panel in the figure plots separate dose-response relationships for the slight or greater level of annoyance. The graphic suggests that the prevalence of at least slight annoyance was lower during the first few days of exposure to low amplitude sonic booms with respect to the prevalence in subsequent time periods.

The middle panel of Figure 21 displays the same analysis for moderate or greater annoyance. A similar temporal pattern of lower initial prevalence of annoyance appears as for slight or greater annoyance judgments. The bottom panel has too few responses in the highly annoyed category to reveal any such pattern. Figure 22 shows the total numbers of responses in each sound level bin, a fraction of which were annoyed at each exposure level.

The causes for the above temporal dependencies are unknown. Conjecturally, they could be due to rapid sensitization to low-amplitude booms over the nine-day course of exposure, or perhaps to a partial confounding of the exposure schedule with daily exposure.

A model for sequence effects is shown in Equation 6. Table 40, Table 41, and Table 42 display the results of single-event multi-level logistic regression models in which sequential exposure event, (equivalent to the number of times a panelist has provided an a prompt annoyance judgment) predicts three degrees of annoyance judgments: (1) HA (high annoyance) includes annoyance judgments indicating that panelists described themselves as very or extremely annoyed; (2) MA (panelists described themselves as at least moderately annoyed), and (3) SA (panelists described themselves as at least slightly annoyed); all after adjustment for Perceived Level. Noise exposure levels associated with events serve as a covariate in this model, so that sequential judgments are adjusted for noise dose.

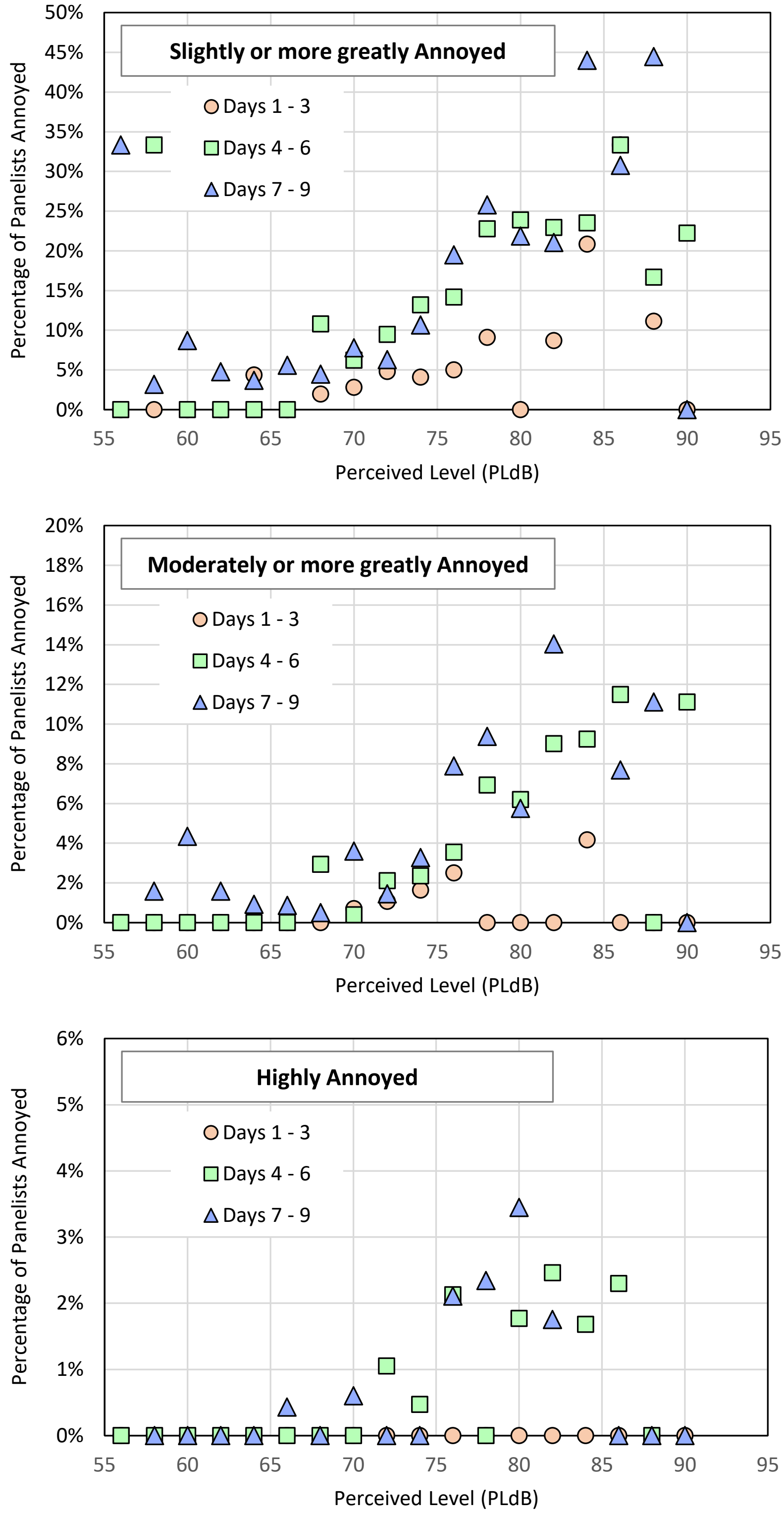

Figure 21. Prevalence of annoyance during three study participation periods

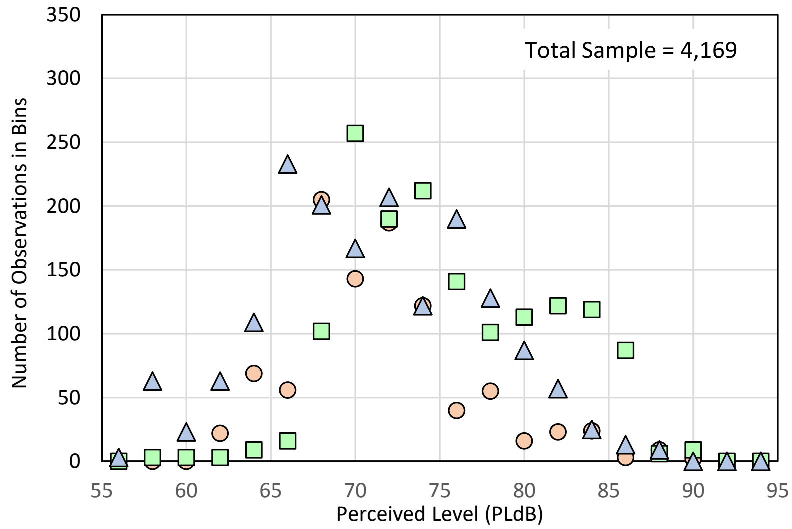

Figure 22. Sample size distribution for three study participation periods of Figure 21

$$\boldsymbol{p(HA = yes) = \frac{e^y}{1+e^y}} \qquad \textbf{Equation 6}$$

where:

$y = \beta_{00} + (\beta_{10})(PL) + (\beta_{20})(Sequence)$

and:

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion level of annoyance;

PL is the noise dose in Perceived Level; and

Sequence is the sequential number of prompt annoyance judgments rendered by panelists (1-52)

Table 40: Prompt response model of high annoyance due to duration of study participation, after adjusting for Perceived Level of Exposure

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-14.593	1.037	-14.06	353	<.001			
NoiseDose (β_{10})	0.128	0.015	8.82	4,166	<.001	1.136	1.105	1.169
Sequence (β_{20})	0.042	0.016	2.57	4,166	.023	1.043	1.010	1.076

Table 41: Prompt response model of at least moderate annoyance due to duration of study participation, after adjusting for Perceived Level of exposure

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-12.450	0.768	-16.20	353	<.001			
NoiseDose (β_{10})	0.121	0.010	12.02	4,166	<.001	1.128	1.106	1.151
Sequence (β_{20})	0.031	0.009	3.60	4,166	.001	1.032	1.014	1.050

Table 42: Prompt response model of at least slight annoyance due to duration of study participation, after adjusting for Perceived Level of exposure

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-9.151	0.503	-18.19	353	<.001			
NoiseDose (β_{10})	0.097	0.006	15.24	4,166	<.001	1.102	1.088	1.115
Sequence (β_{20})	0.015	0.005	2.75	4,166	.006	1.015	1.004	1.026

Annoyance prevalence rates in all judgment categories appeared to increase with study duration, as well as with the noise level of the boom, although the relationship was statistically significant only for moderately or more greatly annoyed judgments. The levels of successive booms were only partially randomized over test days[21] as shown in Figure 23, so the boom exposure schedule partially confounded numbers of exposures with cumulative exposure levels over the nine days of data collection. The six red squares in the figure indicate booms for which no case records were eligible for analysis. Familiarity with boom exposure, of course, increased as data collection proceeded, as did numbers of annoyance judgments made. It is not possible to determine whether an observed sequential bias is more reasonably attributed to growing annoyance with repeated self-reports, or to greater familiarity with booms over the course of data collection.

The effect of sequence was significant for the moderate or greater level of annoyance, as shown in Table 41. Figure 24 plots the percentage of panelists moderately or more greatly annoyed as a function of the number of times they offered annoyance judgments to illustrate the effect. A family of curves shows the effect at various single event sound levels. The solid black curve shows the effect at the average value of Perceived Level across all annoyance judgments.

[21] For pragmatic reasons, higher level booms were avoided on early data collection days.

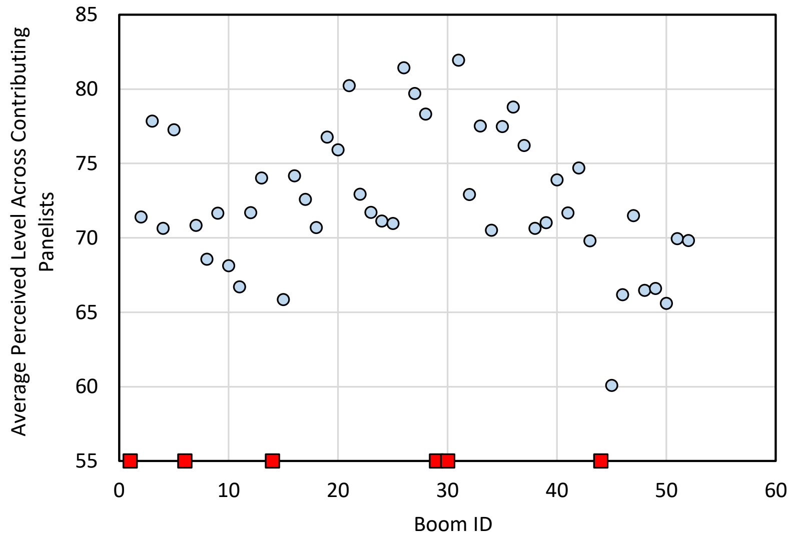

Figure 23. Average Perceived Level across contributing panelists by boom number

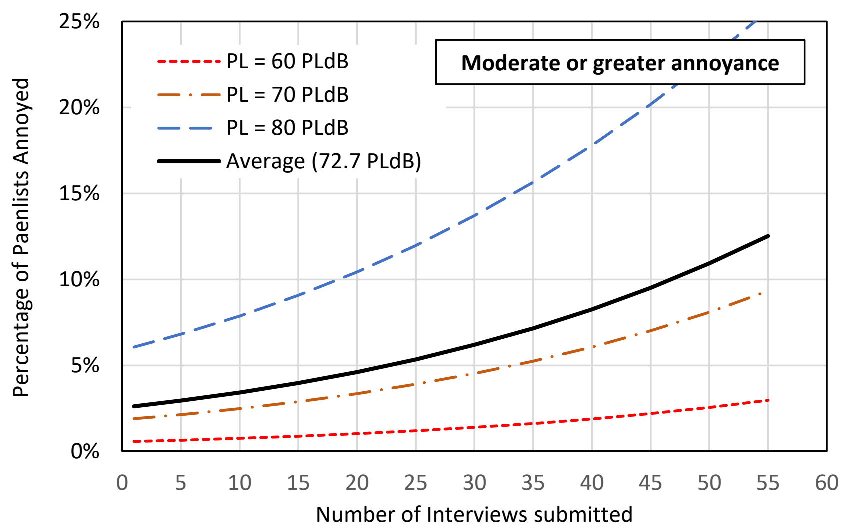

Figure 24. Increase in prevalence of moderate or greater prompt annoyance with increasing number of annoyance judgments offered, parametric in Perceived Level

All of the logistic curve fits in Figure 24 increase at about the same rate *per* doubling of number of annoyance judgments, with only a weak dependence on level. Figure 25 plots this rate of increase for each curve fit. The ordinate of Figure 25 is the factor by which the percentage of annoyed panelists (moderately or more) increased compared with the percentage associated with half the number of annoyance judgments. For example, the 80 PLdB curve fit in Figure 24 shows an annoyance prevalence rate of 17.8% at 40 annoyance judgments. For half that number of judgments (20), the prevalence of annoyance is 10.43%. The ratio (multiplier) is 1.71. The steeper growth rate at low levels behaves much like the rate of growth of loudness with level: that is, the growth rate is greater at low sound levels than it is at higher ones. Combined, the two graphs suggest a potentially important effect of either repeated testing, exposure level, or some combination of the two.

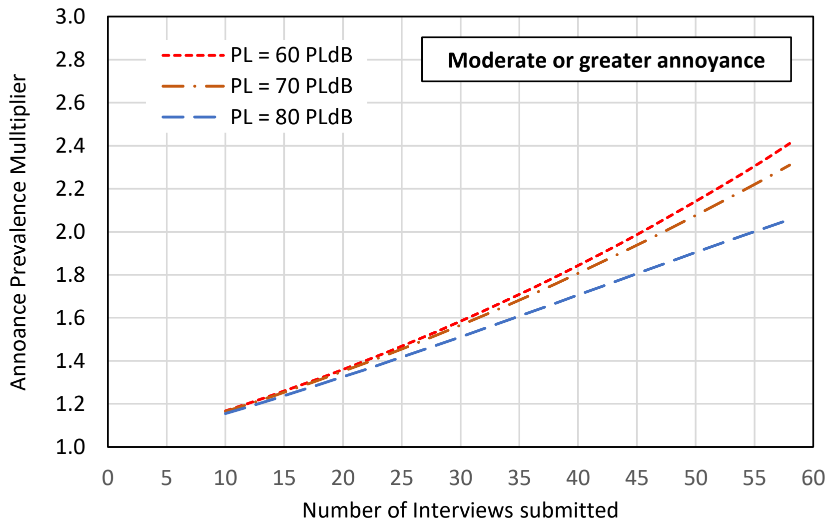

Figure 25. Relative rates of increase of prevalence of prompt annoyance with increasing number of annoyance judgments offered, parametric in Perceived Level

6.3.2 *Sequence effects of length of participation on delayed annoyance judgments*

Table 43, Table 44, and Table 45 show results of delayed response regression models in which sequential day of exposure predicts annoyance as defined by three degrees of annoyance. Panelist judgments are those on the nth day of exposure. The HA category includes panelist annoyance judgments that they were very or extremely annoyed; the MA category includes panelist annoyance judgments that they were at least moderately annoyed; and the SA category includes panelist annoyance judgments of at least slight annoyance; all after adjustments for the single event or Day-Night Average Perceived Sound Level of their exposures.

Table 43: Model of high annoyance due to duration of study participation, after adjusting for average Day-Night Average Perceived Sound Level

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-7.785	0.368	-21.13	385	<.001			
NoiseDose	0.147	0.028	8.27	1,949	<.001	1.088	1.067	1.110
Day of Exposure	0.146	0.028	5.23	1,949	<.001	1.158	1.096	1.242

Table 44: Model of at least moderate annoyance due to duration of study participation, after adjusting for average Day-Night Average Perceived Sound Level

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-7.793	0.454	-17.16	385	<.001			
NoiseDose	0.115	0.012	9.589	1,949	<.001	1.122	1.096	1.149
Day of Exposure	0.229	0.029	7.91	1,949	<.001	1.257	1.188	1.331

Table 45: Model of at least slight annoyance due to duration of study participation, after adjusting for average Day-Night Average Perceived Sound Level

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-5.929	0.413	-14.37	385	<.001			
NoiseDose	0.106	0.01107	96734	1,949	<.001	1.112	1.088	136
Day of Exposure	0.165	0.024	6.88	1,949	<.001	1.180	1.126	1.237

Annoyance in all three degrees increased with duration of participation after controlling for levels of exposure to low-amplitude sonic booms. This sequential bias could be due either to annoyance due to repeated judgments, or to cumulative exposure over the course of the test.

The annoyance parameter estimates in Table 43 were used to plot (in Figure 26) the percentage of panelists highly annoyed as a function of test day to illustrate this sequential effect. The family of curves shows the effect at three PLDNL values. The solid black curve shows the effect at the average PLDNL across all annoyance judgments.

Like the prompt response relationships shown in Figure 24, all of the logistic curve fits in Figure 25 increase at roughly the same rate *per* doubling of number of annoyance judgments, with only a weak dependence on level. Figure 27 plots this rate of increase for each curve fit. The ordinate of Figure 27 is the factor by which the percentage of moderately or more greatly annoyed panelists increased compared with the percentage associated with half the number of annoyance judgments. For example, the 40 PLdB curve fit in Figure 34 shows an annoyance prevalence rate of 3.71% on the 8th test day. At half the number of test days (4), the percentage is 2.10, for a multiplier of 1.76.

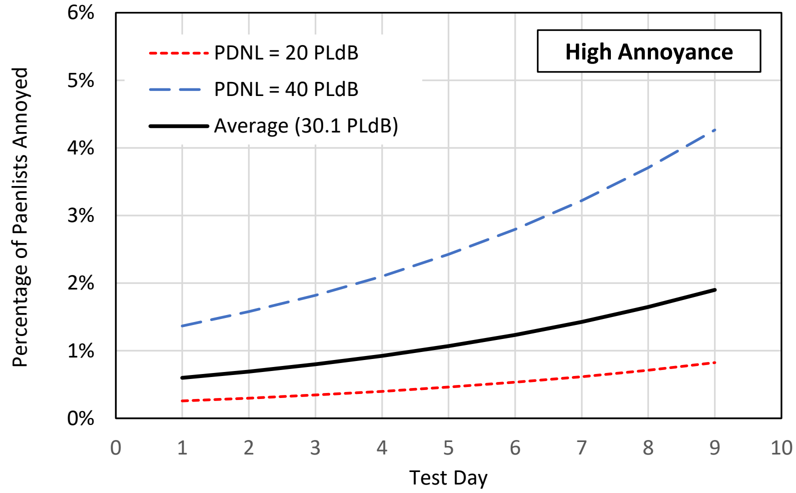

Figure 26. Increase in prevalence of delayed high annoyance judgments with increasing number of test days, parametric in Day-Night Average Perceived Sound Level

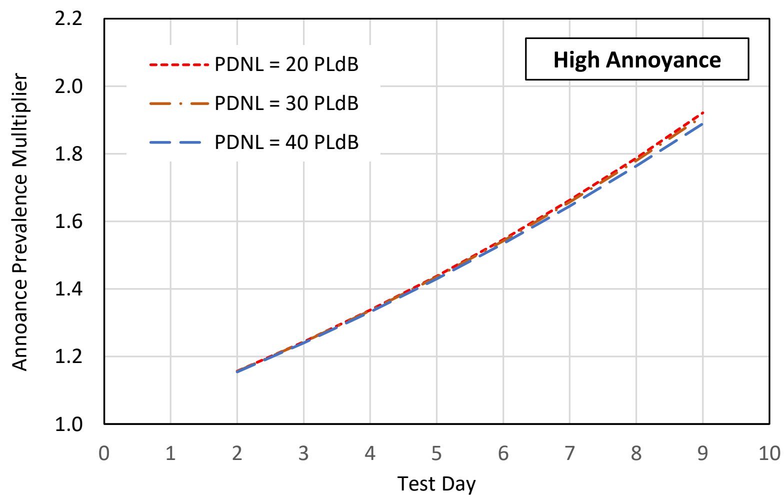

Figure 27. Relative rates of increase of delayed response prevalence of high annoyance with increasing number of test days, parametric in PLDNL

Figure 28 and Figure 29 plot similar results for moderate or greater degrees of annoyance. The annoyance prevalence multipliers in Figure 29 after nine days of participation (*i.e.*, to the end of the study) are at 2.3 to 2.7. These multipliers are similar to the multiplier range of 1.9 to 2.2 for ~50 booms in the prompt or moderate or greater degrees of annoyance (*cf.* Figure 25, page 75).

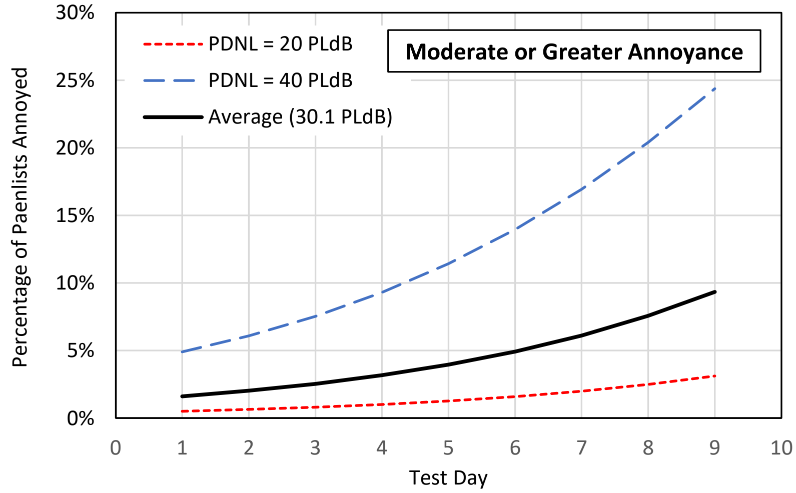

Figure 28. Increase in prevalence of delayed response for moderate or greater annoyance with increasing number of test days, parametric in PLDNL

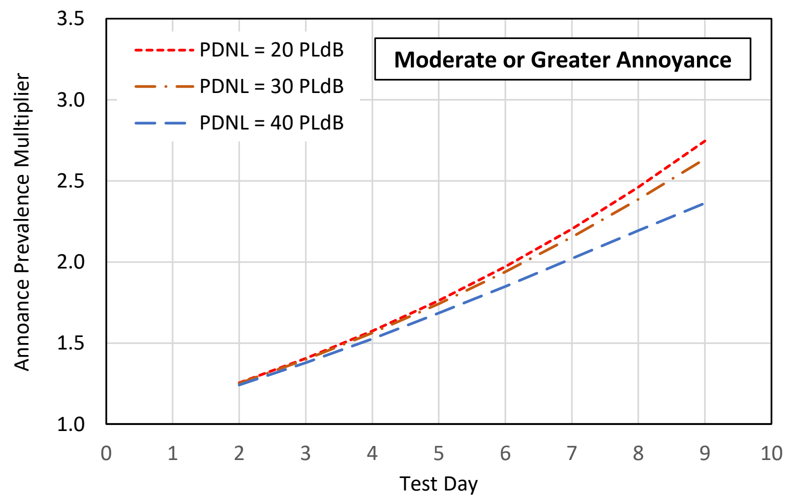

Figure 29. Relative rates of increase of delayed response prevalence of moderate or greater annoyance with increasing number of test days, parametric in PLDNL

Figure 30 and Figure 31 plot comparable results for the slight or greater degrees of annoyance judgments.

The growth rates shown in Figure 27, Figure 29 and Figure 31 are steeper at low levels than at greater ones for all three degrees of annoyance judgments. This growth resembles the growth of loudness as a function of level, because growth rates of loudness with level are greater at low sound levels than at higher ones. In combination, two graphs suggest a potentially important effect of either repeated testing, exposure level, or a combination of the two factors.

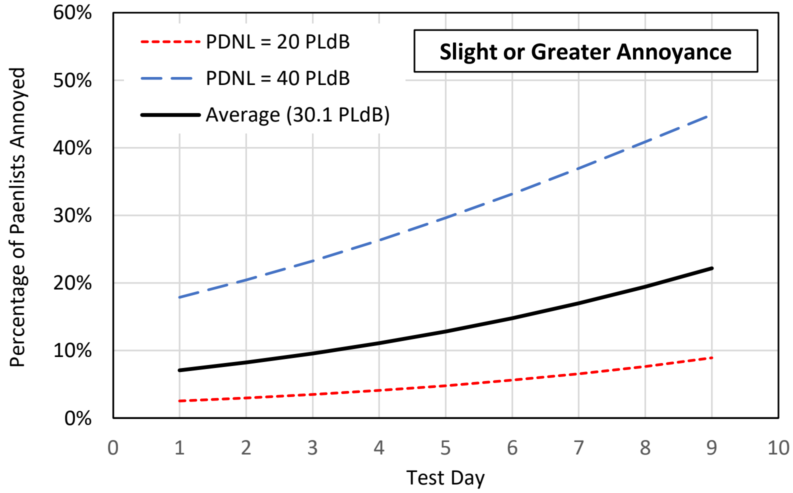

Figure 30. Increase in prevalence of delayed judgments of slight or greater annoyance with increasing number of test days, parametric in PLDNL

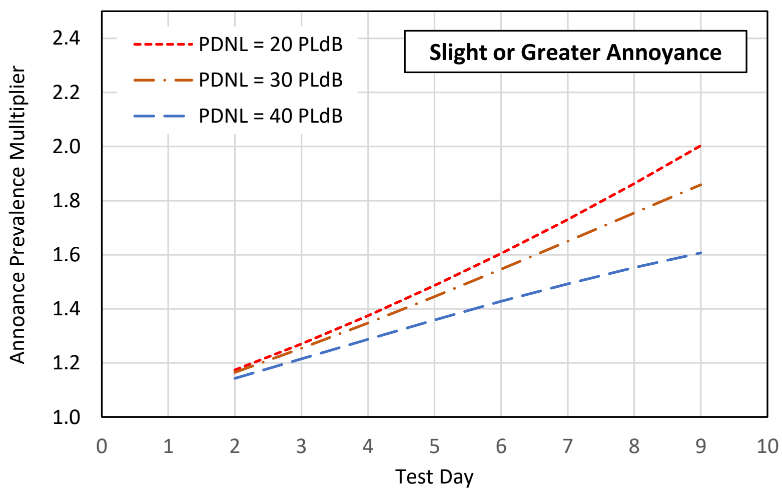

Figure 31. Relative rates of increase of delayed response prevalence of slight or greater annoyance with increasing number of test days, parametric in PLDNL

6.4 Annoyance due to Rattle and Vibration

Annoyance due to rattle or vibration was judged only for single event exposures. Reporting of vibration or rattle was observed at very low Perceived Levels, as shown in Figure 32.[22] The

[22] These judgments are likely to represent annoyance due to secondary emissions of household paraphernalia, rather than annoyance due to structural vibration.

prevalence of reporting vibration or rattle was approximately 10% at a Perceived Level of 65 PLdB, increasing to approximately 50% at 85 PLdB.

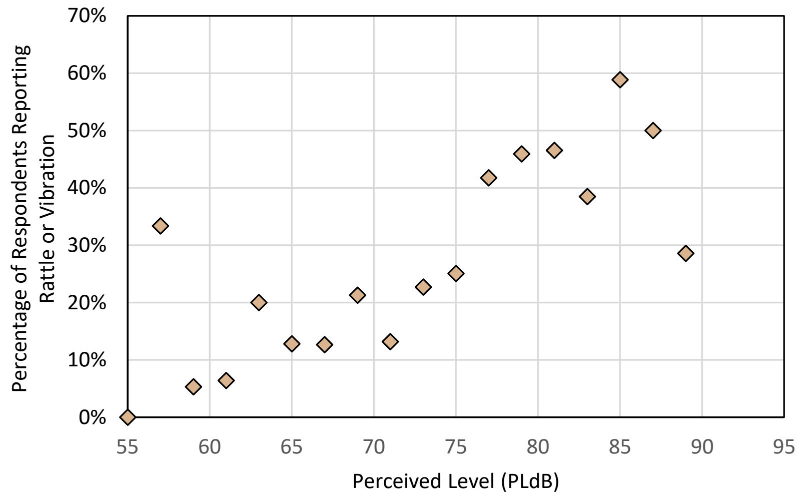

Figure 32. Prevalence of notice of vibration and rattle during study participation period

The same analytic strategy was followed as for all of the above dose-response analyses: multi-level logistic regression for three degrees of annoyance judgments.

Table 46, Table 47, and Table 48 show results of single-event models in which the Perceived Level of exposure and notice of rattle or vibration predict annoyance for three degrees of annoyance judgments: (1) HA (high annoyance), (2) MA (at least moderately annoyed), and (2) SA (at least slightly annoyed). Noticing rattle or vibration significantly increased annoyance for the two lesser degrees of judged annoyance after controlling for the level of noise exposure.

Table 46: Model of high annoyance due to Perceived Level of exposure and notice of rattle or vibration

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-12.400	1.777	-6.978	353	<.001			
NoiseDose	0.071	0.025	2.85	4.166	<.001	1.074	1.022	1.127
Rattle/Vibration	3.545	0.296	11.98	4.166	.005	34.630	19.392	61.839

Table 47: Model of at least moderate annoyance due to Perceived Level of exposure and notice of rattle or vibration

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-10.485	0.984	-10.66	353	<.001			
NoiseDose	0.071	0.014	5.14	4.166	<.001	1.074	1.045	1.103
Rattle/Vibration	3.133	0.184	17.04	4.166	<.001	22.932	15.994	32.879

Table 48: Model of at least slight annoyance due to Perceived Level of exposure and notice of rattle or vibration

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept	-8.103	0.578	-14.01	353	<.001			
NoiseDose	0.066	0.008	8.25	4.166	<.001	1.069	1.052	1.086
Rattle/Vibration	2.377	0.138	17.24	4.166	<.001	10.777	8.225	14.121

The three panels of Figure 33 illustrate the influence of notice of rattle or vibration at varying levels of annoyance. The uppermost panel of Figure 33 shows the influence of vibration or rattle on slight or greater annoyance judgments. Virtually no prevalence of annoyance is reported at Perceived Levels lower than 75 PLdB in the group which did not notice rattle or vibration. In contrast, the group which did report noticing rattle or vibration exhibited an annoyance prevalence rate of 20%, even at Perceived Levels as low as 60 PLdB.

The two lower panels of Figure 33 show that virtually no one who failed to notice vibration reported either moderate or high levels of annoyance, even at boom levels as high as PL = 85 PLdB. Taken together, these panels indicate that of judgments in all degrees of annoyance were strongly influenced by the notice of rattle or vibration at measured noise doses of 85 PLdB or less.

Table 49 breaks down annoyance judgments by degree and notice of rattle or vibration. Of those reporting no annoyance 81.4% did not notice rattle or vibration and 18.6% did. Of all the slight or greater annoyance judgments, 23.5% did not notice rattle or vibration, but 76.5% did. Of the moderate or more greatly annoyed participants, only 8.3% of the annoyance judgments were made by panelists who did not notice rattle or vibration, while the remaining 91.7% came from panelists who did notice rattle or vibration. The influence of rattle or vibration is even more extreme for greater degrees of annoyance. Such findings strongly suggest that notice of rattle or vibration is a consistent component of judgments of all degrees of annoyance, and especially so at the more extreme degrees of annoyance.

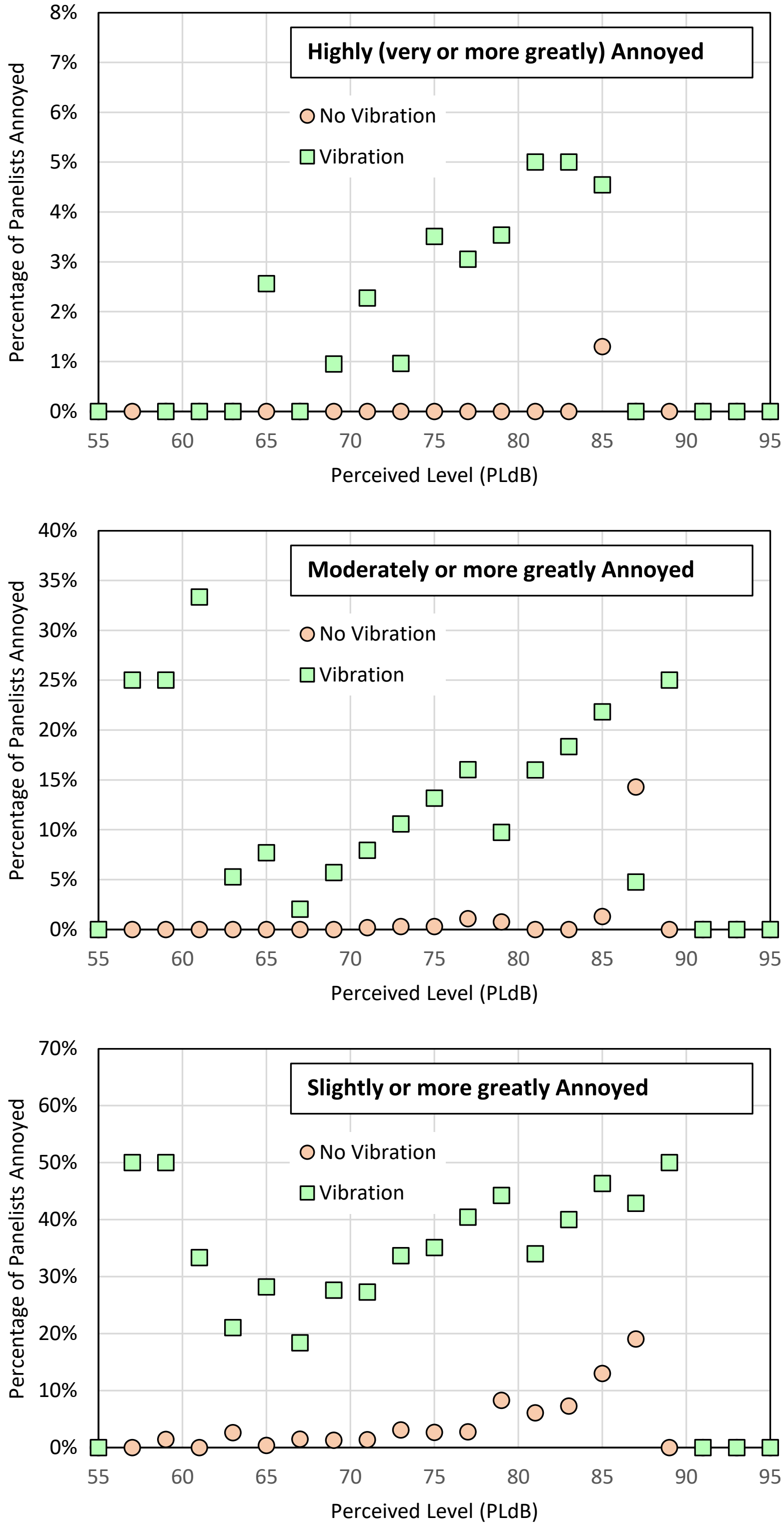

Figure 33. Prevalence of annoyance judgments with and without notice of rattle or vibration

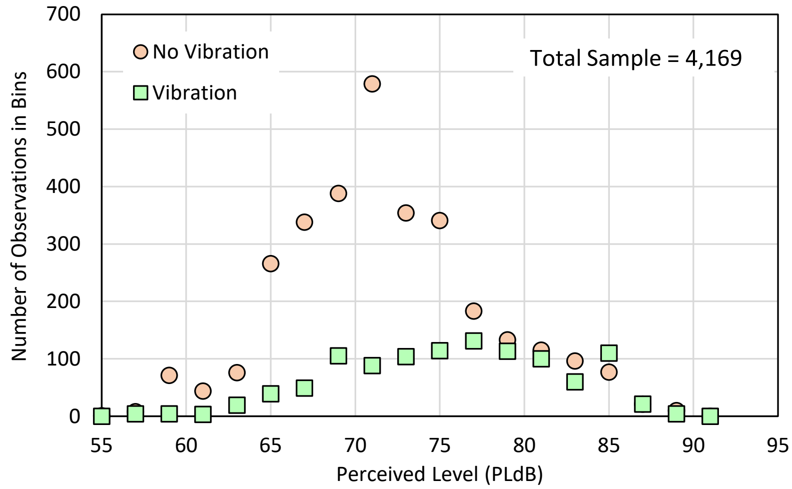

Figure 34. Sample size distribution for notice of vibration or rattle

Table 49: Association between notice of vibration or rattle and annoyance

Annoyance Level	Reporting Vibration	
	No	Yes
Not at all	81.42%	18.58%
Slightly or more	23.46%	76.54%
Moderately or more	8.26%	91.74%
Highly	4.35%	95.65%

Figure 35 plots the two percentage columns of Table 51 as different colored bars. The figure emphasizes the strong association between annoyance and vibration and rattle, especially at moderate or greater degrees of annoyance. The plot raises the basic question of the how the noticeability of rattle of vibration relates to noise dose. Figure 36 illustrates the observed relationship. At 65 PLdB only approximately 10% of panelists noticed vibration or rattle, at 75 PLdB about 30% did and at 85 PLdB 50% did.

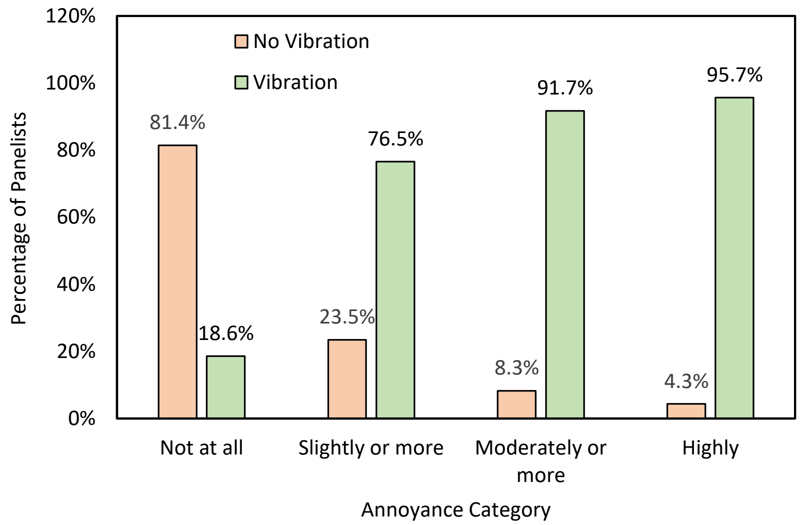

Figure 35. Illustration of strong association between notice of vibration or rattle and severity of annoyance judgments

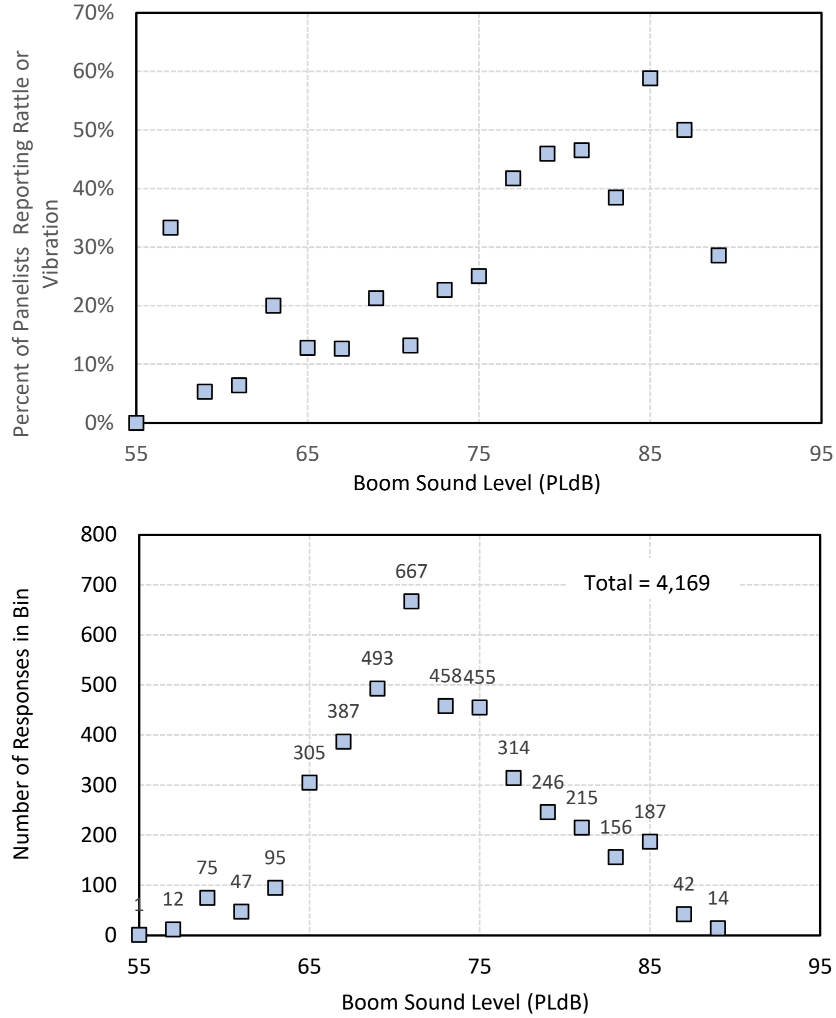

Figure 36. Percentage of panelists noticing rattle or vibration as a function of perceived level

6.5 Range of observed response latencies

Table 50 summarizes the results of the response latency analysis. Although response latency was not a criterion for eligibility for the current analyses, several observations about response latencies may be of interest in X-59 mission planning.

The first column in Table 50 tabulates a number of response latency criteria (in hours). Since no observed latency was greater than five days, a value of 144 hours (6 days x 24 hours) was chosen for the first entry in the table. The second column of Table 50 shows numbers of eligible annoyance judgments with response latencies equal to or less than certain values. The third column shows this number as a percentage of the total number of eligible annoyance judgments (4,148).

The last two columns of Table 50 show numbers of panelists in each latency interval. Column five shows the number of panelists remaining in the pool (*i.e.*, providing one or more annoyance judgments) if a latency criterion had been considered in decisions about eligibility for analysis. The sixth column reports this number as a percentage of the total (364).

Table 50: Numbers of otherwise analysis-eligible annoyance judgments remaining after applying limiting response latency criteria

Maximum Latency (hours)	Number of Eligible Annoyance Judgments	Percentage of Eligible Annoyance Judgments	Number of Panelists	Percentage of Panelists
144	4,168	100.0%	364	100.0%
96	4,167	100.0%	364	100.0%
72	4,163	99.9%	364	100.0%
48	4,161	99.8%	364	100.0%
24	4,130	99.1%	364	100.0%
12	4,058	97.4%	361	99.2%
6	3,939	94.5%	358	98.4%
3	3,774	90.5%	352	96.7%
2.5	3,719	89.2%	352	96.7%
2	3,614	86.7%	349	95.9%
1.5	3,494	83.8%	349	95.9%
1	3,282	78.7%	344	94.5%
0.75	3,095	74.3%	339	93.1%
0.5	2,741	65.8%	325	89.3%
0.333	2,261	54.2%	303	83.2%

Figure 37 shows the effect on the number of analyzable records of restricting eligible records to those meeting varying latency restrictions. The figure plots the two percentage columns shown in Table 50. Note that the number of eligible annoyance judgments falls much more rapidly than the number of panelists as the latency criterion decreases. For example, if a response latency of 30 minutes were established for unambiguous attribution of an annoyance judgment to a particular boom, the number of panelists contributing data would decrease by 11% from the unconstrained percentage, while the numbers of annoyance judgments eligible for analysis would decrease by 34%.

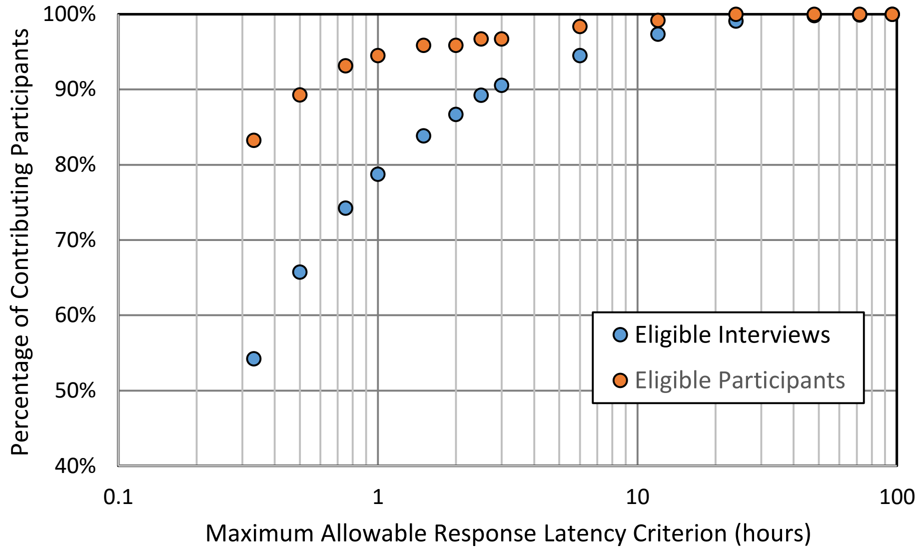

Figure 37. Percentages of annoyance judgments and panelists meeting a range of response latency criteria

The range of response latencies was partitioned into three intervals for this analysis: 0-30 minutes, 30-180 minutes and 180-4,320 minutes (3 days). Separate dose-response plots for these groups are shown in Figure 38, while Figure 39 shows the numbers of observations in each of the 2 PLdB-wide sound level bins). The 0-30 minute latency data points (in the upper panel of Figure 38) trend above those in the longer latency categories. Alternatively, the 0-30 minute data points lie approximately 5 PLdB to the left of the others. Thus, longer latencies are associated with low annoyance prevalence rates. Alternatively, the same annoyance prevalence rate is attained at a lower sound level for the 0-30 minute latency category than for greater latencies.

6.6 Effect of Latency on Prompt Annoyance Judgments

A similar effect may be observed at the moderate or greater level of annoyance, but the scarcity of annoyance judgments to this degree obscures the trend. No such observation may be made at the high level of annoyance due to the scarcity of annoyance judgments in this degree. No inferential analyses were performed on these data. dose-response relationships as a function of response latency may be found in the CTL discussion of §7.2.3.

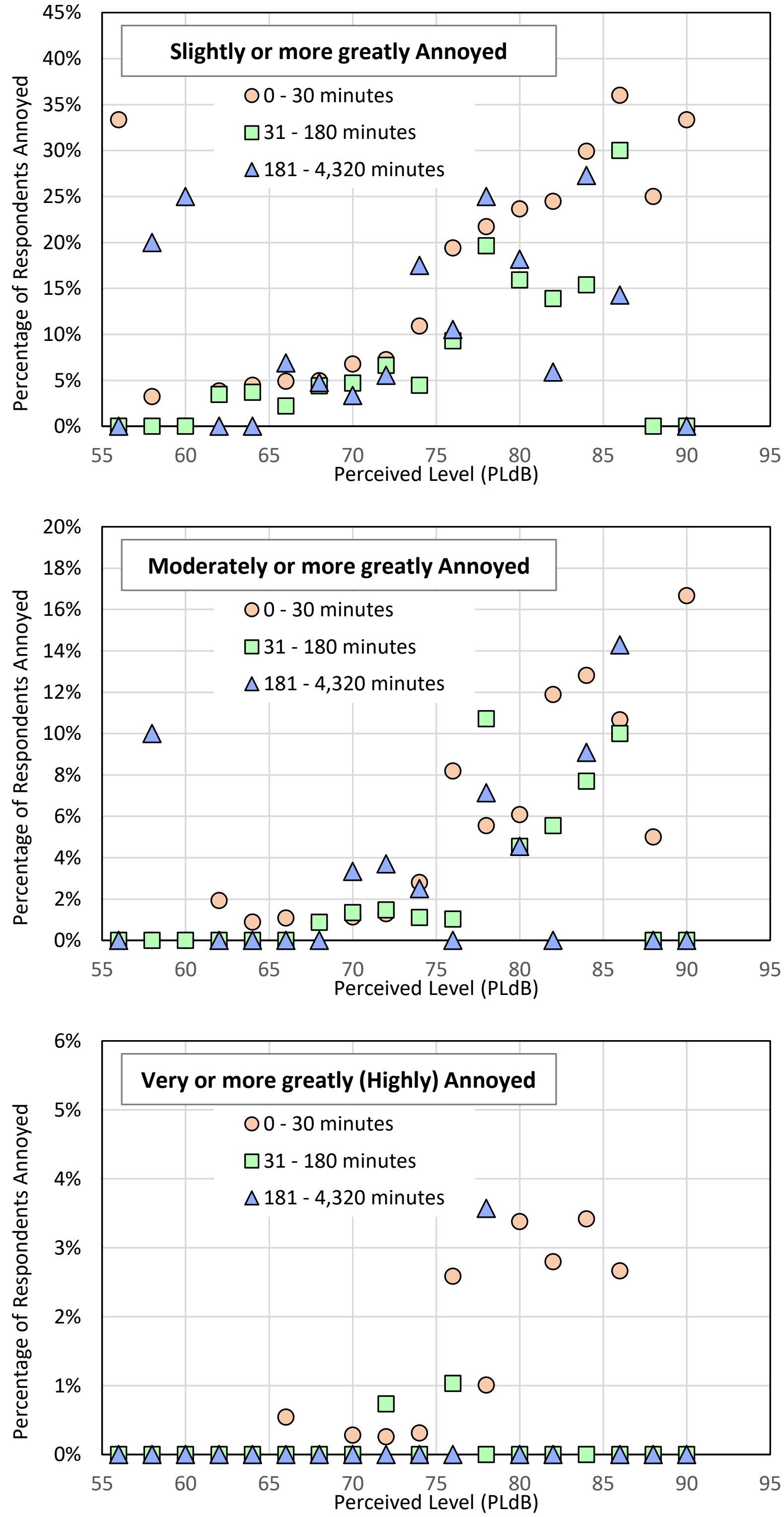

Figure 38. Annoyance judgments as functions of prompt judgment latencies

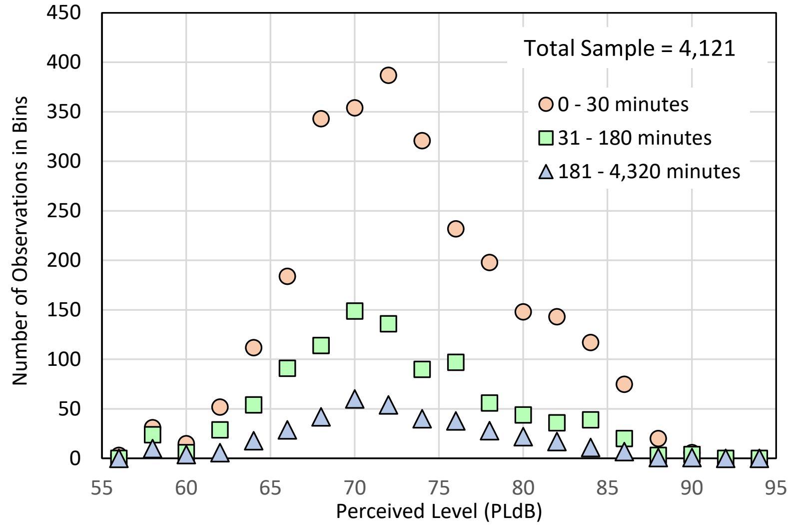

Figure 39. Distribution of sound levels for three classes of prompt response latencies

6.7 Modeling of Effects of Boom Intensity and Reminder Type on Response Latency

A multi-level multiple regression model was constructed to evaluate differences in the time it took to lodge a self-report of annoyance following each boom, adjusted for the Perceived Level of booms. The dependent (predicted) variable of the regression analysis was response latency (in minutes), and the independent (predictor) variables were (1) whether or not reminder messages were sent, and (2) differences between e-mail and text reminder messages for panelists to whom reminder messages were sent. Noise exposure was treated as a covariate. The distribution of latencies was highly positively skewed (z = 145.84), and had excessive kurtosis (z = 247.50), as may be seen in Figure 58 on page 122. Skewness and kurtosis were controlled in multiple regression analyses by truncating their ranges (*i.e.*, re-coding all latencies ≥ 40 minutes to 40 minutes.)

The original e-mail and text message groups receiving no reminders (groups 1 and 3) were once again combined into a single no-reminder group, and the original e-mail and text message groups receiving reminders (groups 2 and 4) were also combined into a single reminder group. This provided for the formation of a single binary variable, *ReminderNone*. In addition, the original two reminder categories (e-mail and text) were contrasted by forming another variable, *EmailText*, in which the no-reminder case records were ignored. Table 51 shows truncated latencies mean values (in minutes) for the reminder message categories.

Table 51: Truncated means and standard deviations for reminder categories

Delivery Method	Statistic	Reminder	No Reminder
E-mail	Mean	28.39	16.70
Text		20.74	18.97
E-mail	Standard Deviation	13.40	16.03
Text		11.04	14.34

Table 52 shows results of the multi-level regression model for predicting annoyance judgment latency. The predictive equation is shown in Equation 7 below.

$$y = \beta_{00} + (\beta_{10})(PL) + (\beta_{20})(ReminderNone) + (\beta_{30})(EmailText) \quad \textbf{Equation 7}$$

where:

y = latency in minutes between boom and annoyance response

β_{00}= intercept for latency

ReminderNone is the code for whether a reminder was sent: 0=No reminder, 1=Reminder;

EmailText is the code for type of reminder (-0.5 for text reminders, 0 for no reminder message, and +0.5 for an e—mail reminder message)

Table 52: Results of a multi-level regression model of effects of reminder messages on latency of annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value
Intercept	34.134	2.401	14.22	353	<.001
NoiseDose	-0.180	0.033	-5.49	4,165	<.001
ReminderNone	2.482	0.769	3.23	4,165	.002
EmailText	3.456	0.528	6.55	4,165	<.001

Annoyance judgment latency significantly decreased with noise level: the greater the level of the boom, the more promptly annoyance judgments were rendered. Events followed by reminder messages were associated with greater annoyance judgment latencies. E-mail reminders were associated with greater annoyance judgment latencies than text reminders. These findings of inferential analyses are further clarified in Figure 59 on page 122. While the combined reminder message group exhibited greater response latencies, the text message group responded just as promptly as the no-reminder group. The e-mail group exhibited much greater response latencies than either the no-reminder or text reminder groups.

Some of these findings may be attributable to the timing and/or notice of reminder messages, rather than to *bona fide* differences attributable to exposure levels *per se*. For example, it might be that panelists had accustomed to notification by text message, or simply have checked their smartphones for text messages more often than they checked for e-mail for messages. As discussed further in §9.1 and §9.2, the issue lacks a simple or definitive explanation two years after the QSF data collection ended.

6.8 Treatment of Missing Data

Missing annoyance judgments of individual exposures may be analyzed in at least two ways: (1) by assuming that no response implies a boom was not noticed (and therefore was logically not at all annoying), and (2) by assuming that the absence of an annoyance judgment justifies imputation of a judgment from a panelist's mean annoyance with other booms of similar amplitude.

No events remained after applying criteria for retention of case records (as detailed in Chapter 3) that failed to elicit a response to any of the questionnaire items. Therefore, no analysis was undertaken of differences between assumptions about those events, nor of the effect of those assumptions on dose-response relationships. Likewise, no analyses were undertaken of any relationship between missing data and panelist attrition.

6.9 Relationships between Reminder Messages and Attrition

Not all panelists provided prompt annoyance judgments for each impulsive exposure, nor on each of the nine study days. Further, the patterns of participation in the study showed little consistency across panelists. Some panelists responded more often at the start of the study, some responded more often later in the study, and others responded sporadically throughout the course of the study period.

Participation attrition rates were therefore evaluated by determining, for each panelist, the number of days of participation, whenever such days occurred during the test period. This practice defines panelist attrition as persistence in study, with larger number of days reflecting greater persistence. Although issuance of reminders after individual events depended on the reminder group assignment (groups 1 and 3 received no reminders; groups 2 and 4 did), all respondents received reminders each morning and evening of participation.

Table 53 dichotomizes the entire panelist group by (1) those who received only the morning and evening reminders and (2) those who received both the morning and evening and the individual event reminders. For each subgroup the number of panelists comprising the group is shown in the second column. The average number of days of participation is shown in the third column, followed by the standard deviation about the average and the lower and upper 95% confidence interval bound on the average. Similar numbers of days of participation may be observed for both groups.

Table 53: Descriptive statistics for comparison of persistence (number of days of study participation) by delivery of reminder messages

Parameter	Number of panelists	Average number of days	Std. Deviation	95% CI for Mean	
				Lower	Upper
Daily reminders only	199	5.14	2.56	4.78	5.50
Daily and single event reminders	187	5.00	2.78	4.60	5.41

Table 54 dichotomizes the daily and single event reminder group in Table 53 by the type of single event reminder message they received. The form of two tables is identical. Again, similar numbers of days of participation may be observed for both groups.

Table 54: Descriptive statistics for comparison of persistence (number of days of study participation) by *form* of reminder messages

Parameter	Number of panelists	Average number of days	Std. Deviation	95% CI for Mean	
				Lower	Upper
E-mail	100	5.20	2.82	4.54	5.76
Text	87	4.78	2.74	4.20	5.37

Planned comparisons implemented through analysis of variance evaluated the significance of effects of reminder messages on persistence. The two planned comparisons were (1) multiple reminder message *vs*. two reminder messages, and (2) for those with multiple reminder messages, e-mail *vs.* text.

Table 55 and Table 56 address the significance of the differences in average values reported in Table 53 and Table 54, respectively. Table 55 and Table 56 are in the conventional format for reporting analyses of variance. The first row of each table shows the source of the variance tested, the number of degrees of freedom it used in the test, the F-statistic (similar to the square of the number of standard deviations the effect was found to lie from the mean), and the *p* value which indicates the likelihood of observing a value of F for a given number of degrees of freedom. In both cases, the F-statistic is too small to indicate a significant difference from the mean, and the *p*-statistic is much greater than .001, thus confirming the visual observations made in Table 53 and Table 54.

Table 55: Planned comparison of effects of reminder messages on persistence

Source of Variance	df	F	*p*
Daily reminders only *vs.* Daily and single event reminders	1	0.248	.62
Error	384		

Table 56: Planned comparison of e-mail *vs.* text reminders on persistence

Source of Variance	df	F	*p*
E-mail *vs.* text message reminder types	1	1.05	.66
Error	186		

No statistically significant relationships were observed between persistence and number or type of reminder messages. In other words, no evidence was found for any relationship between persistence (panelist attrition) and reminder messages.

6.10 Potential Effect of Inter-Boom Intervals on Prompt Annoyance Judgments

"Prompt" self-reports of annoyance judgments were made with latencies greater than 30 minutes for more than a third of analyzable single events.[23] Many of these latencies are several times greater than the duration of the intervals between successive booms. The longer response latencies raise questions about how reliably panelists were able to associate individual annoyance judgments with specific booms. While the experimental protocol prompted participants to supply the approximate time of the boom for each annoyance judgment, no *post hoc* means were available to confirm how reliably they were able to do so, especially when one or more booms intervened between a given boom and the time recorded for the annoyance judgment.

Table 57 lists each boom during the nine-day test period and its time of occurrence. It also lists the elapsed time interval from the previous boom. Hatched cells are all associated with the first boom of the day, for which there were no preceding ones.

[23] Response latencies (some greater than 24 hours) were calculated using data set variables as ReportStart minus the average of BOOM_START_TIME and BOOM_END_TIME. The BOOM_START_TIME and BOOM_END_TIME typically differed by 1 minute so there is an inherent uncertainty of 0.5 minute in this calculation.

Table 57: Boom schedules and inter-boom intervals

Boom ID	Test Day	Boom Start (hh:mm)	Inter-Boom Interval (min)	Boom ID	Test Day	Boom Start (hh:mm)	Inter-Boom Interval (min)
1	1	8:57		29	6	11:17	
2	1	9:22	25	30	6	11:38	21
3	1	9:44	22	31	6	13:03	85
4	1	13:18	214	32	6	13:32	29
5	1	13:47	29	33	6	14:57	85
6	1	15:04	77	34	6	15:27	30
7	1	15:35	31	35	7	15:05	
8	2	10:57		36	7	15:30	25
9	2	11:22	25	37	7	15:56	26
10	2	11:47	25	38	7	16:37	41
11	2	13:57	130	39	7	16:56	19
12	2	14:22	25	40	8	8:58	
13	2	14:48	26	41	8	9:23	25
14	3	8:20		42	8	9:43	20
15	3	8:43	23	43	8	14:01	258
16	3	9:04	21	44	8	14:27	26
17	3	11:58	174	45	8	14:58	31
18	4	11:01		46	8	15:22	24
19	4	11:16	15	47	8	15:47	25
20	4	12:56	100	48	9	8:57	
21	4	13:18	22	49	9	9:22	25
22	5	9:57		50	9	9:47	25
23	5	10:32	35	51	9	10:57	70
24	5	11:58	86	52	9	11:27	30
25	5	12:22	24				
26	5	12:48	26				
27	5	15:58	190				
28	5	16:27	29				

Figure 40 plots the distribution of the inter-boom intervals tabulated in Table 57. The orange-shaded, rightmost bar includes all intervals greater than 180 minutes, of which the largest was 258 minutes (4 hours, 18 minutes). The large number of annoyance judgments made during times from 20 and 30 minutes after a boom underscores the importance of accurately linking the timings of booms and annoyance judgments. The large numbers of booms generated at short intervals required panelists to separately recall the annoyance of more than one boom if their annoyance judgments were delayed until after the occurrence of intervening booms. Their abilities to perform this task have implications for dose-response analyses.

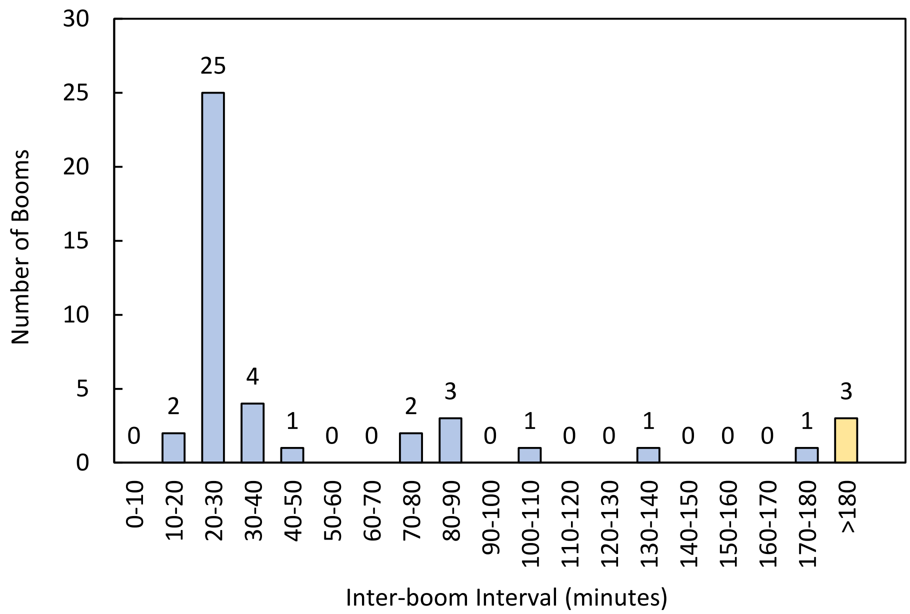

Figure 40. Distribution of inter-boom intervals

Associating an incorrect noise dose with a particular annoyance judgment adds both noise dose and annoyance uncertainty to dose-response relationships. This reduces the precision with which degrees of annoyance may be associated with noise doses. It also lowers the slope of the dose-response relationship, so that estimates of increased annoyance from increased noise dose are underestimated.

6.11 Small Percentages of High Annoyance Judgments

The small percentages of QSF18 panelists who described themselves as highly annoyed by impulsive exposure presents several problems of interpretation of findings:

- The dearth of highly annoyed judgments renders the findings vulnerable to potential methodological criticisms. The most comprehensive defense of the data collection methods would be a demonstration that panelists experiencing higher boom levels would in fact describe themselves as highly annoyed by them.

- The dearth of highly annoyed judgments introduces the need for regulatory analyses that are not directly comparable to those by which FAA and ICAO currently judge the significance of noise exposure.
- The dearth of highly annoyed judgments complicates interpretation of inter-community and inter-noise source differences in tolerance for exposure to sonic booms. Constructing dose-response relationships for degrees of annoyance other than high annoyance also raises questions about whether the slopes of dose-response functions are constant for all degrees of judged annoyance.

6.12 Differential Annoyance of Panelists in Different Locations

Some of the current findings suggest that panelists were more greatly annoyed by low amplitude sonic booms experienced while at home than by booms of comparable level experienced while at work or elsewhere. For judgments of slight or greater annoyance, for example, it appears that panelists were more tolerant of exposure occurring somewhere other than at work or at home (*cf.* Figure 15 on page 57). The evidence is not robust, but if the finding can be further confirmed, it would suggest that the responses of panelists who are at home at the time of exposure can serve as a worst-case estimate of community response to sonic booms.

6.13 Comparison of dose-response functions with CHABA findings

Three pre-requisites for a direct comparison of the QSF18 findings about the annoyance of low-amplitude sonic booms with prior findings about community response to high energy impulsive sounds, as summarized by CHABA (Fidell, 1996), are

- A common definition of impulsive exposure;
- A common metric of exposure; and
- A common metric of annoyance

None of these conditions is met in the present case. First, only a small number of the low-amplitude booms created during the QSF18 data collection exercise meet the CHABA definition of a high energy impulse as one with a CSEL value in excess of 85 dB. Second, the bulk of the analyses reported by Page *et al.* (2020) were conducted in units of Perceived Level, not in C-weighted units (conversion of PL-weighted to C-weighted units is uncertain in the Galveston measurements, as shown in Figure 70 on page 144.) Third, very few annoyance judgments reported by Page *et al.* (2020) were in the "very" or "extremely" annoyed categories.

Under these conditions, no useful comparison of the CHABA and QSF18 findings can be made.

7 CTL-BASED DOSE-RESPONSE RELATIONSHIPS

The dose-response relationships developed in the preceding chapters were all derived by generic correlational techniques as two parameter (slope/intercept) fits of logistic functions to field observations of annoyance prevalence rates. An alternate method for developing dose-response relationships described in this chapter (and in greater detail in Chapter 14) takes a different approach. Whereas logistic regression fits a two-parameter curve to a data set, Community Tolerance Level ("CTL") analysis assumes a fixed slope, and uses a variable intercept to find a maximum likelihood fit (Taraldsen *et al.*, 2016) of a data set to a pre-determined annoyance growth function. In other words, rather than deriving a dose-response relationship between noise exposure and the prevalence of annoyance in communities via a curve fitting procedure without an intrinsic causal rationale, CTL analysis fits a causal model to a set of field observations.

The underlying causal assumptions of CTL analysis are 1) that dose-driven annoyance is caused by the effective (duration-adjusted) loudness of noise exposure, and 2) that any observed deviations from the assumed rate of growth of loudness with exposure are attributable to the aggregate influences on annoyance judgments of non-acoustic factors.[24] Unlike logistic regression (in which the slope of the fitting function in a given community is a free parameter), CTL analysis thus assumes that the *rate* of growth of the prevalence of annoyance with sound exposure is fixed and invariant across communities. The effective loudness of exposure, in turn, is assumed (on the basis of psychoacoustic research such as Stevens, 1972) to grow as the 0.3 power of exposure. CTL analysis thus attributes all observed deviations from the assumed growth rate of annoyance in different communities to differences in the value of the intercept of the fitting function. (Different intercepts cause different ranges of the same, non-linear assumed growth rate function to be displayed.)

CTL analyses of the current data are of interest because dose-response relationships derived by univariate regression offer no causal explanation for observed differences among communities in annoyance prevalence rates. The variable slope, two-parameter fit of regression analysis also blurs the distinction between acoustic and non-acoustic determinants of annoyance.

Multi-level regression modeling that explicitly considers individual effects provides no pragmatically useful explanation for aircraft noise regulatory policy purposes, because no plausible regulatory policies can take individual differences in sensitivities to noise exposure into consideration. Regression analysis offers no useful explanation of correlations between predictor and predicted variables beyond "that's the way the numbers turned out." Indeed, because correlation is symmetric, regression analysis is even indifferent between which variable (exposure or prevalence of annoyance) is considered the predictor and which is considered the predicted.

---

[24] Note that it is only the *aggregate* effects of all non-acoustic influences that are of interest in CTL analysis. Individual non-acoustic influences on annoyance judgments play no role in aviation noise regulatory policy.

The advantages for present purposes of CTL's assumption of a common slope for dose-response relationships in different communities thus include:

(1) that it yields a more parsimonious, single parameter dose-response function than that provided by regression analysis;

(2) that it makes a clear and simple distinction between acoustic and non-acoustic influences on community response to noise exposure;

(3) that it extracts more useful information than regression analysis about the position of an entire dose-response relationship from single data points;

(4) that it can more simply accommodate a non-zero origin for dose-response curves than logistic regression; and

(5) as a matter of practical regulatory utility, CTL analysis provides a simple, direct, and decibel-denominated measure of differences among communities and noise sources in community-level tolerance of noise exposure.

7.1 CTL Dose-response Functions for Delayed Responses

Figure 41 presents CTL analyses of delayed annoyance judgments for the three degrees of annoyance judgments analyzed in other sections of this report. The 95% confidence intervals show the lateral (not ordinal) uncertainty in the position of the curve. The data have been binned for both analysis[25] and presentation purposes.[26] Figure 42 shows numbers of annoyance judgments in each bin. Although the numbers of annoyance judgments in Figure 42 are sparse above Perceived Levels of 40 PLdB, the maximum likelihood method used to estimate CTL values limits the influence of small numbers of observations on the lateral offset of CTL functions for all three degrees of annoyance judgments.

---

[25] The CTL procedure can act on either binned data or individual responses (consisting of ones and zero for annoyed not annoyed). Two decibel wide dose bins were chosen for these illustrations for the sake of consistency with graphic presentations elsewhere in this report.

[26] See discussion in §3.5.

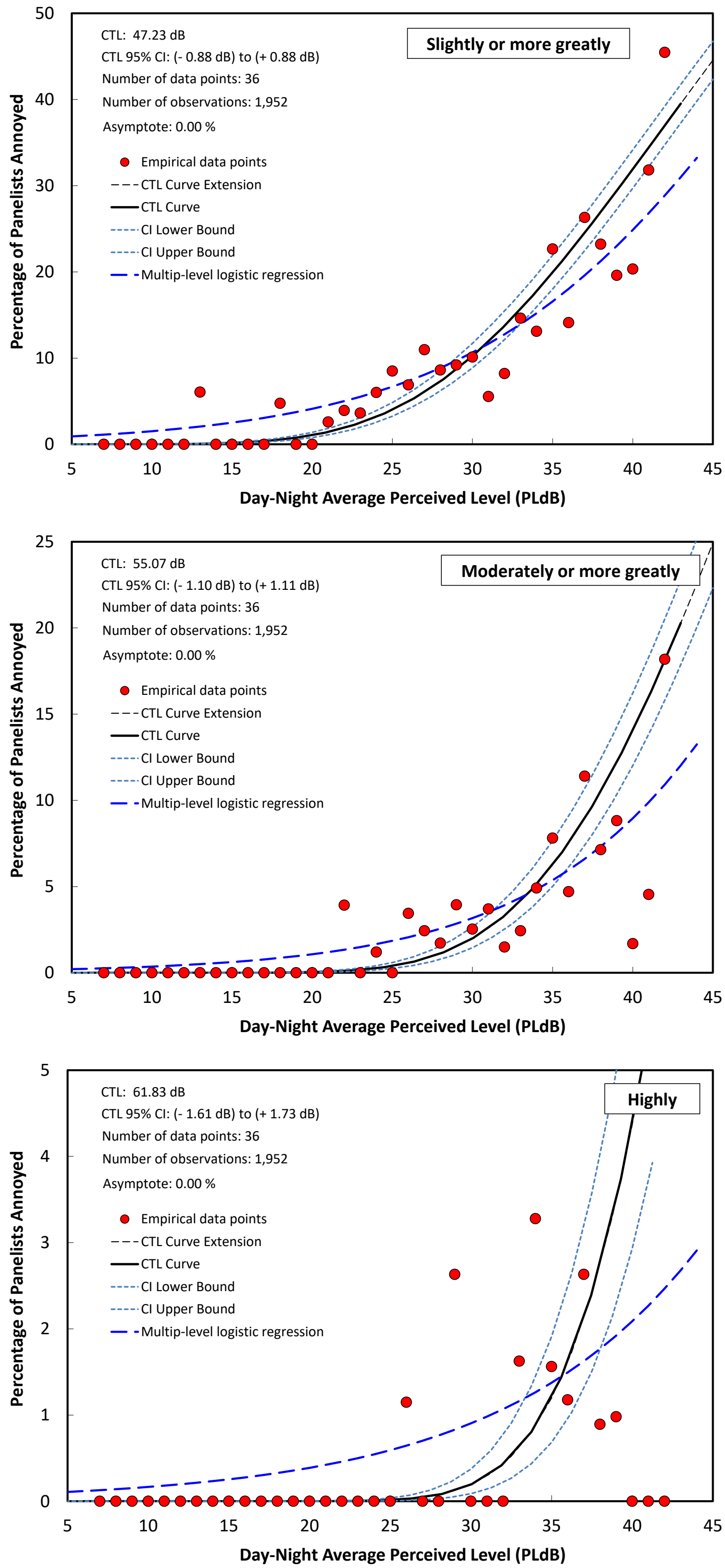

Figure 41. CTL-based dose-response curves of delayed judgments for three degrees of annoyance

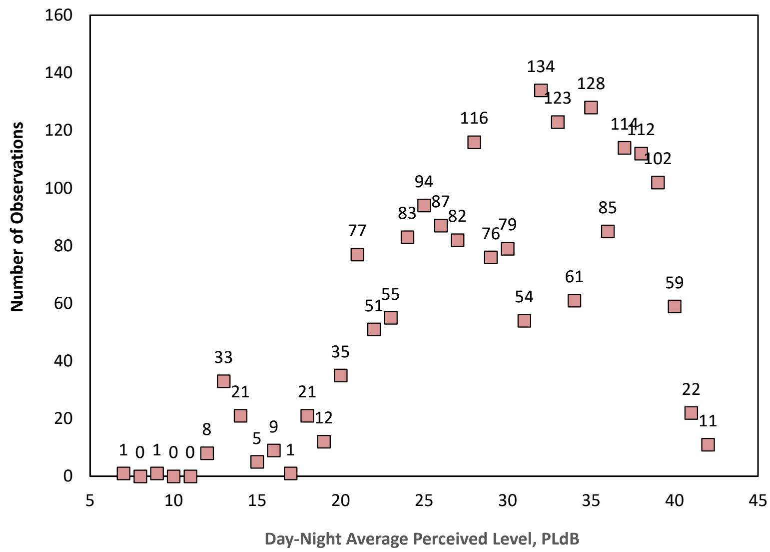

Figure 42. Numbers of delayed response annoyance judgments by sound exposure level

7.2 CTL Dose-response Functions for Prompt Responses

This section applies CTL analysis to the prompt response data set.

7.2.1 *Dose-response functions for all individual event annoyance judgments*

Figure 43, Figure 44, and Figure 45 present CTL analyses applied to the prompt response data set as a whole for the three degrees of annoyance previously analyzed in this report. Within both Figure 43 and Figure 44, the top two panels show CTL-derived dose-response functions based on two different assumptions about the lower asymptote of the data. The bottom panels show the number of annoyance judgments tabulated in 2 PLdB wide bins. Within each bin the percentages of panelists annoyed were computed. Each of the two degrees of annoyance is discussed separately in the paragraphs below.

The upper panel of Figure 43 shows the CTL curves in solid red for all analyzable prompt response judgments of slight or greater annoyance. The curve fit makes the assumption that exposure-driven annoyance begins at 0% annoyed. In other words, it assumes that *all* annoyance with sonic booms is controlled exclusively by its exposure level, no matter how low. The field observations themselves do not fully support this assumption, however, since annoyance prevalence rates of approximately 5% in the 60 to 70 PLdB range are elevated above zero, and do not increase systematically with exposure level. The dashed blue line is the logistic regression fit for slight or greater annoyance judgments, as shown in Figure 6 on page 38.

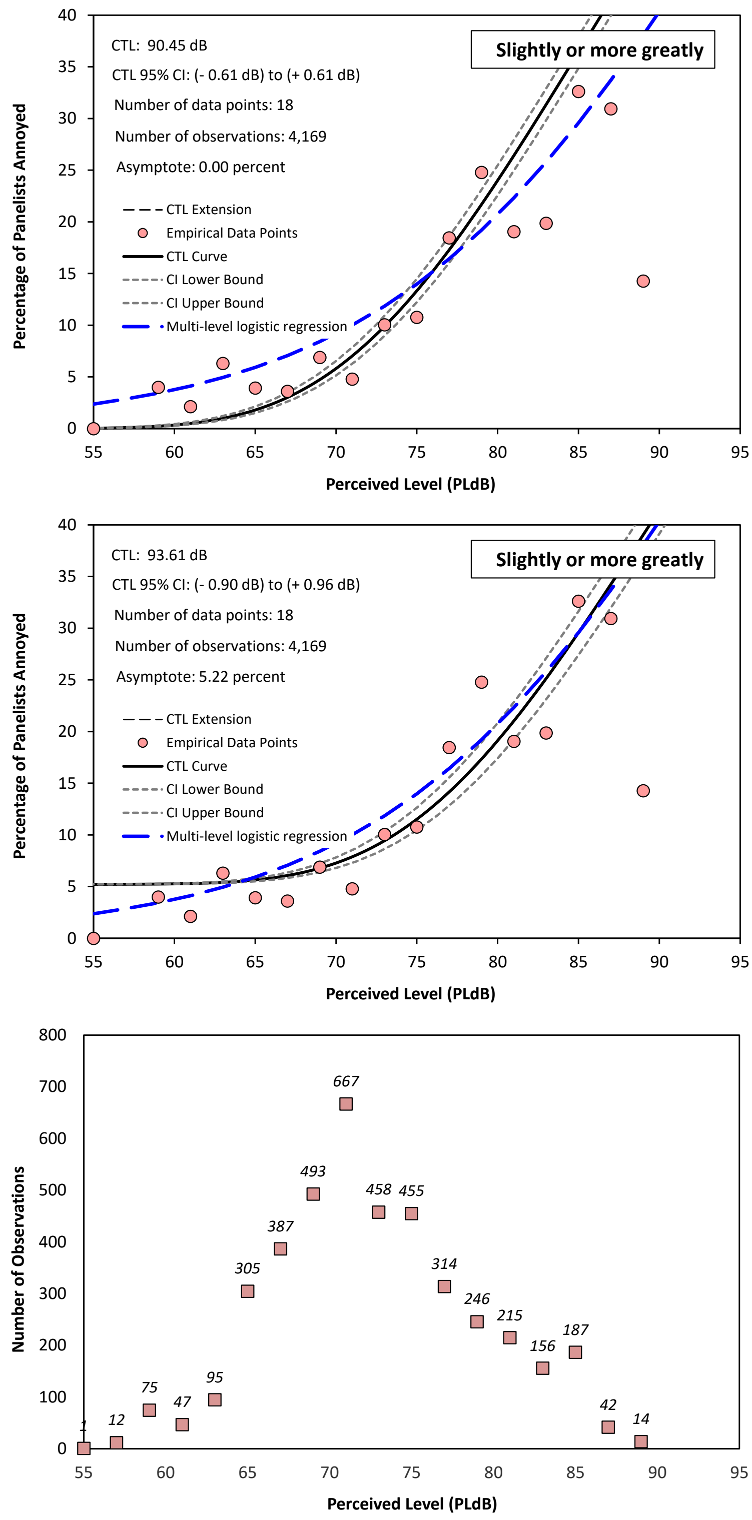

Figure 43. CTL-based dose-response curves for all slight or greater annoyance judgments, with and without a non-zero asymptote

The middle panel of Figure 43 does not make the assumption that exposure-driven annoyance begins at 0% annoyed at low exposure levels. Instead, it assumes that a fraction of the panelists reported slight or greater annoyance with low-amplitude booms independently of their sound levels. (The remainder of the panelists are assumed to base their annoyance judgments on noise dose information.) A lower asymptote for the dose-response curve of 4.1% in this case maximizes the likelihood function[27] that generates the best fit curve. The logistic curve from the top panel is reproduced for comparative purposes. The bottom panel of Figure 43 displays the numbers of observations in each sound level bin.

Figure 44 presents a similar analysis of all analyzable single event judgments of moderate or greater annoyance. Like the top panel of Figure 45, the analysis assumes the curve begins at 0% for low noise doses. The CTL relationship shown in the middle panel assumes an asymptote of 1.1%, which provides the maximum likelihood fit to the annoyance judgments.

Figure 45 shows the CTL curve for all analyzable single event judgments of high (very or greater) annoyance judgments. These data exhibit no evidence of a non-zero asymptote.

7.2.2 *CTL analysis of tolerance for exposure by panelist location*

Figure 46, Figure 47, and Figure 48 plot dose-response relationships between at-home and at-work panelists. Individual annoyance judgment pairs were assigned to 2 PLdB wide bins and the percentage annoyed values computed in each bin. At-home data points are shown as red circles, while at-work points are shown as blue squares. The text block in the figures tabulates the results of the CTL analyses separately for both the at-home and at-work panelists. The solid curves show the positions of the two CTL curves. The dashed curves show the respective confidence intervals about the maximum likelihood estimates of CTL values.

The confidence intervals represent the uncertainty in the lateral positions of the CTL curves, not the uncertainty in predicted percentages of annoyed panelists. Ten or fewer panelists were contained in each exposure bin at Perceived Level values of 64 PLdB and lower. Hence, a difference of a single panelist's annoyance judgment produces a change of 10% or more in a bin's calculated percentage. Since the maximum likelihood fitting process ascribes very little weight to these data points, they do not materially affect the lateral position of the CTL curve.

---

[27] The maximum likelihood procedure iteratively varies the lateral position of CTL curve until the joint probability of all the data points arising out of the curve position is maximized. At each trial curve position the binomial probability, P, of observing exactly k annoyed judgments out of n observations given a value of p determined by the present curve position is computed for each data point:

$P(k, n, p) = nCk * p^k * (1-p)^{(n-k)}$, where $nCk = n! / k! (n-k)!$.

These probabilities are multiplied across all data points to determine the joint likelihood (probability) of the data points arising from that position of the curve. Iterations continue until the search increments become less than a specified criterion value (usually on the order of 0.005 dB).

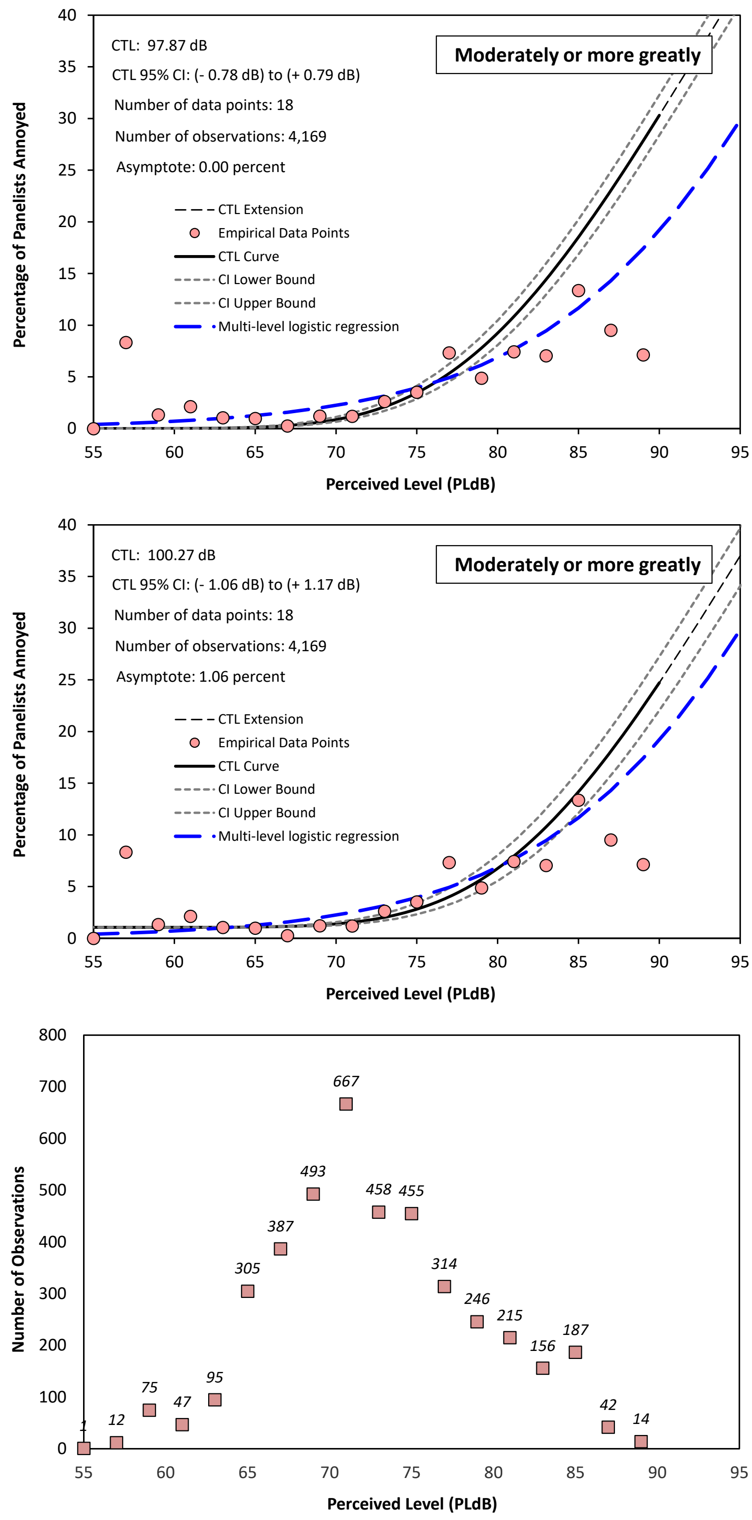

Figure 44. CTL-based dose-response curves for all single event moderate or greater annoyance judgments, with and without a non-zero asymptote

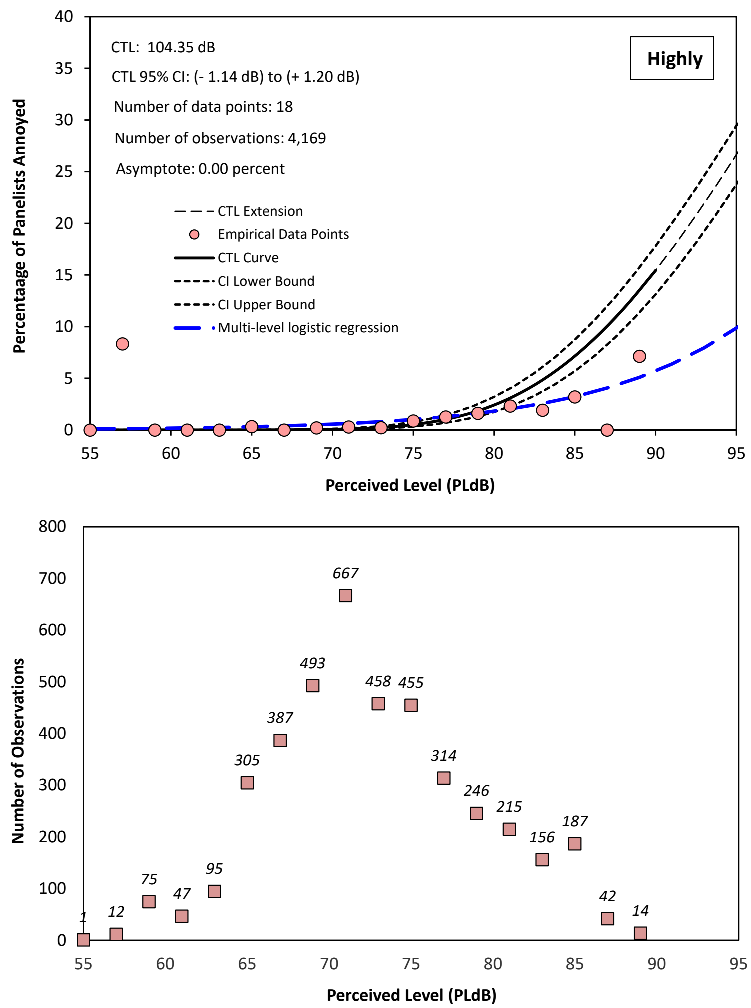

Figure 45. CTL-based dose-response curves for all single event judgments of "high" annoyance

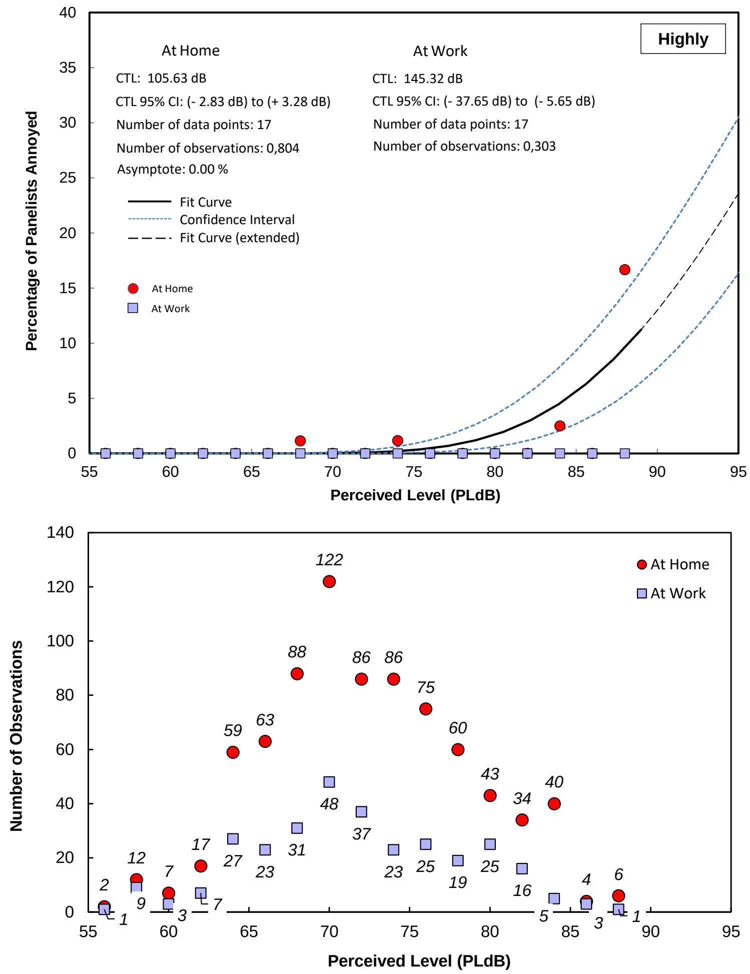

Figure 46. CTL-derived at-home and at-work dose-response relationships for highly annoyed judgments

Figure 47 presents the same type of findings as Figure 46, but for moderate or greater degrees of annoyance judgments. The set of judgments depicted in Figure 47 contains greater numbers of annoyed panelists than the highly annoyed case, allowing the lateral positions of the CTL curves to be computed with less uncertainty. (Note that the red data point at 60 PLdB just below the figure legend represents the opinion of a single panelist.) The 95% confidence interval on the difference of 5.0 PLdB was ±2.4 PLdB wide, implying a difference that is meaningfully different from zero, and is unlikely to have occurred by chance alone.

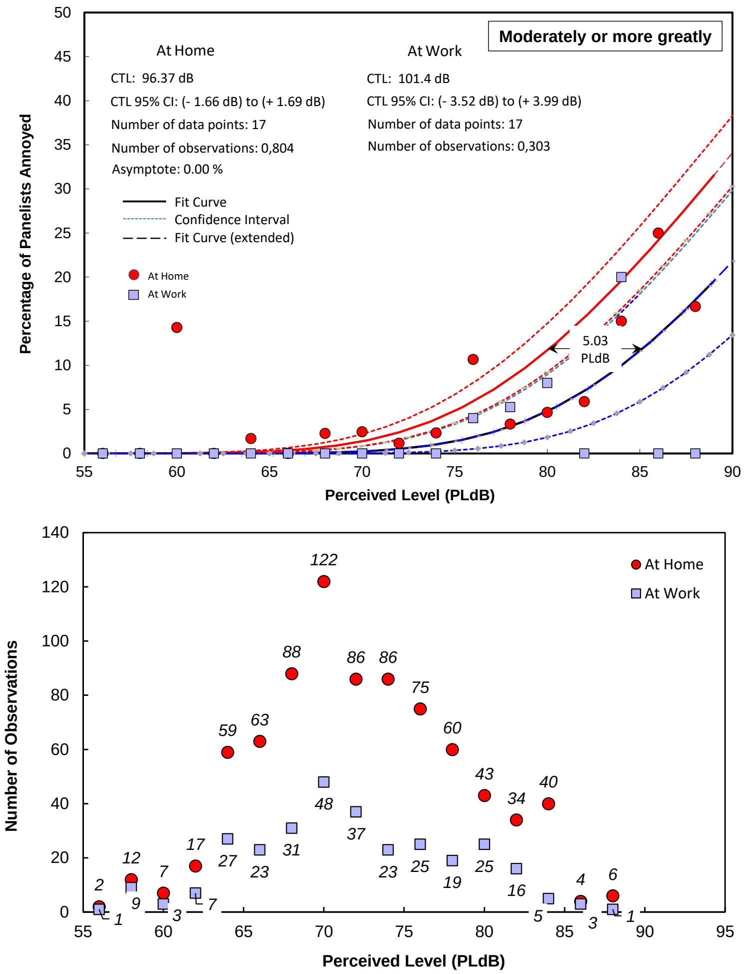

Figure 47. CTL-derived at-home and at-work dose-response relationships for moderate or greater annoyance judgments

Figure 48 presents findings for slight or greater degrees of annoyance judgments. As predictable, the larger numbers of annoyed panelists result in higher annoyance prevalence rates, which in turn allow CTL confidence intervals on the lateral positions of the curves to be computed with greater accuracy. Note that the percentage values for four data points at 62 PLdB or lower were produced by two or fewer panelists. The 95% confidence interval on the difference of 4.7 PLdB was ±2.8 PLdB wide, implying the difference is meaningfully different from zero.

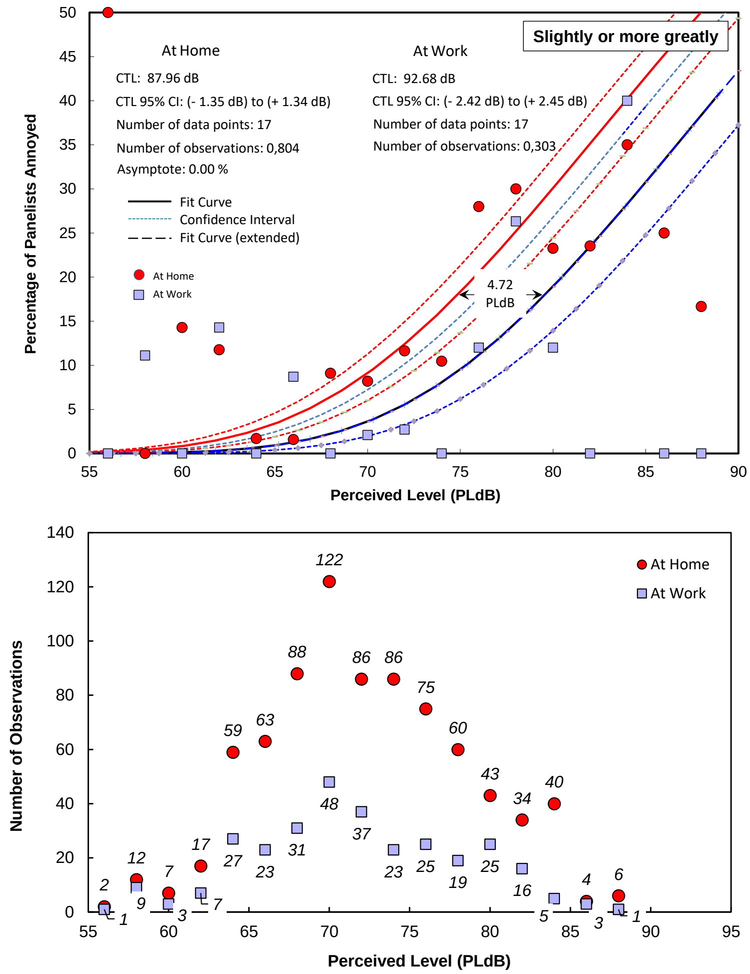

Figure 48. CTL-derived at-home and at-work dose-response relationships for slight or greater annoyance judgments

Table 58 summarizes the results of a CTL analysis of dose-response relationships for the two panelist subgroups discussed in §5.2 of the report. The table shows the findings of Figure 46, Figure 47, and Figure 48 in tabular form. The three annoyance subgroups are identified in the first column. For each of these groups, the mean CTL and its 95% confidence interval are shown separately for the at-home and at-work panelist groups. No high annoyance judgments were found among At Work records (shaded in pink) were found.

Separate vertical panels separate the at-home from the at-work findings. The numbers of panelists providing self-reported annoyance judgments are shown under each location heading. A total of 804 valid dose-response pairs were found for which it could be determined whether the panelist was at home at the time of impulsive exposure. Another 303 valid pairs were found for which it could be determined the panelist was at work. Each panel shows the computed CTL value and its 95% confidence interval for each degree of annoyance judgment, as well as the number of panelists reporting annoyance judgments of that degree on higher. The rightmost vertical panel shows the differences between the two.

Table 58: CTL estimates for At Home and At Work panelists by panelist location

Degree of Reported Annoyance	Statistic	At Home N = 804		N	At Work N = 303		N	Work - Home	
Slightly+	Mean	87.96		84	92.68		14	4.72	
	95% CI	-1.35	1.34		-2.42	2.45		-2.77	-2.79
Moderately+	Mean	96.37		28	101.4		5	5.03	
	95% CI	-1.66	1.69		-3.52	3.99		-3.9	-4.34
Highly	Mean	105.63		2			0		
	95% CI	-2.83	3.28						

The results of the at-home CTL analysis for slight or greater annoyance judgments show a PL-weighted CTL of 88 PLdB, with a 95% confidence interval of ± 1.3 PLdB. The total number of panelists who reported slight or greater annoyance was 84 (across all noise dose levels). Similarly, the CTL for at-work panelists was 92.7 PLdB, with a 95% confidence interval of ± 2.4 PLdB. The confidence interval for the at-work group was greater than for the at-home due to the smaller sample size and smaller number (only 14) of panelists reporting slight or greater annoyance. The difference between at-home and at-work PL-weighted CTL values in community tolerance levels was 4.7 PLdB, with an estimated 95% confidence interval of ± 2.8 PLdB. This difference suggests the at-work panelists were more tolerant of exposure to slightly or more greatly annoying low-amplitude sonic booms than the at-home population.

The same analysis was performed for panelists who described themselves as moderately or more greatly annoyed. The difference in tolerance levels between the at-work and at-home populations was 4.7 PLdB, very little different from the difference in tolerance for exposure to slightly or more greatly annoyed panelists.

A comparison of highly annoyed populations was not possible, since no panelist in the at-work group reported high annoyance with exposure to low-amplitude sonic booms. Taken together, however, the Slight or greater and Moderate or greater annoyance judgments suggest that at-work panelists were approximately 5 PLdB more tolerant of exposure to low-amplitude sonic booms than at-home panelists.

7.2.3 *CTL analysis by response latencies for prompt judgments*

Figure 49 shows dose-response data points and CTL curves for two prompt response latency periods: 0 to 30 minutes, and 31 to 180 minutes. (These are the same data displayed in

Figure 38.) The dashed lines are CTL fits to each data subset. The CTL curves differ by approximately 4 PLdB on the abscissa, and by 5% (at 75 PLdB) on the ordinate.

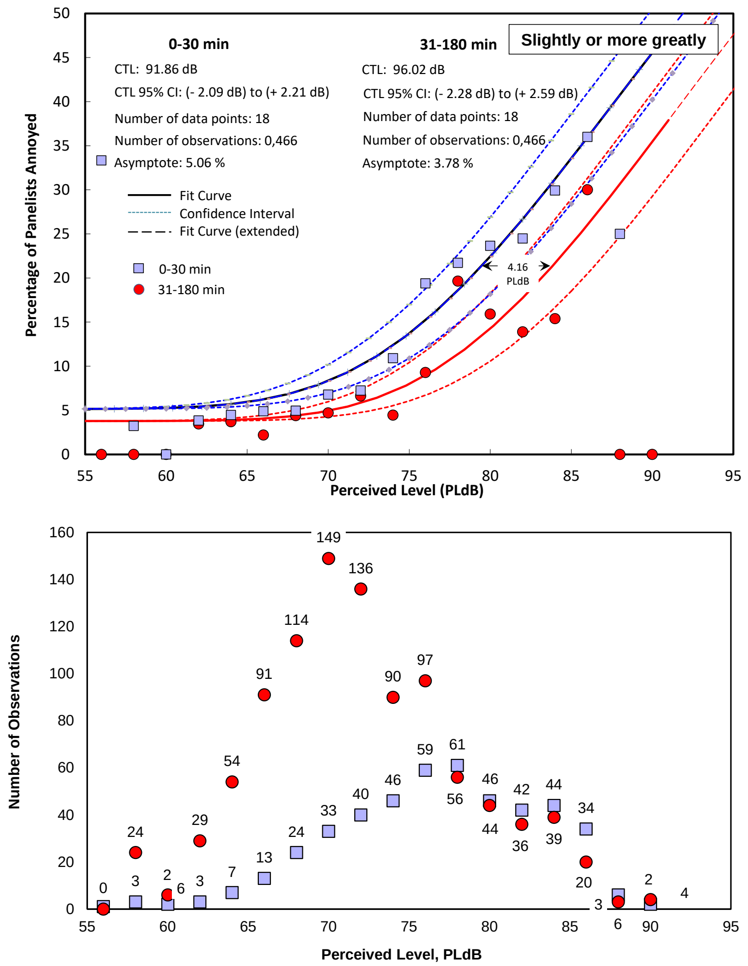

Figure 49. CTL-derived dose-response relationships for slight or greater annoyance judgments for two response latency periods

THIS PAGE INTENTIONALLY LEFT BLANK

8 REPRESENTATIVENESS OF QSF18 PANEL SAMPLE

Both the demographic and geographic representativeness of responses to the QSF18 questionnaire items were examined. Since the address-based sampling frame itself was not provided, the representativeness of the QSF18 data collection was assessed with respect to the demographic characteristics of the corpus of completed annoyance judgments with 2020 census demographic estimates[28]. This Chapter discusses 1) relationships between several demographic characteristics of the Galveston-area population and those of panelists who provided their opinions about low-amplitude sonic booms, and 2) estimated geographic locations of panelists at times of exposure.

8.1 Demographic Representativeness of QSF18 Sample

A limited amount of demographic information (age, education, gender), described in the following subsections, was collected about panelists during the panelist recruitment process. Additional information about number of people in households, numbers of children younger than six year of age in households, and numbers of people 18 years or older in households, was collected, but not analyzed.

8.1.1 *Education*

Responses in the QS18 data set over-represent the Galveston college educated population, and under-represent the Galveston population with a high school (or lower) educational level. Table 59 summarizes comparable categories of educational attainment for the Galveston population and for the panelists. Figure 50 graphically compares percentages of the Galveston population attaining various educational levels and those represented in the Galveston annoyance judgments. This graph illustrates that the educational attainment of panelists who participated in the study did not accurately reflect the Galveston population at large.

8.1.2 *Age*

Panelists whose opinions are reflected in the data set were disproportionately older than the Galveston population. Table 60 and Figure 51 show that two-thirds of the responses in the data set were those of panelists who were at least 55 years of age, whereas only 40% of Galveston residents were similarly aged. Conversely, less than 10% of the responses in the QSF18 data set were between the ages of 18 and 34 years, whereas three times more Galveston residents were of similar ages. Similarly, 22% of the responses in the QSF data set were generated by panelists aged 35 to 54 years, while almost 30% of the Galveston population is of the same age.

[28] Census data is based upon 2019 updated demographic estimates compiled under license from Claritas©.

Table 59: Summary of educational attainment of Galveston population and of panelists

2019 GALVESTON CENSUS			QSF18 PANELISTS			QSF18
Census Education Category	Population 25+ by Education Level*	Percent of 25+ Pop.	Educational Attainment Category	Frequency of Responses	% in QSF18 Responses	
High School or Less	15,529	40.76%	1 = HS or less	1,029	17.32%	-23.44%
Some College	8,261	21.68%	2 = Some College	2,078	34.98%	13.29%
Associates or Bachelor's Degree	9,383	24.63%	3 = Bachelor's Degree; 4 = Some Graduate School	1,814	30.53%	5.91%
Master's Degree or Higher	4,927	12.93%	5 = Graduate degree	1,014	17.07%	4.14%
Unknown	0	0.00%	NA = no response	6	0.10%	0.10%
Totals	**38,100**	**100.00%**		**5,941**	**100.00%**	

* Educational attainment is available only for the percent of population 25 years of age or older. It is unlikely, however, that the observed differences in the level of educational attainment might be accounted for by the seven-year age discrepancy with respect to the 18+ selection criterion of the ABS sampling.

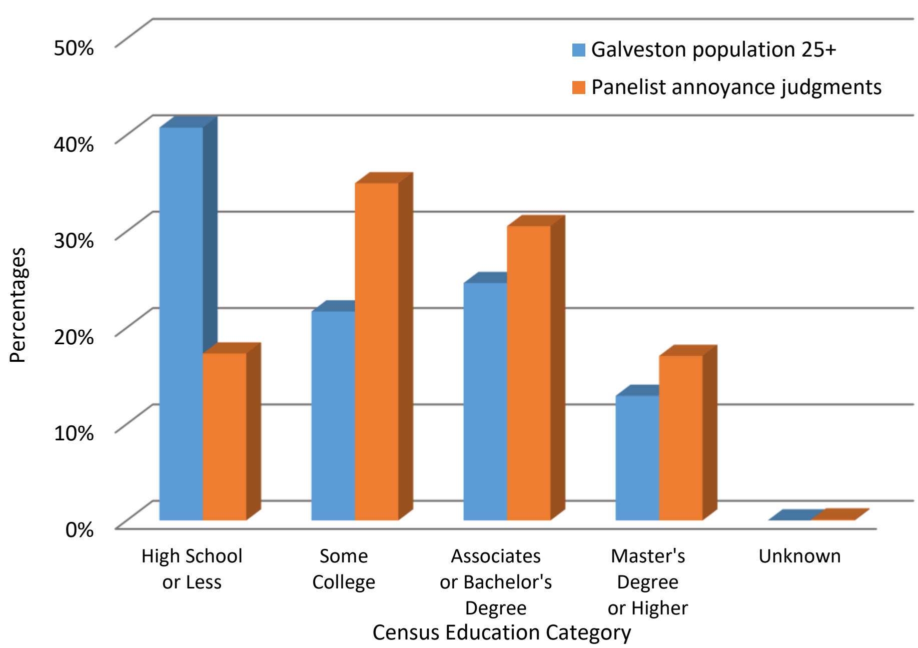

Figure 50. Comparison of educational attainment of QSF18 panelists and Galveston population

Table 60: Comparison of age distributions of QSF panelists and the 2019 age distribution of the population of the City of Galveston

Age in 2019 (years)	2019 US CENSUS		GALVESTON PANELISTS		
	Updated Galveston population*	Percent of population 18+ years of age	Number of responses in data set	Percentage of responses	Ages of panelists in census category groupings
Unknown	NA	NA	71	1.20%	NA
18 to 34	13,100	30.18%	565	9.51%	-20.67%
35 to 54	12,872	29.66%	1,326	22.32%	-7.34%
55+	17,430	40.16%	3,979	66.97%	26.81%
Totals	**43,402**	**100.00%**	**5,941**	**100.00%**	

* Source: Claritas, 2019.

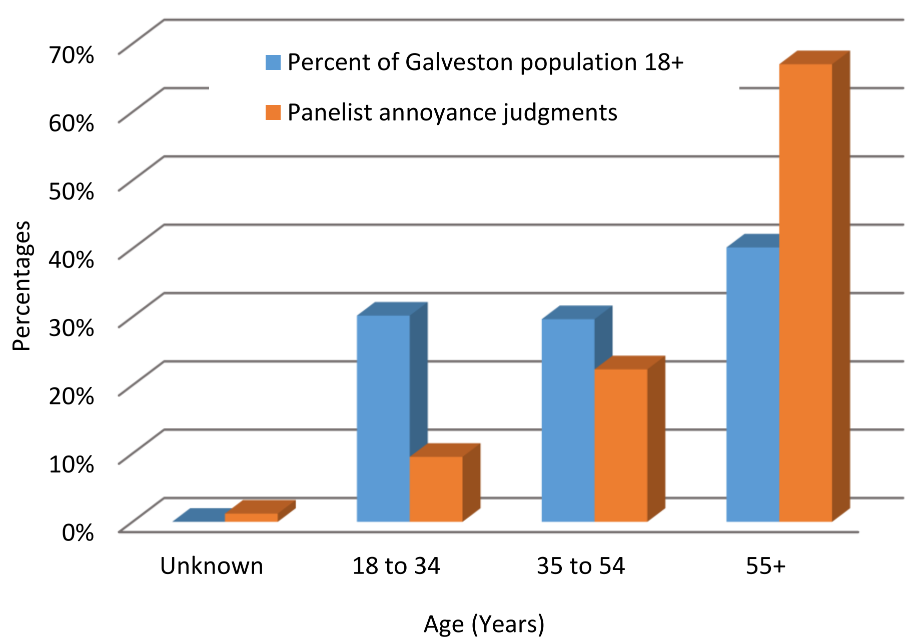

Figure 51. Comparison of ages of panelists contributing opinions to data set and Galveston population

Table 61 provides additional insight into the age demographics of the panelist and Galveston populations. Figure 52 compares the percentage of the Galveston population by age group with that of the panelists' annoyance judgments in finer gradations than in Figure 51. The table and graphic reinforce the experience of the survey industry as a whole: The population between ages 55 and 74 appears more willing (or is more readily available) to participate in social surveys in general.

Table 61: Detailed Comparison of age distributions of QSF panelists and the 2019 age distribution of the population of the City of Galveston

Age	Percent Galveston Population 18 Plus	Percentage of Panelist Annoyance Judgments	Age	Percent Galveston Population 18 Plus	Percentage of Panelist Annoyance Judgments
NA	0.00%	1.20%	55-64	18.23%	33.43%
18-20	5.60%	0.05%	65-74	13.48%	28.92%
21-24	6.62%	0.82%	75-84	6.14%	3.94%
25-34	17.97%	8.63%	85+	2.30%	0.69%
35-44	15.42%	7.88%	**Total**	**100.00%**	**100.00%**
45-54	14.24%	14.44%			

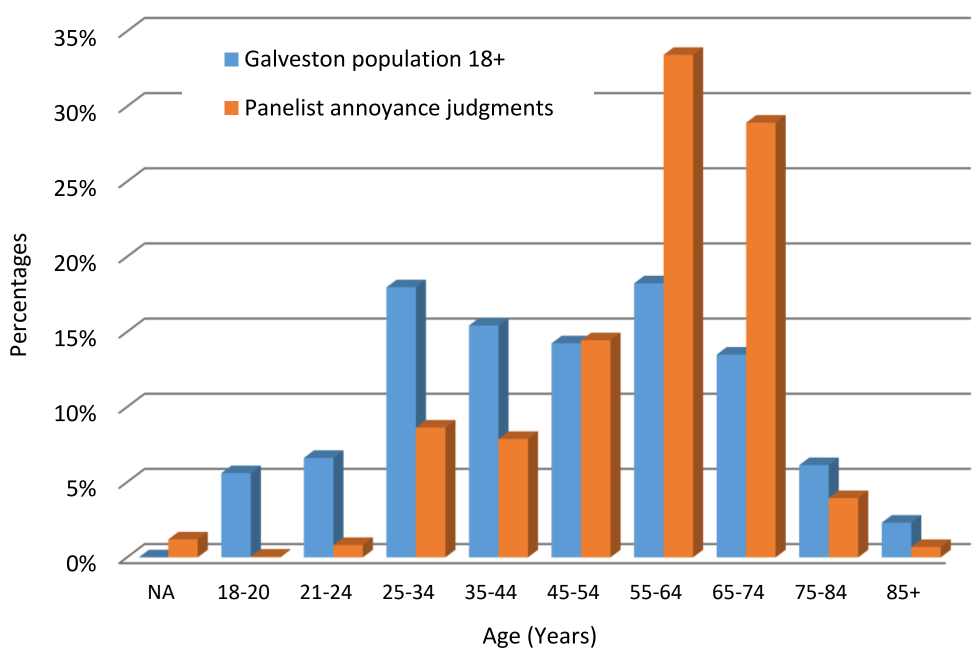

Figure 52. Age distribution of panelists whose opinions are reflected in the data set.

8.1.3 *Gender*

The QSF18 panel data collection yielded reasonably representative proportions of responses by panelist gender to the larger Galveston population. The proportion of female panelists contributing data to the sample under-represented the female population by only 4.13%, and over represented the male population by 4.02%. Table 62 tabulates these findings, while Figure 52 compares the population and sample percentages.

Table 62: Gender splits of Galveston population and panelists

2019 GALVESTON CENSUS		GALVESTON PANELISTS			DIFFERENCE: GALVESTON - CENSUS
Gender	Population Percentage	Gender in QSF18 Data Set	Number of Responses	Percentage of Responses	
Male	51.09%	F	3,274	55.11%	4.02%
Female	48.91%	M	2,660	44.77%	-4.13%
Unknown	0.00%	X	7	0.12%	0.12%
Totals	**100.00%**		**5,941**	**100.00%**	

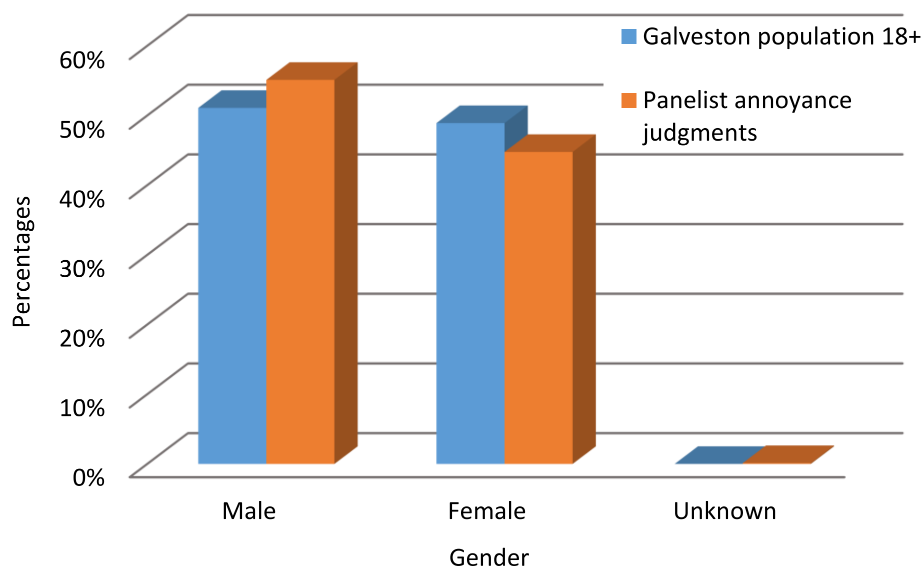

Figure 53. Gender distribution

8.1.4 *Adequacy of QSF18 sampling scheme*

Overall, the QSF18 sampling method yielded a sample that mis-represented young persons, as well as those with a high school education or less (by 23.4%.) The mis-representation was due to a combination of the sample selection mode, first-come-first served recruitment procedure, and lack of stratification. Approximately 29% of adults in the U.S. with a high school education or less do not have internet access, and therefore would be unable to take part in online recruitment efforts.

8.2 Estimated Locations of Panelists in Page *et al.* (2020) Database

Although quite a few case records in the spreadsheet provided by NASA were accompanied by plausible latitude/longitude geographic coordinates, a substantial number of the case records were not. It is unclear whether panelist location estimates in the Page *et al.* (2020) were developed from various estimates of panelist locations at the time of the annoyance judgment, or whether they were modified from original Address Based Sample (ABS) information. For example, if panelists responded at the time of annoyance judgment that they were at home, their locations could have been estimated from known latitude/longitude coordinates of their homes, or from whatever other information may have been available from reporting device IP addresses or smartphone location services at the time of the annoyance judgment.

Figure 54 shows that 10.7% of estimated panelist locations fell outside of Galveston County. If these panelist locations were derived from ABS sampling, then as many as 637 mailing pieces with accompanying participation incentives could have been posted to addresses outside of the Galveston study area.

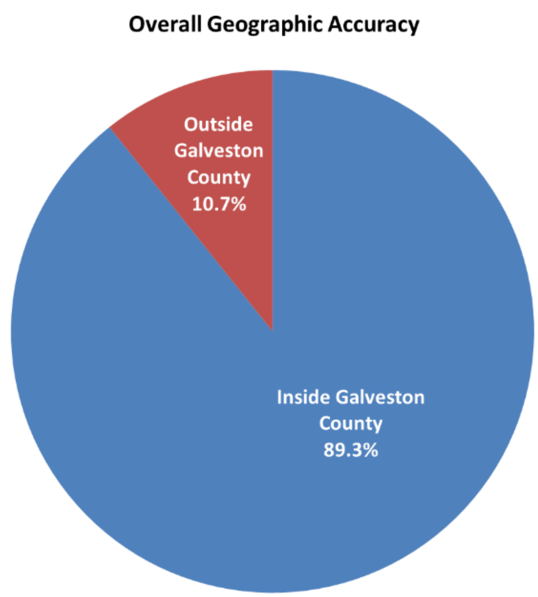

Figure 54. Percentages of panelist locations in the Page *et al.* (2020) database inside and outside Galveston County.

Table 63 shows panelist location summary statistics on the estimated panelist locations shown in Table 10 on page 28 for the QSF18 data collection.

Table 63: Distribution of estimated panelist locations

General Geographic Area of Estimated Panelist Locations	Frequency of Case Records	Percent of Overall Responses
Galveston County, Texas	5304	89.3%
Harris County, Texas	88	1.5%
Other Texas Counties	61	1.0%
California	4	0.1%
Florida	2	0.0%
Louisiana	2	0.0%
Country of Colombia (South America)	4	0.1%
FIPS CODE UNKNOWN	476	8.0%
Totals	**5,941**	**100.0%**

THIS PAGE INTENTIONALLY LEFT BLANK

9 ADDITIONAL OBSERVATIONS

This Chapter discusses several observations made in the course of initial screenings and analyses of the prompt response data set. The discussion is primarily in narrative terms, although one formal multi-level modeling analysis of a subset of the data is also reported.

9.1 Observations about Reminder Messages and Response Latencies

Several interactions were noted among the three reminder message groups and response latencies for prompt annoyance judgments. Figure 53 replots Figure 7 to emphasize major differences in observed response rates among the three reminder groups. The figure shows a 10% rate for the no-reminder message group, and a doubling and tripling of that rate for the e-mail and text reminder message groups, respectively. Reminder messages in the context of self-reporting of annoyance judgments thus appeared to have a direct effect on study participation rates. It is interesting to further note where the additional responses in the e-mail and text message groups came from, and to speculate how they could have affected observed dose-response relationships.

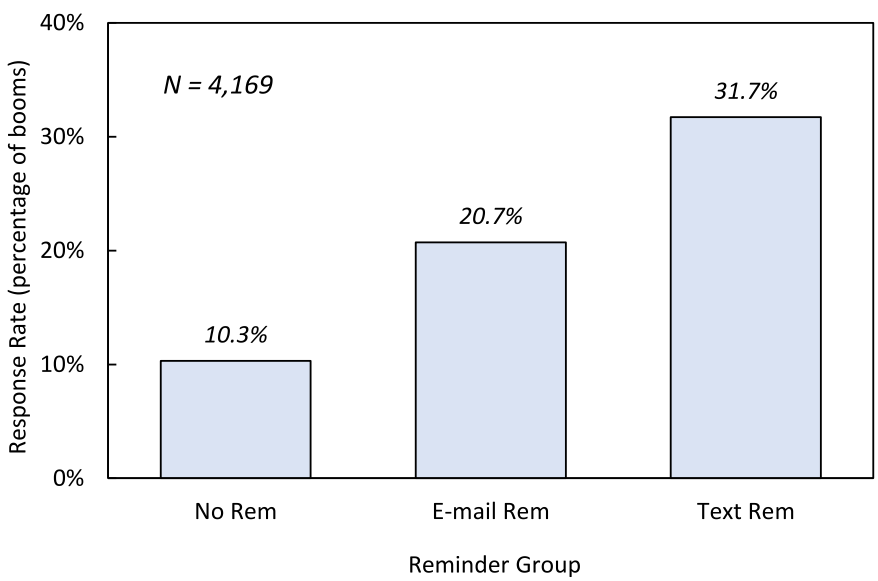

Figure 55. Response rates for three reminder groups

Figure 56 plots the prompt response data shown in Table 11 on page 34 to show the percentage of prompt responses observed in the five annoyance response categories. The figure shows that the great majority of judgments were in the not-at-all annoyed category. These judgments were made 1) by panelists who noticed a boom but were not annoyed by it, and 2) by panelists who did <u>not</u> notice a boom, and therefore could not logically have been annoyed by it.

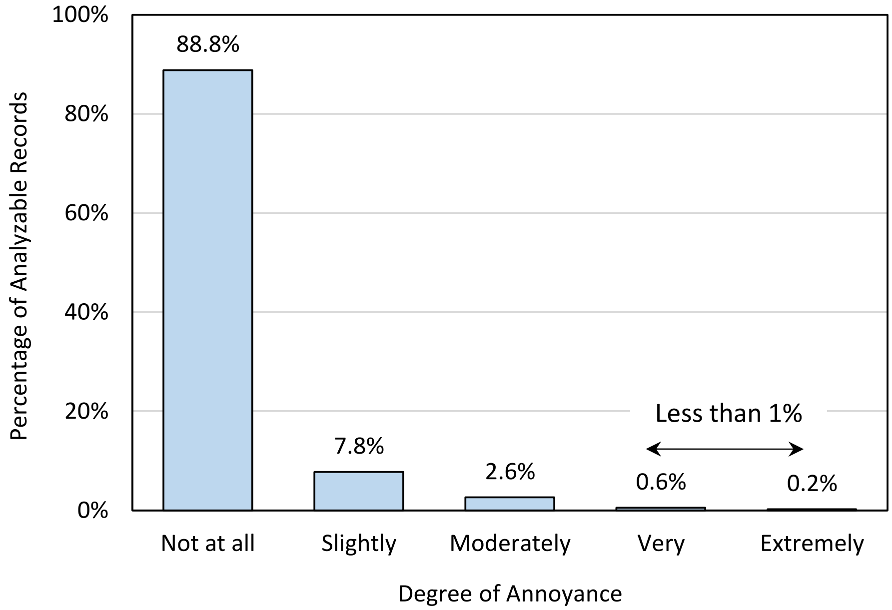

Figure 56. Distribution of prompt response degrees of annoyance

9.1.1 *Differences in annoyance judgment by reminder group*

The annoyance judgments plotted in Figure 56 (across all responses) differ somewhat when examined at the reminder group level. Figure 57 demonstrates this observation by dividing the bars in Figure 54 into three reminder groups (the bars for each reminder subgroup sum to 100%). The order in which the reminder group percentages increase or decrease within each response category is of particular interest in this plot. For the not-at-all degree of annoyance, percentages across the three groups increase systematically from the no-reminder to text reminder group. Conversely, for the higher degrees of annoyance the percentages across the reminder groups decrease systematically. This observation encourages speculation that e-mail messaging and text messaging were increasingly more successful in obtaining responses from panelists who did not in fact notice the booms. Such responses would logically have the effect of increasing the not-at-all annoyance percentages, and decreasing the percentages in the categories of slight or greater annoyance.

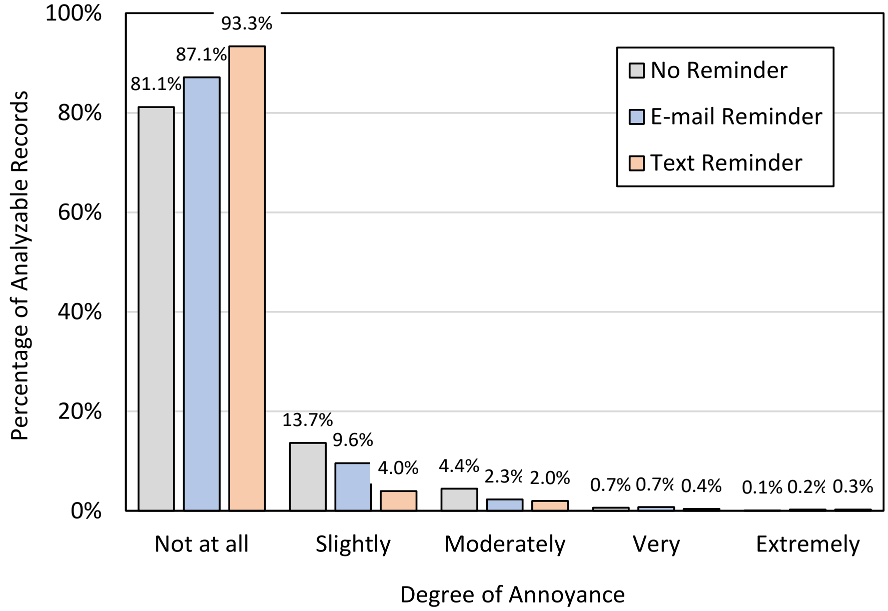

Figure 57. Distributions of annoyance degrees by reminder group

9.1.2 *Differences in response latency by reminder group*

Figure 58 shows the distribution of observed latencies for all prompt response eligible case records. The ordinate plots the number of eligible case records falling in a number of latency ranges. Note that at latencies greater than 180 minutes (3 hours), the distribution is relatively flat out to about 9 hours. At latencies shorter than 180 minutes, the record count falls in an exponential-like manner.

Figure 59 shows the first six bins (0-30 minutes through 150-180 minutes) with the data displayed separately for the three reminder message groups. The ordinate shows percentages of all responses eligible for analysis in the reminder group that fall in the latency range shown on the abscissa. In the case of the no-reminder and text reminder groups, approximately 75% of the responses occurred within the first 30 minutes following the boom. Another 12% of responses were observed in the 30 to 60-minute latency range. The percentages for these two reminder groups are relatively consistent across the entire latency range.

The distribution of response latencies within the e-mail reminder message group is quite different from the other two, however. Only about 45% of responses were observed within the first 30 minutes. Greater percentages than those seen in the other two reminder groups were observed in the longer latency ranges. This pattern of findings suggests that:

1) panelists in the no-reminder message group who noticed booms judged their annoyance fairly promptly;

2) panelists in the text reminder group also responded fairly quickly to text message alerts on their cell phones, even if they did not hear booms; and

3) panelists in the e-mail reminder group who noticed booms may have responded as quickly as those in the no-reminder message group. For booms that went unnoticed, however, an e-mail reminder itself was not as promptly noticed as a text reminder, delaying judgments of the annoyance of unnoticed booms.

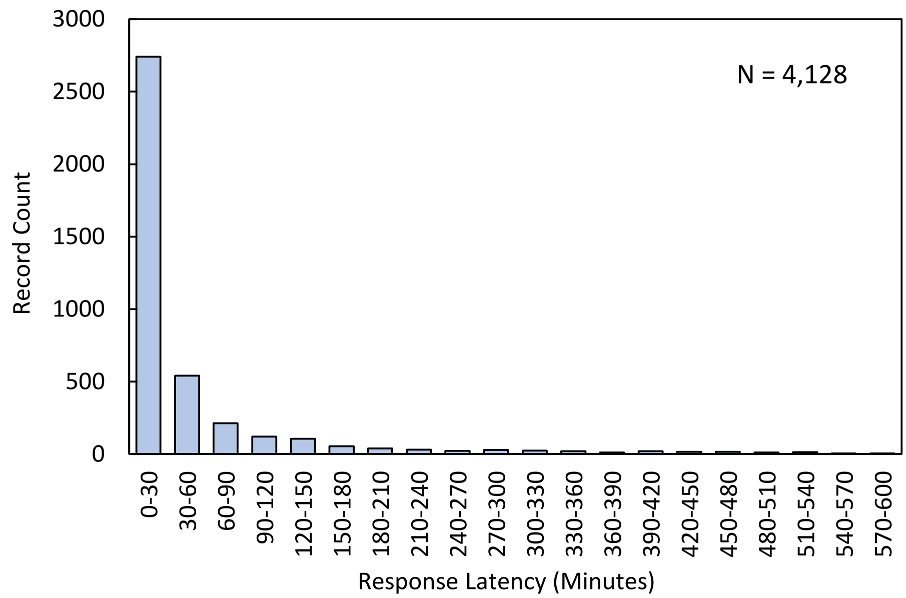

Figure 58. Distribution of response latencies

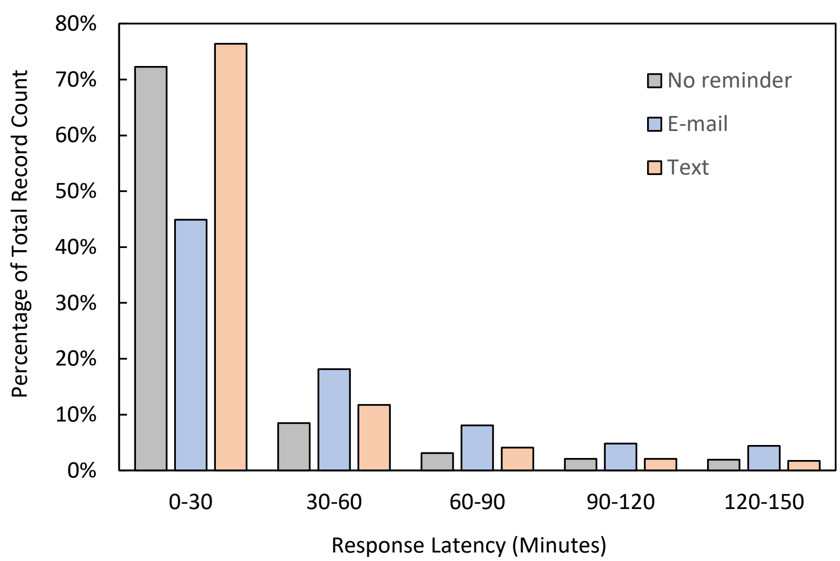

Figure 59. Distribution of response latencies by reminder type

9.1.3 *Differences in "noticed" and "not noticed" responses by reminder group*

Figure 60 examines reminder messages by dichotomizing each of the three reminder message types into "noticed" and "not noticed" groups. The heights of each bar pair sum to 100%. The fewest "not noticed" responses occurred in the no-reminder group, while the greatest number were in the "text message" group. That *any* "not noticed" responses were reported in the no-reminder group is an anomaly, however.

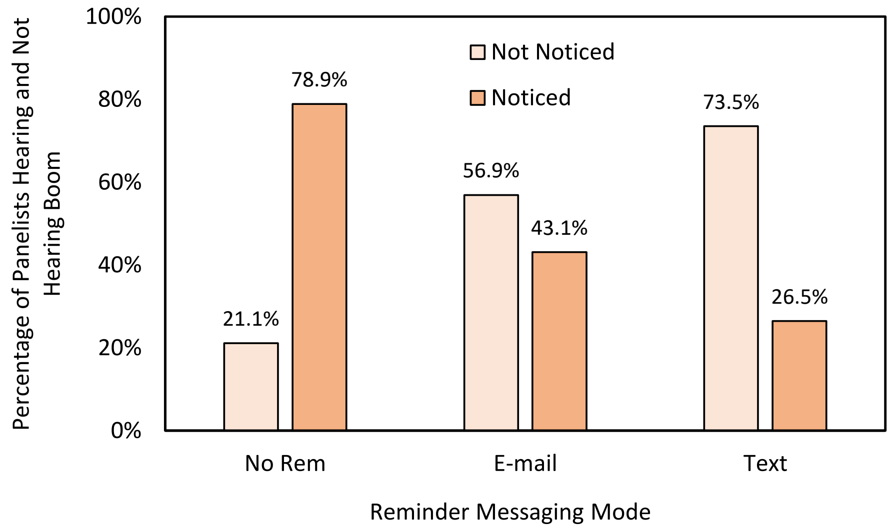

Figure 60. Reminder message categories by reported notice of booms

Figure 61 partitions each of the bars in Figure 59 into "noticed" and "not noticed" groups. The lighter bars represent "not noticed" responses, while the darker ones represent "noticed" responses. The heights of all bars within panels sum to 100%. Slightly more than 70% of all annoyance judgments were made within 30 minutes of the boom for the no-reminder group in the top panel.[29] In the e-mail group, annoyance judgments are much more broadly dispersed in time, and also in notice of booms. Annoyance judgment latencies were more concentrated in the lower exposure ranges for panelists who did not notice booms than for those who did. In the text reminder group, nearly three times as many panelists who did not hear booms responded within the first 30 minutes that those who did notice booms. Thus, both response latency and notice of booms would seem to be linked to reminder type in some (but not necessarily a causal) manner. Such a linkage could lead to unreliable or even misleading conclusions about the annoyance of actual low-amplitude booms.

[29] The dark-shaded bar in the 0-30 minute group may be due to a form of coding error, since the QSF18 data collection and archiving protocols allowed no obvious way for panelists to judge the annoyance of booms that they had not noticed. The anomaly could also reflect the manner in which specific annoyance judgments were assigned to individual booms; *i.e.*, without attempting to identify or make any adjustment for false alarms, in which sounds were interpreted as sonic booms even though no boom had occurred.

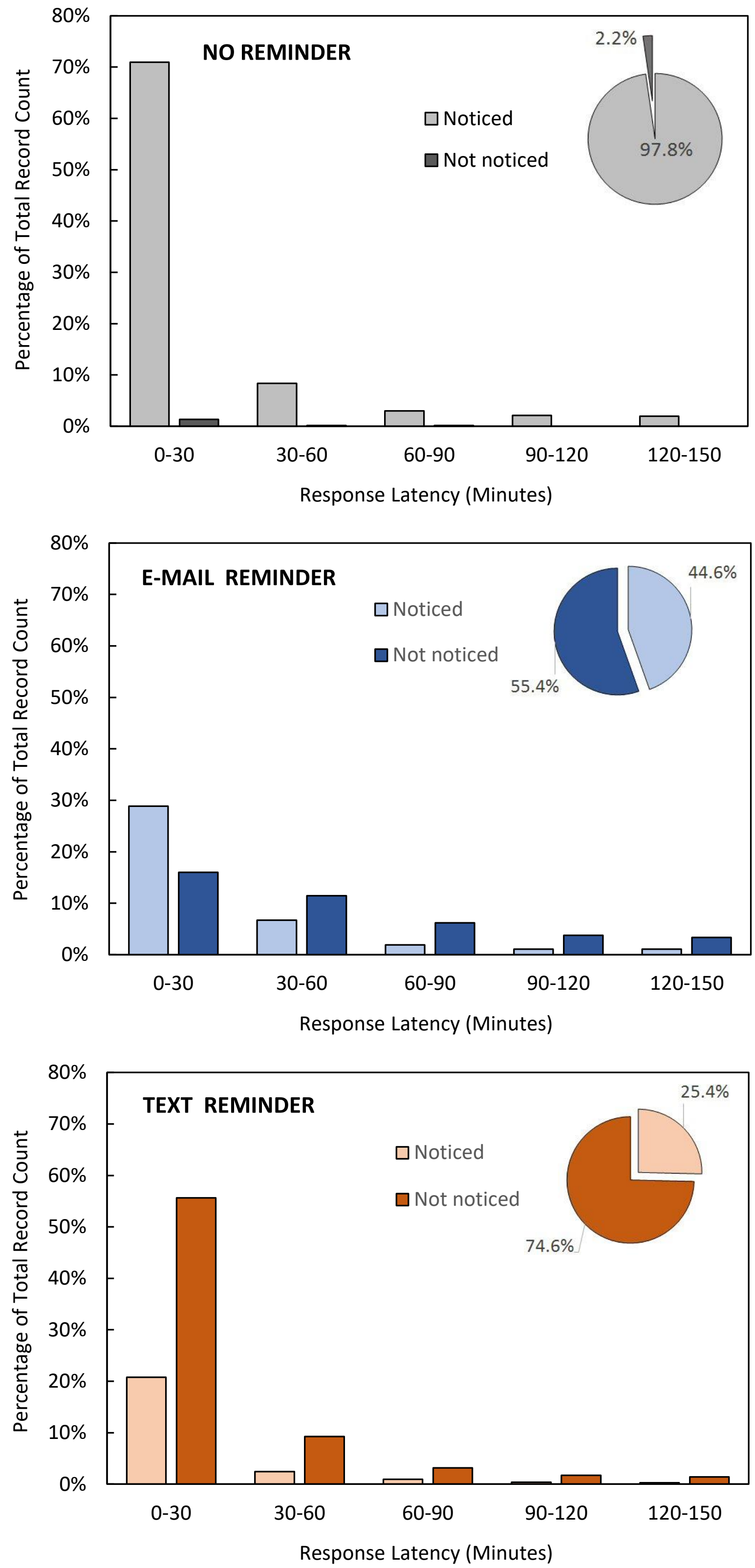

Figure 61. Proportions of "not noticed" responses in reminder message groups

9.2 Absence of Effect of Reminder Messages on Delayed Annoyance Judgments

As described in §9.1, reminder messages (or the type of reminder, if one was issued) might have affected prompt annoyance judgments. A related question arises in the analysis of delayed responses: Could the type of reminder message intended to solicit prompt annoyance judgments have affected panelists' delayed responses? More generally, might the data collection protocol for one set of annoyance judgments have affected the other set?

This subsection speculates about the origins of effects of type and number of reminder messages on delayed annoyance judgments. Recall that all panelists received reminder messages in the morning and evening. Assignment to reminder message groups meant that panelists received additional messages after each noise event. Differences in delayed responses thus refer to differences in annoyance judgments between panelists who received two messages each day, and those who received multiple messages each day.

Table 64 shows the numbers of panelists and annoyance judgments for each of the three reminder message groups. The table also shows the average number of annoyance judgments per panelist in each group and the response rate (the number of annoyance judgments divided by the product of 9 days and the number of panelists). Unlike reminder messages for prompt responses, the table shows the response rates to be about equal at slightly over 50%, independent of the form of prompt reminders. Figure 62 illustrates these response rates. The issuance of single event reminder messages appears to have no effect on the end-of-day response rate.

Table 64: Numbers of panelists and cases associated with delayed response reminders

Quantity	Reminder Group			Totals
	No Rem	E-mail Rem	Text Rem	
# of panelists	201	100	85	386
# of Annoyance judgments	1,019	520	413	1,952
Annoyance judgments *per* panelist	5.07	5.20	4.86	
Response rate[1]	56.3%	57.8%	54.0%	

[1] Number of annoyance judgments divided by (number of panelists x 9 days)

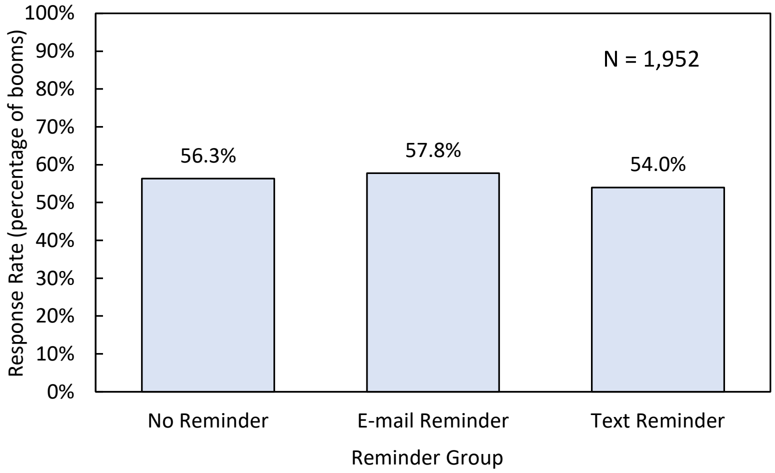

Figure 62. Delayed annoyance judgment rate by reminder message category

Figure 63 plots the prevalence of annoyance in various degrees among panelists in each of the three reminder message groups. In contrast to the prompt response annoyance judgments shown in Figure 9 (page 44), no differences in prevalence of annoyance among the reminder message groups are apparent.

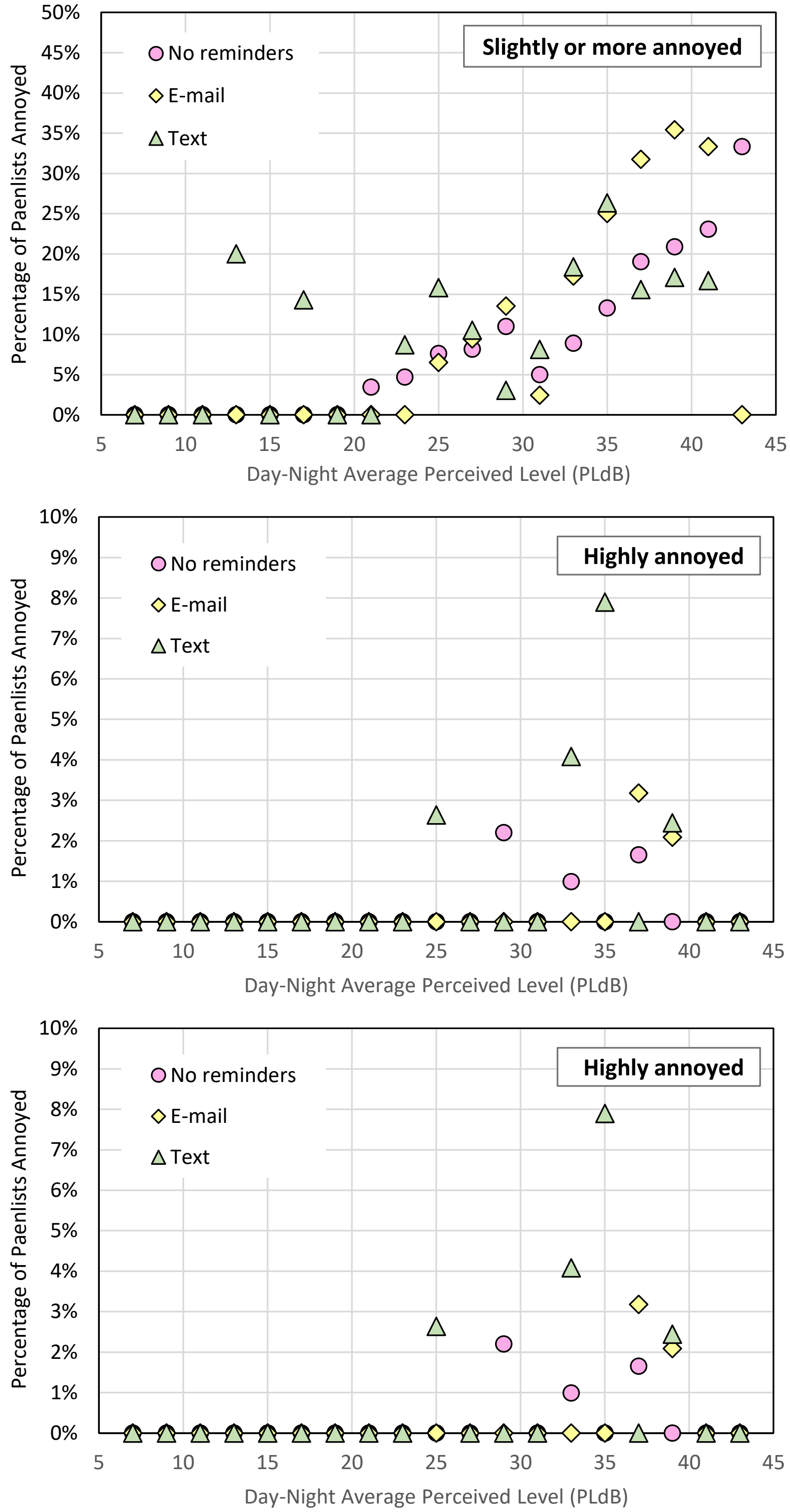

Figure 63. Delayed response annoyance judgments for three reminder messaging modes

Figure 64 shows the distribution of numbers of completed annoyance judgments *per* exposure bin for each reminder message group. *per* the second row of Table 64, the largest numbers of annoyance judgments were made by panelists in the no-reminder group; the next largest from the e-mail group; and the least in the text message group. The possible effects of the reminder message type on observed dose-response relationships are discussed in detail in the following underlined subsections.

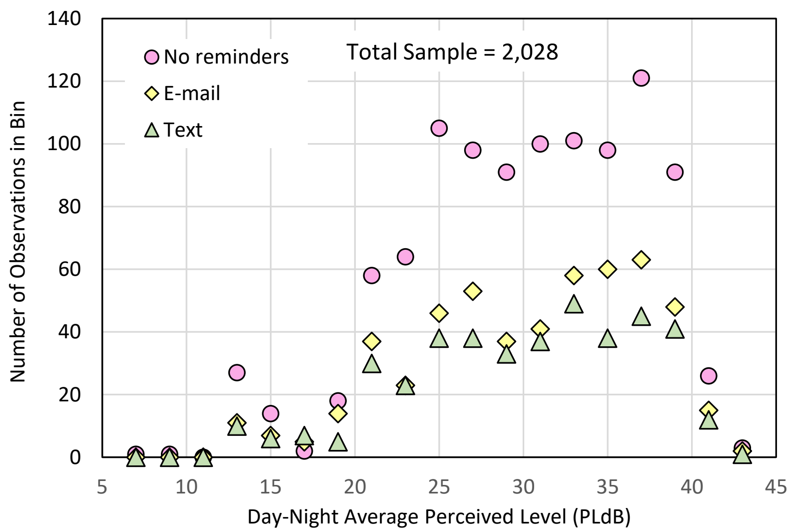

Figure 64. Numbers of delayed response annoyance judgments *per* exposure bin for three reminder messaging modes

9.2.1 *Multi-level Logistic Regression for Reminders for Delayed Responses*

The data set was again dichotomized by reminder and no-reminder to form a single binary variable, named “ReminderNone”, in order to determine whether issuance of reminder messages affected the dose-response relationship. The two reminder message categories were contrasted to form another variable, EmailText, in which no-reminder events were ignored. This permitted two comparisons: one in which issuance of reminders was compared with absence of reminders, and the other in which the two forms of reminder messages, e-mail and text, were compared.

The multi-level logistic regression equation for predicting annoyance from noise exposure, whether a reminder message was sent, type of reminder message, and the interactions between 1) whether a reminder message was sent and noise exposure, and 2) between the form of reminder messages and noise exposure. The generic form of the multi-level logistic regression equation for predicting annoyance from noise exposure, whether a reminder message was sent, the form of reminder message, and the interactions between the sending of reminder messages and noise exposure level, and between type of reminder messages and noise exposure level, is:

$$p(\boldsymbol{annoyed} = \boldsymbol{yes}) = \frac{e^y}{1+e^y}$$ **Equation 8**

where:

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion level of annoyance,

$y=\beta_{00} + (\beta_{10})(PLDNL) + (\beta_{20})(MultipleReminder) + (\beta_{30})(EmailText) + (\beta_{40})(MultipleReminder)(PLDNL) + (\beta_{50})(EmaiText)(PLDNL)$

and

β_{00}= intercept for annoyance

PLDNL is the noise dose in PLdB

MultipleReminder is the code for whether a reminder message was sent after each noise event: 0=Two reminders, 1=Multiple Reminders;

EmailText is the code for type of reminder (-0.5 for text reminders, 0 for no reminder message, and +0.5 for an e--mail reminder message);

MultipleReminder x PLDNL is an interaction term (the product of PLDNL times the *MultipleReminder* reminder message code) which tests the interaction between sending multiple reminder messages and sending two and noise dose; that is, whether the dose-response relationship varies with the presence or absence of multiple reminder messages; and

EmailText x PLDNL is another interaction term (the product of PLDNL times *EmailText* code) which tests the interaction between type of reminder and noise dose, that is, whether the dose-response relationship differs by type of reminder message.

Table 65, Table 66, and Table 67 show Equation 8 parameter values for predicting prevalence of annoyance judgment degrees for delayed response (end-of-day/cumulative exposure) models, in which Day-Night Average Perceived Sound Level (PLdB) and form of reminder message predict annoyance, as defined by the three annoyance judgment levels: HA, MA, and SA. "Reminder x PLDNL" tests the interaction between sending a reminder and the noise dose, that is, whether the dose-response relationship depends on whether reminder messages were sent. "Reminder Type x PLDNL" tests the interaction between type of reminder and the noise dose; that is, whether the dose-response relationship differs by form of reminder message.

Table 65 provides parameter values for Equation 8 for predicting high annoyance (HA) as shown below:

$y=\beta_{00} + (\beta_{10})(PLDNL) + (\beta_{20})(MultipleReminder) + (\beta_{30})(EmailText) + (\beta_{40})(MultipleReminder)(PLDNL) + (\beta_{50})(EmailText)(PLDNL)$

$$y = -6.785 + (0.060)(PLDN) + (-2.260)(Reminder\ v\ None) + (-3.360)(Email\ v\ text) + (0.080)(Reminder\ v\ None)(PLDN) + (0.081)(Reminder\ Type)(PLDN)$$

Table 66 and Table 67 display values for the equation for predicting at least moderate annoyance (MA) and at least slight annoyance (SA).

Table 65: Delayed response model of high annoyance due to Day-Night Average Perceived Sound Level, form of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-6.785	0.410	-16.53	381	<.001			
Noise level (β_{10})	0.060	0.013	4.65	1,946	<.001	1.061	1.035	1.088
MultipleReminder (β_{20})	-2.260	0.631	-4.05	1,946	<.001	0.077	0.022	0.267
EmailText (β_{30})	-3.360	0.480	-7.00	1,946	<.001	0.035	0.014	0.089
MultipleReminder x PLDNL (β_{40})	0.080	0.020	4.44	1,946	<.001	1.094	1.051	1.138
EmailText x PLDNL (β_{50})	0.081	0.156	5.20	1,946	<.001	1.084	1.052	1.118

Table 66: Delayed response model of moderate or greater annoyance due to Day-Night Average Perceived Sound Level, form of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-6.418	0.542	-11.83	381	<.001			
Noise level (β_{10})	0.091	0.169	5.40	1,946	<.001	1.095	1.060	1.132
MultipleReminder (β_{20})	-1.380	0.741	-1.86	1,946	.062	0.252	0.059	1.074
EmailText (β_{30})	-2.195	0.504	-4.95	1,946	<.001	0.082	0.031	0.222
MultipleReminder x PLDNL (β_{40})	0.057	0.023	2.45	1,946	.015	1.058	1.011	1.107
EmailText x PLDNL (β_{50})	0.062	0.157	3.93	1,946	<.001	1.064	1.032	1.097

Dose-response relationships varied by the form of reminder messages for all annoyance judgment criteria. A stronger (that is, steeper slope) relationship between noise exposure levels and annoyance judgments was observed for e-mail reminder messages than for text reminder messages. For the high annoyance criterion, dose-response relationships differed depending on whether multiple reminder messages had been sent: the slope of the relationship was steeper when additional reminder messages had been sent to panelists.

Table 67: Delayed response model of slight or greater annoyance due to Day-Night Average Perceived Sound Level, form of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-5.126	0.481	-10.66	381	<.001			
Noise level (β_{10})	0.095	0.015	6.36	1,946	<.001	1.100	1.068	1.132
MultipleReminder (β_{20})	-0.127	0.701	-0.18	1,946	.857	0.881	0.223	3.480
EmailText (β_{30})	-1.342	0.010	-2.63	1,946	.009	0.261	0.096	0.710
MultipleReminder x PLDNL (β_{40})	0.013	0.021	0.61	1,946	.545	1.013	0.972	1.055
EmailText x PLDNL (β_{50})	0.040	0.144	2.76	1,946	.006	1.040	1.012	1.070

For all annoyance judgment categories, a greater prevalence of annoyance appeared to be associated with text messages than for e-mail, producing odds ratios near zero (*i.e.*, distant from 1). For Highly Annoyed judgments, issuance of multiple reminder messages appeared to increase the degree of judged annoyance. However, all of these main effects are ambiguous in the light of the interactions in which they are included.

9.2.2 *Form of Reminder Message Follow-up Analyses of Significant Interactions*

Separate dose-response relationships were evaluated to explore the statistically significant interactions between reminder type and Day-Night Average Perceived Sound Level for all of the annoyance criteria: high annoyance, moderate+ annoyance, and slight+ annoyance. Equation 9 shows the relationship used in the evaluation.

$$p(annoyed = yes) = \frac{e^y}{1+e^y} \qquad \textbf{Equation 9}$$

where:

$y=\beta_{00} + (\beta_{10})(PLDNL)$

Table 68 and Table 69 provide Equation 9 parameter values for the Highly Annoyed criterion for e-mail and text reminders, respectively. Figure 65 plots the dose-response relationships using Equation 9 and Table 68 and Table 69. It also plots the delayed-response judgments for the e-mail and text reminder groups. Although the relationship between noise and high annoyance among responses associated with text reminder messages is of low magnitude (odds ratio of 1.033, quite close to 1 as seen in Table 69), it is unlikely to have occurred by chance alone. A steeper slope (stronger dose-response relationship) is associated with e-mail (odds ratio = 1.334 in Table 68). At Day-Night Average Perceived Levels lower than about 42 PLdB, Figure 65 shows a greater prevalence of high annoyance being associated with text messages than with e-mail messages. Insufficient numbers of annoyance judgments preclude conclusions at levels greater than 42 PLdB.

The bottom panel of Figure 65 shows the total numbers of annoyance judgments in each sound level bin for the text and e-mail reminder message panelists. The numbers highly annoyed in each bin are used to determine the fraction (prevalence) of highly annoyed panelists. This panel also applies to Figure 66.

Table 68: Delayed response model of high annoyance due to average Day-Night Average Perceived Sound Level for panelists receiving e-mail reminders

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-15.12	2.276	-6.64	99	<.001			
Noise level (β_{10})	0.288	0.059	4.87	518	<.001	1.334	1.187	1.498

Table 69: Delayed response model of high annoyance due to average Day-Night Average Perceived Sound Level for panelists receiving text reminders

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-5.823	0.484	-12.023	84	<.001			
Noise level (β_{10})	0.064	0.017	3.83	411	<.001	1.033	1.032	1.103

Table 70 and Table 71 provide Equation 9 parameter values for the moderate of greater annoyance criterion for e-mail and text reminders, respectively. Figure 66 plots the delayed-response judgments for the e-mail and text reminder groups versus the Day-Night Average Perceived Level. It also plots the dose-response relationships using Equation 9 and Table 70 and Table 71. The curve is steeper for e-mail than text reminder messages, but the two curves cross at a Day-Night Average Perceived Sound Level of about 40 PLdB. At lower levels, greater annoyance is reported when reminders are issued via text messages. At higher levels the data are insufficient to draw any conclusions.

Table 70: Delayed response model of moderate or greater annoyance due to average Day-Night Average Perceived Sound Level of exposure for panelists receiving e-mail reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.072	0.748	-13.47	99	<.001			
Noise level (β_{10})	0.204	0.238	8.56	518	<.001	1.226	1.170	1.584

Table 71: Delayed response model of moderate or greater annoyance due to average Day-Night Average Perceived Sound Level for panelists receiving text reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-5.202	0.006	-9.48	84	<.001			
Noise level (β_{10})	0.083	0.017	4.89	411	<.001	1.087	1.051	1.124

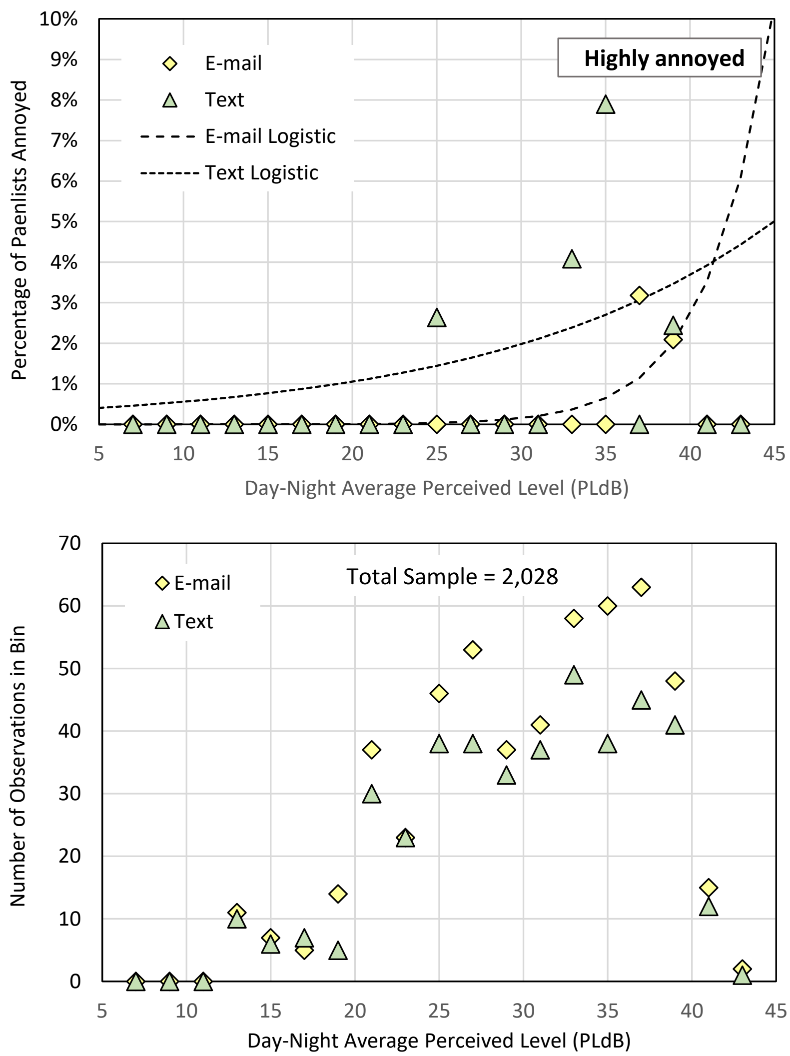

Figure 65. Delayed response annoyance judgments and logistic curves for high annoyance for e-mail and text message reminder groups

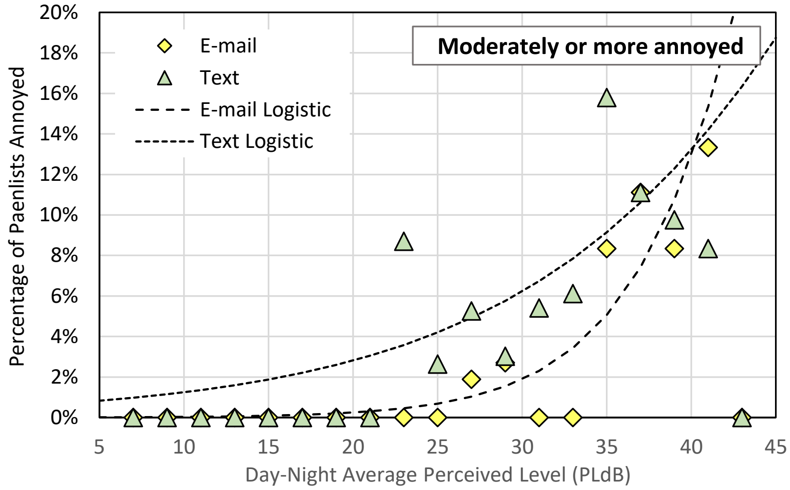

Figure 66. Delayed response annoyance judgments and logistic curves for moderate or greater annoyance for e-mail and text message reminder groups

9.2.3 *Multiple vs. Two Reminder Analyses of Significant Interactions*

Another set of analyses of dose-response relationships was undertaken to explore the statistically significant interactions between Multiple Reminders vs. Two Reminders and PLDNL for high annoyance judgments shown in Table 72 and Table 73, respectively. The dose-response relationship was stronger (*i.e.*, was steeper) when multiple reminder messages were sent.

Table 72 and Table 73 show dose-response relations for the highly annoyed for panelists receiving multiple reminder messages and those receiving two reminder messages. Figure 67 replots the annoyance judgments from the bottom panel of Figure 14 (page 126) to show the dose-response curves generated from these tables.

Table 72: Delayed response model of high annoyance due to Day-Night Average Perceived Sound Level for panelists receiving multiple reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.556	0.505	-14.97	184	<.001			
Noise level (β_{10})	0.102	0.016	6.45	931	<.001	1.110	1.084	1.142

Table 73: Delayed response model of high annoyance due to Day-Night Average Perceived Sound Level for panelists not receiving two reminder messages

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-6.630	0.369	-17.95	200	<.001			
Noise level (β_{10})	0.057	0.012	4.84	1017	<.001	1.058	1.034	1.083

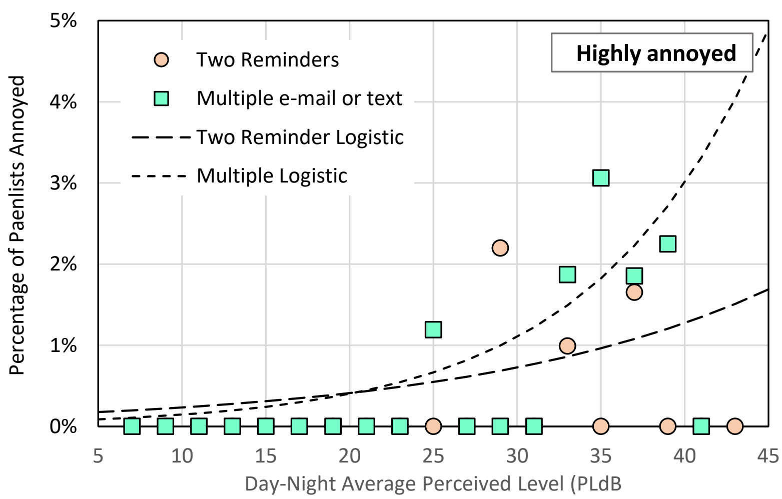

Figure 67. Logistic Dose-Response relationships for high annoyance as a function of reminder message type

The dose-response curve for responses for which multiple reminder messages were issued (either e-mail or text) was steeper than for responses which had only two reminder messages. However, the curves for the two conditions cross at an average Day-Night Average Perceived Level of about 20 PLdB. Below this noise level, days on which two reminders were issued appeared to have slightly more responses of high annoyance than days on which multiple reminders were issued. Above this noise level, responses of high annoyance were more prevalent among days on which multiple reminders were sent than on days on which two reminder messages were sent.

Figure 68 shows numbers of panelists in each of the sound level bins from which annoyance prevalence percentages were calculated. Once again, the majority of the annoyance judgments were made at exposure levels ranging from 20 to 40 PLdB.

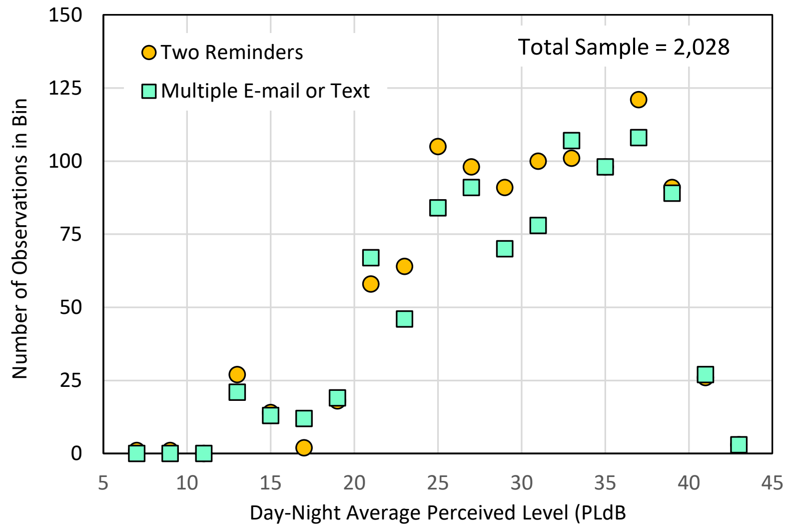

Figure 68. Numbers of delayed response annoyance judgments *per* exposure bin for multiple reminder messages and two reminder message groups

9.3 Possible False Alarms in Prompt Response Data Set

Reliable attribution of a boom to an annoyance judgment in the QSF data set relied on panelists' abilities to estimate the time of occurrence of a boom within a 15-minute window. The precision of this estimation affected all (true, false, and no) reminder message conditions, and contributed uncertainty to the entire association process. The origin of the hundreds of case records with identical combinations of Participant_ID and Boom_ID in the database generated by Page *et al.* (2020) (as described in §3) is unknown. They were excluded from the present analyses as unexplained coding anomalies. Some of these replicated records, however, may have been false alarms: responses incorrectly interpreted as annoyance judgments of particular sonic booms that were in fact annoyance judgments of other impulsive sounds. This is particularly true of self-reports received in the "no-reminder message" condition. If interpreted as correct detections of low-amplitude sonic booms, rather than as false alarms, inclusion of such replicated reports would have systematically *over*estimated the annoyance of low absolute level sonic booms.

The uncertainty of interpretation of self-report/free response data collection methods as correct reports of the presence of a signal when it is in fact present, and as false alarms (incorrect reports of the presence of a signal when it is not in fact present) has been studied and well understood for decades (Lucas, 1967). Such ambiguity of interpretation is inherent in all unstructured, self-report data collection.

Self-report data collection methods, as well as *post hoc, ergo propter hoc* attribution of annoyance judgments received at random times to particular booms, should be avoided in future community response testing involving the X-59 aircraft to avoid such ambiguities. Instead of instructing panelists in future studies to judge the annoyance of any sounds that they believe might be low-amplitude sonic booms, their annoyance judgments should be solicited by live interviews conducted at tightly controlled times. If a clear indication of false alarm rates or annoyance judgments is desired, panelists may be interviewed at times other than shortly after the shockwave of an X-59 overflight reaches their residences.

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

10 IMPLICATIONS FOR X-59 COMMUNITY RESPONSE TESTING

The subsections of this Chapter discuss implications of the present analyses for subsequent X-59 field testing study.

10.1 Sampling and Recruitment

Initial solicitations of interest in panel participation on the basis of address-based sampling should be complemented in future studies by telephone or other qualification interviews of potential panelists. Enrolling panelists on a first-come, first-enrolled basis invites sample representation errors, whereas a qualifying round of annoyance judgment permits demographic stratification of the panel by enforcing population group (census) quotas to avoid sample representation errors. A more comprehensive rationale for excluding Post Office Box addresses (as in the Galveston panelist recruitment) would also be helpful during X-59 testing, particularly in rural areas of the U.S.

Further, an exclusively on-line recruitment/qualification process misses about 10%[30] of U.S. adults who do not routinely use the Internet. (According to ongoing research sponsored by the Pew Research Center, the resulting demographic bias would be even greater among Black, Hispanic, older, and lower income groups.) Note also that in other ABS studies (*e.g.*, Messer and Dillman, 2011), panelists who returned mail surveys in a survey "pushed" from a mailing piece were quite different demographically from those who participated online.

10.2 Generalizability of QSF18 Annoyance Judgments

The findings of the present analyses (summarized in Table 1) revealed relatively few effects that were unlikely to have occurred by chance alone, or that were of sufficient importance to have major implications for the design of future X-59 testing. For the most part, the absence of such effects was not due to insufficient statistical power.

The credibility of the findings of the QSF18 study rests in large part on details of the study design and its execution. For non-technical reasons, for example, the exposure schedule was partially confounded with study duration, rather than fully randomized. In some cases, exclusion of uninterpretable cases from analysis had only minor effects on the current findings. In other cases, it could not be determined whether additional records should also have been excluded from analysis. The extensive grooming of the provided databases that was necessary for quality control purposes raises issues that NASA may wish to further investigate. For example, the occurrence of multiple case records with divergent response coding raises questions about whether coding

---

30 Source: PEW Research Center survey, conducted Jan. 8th through Feb. 7th, 2019; see also https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/.

anomalies were limited to replicate records, or whether they might be present in other records as well.

Likewise,

- the locations of panelists at their times of exposure could not be fully determined by post-processing of the data set provided by NASA;
- the association of individual responses to particular booms could not be verified in some cases;
- the very small numbers of panelists in some categories of comparative analyses limit the usefulness of findings. (The limited range of boom exposure levels is an underlying cause of this "small N" problem, at least for analyses of the prevalence of high annoyance.)

All things (including the limited representativeness of the Galveston sample of the study area population) considered, generalizations of the findings of the current analysis to populations beyond the panel sample recruited in Galveston should be approached with caution.

10.3 Improving the Efficiency of Data Collection

Table 108 of Appendix D documents the inefficiency of the QSF18 data collection effort.

- Complete data was collected for a variety of reasons from only 64 of the 408 panelists over the nine exposure days of the study;
- Only small fractions of the recruited panelists judged the annoyance of most exposure incidents;
- the annoyance judgments of only 16% of panelists could be reliably linked to particular low-amplitude sonic booms; and
- 41% (2899/7068) of all case records did not warrant analysis for various reasons.

The causes of these data losses ranged from

- an inability to unambiguously determine panelist locations at times of exposure (and hence, an inability to determine whether panelists were absent from the study area or simply did not notice low-amplitude impulses);
- to failures of panelists to self-report in a timely manner (or at all);
- to data entry errors and insufficient quality control measures;
- to the boom exposure schedule and to timing of reminder messages;
- to panelist attrition; and
- to combinations of these and other factors.

The net effect of such data losses on the derivation of dose-response functions is that most of the judgments about the annoyance of low-amplitude sonic booms were made by a relatively small number of recruited panelists. This, in turn, reduced the power of statistical analyses intended to detect differences in annoyance associated with exposure to booms. It also introduced uncertainty into dose-response relationships, and particularly into their lower asymptotes – a region of considerable regulatory interest.

Although specific causes of data losses were varied, the study goals and design themselves may have been underlying causes for much of the data loss. The following sub-sections identify and suggest improvements to study design that may improve the data collection efficiency of X-59 field testing.

10.4 Credibility of QSF Estimates of Exposure Levels

Prior analyses of the relationships among common noise dose metrics of low-amplitude sonic boom sound levels (*cf*. Fidell, Horonjeff, and Harris, 2012) showed that most of them differ only by constants or scale factors, and are so highly correlated (r = 0.90 or greater) as to be effectively collinear. High pair-wise correlations among single-event noise dose metrics imply that their abilities to predict values of response variables are nearly identical. This, in turn, implies that resources expended in investigating those additional relationships cannot yield distinct information. Table 74 shows the coefficients of determination (r^2 values) among the noise metrics calculated by Page *et al.* (2020). The coefficients of determination shown in the table are generally lower than those found in similar comparisons in the Edwards AFB dose-response pilot study (Fidell, Horonjeff and Harris, 2012). This outcome appears due in large part to noise dose uncertainty at low boom sound levels, discussed further in the paragraphs below.

Table 74 and Figure 69 illustrate the high correlation between measured values of A-weighted and PL sound exposure levels as reported by Page *et al.* (2020) and Page *et al.* (2014) at Edwards AFB, CA. The black plotting symbols show 4,722 noise dose estimates from the QSF18 data, while the red plotting symbols show average values across all boom monitor stations for each of the 96 Edwards AFB booms reported by Page *et al.* (2014). Figure 69 also shows linear regressions for each data set– the dashed blue line for the QSF18 estimates, and the dashed red line for the Page *et al.* (2014) measurements. The QSF18 regression accounts for 93% of the variance in the relationship, with a linear regression slope of nearly 1. The regression for the Page *et al.* (2014) data accounts for 99% of the variance in the relationship between A-level and PL, with a slope that is again very nearly one.

Similarly, Figure 70 illustrates the correlations between measured values of C-weighted and PL sound exposure levels that were observed by Page *et al.* (2020) and by Page *et al.* (2014). The black plotting symbols show 4,722 noise dose estimates from the QSF18 data, while the red plotting symbols plot the average values across all boom monitor stations for each of the Page *et al.* (2014) booms. Simple linear regressions for each data set are also shown – the dashed blue line for the QSF18 data, and the dashed red line for Page *et al.* (2014). Whereas the Page *et al.* (2014) regression accounts for virtually all (96%) of the variance in the relationship between A-weighted and Perceived Levels, the regression for the Page *et al.* (2020) data accounts for little more than half of the variance in the relationship between C-weighted sound level and Perceived Level values. The slopes of both regression lines are also far lower than 1.

The relationship between A-weighted sound level and Perceived Level measurements in Figure 69 is exceptionally tight, with the exception of ~35 data points lying below and to the right of the general trend. These outliers divide into clusters which for the most part run either horizontally or parallel to the main trend. The clusters running parallel to the trend line are separated from the main trend by very nearly 10 decibels on both the x and y axes. This makes it difficult to determine whether the outliers are caused by erroneous values in one or the other variable. It is nonetheless likely that the deviant values are due to incorrect estimates on one or the other axis, but not on both axes. The horizontal cluster of points in Figure 69 implies that a single value was assigned to the A-level of a small number of differing PL values.

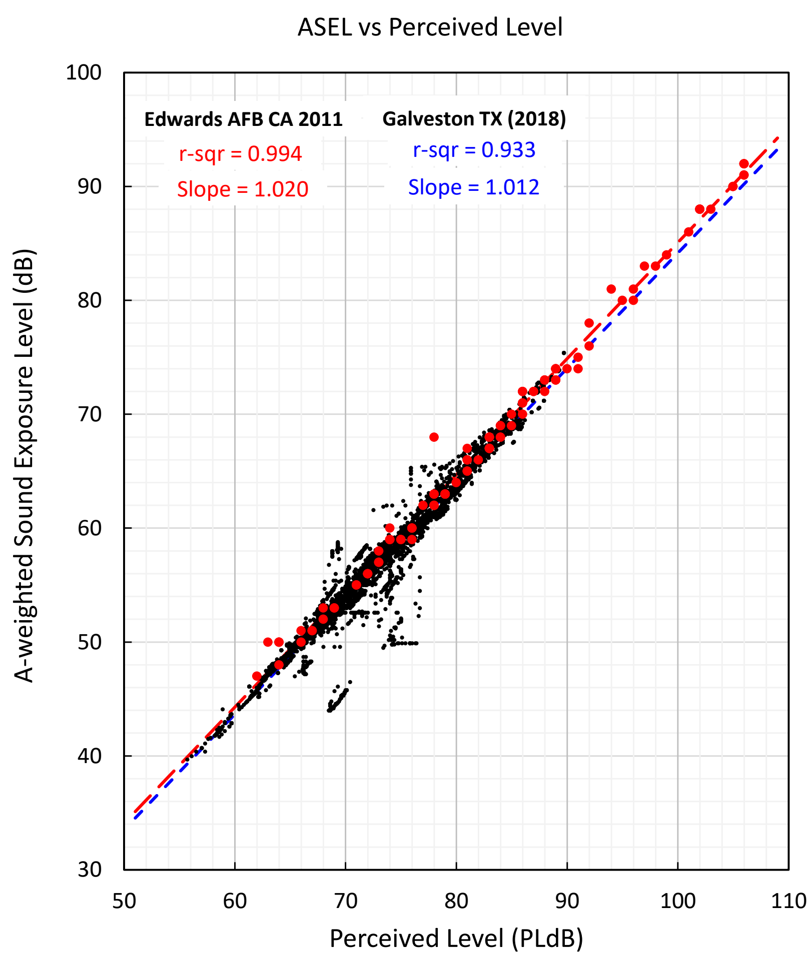

Figure 69. QSF18 and Edwards AFB pairs of A- *vs.* PL-weighted sound levels

The relationship between C-weighted sound level and Perceived Level in Figure 70, on the other hand, shows many more vertical and horizontal clusters of point. The underlying cause is not apparent from the data themselves, but the additional scatter most certainly affects the coefficient of determination between the two metrics.

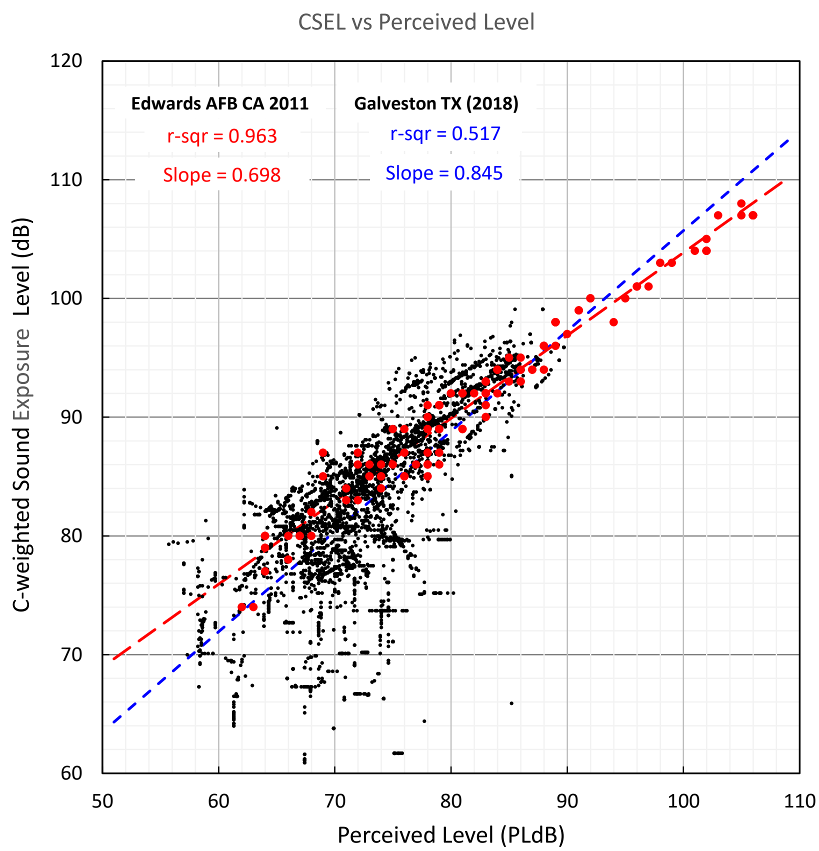

Figure 70. QSF18 and Edwards AFB pairs of C- *vs.* PL-weighted sound levels

Table 74: Coefficients of determination among alternate metrics of low-amplitude sonic booms in QSF18 noise dose estimates

	PL	ASEL	BSEL	CSEL	DSEL	ESEL	FSEL	ISBAP	LLZf	LLZd	PNL	10 Log MaxPSF
PL	1.000	0.933	0.697	0.517	0.599	0.766	0.292	0.702	0.829	0.838	0.871	0.386
ASEL		1.000	0.705	0.514	0.602	0.780	0.296	0.689	0.814	0.823	0.876	0.386
BSEL			1.000	0.782	0.871	0.897	0.491	0.907	0.732	0.721	0.785	0.592
CSEL				1.000	0.916	0.692	0.642	0.834	0.573	0.559	0.555	0.751
DSEL					1.000	0.777	0.629	0.907	0.654	0.639	0.653	0.724
ESEL						1.000	0.432	0.844	0.775	0.760	0.851	0.540
FSEL							1.000	0.517	0.304	0.294	0.295	0.923
ISBAP								1.000	0.751	0.739	0.764	0.623
LLZf									1.000	0.990	0.910	0.399
LLZd										1.000	0.908	0.388
PNL											1.000	0.395
MaxPSF												1.000

Note: Shaded cells denote correlations of 0.9 or greater.

Integrated noise metrics – ASEL: A-weighted sound exposure level; BSEL: B-weighted sound exposure level; CSEL: C-weighted sound exposure level; DSEL: D-weighted sound exposure level; ESEL: E-weighted sound exposure level; FSEL: F-weighted sound exposure level.

Non-integrated noise metrics – PL: Perceived Level; ISBAP: Indoor sonic boom annoyance predictor; LLZf: Zwicker loudness level; LLZd; PNL: Perceived Noise Level; 10 log MAXPSF: 10 (log(maximum overpressure in pounds *per* square foot)).

10.5 Relative Importance of Short- and Long-Term Effects of Exposure

Much of the complexity and expense of the QSF data acquisition study was due to the goal of characterizing both prompt (single event) and long term (cumulative) response to exposure to low-amplitude booms. Subsonic aircraft noise regulatory policy deals exclusively with the prevalence of a consequential degree of residential annoyance with long-term, cumulative exposure. If detailed discussions with FAA and ICAO prior to the design of X-59 community response testing could confirm the lesser regulatory relevance of quantifying prompt responses to individual low-amplitude sonic booms, such confirmation could result in less costly, more efficient, and larger scale community response studies during X-59 field deployments. Similar discussions could also be held concerning regulatory agencies' interests in quantifying annoyance in non-residential settings.

Even if pre-deployment discussions with regulatory agencies were to indicate that they had interests in information about both single-event and cumulative annoyance of sonic booms, they might also confirm that the regulatory agencies were interested to a greater degree in cumulative than in single-event reactions. If so, it would be possible to design studies which yielded more information about delayed (cumulative) than about prompt (single-event) reactions. For instance, panelists could be interviewed shortly after the last X-59 flight of a day about their prompt annoyance judgments of a particular boom event, and at the same time, about their judgments of the annoyance of the day's worth of impulsive exposures.

10.5.1 *Self-report vs. solicited interviews*

Given the large numbers of long latency self-reports of prompt annoyance judgments, reliance on self-reporting of reactions to sonic booms complicates attribution of specific annoyance judgments to specific booms. This unreliability also complicates distinction of hits (correct detections of low-amplitude booms) from false alarms (incorrect detections of false alarms when they are not present), and hence, accurate estimation of annoyance prevalence rates. Self-reporting of annoyance judgments further requires administrative effort to schedule and send reminder messages. Sending of reminder messages also blurs the distinction between self-report (complaint-like behavior) and solicited annoyance judgments.

One potential remedy for the ambiguity between "didn't notice a boom" and "wasn't annoyed by a boom" self-reports might be a study design which waits, say, ten minutes after a boom event for a spontaneous self-report, but if none is received, actively solicits (via text message or other means) a response to a questionnaire item concerning whether a panelist was at-or near-home at the time of exposure.

However, open-ended – that is, one way – contact attempts with panelists via text or e-mail reminder messages offer no confirmation that reminder messages have been timely received, nor do they guarantee prompt responses. Soliciting interviews at known times relative to the occurrence of exposure events, either by interactive voice response or by live agent interviewing, could reduce ambiguities of association of prompt annoyance judgments with particular booms. A live agent interview could also reinforce panelists' understandings of questionnaire items and study procedures, while encouraging continued study participation.

About 50 centrally-supervised CATI interviewers would suffice to make initial contact attempts and one or two prompt callbacks at a given interview site. If interviews are to be conducted in future X-59 field testing shortly after a carpet boom reaches half a dozen communities *per* flight, then several hundred interviewers would need to be trained. Given that panelists will have been recruited and provided with participation incentives, the response rate should be far higher than those observed by Fidell *et al.* (2019) in prior NASA-sponsored research

10.5.2 *Implications of differences between inverted dive and carpet booms*

Differences between impulsive exposure produced by supersonic diving maneuvers and carpet booms produced in straight-and-level flight, and between relatively short and lengthier duration test periods, carry several implications for future X-59 field testing. These implications include diminished needs for highly detailed local (single site) meteorological measurements in aircraft operating and interviewing areas immediately prior to test flights; greater utility for adaptive (rather than fixed exposure schedule) study designs; and reduced needs for a dense, small-scale network of boom monitors (albeit greater needs for wide-area monitoring.) Adaptive study designs of responses to carpet booms would be facilitated if acoustic data collection and analysis were more highly automated and more rapidly accomplished than in the QSF18 data collection exercises.

10.5.3 *Utility of very low-amplitude exposure in future testing*

The "no response/out of study area" ambiguity can be reduced by confining planned boom exposure levels to a more noticeable range of levels, over which noise metrics are also more readily calculable. Because annoyance with noise exposure is not caused solely by noise exposure, some respondents may continue to report annoyance even at very low exposure levels, for entirely non-acoustic reasons. As a separate matter, it is also unlikely that regulatory agencies would endorse definitions of the threshold of significant impulsive noise exposure that differed greatly from the values implied by the FICON (1992) dose-response function. To insist on a zero prevalence of high annoyance as a threshold of significance for impulsive noise exposure, but to tolerate a 12.3% prevalence of annoyance for subsonic aircraft noise, would beg questions about the established (FICON, 1992) rationale for regulating aircraft noise.

One disadvantage of a study design that minimizes the amount of scheduled exposure to very low levels of low-amplitude booms would be an inability to apply logistic regression techniques to infer an empirical "zero annoyance" point.[31] This disadvantage can be mitigated (parsimoniously) by relying on CTL analysis, or (less parsimoniously) on the four-parameter logistic Hill equation as alternatives to conventional logistic regression. CTL analysis, which fits field observations to an exponential function with a fixed slope, can yield a dose-response function with a non-zero asymptote that satisfies a maximum likelihood criterion (Mestre *et al.*, 2016, 2017). The lower asymptote of the CTL function could then serve as a more accurate estimate of a minimal noise impact threshold than a logistic function.

10.6 Initial Lesser Prevalence of Annoyance

Both the prompt and delayed response results of §6.3 suggest that the prevalence of annoyance with sonic booms in the QSF18 data collection was lower during the first few days of testing for all degrees of annoyance judgments. The reasons for this sequential effect are unknown, but may have had something to do with the partial confounding of the exposure schedule with study duration.

As described in §4.3, Table 4-3 and Appendix T of Page *et al.* (2020), daily exposures to low-amplitude sonic booms in Galveston resembled an increasing method of limits schedule, with some within-day randomness. Per-event impulsive noise exposure was planned to be limited for the first few days to 80 PLdB, after which it increased to 85, 90, and 95 PLdB on successive days. This partial confounding of level of exposure with the duration of the test complicates distinguishing sequential effects from effects of exposure level *per se*.

10.7 Panelist Attrition

About a quarter of the recruited and confirmed panelists were lost to attrition over the nine exposure days of the QSF18 data collection exercise, resulting in a 28% reduction of prompt responses and 23% delayed of delayed responses. An attrition rate this great over only nine days of exposure suggests either a need for greater over-sampling in longer duration X-59 deployments, and/or a need for more active and sustained efforts to retain panelists.

Plans should be made to accommodate an attrition rate of at least a third, if not half, of recruited panelists in weeks-long X-59 field testing. Such over-recruitment would be particularly appropriate in field studies whose designs confound study duration with boom exposure levels (*e.g.*, ascending method-of-limits exposure schedules) to maintain adequate numbers of participants at higher exposure levels.

---

[31] Logistic regression forces the dose-response relationship to a lower asymptote of zero, even if the prevalence of annoyance does not empirically asymptote at zero. This can lead to over-estimation of the prevalence of annoyance due to exposure to low-amplitude sonic booms.

10.8 Future Needs for High Precision of Measurement at Low Signal-to-Noise Ratios

The most challenging aspects of measuring exposure to low-amplitude sonic booms produced by the inverted supersonic dive maneuver is distinguishing their waveforms from ambient sound levels at very low signal-to-noise ratios. Doing so presents much less of a challenge for carpet booms produced by the X-59 aircraft, for two primary reasons. First, the width of the carpet boom corridor is far greater than that produced by the diving maneuver. Second, as described by Morgenstern *et al.* (2012) and Maglieri (2019), the cross-track homogeneity of carpet boom exposure is considerably greater than that of booms produced by the diving maneuver. In combination, the two phenomena permit a wider choice of low ambient-level measurement sites High ambient noise is the greatest impediment to obtaining useful measurements at Perceived Levels lower than about 75 PLdB.

Further, the payoff for extreme precision of measurement of low-amplitude carpet booms at low signal-to-noise ratios is minimal. As Figure 7 on page 40 illustrates, the prevalence of high annoyance with sonic booms at maximum sound levels below about 70 PLdB is so close to zero as to be difficult to distinguish from zero. Estimates of boom magnitudes at levels lower than about 70 PLdB could be in error by as much as ± 2 PLdB without affecting characterization of community response to sonic booms. Given that the whole purpose of the X-59 program is to collect information to be presented to FAA and ICAO to reconsider their prohibitions on overland supersonic flight, little purpose is served by incurring high costs to measure boom exposure at very low signal-to-noise ratios.

Should an annoyance criterion lower than "highly" be of interest to regulators, the issue of ambient noise interference could become acute. Table 75 shows the Perceived Levels at which the dose-response curves in Figure 7 reach selected percentages of panelists annoyed. The first column lists three low degrees of annoyance at which the dose-response curves begin to depart the abscissa. The remaining columns show the Perceived Levels at which these percentages occur for three different degrees of annoyance.

Table 75: Perceived Levels (dB) at which annoyance judgments reach various prevalence rates

Percentage of population annoyed	Degree of Annoyance		
	Slightly or more	Moderately or more	Highly (very or more)
1%	--	63	75
2%	53	69	81
5%	63	77	89

The current analyses indicate that typical suburban environments ambient levels can create difficulties in measuring carpet booms at Perceived Levels lower than 75 PLdB. The pink-highlighted cells in the table indicate points on the dose-response curve where exposure noise metrics would be difficult to measure for lack of sufficient signal to noise ratio.

10.9 Determining Panelist Locations at Times of Exposure

The effort expended in determining panelist locations at the times of their exposures to low-amplitude sonic booms, and in detailed estimation of panelists' personal exposures at such times, increased the costs of the QSF18 data collection and analyses. It may not be necessary to incur similar costs in future field tests of community response to X-59 carpet booms, however. To be as easily understandable and persuasive as possible, evidence about community response to low-amplitude carpet booms should be of a form and nature as consistent as possible with regulatory agencies' long-established aircraft noise regulatory policies.

As noted in §10.5, these regulatory policies limits are based exclusively on long-term, outdoor residential noise exposure levels, rather than on indoor, personal, occupational, recreational, or other forms of aircraft noise exposure. Thus, for example, FAA's definition of the significance of community response to aircraft noise exposure is conditioned on a hypothetical annual average day's exposure to cumulative, outdoor residential neighborhood aircraft noise levels. Although different residents of airport-vicinity neighborhoods may spend very different amounts of time at home, and their waking hours may be spent in noise environments very different from those of their homes, their aircraft noise exposure for regulatory policy analyses is that outside their residences.

In other words, federal regulation is blind to distinctions between indoor and outdoor exposure of household residents to aircraft noise, even though the two forms of exposure may differ by ~20 dB.[32] The long-established practice of basing regulatory policy on outdoor residential rather than indoor personal noise exposure is necessary because the total personal exposure of individuals to aircraft noise, whether created by subsonic or supersonic overflights, cannot be cost-effectively or reliably estimated (Fidell, 2015). Thus, current aircraft noise regulation is based on nominal exposure contours, and FAA discourages comparisons of measured *vs.* modeled exposure. Further, regulatory policy points are defined only for 5 dB-wide exposure intervals, and noise exposure contours are prepared only in 5 dB increments.

Further, the limited accuracy and availability of estimates of panelist location obtained from cell phones may be further compromised in the future by FCC actions to restrict cell carriers from entering into location-sharing agreements.[33] Location information in future studies of

---

[32] It may seem that outdoor residential noise exposure differs from personal residential noise exposure only by the sound transmission loss of a residence. This is not the case for sonic booms. The low-frequency content of sonic booms can excite structural resonances and secondary emissions to create potentially annoying rattle that is audible only indoors.

[33] According to the Wall Street Journal of 28 February 2020, the Federal Communications Commission "is seeking hundreds of millions of dollars in fines from the country's top cellphone carriers after officials found the companies failed to safeguard information about customers' real-time locations."

community response to X-59 carpet booms should be obtained by direct questioning of study participants.[34]

10.10 Extension of Cumulative Exposure Range

FAA regulatory policy assesses the significance (*i.e.*, warranting of mitigation) of aircraft noise impacts on communities in terms of the A-weighted Day-Night Average sound level (DNL) value that such exposure creates on an annual average day. FAA is therefore likely to consider information that NASA produces about community response to low-amplitude sonic booms in similar cumulative noise exposure units in the context of its existing regulations. The agency's current L_{dn}= 65 dB definition of a threshold of significance of aircraft noise exposure implies that it regards a prevalence of high annoyance of 12.3% as indicating a significant degree of aircraft noise impact.[35]

Figure 6-47 of Page *et al.* (2020) shows that the observed prevalence of cumulative high annoyance judgments in the QSF18 data at daily exposures of $L_{Cdn} < 50$ dB did not exceed 1%. On its face, this finding suggests that the prevalence of high annoyance with low-amplitude sonic booms is notably lower than that which FAA policy currently recognizes as "significant" for regulatory purposes. However, given some of the ambiguities in both physical and social measurements in the Galveston data collection, it is difficult to dismiss speculation that the observed very low prevalence rates of high annoyance might be due to methodological artifacts, rather than to substantive annoyance judgments.

The most direct precaution against such speculation in future field studies of community response to low-amplitude carpet booms is to include exposure conditions expected to create higher annoyance prevalence rates. If such higher levels of exposure to low-amplitude sonic booms are found to be associated with proportionately higher annoyance prevalence rates, it can be persuasively argued that the very low annoyance prevalence rates observed in Galveston were not due to the insensitivity of the research methods to actual annoyance created by impulsive noise exposure, but that the findings are methodologically credible.

10.11 Instrumentation Improvements to Increase Data Collection Efficiency

Although the unattended field instruments deployed by Page *et al.* (2020) to make acoustic measurements of sonic boom levels were capable of remote downloading, they did not support

---

[34] Unless such information is sought in a separate interview, this tactic does not resolve whether non-responding panelists simply failed to notice a low-amplitude boom, or were not present within a designated exposure area at the time of an overflight.

[35] FAA's policy is based on its reliance on the obsolete and demonstrably incorrect FICON (1992) dose-response relationship, as well as its former charter – rescinded by Congress in 1996 – to promote civil aviation. The latest international technical consensus standard, ISO 1996-1:2016, indicates that roughly 28% (not 12.3%) of the residential population, is highly annoyed by aircraft noise exposure at a DNL value of 65 dB.

rapid or geographically-distributed processing of acoustic field measurements. Such instrumentation limits NASA's ability to employ adaptive experimental designs in larger scale studies of community response to X-59 overflights.

Adaptive study designs can improve the efficiency and cost-effectiveness of data collection, but require near-real time decisions to be made about sequential exposure levels. Such decisions, in turn, typically require a quick evaluation of sound levels produced during flyovers, and associated annoyance prevalence rates, before the next scheduled overflight occurs. Pragmatically, such decisions are only possible (and affordable) with a greater degree of automation of data processing, and with the capability for multiple, geographically-dispersed analysts to participate in the decision making.

10.12 Acoustic Measurement Uncertainty Appropriate for X-59 Deployments

Extreme precision in measurement of low-amplitude booms is not essential for purposes of inferring a dose-response relationship that is relevant for regulatory analyses of the prevalence of annoyance associated with low-amplitude carpet booms. Precision of measurement to the nearest 1 or 2 decibels is more than ample for this purpose. Over much of the range of the dose-response function for prediction of the prevalence of high annoyance, differences in sound levels of such magnitude have little or no meaningful effect on predicted annoyance prevalence rates.

It follows that resources need not be expended on greater precision of measurement for X-59 field testing purposes. Greater precision of boom measurement may be of interest for other purposes (such as quantifying the effects of short-range atmospheric turbulence on the uncertainty of boom levels), but is unlikely to affect regulatory decisions on tolerable annual average levels of cumulative boom exposure.

10.13 Likelihood of Regional Differences in X-59 Community Response Tests

Regional differences in dose-response relationships that are likely to be observed during X-59 deployments to differing geographic regions are readily anticipatable. If such differences are characterized solely by regression analyses, they may impede straightforward interpretation of findings for regulatory purposes. Perspective on such regional differences in tolerance for low amplitude sonic booms may be gained through CTL analysis, however. Figure 71 (from Fidell *et al.*, 2014), for example, shows that the standard deviation of the distribution of CTL values for subsonic aircraft noise exposure is 7.0 dB. Differences of this magnitude in tolerance for low-amplitude sonic booms should no more deter regulatory agencies from formulating policy for low-amplitude sonic booms than they do for regulation of subsonic aircraft noise.

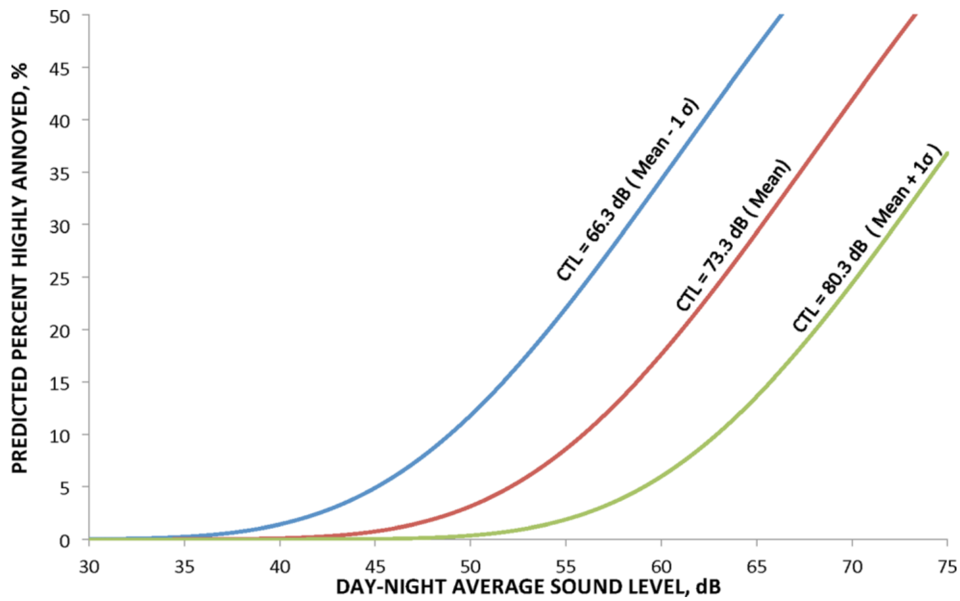

Figure 71. Illustration of effects of ± 1σ differences in CTL values on Dose-Response relationships for subsonic aircraft noise (Fidell *et al.*, 2014)

11 GLOSSARY OF ABBREVIATIONS AND OTHER TERMINOLOGY

ABS: Address-based sampling

Active panelist: A member of the original recruitment pool who contributed one or more analyzable annoyance judgments to the analysis at the conclusion of the field study.

Analyzable panelist: A member of the original recruitment pool who contributed one or more analyzable annoyance judgments to the analysis at the conclusion of the data screening process.

ADS-B: Automatic Dependent Surveillance – Broadcast

BDNL: B-weighted Day-Night average sound level (in decibels)

BSEL: Letter abbreviation for B-weighted sound exposure level (represented in mathematical expressions by the symbol L_{BE}, *per* ANSI standard S1.1-1994 *et seq.*) (in decibels)

CATI: Computer-assisted telephone interviewing

CHABA: Committee on Hearing, Bioacoustics and Biomechanics of the U.S. National Academy of Sciences

CSEL: Letter abbreviation for C-weighted sound exposure level (represented in mathematical expressions by the symbol L_{CE}, *per* ANSI standard S1.1-1994 *et seq.*) (in decibels)

CTL: Community Tolerance Level (in decibels)

Decibel: The measurement unit of a physical sound level (*e.g.*, A-weighted sound level or C-weighted sound level)

Delayed annoyance judgment: A response made at the end of the day to the day's worth of sonic booms as a whole.

DNL: Abbreviation for Day-Night Average Sound Level (in decibels)

DDNL: Abbreviation for D-weighted Day-Night Average Sound Level (in decibels)

DSEL: Letter abbreviation for D-weighted sound exposure level (represented in mathematical expressions by the symbol L_{DE}, *per* ANSI standard S1.1-1994 *et seq.*) (in decibels)

FIPS: Federal Information Processing Standards

GIS: Geo-Information System

HA: Abbreviation and symbol for "high annoyance" judgments (composed of self-description of panelists as either very or extremely annoyed

HLM: Hierarchical linear modeling

LBFD: Low-boom flight demonstrator (aircraft)

MLM: Multi-level (regression) modeling

PL: Abbreviation for Perceived Loudness Level (or Perceived Level)

PLdB: The measurement unit for the Mark VII perceived loudness level (Stevens, 1971). The initial "PL" indicates that this measure of perceived loudness differs from measures of physical sound level. In both cases, these noise metrics express a measured quantity in terms of the logarithm of its ratio to a standard reference value. The measured quantity is first converted to a ratio (division by a standardized reference quantity), then the common logarithm is taken of this ratio, and finally the logarithm is multiplied by a constant (C).

The term "decibel" is applied when the multiplier is 10. For simple frequency-weighted and unweighted sound level measurements, the physical quantity is sound intensity (proportional to pressure squared), expressed in Pascals, and referred to 20 microPascals. The multiplier (C) on the logarithm is 10, the SI definition of the decibel.

For Perceived Level measurements, however, the quantity of concern is loudness (measured in sones), with respect to an implied reference quantity of 1 sone. Further, the multiplier of the common logarithm is not 10, but 9 (for reasons discussed by Stevens, 1971). This multiplier does not meet the standardized (SI) definition of a decibel, and must therefore be called something else. The calculation of Perceived Level addresses the well-documented non-linear growth of loudness with increasing sound intensity, especially pronounced at frequencies below about 400 Hz. Therefore, the perceived loudness expressed in a decibel-like manner behaves differently from sound intensity expressed in decibels. The decision made by Stevens (1971) to use Perceived Level (or loudness) decibels (PLdB) as his unit of measurement remains in common use today.

PLDNL: Abbreviation for Day-Night Average Perceived Sound Level

Prompt annoyance judgment: A judgment of the annoyance of a single sonic boom.

SPSS: Statistical Package for the Social Sciences (see IBM SPSS Statistics for Windows in the Reference section

Table of data set variable names used in current analyses: (See Table 76 below.)

Table 76: Data set variables used in analyses

Variable Mnemonic	Description	Values
PARTICIPANT_ID	Participant identification number	Integers ranging from 100258 to 892998 (408 unique values)

BOOM_ID	Boom identification number	1, 2, 3, …, 52 (integers)
Group	Reminder group to which the participant was assigned	1, 2, 3, 4 (integers)
E6	Response to single-event survey question E6: *Did you hear a sonic thump?*	0, 1 (integers), NA
Exc_Future	Should we exclude the case because it refers to an event too far in the future from when the SE report was initiated? Exclusion criteria is anything more than 900 seconds in the future	0, 1 (integers)
HA	Binary for whether response was highly annoyed or not	0, 1 (integers)
Annoy	Response to single-event survey question E7: *How much did the sonic thump bother, disturb, or annoy you?*	1, 2, 3, 4, 5 (integers)
Vibration	Response to single-event survey question E10: *Vibration is a motion. The motion may be seen, felt or heard. Rattle is a type of noise that can occur when objects move due to a vibration. Did you see, hear, or feel vibration or rattle?*	0, 1 (integers)
Startle	Response to single-event survey question E11: *Did the sonic thump startle you?*	0, 1 (integers)
BOOM_DAY	Date of thump event (11/5-11/8, 11/10-11/11, 11/13-11/14)	1, 2, …, 9 (integers)
BOOM_NUMBER	Boom sequence number	1 to 52, integers
BOOM_START_TIME	Start time of boom	MM/DD/YYYY HH::MM format
BOOM_END_TIME	End time of boom	MM/DD/YYYY HH::MM format
PARTICIPANT_GRP	Reminder group for which the participant was assigned to	1, 2, 3, 4 (integers)
HEARD	Whether participant heard thump or not. But I suspect not 100% match with the E6 column[2]	0, 1 (integers)
LAT	Latitude of participant	Continuous: 0, 10 to ~33
LON	Longitude of participant	Continuous: 0, -75 to -117
MATCH_TYPE	Matching location for calculating metric?[1]	1 to 9 (integers)
PL	Perceived Level (dB)	Continuous (rounded to nearest tenths)
USED_IN_ANALYSIS_PL	Binary for whether response was included in dose-response analysis using each metric in QSF18 APS report	0 or 1 (integers)

[1] Verbatim from data set coding key. Analysis revealed this to be true.
[2] Verbatim from data set coding key. Presume the "?" refers to any dose metric of choice.

THIS PAGE INTENTIONALLY LEFT BLANK

12 REFERENCES

Borsky, P. (**1965**) "Community reactions to sonic booms in the Oklahoma City area", USAF Aerospace Medical Research Laboratory, Report AMRL-TR-65-37, Wright-Patterson Air Force Base, OH.

du Toit, M. (**2020**) Personal communication, May 18, 2020

Federal Interagency Committee on Noise (FICON) (**1992**). "Federal Agency Review of Selected Airport Noise Analysis Issues," Report for the Department of Defense, Washington, DC.

Fidell, S. (ed.) (**1996**) "Assessment of Community Response to High Energy Impulsive Sounds," National Research Council, National Academy Press, Washington, D.C.

Fidell, S., Mestre, V., Schomer, P., Berry, B., Gjestland, T., Vallet, M., and Reid, T. (**2011**). "A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure," J. Acoust. Soc. Am. 130(2), 791–806.

Fidell, S., Horonjeff, R., and Harris, M., (**2012**) "Pilot Test of a Novel Method for Assessing Community Response to Low-Amplitude Sonic Booms," NASA/CR-2012-217767.

Fidell, S., Schultz, T. J., and Green, D. M. (**1988**). "A theoretical interpretation of the prevalence rate of noise-induced annoyance in residential populations." Journal of the Acoustical Society of America 84, 2109-2113.

Fidell, S., Mestre, V., Schomer, P., Horonjeff, R., and Reid, T., (**2014**) "A systematic rationale for defining the significance of aircraft noise impacts," J. Acoust. Soc. Am., 136 (3), 1129-1138.

Fidell, S., and Horonjeff, R., (**2019**) "Field Evaluations of Sampling, Interviewing, and Flight Tracking of NASA's Low Boom Flight Demonstrator Aircraft", NASA/CR -2019-22057, March, 2019.

Fields, J.M. (**1997**) "Reactions of Residents to Long-Term Sonic Boom Noise Environments," NASA Contractor Report 201704, Contract NAS 1-20103, NASA Contractor Report 201704.

IBM Corp. (**2019**) IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp.

International Organization for Standardization (ISO) (**2016**). Acoustics – Description, Measurement and Assessment of Environmental Noise - Part 1: Basic Quantities and Assessment Procedures, ISO 1996-1, Geneva, Switzerland.

Lee, J., Rathsam, J., and Wilson, A. (**2019**) Statistical Modeling of Quiet Sonic Boom Community Response Survey Data, NASA/TM-2019-220427.

Maglieri, D. (**2019**) Personal communication.

Messer, B., and Dillman, D., (**2011**) “Surveying the General Public over the Internet Using Address-Based Sampling and Mail Contact Procedures”, Public Opinion Quarterly, Volume 75, Issue 3, pp. 429–457.

Mestre, V., Fidell, S. Horonjeff, R., Schomer, P., Hastings, A., Tabachnick, B, and Schmitz, F. (**2017**) "Assessing Community Annoyance of Helicopter Noise," Research Report 181, Transportation Research Board of the National Academies.

Mestre, V., Horonjeff, R., and Fidell, S. (**2016**) "Accounting for non-zero asymptotic annoyance prevalence rates in community tolerance level analyses," presentation at the 171st Meeting of the Acoustical Society of America, Salt Lake City, Utah, Session 1pNS.

Morgenstern, J., Buonanno, M., Nordstrud, N. (**2012**) “N+2 low boom wind tunnel model design and validation”, AIAA 2012-3217, 30th AIAA Applied Aerodynamics Conference, New Orleans, LA.

Page, J., Hogdon, K., Hunte, R., Davis, D., Gaugler, T., Downs, R., Cowart, R., Maglieri, D., Hobbs, C., Baker, G., Collmar, M., Bradley, K., Sonak, B., Crom, D., and Cutler, C. (**2020**) “Quiet Supersonic Flights 2018 (QSF18): Galveston, Texas”, NASA Langley Research Center, Final Technical Report, NASA CR-2014-218180; see also Final Technical Report, NASA/CR–2020-220589. See also NASA/CR–2020-220589/Appendices/Volume II.

Raudenbush, S.W., Bryk, A.S, & Congdon, R. (**2004**). “HLM 6 for Windows” Skokie, IL: Scientific Software International, Inc.

Stevens, S.S. (**1972**) “Perceived Level of Noise by Mark VII and Decibels (E.), J. Acoust. Soc. Am. 51, 575-601.

Tabachnick, B. G., and Fidell, L. S. (**2019**). “Using Multivariate Statistics,” 7th ed. Boston: Pearson.

Taraldsen, G., Gelderblom, F., and Gjestland, T. (**2016**) “Measurement of aircraft noise Community tolerance levels” J. Acous. Soc. Am 140, 692.

13 APPENDIX A: DOSE-RESPONSE ANALYSES IN UNITS OTHER THAN PERCEIVED LEVEL

This Appendix repeats analyses of Chapters 4 and 5 (conducted in units of Perceived Level) in alternate units, for the sake of completeness. For delayed response annoyance judgments, the alternate units are BDNL and DDNL; for the single event (prompt response) annoyance judgments, the alternate units are BSEL and DSEL. These analyses are not further interpreted, as they are based on sound level measurements optimized for the circumstances of exposure to low-amplitude sonic booms produced by an F-18 inverted dive maneuver in the QSF18 data collection, and may not be directly applicable to estimation of exposure levels created by carpet booms that the X-59 aircraft will produce in straight and level supersonic flight.

Parameters whose *p* value is less than .001 (other than the *Intercept* of *Noise level*) have been highlighted in light green. Parameters whose *p* value are less than .001 in the Perceived Level analyses, but greater than .001 with B- or D-weighted analyses, have been highlighted in light blue. No B- or D-weighted parameters had *p* values less .001 where Perceived Level parameters did not.

13.1 Dose-Response Relationships

13.1.1 *Dose-Response relationships for prompt responses*

Table 77 through Table 82 show MLM logistic fit parameter values to Equation 1. The subscripted β coefficients correspond to those in the expression for y in the equation. BSEL and DSEL replace PL in the expression for y.

Table 77: Logistic multi-level model of prompt Dose-Response relationship between BSEL and high annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.506	0.916	-14.75	352	<.001			
NoiseDose (β_{10})	0.121	0.012	9.84	4,162	<.001	1.129	1.102	1.156

Table 78: Logistic multi-level model of prompt Dose-Response relationship between BSEL and at least moderate annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-12.438	0.681	-18.26	352	<.001			
NoiseDose (β_{10})	0.126	0.009	13.53	4,162	<.001	1.134	1.114	1.155

Table 79: Logistic multi-level model of prompt Dose-Response relationship between BSEL and at least slight annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	Upper
Intercept (β_{00})	-9.245	0.419	-22.08	352	<.001			
NoiseDose (β_{10})	0.104	0.006	18.30	4,162	<.001	1.107	1.095	1.119

Table 80: Logistic multi-level model of prompt Dose-Response relationship between DSEL and high annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	Upper
Intercept (β_{00})	-16.031	0.971	-16.518	352	<.001			
NoiseDose (β_{10})	0.152	0.013	11.87	4,162	<.001	1.164	1.135	1.194

Table 81: Logistic multi-level model of prompt Dose-Response relationship between DSEL and at least moderate annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	Upper
Intercept (β_{00})	-15.012	0.756	-19.85	352	<.001			
NoiseDose (β_{10})	0.158	0.010	15.59	4,162	<.001	1.171	1.148	1.194

Table 82: Logistic multi-level model of prompt Dose-Response relationship between DSEL and at least slight annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	Upper
Intercept (β_{00})	-10.904	0.456	-23.89	352	<.001			
NoiseDose (β_{10})	0.122	0.006	20.30	4,162	<.001	1.129	1.116	1.143

13.1.2 *Dose-Response relationships for delayed responses*

Table 83 through Table 88 show MLM logistic fit parameter values to Equation 1. The subscripted β coefficients correspond to those in the expression for y in the equation. BDNL and DDNL replace PL in the expression for y.

Table 83: Logistic multi-level model of delayed Dose-Response relationship between BDNL and high annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.783	0.277	-28.08	383	<.001			
NoiseDose (β_{10})	0.103	0.010	10.39	1,947	<.001	1.109	1.088	1.129

Table 84: Logistic multi-level model of delayed Dose-Response relationship between BDNL and at least moderate annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-6.937	0.342	-20.29	383	<.001			
NoiseDose (β_{10})	0.121	0.011	11.02	1,947	<.001	1.128	1.104	1.153

Table 85: Logistic multi-level model of delayed Dose-Response relationship between BDNL and at least slight annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-5.141	0.371	-13.88	383	<.001			
NoiseDose (β_{10})	0.104	0.011	9.22	1,947	<.001	1.110	1.086	1.135

Table 86: Logistic multi-level model of delayed Dose-Response relationship between DDNL and high annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-9.043	0.400	-22.55	383	<.001			
NoiseDose (β_{10}))	0.138	0.013	10.46	1,947	<.001	1.148	1.119	1.178

Table 87: Logistic multi-level model of delayed Dose-Response relationship between BDNL and at least moderate annoyance judgments

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.994	0.422	-18.96	383	<.001			
NoiseDose (β_{10})	0.149	0.013	11.26	1,947	<.001	1.161	1.131	1.192

Table 88: Logistic multi-level model of delayed Dose-Response relationship between BDNL and at least slight annoyance

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.994	0.422	-18.96	383	<.001			
NoiseDose (β_{10})	0.149	0.013	11.26	1,947	<.001	1.161	1.131	1.192

13.2 Other Dose-Response Inferential Analyses

13.2.1 *Effects of reminder messages on prompt Dose-Response relationships*

Table 89 through Table 94 show MLM logistic fit parameter values to Equation 2. The subscripted β coefficients correspond to those in the expression for *y* in the equation. BSEL and DSEL replace PL in the expression for *y*.

Table 89: Model of prompt high annoyance due to BSEL, type of reminder message, and interactions[36]

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-14.017	1.472	-9.53	352	<.001			
NoiseDose (β_{10})	0.125	0.019	6.44	4,158	<.001	1.133	1.091	1.177
ReminderNone (β_{20})	0.393	1.818	0.22	4,158	.829	1.482	0.042	52.23
EmailText(β_{30})	-2.725	2.113	-1.29	4,158	.198	0.066	.001	4.146
ReminderNone x BSEL (β_{40})	-.001	0.024	-0.05	4,158	.960	0.999	0.952	1.047
EmailText x BSEL(β_{30})	0.039	0.029	1.35	4,158	.179	1.040	0.982	1.101

Table 90: Model of prompt moderate or greater annoyance due to BSEL, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-11.578	1.479	-7.83	352	<.001			
NoiseDose (β_{10})	0.116	0.020	5.88	4,158	<.001	1.123	1.081	1.168
ReminderNone (β_{20})	-1.266	1.723	-0.74	4,158	.462	0.282	0.010	8.249
EmailText(β_{30})	0.077	1.735	0.04	4,158	.965	1.079	0.036	32.399
ReminderNone x BSEL (β_{40})	0.014	0.023	0.62	4,158	.532	1.015	0.970	1.061

[36] Parameters whose *p* values are less than .001 in the Perceived Level analyses, but not in B- or D-weighted sound level units, have been highlighted in light blue.

							95% CI	
EmailText x BSEL(β_{30})	-0.000	0.024	-0.02	4,158	.179	1.000	0.970	1.161

Table 91: Model of prompt slight or greater annoyance due to BSEL, type of reminder message, and interactions

							95% CI	
Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	Lower	Upper
Intercept (β_{00})	-8.353	0.822	-10.16	352	<.001			
NoiseDose (β_{10})	0.091	0.011	8.34	4,158	<.001	1.095	1.072	1.118
ReminderNone (β_{20})	-1.403	0.918	-1.53	4,158	.126	0.246	0.041	1.487
EmailText(β_{30})	2.476	0.974	2.54	4,158	.011	11.893	1.762	80.269
ReminderNone x BSEL (β_{40})	0.017	0.012	1.37	4,158	.169	1.017	0.993	1.042
EmailText x BSEL(β_{30})	-0.029	0.013	-2.19	4,158	.029	0.971	0.047	1.997

Table 92: Model of prompt high annoyance due to DSEL, type of reminder message, and interactions[37]

							95% CI	
Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	Lower	Upper
Intercept (β_{00})	-20.516	2.181	-9.41	352	<.001			
NoiseDose (β_{10})	0.206	0.027	7.56	4,158	<.001	1.228	1.165	1.296
ReminderNone (β_{20})	4.541	2.316	1.96	4,158	.050	93.825	1.003	8776.44
EmailText(β_{30})	-2.596	2.094	-1.24	4,158	.216	0.075	1.001	4.519
ReminderNone x DSEL (β_{40})	-0.054	0.029	-1.84	4,158	.065	0.947	0.894	1.003
EmailText x DSEL(β_{30})	0.037	0.013	-2.19	4,158	.029	1.038	0.981	1.098

Table 93: Model of prompt moderate or greater annoyance due to DSEL, type of reminder message, and interactions

							95% CI	
Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	Lower	Upper
Intercept (β_{00})	-16.125	1.532	-10.53	352	<.001			
NoiseDose (β_{10})	0.174	0.020	8.83	4,158	<.001	1.190	1.145	1.236
ReminderNone (β_{20})	1.330	1.577	0.84	4,158	.399	3.782	0.172	83.238
EmailText(β_{30})	0.121	2.094	-1.24	4,158	.216	0.075	1.001	4.519

[37] Parameters whose *p* values are less than .001 in the Perceived Level analyses, but not in B- or D-weighted sound level units, have been highlighted in light blue.

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
ReminderNone x DSEL (β_{40})	-0.054	1.687	0.07	4,158	.943	0.947	0.894	1.003
EmailText x DSEL(β_{30})	0.037	0.013	-2.19	4,158	.029	1.129	0.041	30.806

Table 94: Model of prompt slight or greater annoyance due to DSEL, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.529	0.842	-12.50	352	<.001			
NoiseDose (β_{10})	0.118	0.011	10.76	4,158	<.001	1.125	1.102	1.150
ReminderNone (β_{20})	-0.677	0.857	-0.79	4,158	.430	0.508	0.095	2.724
EmailText(β_{30})	2.384	0.967	2.47	4,158	.014	10.849	1.630	72.226
ReminderNone x DSEL (β_{40})	0.007	0.115	0.56	4,158	.943	1.007	0.984	1.029
EmailText x DSEL(β_{30})	-0.027	0.013	-2.09	4,158	.036	0.973	0.948	0.998

13.2.2 *Effects of home vs. work location on prompt Dose-Response relationships*

Table 95 through Table 100 show MLM logistic fit parameter values to Equation 4. The subscripted β coefficients correspond to those in the expression for *y* in the equation. BSEL and DSEL replace PL in the expression for *y*.

Table 95: Model of prompt high annoyance due to BSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.361	0.892	-14.98	352	<.001			
NoiseDose (β_{10})	0.119	0.012	9.91	4,160	<.001	1.127	1.100	1.153
HomeWork (β_{20})	-1.411	1.338	-1.05	4,160	.292	0.244	0.018	3.361
HomeWork x BSEL (β_{30})	0.018	0.019	0.96	4,160	.340	1.018	0.981	1.056

Table 96: Model of prompt moderate or greater annoyance due to BSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-12.539	0.711	-17.62	352	<.001			
NoiseDose (β_{10})	0.127	0.010	12.93	4,160	<.001	1.135	1.113	1.157
HomeWork (β_{20})	0.849	0.942	0.90	4,160	.368	2.336	0.369	14.797
HomeWork x BSEL (β_{30})	-0.006	0.013	-0.48	4,160	.633	0.994	0.696	1.019

Table 97: Model of prompt slight or greater annoyance due to BSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-9.198	0.434	-21.17	352	<.001			
NoiseDose (β_{10})	0.100	0.006	17.54	4,160	<.001	1.102	1.093	1.118
HomeWork (β_{20})	-0.240	0.637	-0.38	4,160	.707	0.787	0.226	2.744
HomeWork x BSEL (β_{30})	0.008	0.009	0.824	4,160	.410	1.008	0.989	

Table 98: Model of prompt high annoyance due to DSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-15.770	0.935	-16.86	352	<.001			
NoiseDose (β_{10})	0.149	0.012	11.99	4,160	<.001	1.160	1.132	1.189
HomeWork (β_{20})	-2.973	1.375	-2.16	4,160	.030	0.051	0.003	0.757
HomeWork x DSEL (β_{30})	0.037	0.019	1.96	4,160	.049	1.038	1.000	1.078

Table 99: Model of prompt moderate or greater annoyance due to DSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-15.104	0.792	-19.07	352	<.001			
NoiseDose (β_{10})	0.158	0.011	14.88	4,160	<.001	1.171	1.147	1.196
HomeWork (β_{20})	0.800	1.111	0.72	4,160	.472	2.224	0.252	19.616
HomeWork x DSEL (β_{30})	0.005	0.148	-0.36	4,160	.717	0.995	00.966	1.024

Table 100: Model of prompt slight or greater annoyance due to DSEL, home *vs.* work location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.951	0.471	-23.23	352	<.001			
NoiseDose (β_{10})	0.122	0.006	19.79	4,160	<.001	1.129	1.116	1.143
HomeWork (β_{20})	0.306	0.682	0.45	4,160	.653	1.358	0.357	5.169
HomeWork x DSEL (β_{30})	<.001	0.010	0.03	4,160	.977	1.000	0.981	1.020

13.2.3 *Effects of determination of panelist locations*

Table 101 through Table 106 show MLM logistic fit parameter values to Equation 5. The subscripted β coefficients correspond to those in the expression for *y* in the equation. BSEL and DSEL replace PL in the expression for *y*.

Table 101: Model of prompt high annoyance due to BSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-13.123	0.970	-13.53	352	<.001			
NoiseDose (β_{10})	0.118	0.131	9.00	4,160	<.001	1.125	1.097	1.155
Determined (β_{20})	-5.113	3.307	-1.55	4,160	.122	0.006	0.000	1.148
Determined x BSEL (β_{30})	0.053	0.043	1.24	4,160	.217	1.055	0.969	1.148

Table 102: Model of prompt moderate or greater annoyance due to BSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-12.113	0.788	-15.36	352	<.001			
NoiseDose (β_{10})	0.122	0.011	11.21	4,160	<.001	1.130	1.106	1.154
Determined (β_{20})	-1.345	1.132	-1.19	4,160	.235	0.261	0.028	2.396
Determined x Determined x BSEL (β_{30})	0.018	0.016	1.13	4,160	.259	1.018	0.987	1.050

Table 103: Model of prompt slight or greater annoyance due to BSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-8.780	0.479	-18.33	352	<.001			
NoiseDose (β_{10})	0.096	0.006	15.19	4,160	<.001	1.100	1.087	1.114
Determined (β_{20})	-1.809	0.813	-2.22	4,160	.026	0.164	0.033	0.807
Determined x Determined x BSEL (β_{30})	0.022	0.011	1.98	4,160	.047	1.023	1.000	1.046

Table 104: Model of prompt high annoyance due to DSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-15.393	1.010	-15.25	352	<.001			
NoiseDose (β_{10})	0.146	0.013	10.84	4,160	<.001	1.157	1.127	1.188
Determined (β_{20})	-10.373	4.229	-2.45	4,160	.014	0.000	0.000	0.124
Determined x DSEL (β_{30})	0.118	0.054	2.20	4,160	.028	1.125	1.013	1.249

Table 105: Model of prompt moderate or greater annoyance due to DSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-14.618	0.874	-16.73	352	<.001			
NoiseDose (β_{10})	0.153	0.012	13.01	4,160	<.001	1.165	1.139	1.192
Determined (β_{20})	-1.171	1.354	-1.26	4,160	.206	0.181	0.013	2.564
Determined x DSEL (β_{30})	0.022	0.182	1.21	4,160	.227	1.022	0.986	1.060

Table 106: Model of prompt slight or greater annoyance due to DSEL, determined *vs.* Undetermined location, and interaction

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-10.435	0.514	-20.28	352	<.001			
NoiseDose (β_{10})	0.116	0.007	17.21	4,160	<.001	1.123	1.108	1.138
Determined (β_{20})	-1.843	0.987	-1.87	4,160	.062	0.158	0.023	1.096
Determined x DSEL (β_{30})	0.022	0.013	1.68	4,160	.093	1.023	0.996	1.050

14 APPENDIX B: INTRODUCTION TO CTL ANALYSIS

This appendix, adapted in part from Fidell *et al*. (2014), summarizes mathematical relationships among the parameters central to CTL analysis. Greater detail about the development of these relationships may be found in Fidell *et al.* (2011).

14.1 CTL *vs.* Regression-Based Dose-Response Analysis

Univariate logistic regression has been for several decades the conventional statistical technique used to develop dose-response relationships between aircraft noise exposure and the prevalence of high annoyance in communities.[38] Fidell *et al.* (2011) introduced an alternate approach to developing dose-response relationships which accounts for notably more of the variance in community response data than regression based on noise exposure alone. The additional variance is accounted for by a single explanatory variable in addition to noise exposure: the Community Tolerance Level (abbreviated CTL, and represented in mathematical expressions as L_{ct}). CTL values are estimates of a DNL value at which half of a community describes itself as highly annoyed by transportation noise exposure (and half does not.) The preferred estimates of CTL values are derived as maximum likelihood ratio fits of empirical observations of the prevalence of high annoyance at various DNL values to a growth function with a pre-specified slope, rather than as a simple two parameter (slope/intercept) curve fit.

CTL-based predictions of annoyance prevalence rates are based on the observation that annoyance with transportation noise exposure grows at a rate very similar to the rate of growth of duration-adjusted ("effective") loudness with sound level (Stevens, 1972). Fidell *et al.* (2011) and Schomer *et al.* (2012) have shown that the fits of social survey data sets to effective loudness functions can be found by first converting DNL values for interviewing sites in the same community into a noise level, *m*, calculated as $m = (10^{(\mathrm{DNL}/10)})^{0.3}$.

Annoyance prevalence rates for the calculated noise dose are then predicted as $p(\mathrm{HA}) = e^{-(A/m)}$, where *A* is a scalar, non-acoustic decision criterion originally defined by Fidell, Schultz and Green (1988). The dose parameter, *m*, controls the rate of growth of annoyance on the ordinate of a dose-response relationship, while the decision criterion parameter, *A*, translates the growth function along the abscissa.

[38] Various forms of multivariate regression analysis have more recently been applied to infer relationships between noise exposure as well as other variables on the prevalence of noise-induced annoyance. Predictor variables other than noise exposure have little relevance to regulatory decision making, however. FAA, for example, does not recognize any formal role in aircraft noise regulatory policy for individual-level influences on annoyance prevalence rates, including demographic characteristics, personality traits, ability to habituate to noise, beliefs about misfeasance and malfeasance, economic dependence on noise sources, fear of aircraft crashes, and sensitivity to other noise sources (street traffic, commercial, and military aircraft, *etc.*)

14.2 Basic relationships

Equations B-1 through B-6 describe relationships among the parameters needed to infer the quantities *A* (a community's annoyance decision criterion) and Community Tolerance Level from a set of social survey findings on the prevalence of high annoyance with noise exposure. Equation B-1 shows the basic relationship among these parameters:

$$p(HA) = e^{(-A/m)} \qquad \text{Equation B-1}$$

where:

p(HA) is the proportion of the population highly annoyed by noise exposure,
A is a community's annoyance decision criterion, and
m is the noise level, defined as:

$$m = \left(10^{\left(\frac{L_{\mathrm{dn}}}{10}\right)}\right)^{0.3} \qquad \text{Equation B-2}$$

14.2.1 *Functional form of effective loudness function*

The functional relationship specified by Eq. B-1 is a sigmoid. The inferred value of *A* is the one that yields the maximum likelihood between the actual paired observations of noise exposure and annoyance prevalence rates (L_{dn}, %HA) in the empirical data set. The sigmoid anchored to the abscissa by a particular value of the scalar variable, *A*.

The value of DNL at which 50% of the population is highly annoyed by noise exposure is calculated from the value of *A* that produces the maximum likelihood estimator (MLE).

14.2.2 *Relationship between A and CTL*

The relationship between CTL and A is shown in Equation B-3.

$$L_{ct} = 33.3 \log_{10}(A) + 5.32 \qquad \text{Equation B-3}$$

If CTL for a community is known, then the annoyance decision criterion, *A*, may be calculated from CTL as shown in Equation B-4:

$$A = 10^{\left(\frac{L_{ct} - 5.32}{33.3}\right)} \qquad \text{Equation B-4}$$

Likewise, the proportion of the population highly annoyed at a specific DNL may be computed from Equation B-5:

$$p(HA) = e^{\frac{-10^{\left[{Lct - 5.32}/{33.3}\right]}}{\left(10^{\left({Ldn}/{10}\right)}\right)^{0.3}}} \qquad \text{Equation B-5}$$

14.3 Potential adjustment for non-zero asymptotic annoyance rates

In some cases, empirically measured annoyance prevalence rates may not approach zero percent annoyance at low exposure levels. Instead, the prevalence of high annoyance approaches a non-zero residual value, and does not decrease further at lower levels of noise exposure. A minor modification to Eq. B-1, shown below in Eq. B-6, can account for a non-zero annoyance prevalence rate at low exposure levels.

$$p(HA) = r + [1 - r] \bullet [e^{(-A/m)}] \qquad \text{Equation B-6}$$

The parameter r represents the proportion of highly annoyed panelists whose responses are not level-dependent. The term (1-r) acts as a compression factor on the sigmoidal exponent term, so that the sum of the r and (1-r) terms ranges from r to 1.

THIS PAGE INTENTIONALLY LEFT BLANK

15 APPENDIX C: INTERPRETING TABLES SUMMARIZING MULTI-LEVEL LOGISTIC REGRESSION ANALYSES

This Appendix is intended as an aid to interpreting tables elsewhere in this report describing the results of multi-level regression models. Readers interested in further detail may find it in Tabachnick and Fidell (2019).

15.1 Parameter Estimates

$$\boldsymbol{p(annoyed = yes) = \frac{e^y}{1+e^y}} \qquad \textbf{Equation C-1}$$

where:

$y = \beta_{00} + \beta_{10}$ *(NoiseDose)* $+ \beta_{20}$ *(another level 1 predictor)* $+ \cdots + \beta_{01}$ *(level 2 predictor)* $+ \beta_{02}$ *(another level 2 predictor)* $+ \cdots + \beta_{0k}$

and:

p(annoyed = yes) is the predicted fraction of annoyance judgments in which a boom is judged to be annoying to a degree meeting or exceeding a criterion level of annoyance: high annoyance (HA), at least moderate annoyance (MA), or at least slight annoyance (SA).

β_{00} = intercept for predicted fraction of *annoyance* (defined as either HA, MA, or SA)

β_{10} = average dose-response relationship *i.e.*, slope for *NoiseDose* (in PLdB): either PLDNL (delayed responses) or PL (prompt responses, averaged over all panelists

NoiseDose = noise dose in metric-appropriate units (dB or PLdB)

β_{20} = effect of another level 1 predictor

$\beta_{03\ldots0k}$ = effect of additional level 1 predictors, *k* = number of level 1 predictors

β_{01} = effect of first level 2 predictor

$\beta_{02\ldots0l}$ = effect of additional level 2 predictors, *l* = number of level 2 predictors, and

Table 20 is reproduced below as Table 107 as an example of interpretation of tables showing results of inferential dose-response relationships derived through multi-level logistic regression in a two-level model with one level1 and four level 2 predictors.

Table 107: Model of high annoyance due to Perceived Level of single event exposure, type of reminder message, and interactions

Parameter	Parameter estimate	Std Error	*t*-ratio	Approx. df	*p*-value	Odds Ratio	95% CI Lower	95% CI Upper
Intercept (β_{00})	-7.603	0.405	-18.76	381	<.001			
NoiseDose(β_{10})	0.085	0.011	7.87	1,946	<.001	1.088	1.066	1.111
Reminder v Not (β_{01})	-0.328	0.423	-0.78	381	.438	0.720	0.314	1.651
EmailText(β_{02})	-2.078	0.351	-5.92	381	<.001	0.125	0.063	0.250
Reminder x NoiseDose(β_{03})	0.027	0.013	2.00	381	.045	1.027	1.001	1.055
Type x NoiseDose(β_{04})	0.048	0.013	3.70	381	<.001	1.088	1.023	1.077

When a second level 1 predictor (*e.g.*, day of exposure) is present, β_{20} = Day of exposure parameter estimate

15.1.1 *Interpretation of parameter estimates*

Interpretation of any parameter estimate implies that all other predictors in the model are held constant at their own means. For example, noise exposure level is held constant at its mean value, and so on. Likewise, interpretation of the parameter estimate for *ReminderNone* in Table 107 includes adjustments for *NoiseDose* and *EmailText* and both interactions. In other words, all other variables are held constant.

Likewise, interpretation of the sign of a parameter estimate is straightforward. A positive sign for the parameter estimate of a continuous variable means that the probability of annoyance increases as the value of the predictor variable (*e.g.*, noise level) increases. For a categorical variable (*e.g.*, ReminderNone) the interpretation of the sign depends on the coding of the variable. For example, if reminder is coded 1 if present and 0 if absent, a positive coefficient means that annoyance is greater for those who are reminded than for those who are not sent reminder messages.

15.1.2 *Interpretation of odds ratios*

Odds ratios contribute to the interpretation of the magnitude of a predictor variable's influence. They vary around one, which signifies a chance effect. The farther the odds ratio is from one (either smaller or larger, varying from zero to infinity) the stronger is the effect. For example, an odds ratio of 2 for noise level would indicate that the odds of responding "highly annoyed" would double with each 1dB increase in noise level. The actual odds of 1.088 indicate a much smaller increase in likelihood of high annoyance: each single decibel increase in noise level increases the likelihood of a high annoyance response by about 9%. An odds ratio less than one indicates a negative relationship. For the effect of e-mail *vs.* text reminder messages, the odds ratio of 0.125 indicates that a response of high annoyance is one-eighth as likely to be associated with an e-mail reminder message than with a text message one.

Odds ratios are highly dependent on units of measurement and coding of categorical variables. For example, the coding of reminder messages *vs.* none is (0, 1), so that a one-unit change indicates the difference between reminders and no reminders. The coding of e-mail *vs.* text reminder messages is (-.5, 0, .5) with no-reminder panelists coded zero, so that panelists who received no-reminder messages are included in the analysis, but do not contribute to the test of reminder message type. Therefore, the difference between e-mail and text reminder messages is also a one-unit change in the coding category.

15.2 Definitions of Other Inferential Statistics

Std Error is the standard deviation of the effect (*e.g.*, Noise level) divided by $\sqrt{N}$, where N is the total number of events or days.

***t*-ratio** is the parameter estimate divided by the Std Error (carried to more decimal places than reported in the tables of Chapter 3.)

Approx. df is the sample size (N) minus the number of parameters estimated (not all of which appear in the tables of Chapters 4, 5, 0, 9 and Appendix A.)

***p* value** is the probability of achieving a parameter estimate that much different from zero if, in reality, the true value is zero.

Odds ratio is the change in odds of inclusion in one of the categories of outcome (*e.g.* HA) when the value of a predictor (*e.g.*, noise exposure) increases by one unit.

15.3 Interaction Terms

An interaction considers the relationship among (at least) three variables: a dependent variable (here, prevalence of annoyance), and two predictors (here, noise level and one other predictor). An interaction tests whether the relationship between the dependent variable and one predictor (here, annoyance and noise level) is different when the other predictor (*e.g.*, e-mail *vs.* text reminder messages) changes. Consider, for example, whether the relationship between prevalence of annoyance and noise level (*i.e.*, the dose-response relationship) differs for e-mail and text reminder messages. The finding of a statistically significant interaction indicates that the dose-response relationship is different for the two types of messages. Conversely, the relationship between prevalence of annoyance and a predictor (*e.g.*, e-mail *vs.* text messages) depends on the other predictor (*i.e.*, noise level).

When an interaction effect is statistically significant, the associated main effects cannot be interpreted in isolation. For example, a finding of a significant difference in prevalence of annoyance between e-mail and text reminder messages cannot be directly interpreted because that difference will depend on noise level. Further, the finding of a significant effect of noise level cannot be directly interpreted because the effect of noise level will depend on whether e-mail or text reminder messages were sent.

15.4 Confidence Intervals

Calculation and interpretation of confidence intervals around dose-response functions derived by logistic multi-level regression is neither as simple nor straightforward as is the case for simple linear regression. The nature of the units in which the predicted (dependent) variable is plotted on the ordinate is part of the complexity. The dependent variable is not expressed as a linear quantity, but rather as the natural logarithm of an odds ratio. The bounds of confidence intervals must therefore be interpreted in terms of odds ratios.

Confidence bounds for logistic regression are not simply expressed as a likelihood that infinite numbers of repetitions of similar analyses would yield dose-response functions within specifiable limits. Further, it is not uncommon for confidence intervals for dose-response functions derived by logistic multi-level regression to be so wide that they are unhelpful for most purposes, even with large numbers of degrees of freedom. Worse yet, alternate methods for calculating such intervals do not necessarily yield the same bounds.

A method suggested by du Toit[39] for calculating such bounds is described below. Note that the method does not necessarily yield the smallest confidence intervals.

$y = \alpha + \beta \cdot x$ **Equation C-2**

$Var(y) = Var(\alpha) + x^2 \cdot Var(\beta)$ **Equation C-3**

$Var(y) = [StdErr(\alpha)]^2 + x^2 \cdot [StdErr(\beta)]^2$ **Equation C-4**

$CI_{95} = \pm 1.96 \cdot sqrt[Var(y)]$ **Equation C-5**

Hence, the 95% CI is a function of x^2 (*i.e.*, PL^2)

$P(Annoy = 1) = e^y / (1 + e^y)$ **Equation C-5**

Therefore, the confidence bound curves are:

$P_{upper} = e^{(y + CI95)} / [1 + e^{(y + CI95)}]$ **Equation C-6**

and

[39] Mathilda du Toit, Vector Psychometric Group, LLC, personal correspondence, 18 May 2020.

$$P_{lower} = e^{(y\ -\ CI95)} / [1 + e^{(y\ -\ CI95)}]$$ **Equation C-7**

Application of the above equations to Table 14 (for slight or greater annoyance judgments in the entire set of prompt annoyance judgments eligible for analysis) yields the confidence intervals shown by the short-dashed curves shown below. (Note that the lower bound of the 95% confidence interval hugs the abscissa over a 40 dB range of exposure values; and that the width of the confidence interval at 75 PLdB extends from roughly zero to nearly 90% prevalence of annoyance).

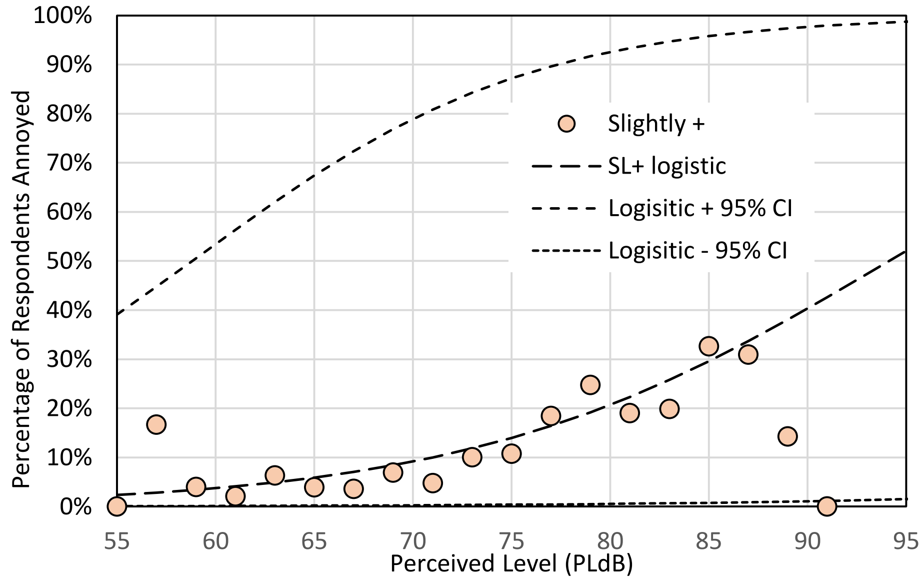

Figure 72. 95% confidence interval bounds for dose-response function of Table 14

THIS PAGE INTENTIONALLY LEFT BLANK

16 APPENDIX D: ATTRITION IN PANELIST PARTICIPATION AND NUMBERS OF CASES BY PARTICIPANT ID

This appendix shows the reductions (losses) in numbers of panelists and numbers of analyzable records at various stages in the data grooming process. The data screening process was conducted in an orderly progression in order to observe the reduction in remaining numbers of contributing panelists and analyzable records following each stage in the process. The prompt response and delayed response data sets are discussed separately in the two subsections below.

16.1 Prompt Responses

Table 108 lists each of the panelists found in the single event database along with the reductions in numbers of analyzable records for each following a series of stratified data screening stages. Column numbers are shown at the head of each column to facilitate the discussion below. The second row shows four stages in the data screening process. The third row identifies the screening criterion being applied at each stage (discussed in some detail below). The fourth and fifth rows show totals for each column. The number of participants is a count of those where one or more records remain after the present and all previous screening stages have been completed. The number of records is a count of the number remaining following the screening stages. These are shown in bold type face.

Column 1 lists all of the unique panelist identifications found in the NASA-supplied data base. In the second column the numbers of records associated with each panelist is shown. The two header rows above these numbers labeled "Panelists" and "Records" show the numbers of participants and numbers of cases remaining follow each stage in the screening process.

Stage 1 of the screening process is shown in Columns 3 and 4. Column 3 tabulates the number of unique cases found for each panelist (that is the number of unique Boom_ID values). Several instances were found where more than one record existed with the same Boom_ID. Since only one combination or Participant_ID and Boom_ID one could be considered eligible for analysis, all but one of the records required elimination. The numbers of such additional records beyond the first is shown in Column 4 (the difference between columns 2 and 3). Non-zero values are shown in red for visual identification purposes.

Stage 2 of the screening process is shown in Columns 5 and 6. Multiple records of a Participant_ID and Boom_ID were found in which the response categories codings did not match across them. That created uncertainty regarding which ones were valid and which were not. A decision was made to remove all such records to avoid adding uncertainty noise to the analyses. Column 5 shows the number of cases remaining after inconsistent multiples had been eliminated. Column 6 shows the reduction in number of unique case records due to inconsistent responses between the multiples.

Stage 3 in the screening process is shown in columns 7 and 8. Column 7 shows the number of cases remaining after the participant's location (Match_Type) was screened for analyzable locations. Column 8 shows the reduction in number of unique case records due to unanalyzable match types.

Several additional parameters were investigated in Stage 4 of the screening process. These included (a) a noise dose (PL) of zero, (b) an annoyance rating greater than "not at all" when the boom was identified as not noticed, (c) the latitude and longitude coded as zero, (d) reminders being sent *prior to* the occurrence of the boom, (e) the boom occurring too far in the future[40] from when the response began, and (f) disagreement being the two variables ("E6" and "Heard"). Column 9 tabulates numbers of cases remaining after these screening criteria were applied. The blue-highlighted cells indicate panelists eliminated from the panelist pool because all of their records were removed during the various stages of screening. Column 10 shows the reduction in number of unique case records due to these additional screenings.

Column 11 shows the reduction in analyzable records across all screening stages. Column 12 shows this loss as a percentage of the original number of records in the database. Pink highlighting in these cells means a reduction at one screening stage or another occurred. The preponderance of pink-highlighted cells indicates record reductions occurred with most participants for one reason or another. Only 64 of the original 408 unique Participant_IDs in the database experienced no reductions in numbers of records.

Table 108: Effects of single event (prompt response) database screening stages by panelist

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
Panelist	408	408	0	406	2	388	18	354	34	54	
Records	7,068	6,312	756	6,041	271	5,530	511	**4,169**	1,361	2,899	
100258	19	17	2	17	0	17	0	13	4	6	32%
100948	16	14	2	12	2	11	1	8	3	8	50%
101814	9	9	0	9	0	9	0	6	3	3	33%
109776	1	1	0	1	0	1	0	0	1	1	100%
112870	34	29	5	29	0	29	0	18	11	16	47%
115081	37	30	7	25	5	22	3	13	9	24	65%
116471	39	35	4	34	1	34	0	27	7	12	31%
120917	1	1	0	1	0	1	0	1	0	0	0%
121398	7	7	0	7	0	7	0	5	2	2	29%

[40] "Too far in the future" was a determination made by the QSF18 contractor during the creation of the single event data set. The data set variable Exc_Future (coded as "0" for inclusion and "1" for exclusion) was the sole criterion used for this purpose in the current analysis. Please refer to the contractor report for further explanation and discussion of this variable.

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		**Stage 1**		**Stage 2**		**Stage 3**		**Stage 4**			
Panelist ID	**Total Records**	**Unique Cases**	**Loss**	**Agreeing Cases**	**Incr. Loss**	**Valid MT**	**Incr. Loss**	**Further Screen**	**Incr. Loss**	**Total Loss N**	**%**
126248	8	8	0	8	0	8	0	7	1	1	13%
126677	4	4	0	4	0	4	0	2	2	2	50%
126947	1	1	0	1	0	1	0	1	0	0	0%
128233	9	9	0	9	0	8	1	6	2	3	33%
130926	7	6	1	5	1	4	1	0	4	7	100%
131930	20	18	2	16	2	16	0	12	4	8	40%
131994	2	2	0	2	0	2	0	2	0	0	0%
137709	25	22	3	20	2	20	0	16	4	9	36%
139826	6	6	0	6	0	6	0	0	6	6	100%
141784	3	3	0	3	0	3	0	3	0	0	0%
142241	47	41	6	39	2	39	0	33	6	14	30%
144270	30	25	5	22	3	22	0	16	6	14	47%
144866	10	10	0	10	0	10	0	8	2	2	20%
150290	3	3	0	3	0	3	0	2	1	1	33%
150597	38	32	6	29	3	29	0	21	8	17	45%
152896	4	4	0	4	0	4	0	4	0	0	0%
155247	21	19	2	17	2	17	0	14	3	7	33%
160529	9	8	1	7	1	7	0	7	0	2	22%
161151	1	1	0	1	0	1	0	1	0	0	0%
161244	4	4	0	4	0	4	0	2	2	2	50%
161802	2	2	0	2	0	2	0	2	0	0	0%
162170	11	10	1	9	1	9	0	8	1	3	27%
162581	42	37	5	34	3	24	10	24	0	18	43%
163112	4	4	0	4	0	0	4	3	-3	1	25%
163161	4	4	0	4	0	3	1	2	1	2	50%
163612	30	26	4	26	0	26	0	5	21	25	83%
164112	39	32	7	28	4	28	0	23	5	16	41%
164562	2	1	1	0	1	0	0	0	0	2	100%
165139	10	10	0	10	0	10	0	8	2	2	20%
165666	17	16	1	16	0	13	3	7	6	10	59%
166343	6	6	0	6	0	6	0	4	2	2	33%
167712	3	3	0	3	0	3	0	2	1	1	33%
169109	43	37	6	32	5	31	1	2	29	41	95%
169512	54	46	8	43	3	32	11	26	6	28	52%
172915	49	34	15	31	3	30	1	16	14	33	67%
173288	1	1	0	1	0	1	0	1	0	0	0%
173820	26	24	2	22	2	22	0	19	3	7	27%
177516	1	1	0	1	0	1	0	1	0	0	0%
179570	38	36	2	36	0	36	0	19	17	19	50%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
182787	39	36	3	35	1	34	1	25	9	14	36%
183378	13	10	3	10	0	10	0	7	3	6	46%
183539	16	14	2	12	2	12	0	10	2	6	38%
186013	12	10	2	9	1	9	0	8	1	4	33%
186204	2	2	0	2	0	2	0	1	1	1	50%
190332	3	3	0	3	0	3	0	3	0	0	0%
190531	7	7	0	7	0	6	1	5	1	2	29%
192410	8	6	2	4	2	3	1	3	0	5	63%
194099	42	34	8	31	3	31	0	26	5	16	38%
198212	32	26	6	24	2	24	0	17	7	15	47%
198863	17	15	2	15	0	14	1	14	0	3	18%
200664	1	1	0	1	0	1	0	0	1	1	100%
203019	11	11	0	11	0	11	0	8	3	3	27%
204676	27	24	3	22	2	22	0	20	2	7	26%
206578	23	19	4	16	3	14	2	14	0	9	39%
206620	37	35	2	34	1	32	2	26	6	11	30%
206913	8	8	0	8	0	8	0	7	1	1	13%
210469	22	18	4	17	1	17	0	9	8	13	59%
210659	22	21	1	21	0	21	0	15	6	7	32%
211611	17	17	0	17	0	17	0	15	2	2	12%
212335	24	24	0	24	0	24	0	21	3	3	13%
213461	1	1	0	1	0	1	0	0	1	1	100%
213700	37	32	5	29	3	29	0	24	5	13	35%
214043	46	40	6	37	3	35	2	27	8	19	41%
215176	49	46	3	45	1	45	0	32	13	17	35%
216486	3	3	0	3	0	3	0	3	0	0	0%
222695	11	11	0	11	0	11	0	7	4	4	36%
222736	1	1	0	1	0	1	0	1	0	0	0%
225045	49	39	10	33	6	32	1	26	6	23	47%
225650	22	19	3	18	1	18	0	16	2	6	27%
225999	36	32	4	31	1	29	2	9	20	27	75%
228288	2	2	0	2	0	2	0	0	2	2	100%
228940	11	9	2	8	1	8	0	6	2	5	45%
229287	44	41	3	40	1	39	1	24	15	20	45%
233227	29	25	4	23	2	22	1	0	22	29	100%
234309	12	12	0	12	0	12	0	11	1	1	8%
234933	1	1	0	1	0	1	0	0	1	1	100%
235184	2	2	0	2	0	2	0	1	1	1	50%
235646	15	14	1	13	1	12	1	7	5	8	53%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
236260	2	2	0	2	0	2	0	1	1	1	50%
244259	21	19	2	18	1	18	0	15	3	6	29%
244285	2	2	0	2	0	2	0	1	1	1	50%
245375	3	3	0	3	0	3	0	0	3	3	100%
245721	5	5	0	5	0	5	0	4	1	1	20%
250693	26	23	3	23	0	22	1	18	4	8	31%
250994	25	20	5	20	0	19	1	8	11	17	68%
257056	41	35	6	34	1	34	0	25	9	16	39%
257310	2	2	0	2	0	2	0	2	0	0	0%
258139	1	1	0	1	0	1	0	0	1	1	100%
259435	9	9	0	9	0	9	0	9	0	0	0%
260422	4	3	1	3	0	3	0	2	1	2	50%
260960	1	1	0	1	0	1	0	1	0	0	0%
262147	18	18	0	18	0	17	1	10	7	8	44%
266716	32	28	4	25	3	24	1	21	3	11	34%
267495	2	2	0	2	0	2	0	1	1	1	50%
267528	10	10	0	10	0	10	0	10	0	0	0%
268706	3	3	0	3	0	3	0	3	0	0	0%
269230	4	4	0	4	0	4	0	4	0	0	0%
269876	23	20	3	17	3	17	0	15	2	8	35%
270183	24	19	5	14	5	14	0	10	4	14	58%
270554	8	8	0	8	0	8	0	0	8	8	100%
270850	27	25	2	24	1	24	0	15	9	12	44%
271187	36	33	3	31	2	31	0	18	13	18	50%
272935	8	8	0	8	0	7	1	7	0	1	13%
273014	33	30	3	30	0	30	0	27	3	6	18%
273623	7	7	0	7	0	2	5	6	-4	1	14%
274065	45	43	2	43	0	39	4	35	4	10	22%
274718	10	10	0	10	0	10	0	8	2	2	20%
276216	6	6	0	6	0	6	0	6	0	0	0%
277473	1	1	0	1	0	1	0	1	0	0	0%
279254	41	40	1	40	0	40	0	30	10	11	27%
282811	2	2	0	2	0	2	0	1	1	1	50%
283495	17	15	2	13	2	13	0	5	8	12	71%
283938	46	46	0	46	0	44	2	36	8	10	22%
285341	34	31	3	29	2	24	5	18	6	16	47%
285357	2	1	1	0	1	0	0	0	0	2	100%
286177	1	1	0	1	0	1	0	1	0	0	0%
287362	41	38	3	38	0	38	0	33	5	8	20%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
287759	49	40	9	37	3	37	0	28	9	21	43%
288613	6	6	0	6	0	6	0	0	6	6	100%
289707	3	3	0	3	0	3	0	3	0	0	0%
290590	16	14	2	13	1	13	0	12	1	4	25%
290783	20	17	3	16	1	16	0	14	2	6	30%
290975	1	1	0	1	0	0	1	1	-1	0	0%
293790	46	44	2	43	1	40	3	37	3	9	20%
296283	17	17	0	17	0	17	0	9	8	8	47%
298398	8	7	1	6	1	6	0	4	2	4	50%
302663	2	2	0	2	0	2	0	2	0	0	0%
307193	11	11	0	11	0	11	0	6	5	5	45%
308089	42	40	2	40	0	37	3	25	12	17	40%
309007	15	13	2	13	0	12	1	8	4	7	47%
309576	4	3	1	3	0	3	0	3	0	1	25%
311614	5	5	0	5	0	5	0	3	2	2	40%
314445	28	28	0	28	0	28	0	17	11	11	39%
315444	1	1	0	1	0	1	0	1	0	0	0%
315462	2	2	0	2	0	2	0	0	2	2	100%
319504	39	36	3	36	0	36	0	30	6	9	23%
322616	2	2	0	2	0	2	0	1	1	1	50%
325466	32	30	2	28	2	27	1	21	6	11	34%
325816	1	1	0	1	0	1	0	1	0	0	0%
326690	3	2	1	2	0	2	0	2	0	1	33%
329099	13	13	0	13	0	12	1	10	2	3	23%
331385	7	6	1	6	0	6	0	4	2	3	43%
332655	1	1	0	1	0	1	0	1	0	0	0%
335236	5	4	1	3	1	3	0	1	2	4	80%
336443	1	1	0	1	0	1	0	1	0	0	0%
337845	15	14	1	14	0	14	0	12	2	3	20%
344197	1	1	0	1	0	1	0	1	0	0	0%
348952	2	2	0	2	0	2	0	0	2	2	100%
349820	2	2	0	2	0	2	0	2	0	0	0%
351401	1	1	0	1	0	1	0	1	0	0	0%
351940	1	1	0	1	0	1	0	1	0	0	0%
353000	3	3	0	3	0	2	1	2	0	1	33%
353406	51	46	5	45	1	43	2	39	4	12	24%
353704	1	1	0	1	0	1	0	1	0	0	0%
353869	1	1	0	1	0	1	0	1	0	0	0%
355419	1	1	0	1	0	1	0	0	1	1	100%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
360360	50	43	7	42	1	40	2	30	10	20	40%
361987	1	1	0	1	0	1	0	0	1	1	100%
361997	35	30	5	27	3	26	1	22	4	13	37%
362478	2	2	0	2	0	2	0	2	0	0	0%
364113	2	2	0	2	0	2	0	2	0	0	0%
364231	34	32	2	32	0	30	2	17	13	17	50%
365571	6	6	0	6	0	6	0	2	4	4	67%
366877	39	33	6	32	1	32	0	17	15	22	56%
366915	12	12	0	12	0	12	0	4	8	8	67%
369023	49	44	5	44	0	44	0	22	22	27	55%
369160	8	8	0	8	0	8	0	7	1	1	13%
369737	2	2	0	2	0	1	1	0	1	2	100%
372391	2	2	0	2	0	2	0	0	2	2	100%
379554	2	2	0	2	0	2	0	1	1	1	50%
380161	42	38	4	38	0	37	1	26	11	16	38%
384523	1	1	0	1	0	1	0	0	1	1	100%
384588	29	28	1	28	0	28	0	21	7	8	28%
384720	20	18	2	18	0	18	0	10	8	10	50%
391171	10	8	2	8	0	8	0	4	4	6	60%
394057	10	10	0	10	0	9	1	10	-1	0	0%
399729	3	2	1	1	1	1	0	0	1	3	100%
403686	2	2	0	2	0	2	0	0	2	2	100%
404904	31	25	6	24	1	24	0	11	13	20	65%
409897	44	38	6	37	1	37	0	24	13	20	45%
413188	4	4	0	4	0	4	0	2	2	2	50%
414687	6	6	0	6	0	6	0	5	1	1	17%
417005	43	39	4	39	0	39	0	33	6	10	23%
417587	10	4	6	4	0	4	0	0	4	10	100%
418994	1	1	0	1	0	1	0	0	1	1	100%
420438	31	25	6	25	0	24	1	7	17	24	77%
424160	2	2	0	2	0	2	0	0	2	2	100%
426190	33	28	5	26	2	26	0	22	4	11	33%
429660	3	3	0	3	0	3	0	2	1	1	33%
431928	1	1	0	1	0	1	0	1	0	0	0%
434485	24	19	5	19	0	18	1	14	4	10	42%
436779	50	46	4	43	3	43	0	28	15	22	44%
438184	40	35	5	34	1	33	1	22	11	18	45%
442466	33	32	1	31	1	30	1	27	3	6	18%
443494	10	8	2	6	2	6	0	3	3	7	70%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
443892	7	7	0	7	0	7	0	4	3	3	43%
446633	37	35	2	35	0	34	1	33	1	4	11%
447455	1	1	0	1	0	1	0	0	1	1	100%
449813	41	32	9	30	2	29	1	0	29	41	100%
450493	1	1	0	1	0	1	0	0	1	1	100%
450562	5	5	0	5	0	5	0	0	5	5	100%
450743	4	4	0	4	0	4	0	4	0	0	0%
456502	44	39	5	38	1	35	3	0	35	44	100%
457137	47	44	3	44	0	42	2	34	8	13	28%
458610	47	40	7	39	1	38	1	23	15	24	51%
460803	1	1	0	1	0	1	0	1	0	0	0%
461413	1	1	0	1	0	0	1	0	0	1	100%
463052	1	1	0	1	0	1	0	0	1	1	100%
463932	44	41	3	40	1	40	0	33	7	11	25%
467846	14	13	1	13	0	13	0	10	3	4	29%
477668	34	31	3	30	1	29	1	18	11	16	47%
478991	35	30	5	27	3	26	1	7	19	28	80%
483756	3	3	0	3	0	1	2	2	-1	1	33%
483903	2	2	0	2	0	2	0	0	2	2	100%
487753	31	28	3	28	0	25	3	19	6	12	39%
488164	4	4	0	4	0	2	2	3	-1	1	25%
504612	4	3	1	2	1	2	0	2	0	2	50%
507188	41	36	5	35	1	33	2	23	10	18	44%
510037	2	2	0	2	0	1	1	1	0	1	50%
511423	1	1	0	1	0	1	0	0	1	1	100%
512334	41	35	6	31	4	31	0	20	11	21	51%
514026	2	2	0	2	0	2	0	2	0	0	0%
514147	1	1	0	1	0	1	0	0	1	1	100%
515366	14	13	1	13	0	13	0	12	1	2	14%
515653	1	1	0	1	0	1	0	0	1	1	100%
516450	34	31	3	30	1	30	0	18	12	16	47%
517748	1	1	0	1	0	1	0	1	0	0	0%
517762	32	30	2	28	2	28	0	27	1	5	16%
520615	11	10	1	10	0	9	1	4	5	7	64%
524371	16	15	1	15	0	14	1	9	5	7	44%
525392	1	1	0	1	0	1	0	1	0	0	0%
525562	1	1	0	1	0	1	0	0	1	1	100%
526001	4	3	1	3	0	3	0	3	0	1	25%
526329	1	1	0	1	0	1	0	1	0	0	0%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
528236	8	6	2	5	1	5	0	2	3	6	75%
528620	29	25	4	24	1	24	0	9	15	20	69%
530102	1	1	0	1	0	1	0	0	1	1	100%
531555	2	2	0	2	0	2	0	0	2	2	100%
533120	4	3	1	3	0	3	0	2	1	2	50%
533365	1	1	0	1	0	1	0	1	0	0	0%
536215	3	3	0	3	0	3	0	0	3	3	100%
539515	3	3	0	3	0	3	0	3	0	0	0%
542868	1	1	0	1	0	1	0	0	1	1	100%
546908	8	7	1	6	1	5	1	3	2	5	63%
553850	23	22	1	21	1	21	0	17	4	6	26%
556639	4	4	0	4	0	4	0	4	0	0	0%
562500	1	1	0	1	0	1	0	1	0	0	0%
566024	38	36	2	35	1	34	1	30	4	8	21%
568255	41	37	4	37	0	37	0	29	8	12	29%
570134	4	3	1	3	0	3	0	3	0	1	25%
570350	35	32	3	32	0	32	0	6	26	29	83%
571394	28	24	4	23	1	22	1	14	8	14	50%
576289	51	47	4	47	0	47	0	40	7	11	22%
576779	1	1	0	1	0	1	0	0	1	1	100%
579917	14	11	3	10	1	10	0	7	3	7	50%
582052	5	5	0	5	0	5	0	3	2	2	40%
584666	20	17	3	17	0	17	0	7	10	13	65%
592272	7	7	0	7	0	7	0	5	2	2	29%
593581	14	7	7	7	0	7	0	4	3	10	71%
595751	5	5	0	5	0	5	0	3	2	2	40%
595994	31	30	1	30	0	29	1	25	4	6	19%
596668	9	7	2	7	0	7	0	6	1	3	33%
597827	4	4	0	4	0	4	0	0	4	4	100%
599238	3	3	0	3	0	2	1	3	-1	0	0%
603420	8	7	1	7	0	7	0	5	2	3	38%
605377	27	19	8	19	0	17	2	13	4	14	52%
607325	35	28	7	27	1	25	2	22	3	13	37%
607674	4	4	0	4	0	4	0	2	2	2	50%
610353	42	24	18	24	0	22	2	13	9	29	69%
613051	33	30	3	30	0	26	4	25	1	8	24%
613148	43	40	3	39	1	39	0	30	9	13	30%
617734	3	3	0	3	0	3	0	2	1	1	33%
623599	37	33	4	33	0	32	1	26	6	11	30%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
624206	1	1	0	1	0	1	0	1	0	0	0%
624897	32	29	3	27	2	27	0	17	10	15	47%
625114	4	4	0	4	0	4	0	4	0	0	0%
626076	3	3	0	3	0	3	0	2	1	1	33%
633249	6	5	1	4	1	4	0	4	0	2	33%
636171	3	3	0	3	0	3	0	2	1	1	33%
641282	14	13	1	13	0	13	0	5	8	9	64%
654323	16	15	1	15	0	13	2	12	1	4	25%
662507	38	34	4	33	1	33	0	24	9	14	37%
665075	1	1	0	1	0	1	0	1	0	0	0%
668572	1	1	0	1	0	1	0	1	0	0	0%
672671	15	15	0	15	0	15	0	7	8	8	53%
673869	32	27	5	26	1	26	0	22	4	10	31%
674103	4	4	0	4	0	4	0	1	3	3	75%
682401	6	6	0	6	0	6	0	5	1	1	17%
682496	2	2	0	2	0	2	0	0	2	2	100%
684915	10	9	1	9	0	9	0	3	6	7	70%
698939	4	4	0	4	0	4	0	4	0	0	0%
705267	32	30	2	30	0	30	0	23	7	9	28%
708789	13	13	0	13	0	13	0	12	1	1	8%
709289	9	9	0	9	0	9	0	8	1	1	11%
710574	27	24	3	24	0	23	1	10	13	17	63%
711692	7	7	0	7	0	7	0	6	1	1	14%
712711	19	17	2	16	1	15	1	13	2	6	32%
712921	50	46	4	44	2	44	0	36	8	14	28%
718045	12	11	1	11	0	11	0	9	2	3	25%
721778	7	7	0	7	0	7	0	5	2	2	29%
723830	41	34	7	31	3	31	0	21	10	20	49%
727478	2	2	0	2	0	2	0	2	0	0	0%
727893	41	38	3	35	3	33	2	31	2	10	24%
730598	6	5	1	4	1	4	0	3	1	3	50%
732344	23	21	2	21	0	20	1	3	17	20	87%
733260	2	2	0	2	0	1	1	1	0	1	50%
734586	48	45	3	44	1	43	1	35	8	13	27%
737109	11	10	1	9	1	9	0	8	1	3	27%
739713	18	17	1	17	0	17	0	14	3	4	22%
742019	28	24	4	21	3	19	2	14	5	14	50%
744183	9	9	0	9	0	9	0	8	1	1	11%
744732	53	42	11	32	10	32	0	23	9	30	57%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		**Stage 1**		**Stage 2**		**Stage 3**		**Stage 4**			
Panelist ID	**Total Records**	**Unique Cases**	**Loss**	**Agreeing Cases**	**Incr. Loss**	**Valid MT**	**Incr. Loss**	**Further Screen**	**Incr. Loss**	**Total Loss N**	**Total Loss %**
747832	6	6	0	6	0	6	0	5	1	1	17%
753190	18	17	1	17	0	17	0	15	2	3	17%
755415	44	38	6	37	1	36	1	28	8	16	36%
756496	52	40	12	35	5	35	0	0	35	52	100%
757711	8	7	1	7	0	6	1	7	-1	1	13%
758847	7	7	0	7	0	7	0	7	0	0	0%
760231	7	7	0	7	0	7	0	5	2	2	29%
765789	36	31	5	31	0	31	0	9	22	27	75%
766541	15	13	2	12	1	11	1	11	0	4	27%
768270	30	24	6	22	2	22	0	7	15	23	77%
769354	3	3	0	3	0	3	0	3	0	0	0%
770424	37	33	4	31	2	29	2	18	11	19	51%
770742	45	39	6	36	3	36	0	28	8	17	38%
772257	11	11	0	11	0	11	0	9	2	2	18%
773134	12	10	2	8	2	8	0	0	8	12	100%
773289	33	32	1	31	1	30	1	22	8	11	33%
774133	39	34	5	31	3	29	2	26	3	13	33%
775161	7	6	1	5	1	5	0	5	0	2	29%
775264	36	32	4	30	2	30	0	26	4	10	28%
775615	47	45	2	44	1	44	0	34	10	13	28%
779314	39	36	3	36	0	35	1	11	24	28	72%
781032	12	11	1	10	1	10	0	10	0	2	17%
781468	5	5	0	5	0	5	0	5	0	0	0%
783080	31	29	2	29	0	27	2	23	4	8	26%
783792	36	34	2	33	1	33	0	27	6	9	25%
787247	1	1	0	1	0	1	0	0	1	1	100%
789717	7	7	0	7	0	7	0	6	1	1	14%
793598	1	1	0	1	0	1	0	0	1	1	100%
795999	33	32	1	32	0	29	3	27	2	6	18%
796277	44	39	5	36	3	36	0	32	4	12	27%
797406	31	30	1	29	1	28	1	21	7	10	32%
799650	44	36	8	33	3	33	0	26	7	18	41%
799921	2	2	0	2	0	2	0	1	1	1	50%
800651	44	35	9	34	1	34	0	19	15	25	57%
801257	1	1	0	1	0	1	0	1	0	0	0%
802904	17	16	1	16	0	16	0	14	2	3	18%
803572	10	10	0	10	0	10	0	4	6	6	60%
808476	22	21	1	20	1	20	0	16	4	6	27%
809028	20	17	3	16	1	14	2	13	1	7	35%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
809189	5	5	0	5	0	5	0	5	0	0	0%
811059	1	1	0	1	0	1	0	0	1	1	100%
811339	15	13	2	12	1	12	0	10	2	5	33%
812339	51	49	2	48	1	37	11	37	0	14	27%
813795	4	1	3	1	0	1	0	1	0	3	75%
815397	7	7	0	7	0	7	0	6	1	1	14%
817759	8	7	1	6	1	6	0	4	2	4	50%
818062	6	6	0	6	0	6	0	6	0	0	0%
823119	3	2	1	1	1	1	0	0	1	3	100%
831746	9	8	1	7	1	7	0	6	1	3	33%
836494	32	30	2	29	1	29	0	24	5	8	25%
836818	48	39	9	37	2	36	1	26	10	22	46%
837719	16	16	0	16	0	16	0	14	2	2	13%
838579	28	25	3	23	2	23	0	20	3	8	29%
840686	1	1	0	1	0	1	0	0	1	1	100%
844512	14	14	0	14	0	14	0	12	2	2	14%
844519	37	34	3	33	1	33	0	24	9	13	35%
844926	41	37	4	36	1	35	1	28	7	13	32%
845029	40	32	8	29	3	29	0	24	5	16	40%
847095	2	2	0	2	0	2	0	1	1	1	50%
848379	31	27	4	23	4	23	0	18	5	13	42%
850567	12	11	1	10	1	7	3	1	6	11	92%
853607	9	9	0	9	0	9	0	7	2	2	22%
854051	3	3	0	3	0	3	0	2	1	1	33%
858711	11	9	2	8	1	8	0	7	1	4	36%
858998	36	33	3	32	1	31	1	25	6	11	31%
862198	20	17	3	14	3	14	0	11	3	9	45%
865261	1	1	0	1	0	1	0	1	0	0	0%
867371	7	7	0	7	0	7	0	4	3	3	43%
869364	1	1	0	1	0	1	0	1	0	0	0%
870057	3	3	0	3	0	3	0	1	2	2	67%
873493	17	15	2	14	1	14	0	9	5	8	47%
874161	12	11	1	11	0	11	0	0	11	12	100%
874599	36	30	6	24	6	0	24	21	-21	15	42%
875638	41	36	5	34	2	0	34	28	-28	13	32%
876331	14	13	1	13	0	0	13	8	-8	6	43%
876826	43	39	4	38	1	0	38	31	-31	12	28%
877408	13	13	0	13	0	0	13	10	-10	3	23%
878941	5	5	0	5	0	0	5	2	-2	3	60%

1	2	3	4	5	6	7	8	9	10	11	12
Screening Stage		Stage 1		Stage 2		Stage 3		Stage 4			
Panelist ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Valid MT	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
879927	3	2	1	2	0	0	2	0	0	3	100%
882307	46	37	9	32	5	0	32	27	-27	19	41%
884555	15	15	0	15	0	0	15	12	-12	3	20%
884621	41	38	3	38	0	0	38	31	-31	10	24%
888380	7	5	2	5	0	0	5	3	-3	4	57%
891304	2	2	0	2	0	0	2	2	-2	0	0%
891806	47	45	2	44	1	0	44	29	-29	18	38%
892937	29	25	4	25	0	0	25	20	-20	9	31%
892998	50	42	8	38	4	0	38	31	-31	19	38%

16.2 Delayed responses

Table 109 lists each of the panelists found in the daily database along with the reductions in numbers of analyzable records for each following a series of stratified data screening stages. Column numbers are shown at the head of each column to facilitate the discussion below. The second row shows four stages in the data screening process. The third row identifies the screening criterion being applied at each stage (discussed in some detail below). The fourth and fifth rows show totals for each column. The number of participants is a count of those where one or more records remain after the present and all previous screening stages have been completed. The number of records is a count of the number remaining following the screening stages. These are shown in bold type face.

Column 1 lists all of the unique panelist identifications found in the NASA-supplied data base. In the second column the numbers of records associated with each panelist is shown. The two header rows above these numbers labeled "Partic's" and "Records" show the numbers of participants and numbers of cases remaining follow each stage in the screening process.

Stage 1of the screening process is shown in Columns 3 and 4. Column 3 tabulates the number of unique cases found for each participant (that is the number of unique test days). Several instances were found where more than one record existed with the same numeric test day. Since only one combination or Participant_ID and Test_day could be considered eligible for analysis all but one of the records required elimination. The numbers of such additional records beyond the first occurrence is shown in Column 4 (the difference between columns 2 and 3). Non-zero values are shown in red for visual identification purposes.

Stage 2 in the screening process is shown in Columns 5 and 6. During visual inspection of the data set it was noticed that when multiple records of a Participant_ID and Test_day were found the response categories did not always match across them. That created uncertainty regarding which ones were valid and which were not. A decision was made to remove all such records to avoid adding uncertainty noise to the analyses. Column 5 shows the number of cases remaining after inconsistent multiples had been eliminated. Column 6 shows the reduction in number of unique case records due to inconsistent responses between the multiples.

Stage 3 of screening several additional parameters were investigated. These included (a) the noise dose (PL) being zero and (b) the "USED_IN_ANALYSIS_PLDN" parameter not being set to "1." The blue-highlighted cells indicate panelists eliminated from the pool because all of their records were removed during the various stages of screening. Column 7 tabulates the number of cases remaining after these screening procedures were applied. Column 8 shows the reduction in number of unique case records due to these additional screenings.

Column 9 shows the reduction in analyzable records across all screening stages. Column 10 shows this loss as a percentage of the original number of records in the database. Pink highlighting in these cells means a reduction at one screening stage or another occurred. The preponderance of pink-highlighted cells indicates record reductions occurred with most participants for one reason or another. Only 64 of the original 408 unique Participant_IDs in the database experienced no reductions in numbers of records.

Table 109: Effects of daily (delayed response) database screening stages by participant

1	2	3	4	5	6	7	8	9	10
Screening Stage		**Stage 1**		**Stage 2**		**Stage 3**			
Particip't ID	**Total Records**	**Unique Cases**	**Loss**	**Agreeing Cases**	**Incr. Loss**	**Further Screen**	**Incr. Loss**	**Total Loss N**	**%**
Particip'ts	**408**	**408**	0	**408**	0	**368**	40	40	
Records	**2,855**	**2,695**	160	**2,681**	14	**1,952**	729	903	
100258	5	5	0	5	0	5	0	0	0%
100948	1	1	0	1	0	1	0	0	0%
101814	9	9	0	9	0	9	0	0	0%
109776	1	1	0	1	0	0	1	1	100%
109942	9	9	0	9	0	8	1	1	11%
112870	1	1	0	1	0	1	0	0	0%
115081	1	1	0	1	0	1	0	0	0%
116471	9	9	0	9	0	9	0	0	0%
121398	9	9	0	9	0	9	0	0	0%
126248	6	6	0	6	0	5	1	1	17%
126677	10	9	1	9	0	5	4	5	50%
126947	7	5	2	4	1	4	0	3	43%
128233	2	2	0	2	0	2	0	0	0%
130926	9	8	1	8	0	0	8	9	100%
131930	5	2	3	1	1	1	0	4	80%
131994	4	3	1	3	0	3	0	1	25%
137709	9	9	0	9	0	8	1	1	11%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
139338	3	3	0	3	0	3	0	0	0%
139826	3	3	0	3	0	0	3	3	100%
141784	9	8	1	8	0	8	0	1	11%
142241	6	6	0	6	0	6	0	0	0%
144270	9	9	0	9	0	6	3	3	33%
144866	6	6	0	6	0	6	0	0	0%
150290	6	6	0	6	0	6	0	0	0%
150597	8	8	0	8	0	8	0	0	0%
152896	9	9	0	9	0	9	0	0	0%
155247	9	9	0	9	0	8	1	1	11%
160529	10	9	1	9	0	8	1	2	20%
161151	3	3	0	3	0	2	1	1	33%
161244	8	7	1	7	0	7	0	1	13%
161802	1	1	0	1	0	1	0	0	0%
162170	10	9	1	9	0	8	1	2	20%
162581	9	9	0	9	0	9	0	0	0%
163112	6	6	0	6	0	6	0	0	0%
163161	8	8	0	8	0	8	0	0	0%
163612	9	8	1	8	0	2	6	7	78%
164112	7	7	0	7	0	7	0	0	0%
165139	10	9	1	9	0	9	0	1	10%
165666	1	1	0	1	0	1	0	0	0%
166343	1	1	0	1	0	0	1	1	100%
167194	9	9	0	9	0	9	0	0	0%
167314	9	9	0	9	0	7	2	2	22%
167712	9	9	0	9	0	7	2	2	22%
169109	9	8	1	8	0	2	6	7	78%
169512	3	3	0	3	0	2	1	1	33%
170034	3	3	0	3	0	3	0	0	0%
172915	10	9	1	9	0	8	1	2	20%
173288	4	4	0	4	0	4	0	0	0%
173820	9	8	1	8	0	6	2	3	33%
177516	7	7	0	7	0	7	0	0	0%
179570	9	9	0	9	0	9	0	0	0%
179648	9	9	0	9	0	9	0	0	0%
182787	7	7	0	7	0	7	0	0	0%
183378	4	3	1	2	1	2	0	2	50%
183461	4	4	0	4	0	4	0	0	0%
186013	10	8	2	8	0	8	0	2	20%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
186204	1	1	0	1	0	0	1	1	100%
190332	9	9	0	9	0	9	0	0	0%
190531	7	7	0	7	0	5	2	2	29%
192410	4	4	0	4	0	2	2	2	50%
194099	8	5	3	5	0	5	0	3	38%
198212	7	7	0	7	0	7	0	0	0%
198863	4	4	0	4	0	4	0	0	0%
203019	8	8	0	8	0	8	0	0	0%
204676	9	9	0	9	0	8	1	1	11%
206578	2	2	0	2	0	2	0	0	0%
206620	7	7	0	7	0	7	0	0	0%
206913	7	7	0	7	0	6	1	1	14%
210469	9	9	0	9	0	8	1	1	11%
210841	5	4	1	4	0	0	4	5	100%
211611	11	9	2	8	1	8	0	3	27%
212335	9	8	1	8	0	8	0	1	11%
213461	9	9	0	9	0	7	2	2	22%
213700	9	9	0	9	0	9	0	0	0%
214043	10	9	1	9	0	9	0	1	10%
215176	6	6	0	6	0	6	0	0	0%
216486	2	2	0	2	0	2	0	0	0%
217431	7	7	0	7	0	5	2	2	29%
222695	9	8	1	8	0	7	1	2	22%
222736	4	4	0	4	0	4	0	0	0%
225045	9	9	0	9	0	9	0	0	0%
225650	9	7	2	5	2	5	0	4	44%
225999	6	5	1	5	0	5	0	1	17%
228288	7	7	0	7	0	5	2	2	29%
228940	9	9	0	9	0	9	0	0	0%
229287	9	9	0	9	0	8	1	1	11%
233227	6	6	0	6	0	0	6	6	100%
234309	1	1	0	1	0	1	0	0	0%
234933	5	5	0	5	0	5	0	0	0%
235646	11	8	3	8	0	8	0	3	27%
236260	9	6	3	5	1	5	0	4	44%
236449	2	2	0	2	0	2	0	0	0%
244259	9	9	0	9	0	8	1	1	11%
244285	1	1	0	1	0	1	0	0	0%
245721	10	9	1	8	1	8	0	2	20%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
250693	9	9	0	9	0	9	0	0	0%
250994	8	5	3	5	0	2	3	6	75%
257056	8	8	0	8	0	8	0	0	0%
257310	8	7	1	7	0	5	2	3	38%
258139	9	9	0	9	0	8	1	1	11%
259435	1	1	0	1	0	1	0	0	0%
260422	2	2	0	2	0	2	0	0	0%
260960	10	8	2	8	0	7	1	3	30%
262147	1	1	0	1	0	1	0	0	0%
266549	6	6	0	6	0	6	0	0	0%
266716	7	7	0	7	0	7	0	0	0%
267489	8	8	0	8	0	2	6	6	75%
267495	2	2	0	2	0	1	1	1	50%
267528	6	6	0	6	0	6	0	0	0%
268706	3	3	0	3	0	3	0	0	0%
269230	9	9	0	9	0	9	0	0	0%
269876	10	9	1	8	1	8	0	2	20%
270183	7	6	1	6	0	6	0	1	14%
270554	7	7	0	7	0	0	7	7	100%
270850	6	6	0	6	0	6	0	0	0%
271187	9	9	0	9	0	7	2	2	22%
271843	9	9	0	9	0	9	0	0	0%
272935	8	8	0	8	0	8	0	0	0%
273623	8	8	0	8	0	8	0	0	0%
274065	9	9	0	9	0	8	1	1	11%
274718	3	3	0	3	0	3	0	0	0%
276216	2	2	0	2	0	2	0	0	0%
279254	2	1	1	1	0	1	0	1	50%
285341	8	8	0	8	0	7	1	1	13%
285357	11	8	3	8	0	6	2	5	45%
286177	9	9	0	9	0	9	0	0	0%
287362	8	8	0	8	0	7	1	1	13%
287759	5	5	0	5	0	5	0	0	0%
289707	5	4	1	4	0	4	0	1	20%
290590	9	9	0	9	0	9	0	0	0%
290783	2	2	0	2	0	2	0	0	0%
290975	1	1	0	1	0	1	0	0	0%
292434	7	7	0	7	0	6	1	1	14%
293790	2	2	0	2	0	1	1	1	50%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	Total Loss %
298398	9	9	0	9	0	5	4	4	44%
298649	8	7	1	7	0	4	3	4	50%
308089	6	6	0	6	0	6	0	0	0%
308272	8	8	0	8	0	2	6	6	75%
309007	11	9	2	9	0	2	7	9	82%
309576	5	5	0	5	0	5	0	0	0%
311119	4	1	3	1	0	0	1	4	100%
311614	5	4	1	4	0	2	2	3	60%
314445	2	2	0	2	0	2	0	0	0%
315462	2	2	0	2	0	0	2	2	100%
319504	9	9	0	9	0	9	0	0	0%
320676	9	9	0	9	0	7	2	2	22%
322616	2	2	0	2	0	0	2	2	100%
324366	8	8	0	8	0	3	5	5	63%
325466	11	8	3	8	0	4	4	7	64%
325816	9	9	0	9	0	3	6	6	67%
325911	2	1	1	1	0	0	1	2	100%
326430	4	4	0	4	0	3	1	1	25%
326690	4	3	1	3	0	1	2	3	75%
329099	2	2	0	2	0	1	1	1	50%
329927	5	5	0	5	0	2	3	3	60%
331385	2	2	0	2	0	2	0	0	0%
332655	8	8	0	8	0	5	3	3	38%
335236	8	7	1	7	0	6	1	2	25%
336443	9	9	0	9	0	5	4	4	44%
337845	7	7	0	7	0	7	0	0	0%
342465	9	9	0	9	0	4	5	5	56%
348952	6	5	1	5	0	4	1	2	33%
349820	10	9	1	9	0	6	3	4	40%
351401	10	8	2	8	0	3	5	7	70%
351940	8	7	1	7	0	7	0	1	13%
353000	9	9	0	9	0	6	3	3	33%
353406	9	9	0	9	0	5	4	4	44%
353704	8	8	0	8	0	6	2	2	25%
353869	7	7	0	7	0	6	1	1	14%
355419	5	4	1	4	0	3	1	2	40%
361987	5	5	0	5	0	3	2	2	40%
361997	9	8	1	8	0	5	3	4	44%
364113	8	8	0	8	0	4	4	4	50%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
364231	9	9	0	9	0	4	5	5	56%
365571	9	9	0	9	0	3	6	6	67%
366915	3	3	0	3	0	3	0	0	0%
368889	10	9	1	9	0	3	6	7	70%
369160	9	9	0	9	0	5	4	4	44%
372391	8	8	0	8	0	4	4	4	50%
379554	1	1	0	1	0	0	1	1	100%
379683	3	3	0	3	0	0	3	3	100%
380161	9	9	0	9	0	6	3	3	33%
383706	9	9	0	9	0	8	1	1	11%
384523	7	7	0	7	0	7	0	0	0%
384588	6	6	0	6	0	3	3	3	50%
384720	9	9	0	9	0	9	0	0	0%
391171	9	9	0	9	0	5	4	4	44%
394057	10	7	3	7	0	3	4	7	70%
402440	1	1	0	1	0	1	0	0	0%
403686	6	6	0	6	0	0	6	6	100%
404904	9	8	1	8	0	3	5	6	67%
406330	11	9	2	9	0	5	4	6	55%
409897	4	4	0	4	0	1	3	3	75%
412345	3	3	0	3	0	0	3	3	100%
413188	1	1	0	1	0	1	0	0	0%
413495	9	6	3	6	0	0	6	9	100%
414687	8	8	0	8	0	4	4	4	50%
417005	7	7	0	7	0	7	0	0	0%
417459	1	1	0	1	0	1	0	0	0%
417587	9	9	0	9	0	6	3	3	33%
418222	9	9	0	9	0	6	3	3	33%
418994	5	5	0	5	0	4	1	1	20%
420160	4	4	0	4	0	2	2	2	50%
420438	9	9	0	9	0	8	1	1	11%
424160	4	4	0	4	0	0	4	4	100%
426190	4	4	0	4	0	4	0	0	0%
429660	2	2	0	2	0	1	1	1	50%
431928	10	9	1	9	0	4	5	6	60%
433165	1	1	0	1	0	0	1	1	100%
434485	7	7	0	7	0	4	3	3	43%
436278	6	6	0	6	0	3	3	3	50%
436779	9	9	0	9	0	8	1	1	11%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
438184	4	4	0	4	0	2	2	2	50%
442466	10	8	2	7	1	6	1	4	40%
443494	10	9	1	9	0	8	1	2	20%
443892	9	9	0	9	0	5	4	4	44%
446633	9	9	0	9	0	6	3	3	33%
447455	3	3	0	3	0	2	1	1	33%
448712	10	9	1	9	0	4	5	6	60%
449813	9	9	0	9	0	0	9	9	100%
450493	8	8	0	8	0	5	3	3	38%
450562	2	2	0	2	0	0	2	2	100%
450743	5	4	1	4	0	1	3	4	80%
455695	5	4	1	4	0	3	1	2	40%
456502	10	9	1	9	0	0	9	10	100%
457137	7	7	0	7	0	5	2	2	29%
458610	9	9	0	9	0	6	3	3	33%
460803	2	2	0	2	0	2	0	0	0%
461413	9	8	1	8	0	0	8	9	100%
463052	3	3	0	3	0	1	2	2	67%
463932	7	7	0	7	0	4	3	3	43%
467846	5	5	0	5	0	3	2	2	40%
473992	2	2	0	2	0	2	0	0	0%
477668	9	9	0	9	0	7	2	2	22%
478991	12	8	4	8	0	3	5	9	75%
483188	2	2	0	2	0	0	2	2	100%
483756	4	4	0	4	0	2	2	2	50%
483903	9	9	0	9	0	8	1	1	11%
488164	9	9	0	9	0	9	0	0	0%
488788	5	5	0	5	0	3	2	2	40%
489120	11	9	2	9	0	7	2	4	36%
497113	2	2	0	2	0	0	2	2	100%
497553	9	9	0	9	0	6	3	3	33%
504612	10	9	1	9	0	7	2	3	30%
507188	2	2	0	2	0	1	1	1	50%
510037	8	7	1	7	0	1	6	7	88%
511423	2	2	0	2	0	2	0	0	0%
512334	9	9	0	9	0	8	1	1	11%
514026	7	6	1	6	0	5	1	2	29%
514147	5	5	0	5	0	0	5	5	100%
515366	8	8	0	8	0	6	2	2	25%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
517762	7	7	0	7	0	4	3	3	43%
520615	5	5	0	5	0	1	4	4	80%
525392	9	8	1	8	0	4	4	5	56%
525562	1	1	0	1	0	0	1	1	100%
526001	3	3	0	3	0	0	3	3	100%
526329	4	4	0	4	0	2	2	2	50%
528236	4	4	0	4	0	2	2	2	50%
530102	11	9	2	9	0	4	5	7	64%
531555	3	3	0	3	0	2	1	1	33%
533120	3	3	0	3	0	2	1	1	33%
533365	7	7	0	7	0	3	4	4	57%
536215	2	2	0	2	0	1	1	1	50%
542868	5	5	0	5	0	2	3	3	60%
546908	1	1	0	1	0	0	1	1	100%
553850	8	8	0	8	0	5	3	3	38%
556639	9	9	0	9	0	6	3	3	33%
562500	6	5	1	5	0	5	0	1	17%
566024	5	5	0	5	0	5	0	0	0%
568255	9	9	0	9	0	5	4	4	44%
570134	8	6	2	6	0	2	4	6	75%
570350	3	3	0	3	0	1	2	2	67%
570468	10	9	1	9	0	3	6	7	70%
571394	3	3	0	3	0	2	1	1	33%
576289	9	9	0	9	0	5	4	4	44%
576779	5	5	0	5	0	2	3	3	60%
576901	5	5	0	5	0	1	4	4	80%
579917	4	3	1	3	0	1	2	3	75%
584666	9	8	1	8	0	3	5	6	67%
585235	8	8	0	8	0	1	7	7	88%
588927	9	9	0	9	0	6	3	3	33%
593581	8	7	1	7	0	3	4	5	63%
594435	9	9	0	9	0	5	4	4	44%
595751	6	6	0	6	0	5	1	1	17%
595796	1	1	0	1	0	0	1	1	100%
596668	9	9	0	9	0	5	4	4	44%
597827	7	7	0	7	0	1	6	6	86%
599238	3	3	0	3	0	0	3	3	100%
603420	10	9	1	9	0	6	3	4	40%
605377	4	4	0	4	0	3	1	1	25%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
607325	8	8	0	8	0	8	0	0	0%
607674	3	3	0	3	0	3	0	0	0%
610353	7	7	0	7	0	4	3	3	43%
613148	9	8	1	8	0	5	3	4	44%
614830	9	9	0	9	0	7	2	2	22%
617734	10	8	2	8	0	6	2	4	40%
620503	9	9	0	9	0	4	5	5	56%
623544	9	9	0	9	0	6	3	3	33%
623599	9	9	0	9	0	5	4	4	44%
624206	8	8	0	8	0	4	4	4	50%
625114	3	3	0	3	0	2	1	1	33%
626076	9	9	0	9	0	4	5	5	56%
633249	3	3	0	3	0	1	2	2	67%
636171	9	9	0	9	0	7	2	2	22%
640500	2	2	0	2	0	0	2	2	100%
641282	2	2	0	2	0	1	1	1	50%
654323	10	8	2	8	0	6	2	4	40%
660923	9	8	1	8	0	5	3	4	44%
660992	3	3	0	3	0	1	2	2	67%
662507	4	4	0	4	0	2	2	2	50%
665075	3	3	0	3	0	1	2	2	67%
667812	8	7	1	6	1	0	6	8	100%
668572	8	8	0	8	0	1	7	7	88%
672245	9	9	0	9	0	5	4	4	44%
672671	1	1	0	1	0	0	1	1	100%
673869	9	8	1	8	0	5	3	4	44%
674103	9	8	1	7	1	4	3	5	56%
682166	10	9	1	9	0	5	4	5	50%
682401	10	9	1	9	0	4	5	6	60%
682496	2	2	0	2	0	0	2	2	100%
697676	9	8	1	8	0	7	1	2	22%
698805	9	9	0	9	0	4	5	5	56%
698939	2	2	0	2	0	0	2	2	100%
705267	8	8	0	8	0	5	3	3	38%
708789	8	8	0	8	0	8	0	0	0%
709289	8	8	0	8	0	8	0	0	0%
710574	9	8	1	8	0	3	5	6	67%
711692	10	9	1	9	0	9	0	1	10%
712711	5	5	0	5	0	3	2	2	40%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
718045	2	2	0	2	0	2	0	0	0%
723830	1	1	0	1	0	1	0	0	0%
727478	6	6	0	6	0	4	2	2	33%
727893	9	9	0	9	0	9	0	0	0%
730598	9	9	0	9	0	5	4	4	44%
732344	8	8	0	8	0	1	7	7	88%
733260	9	9	0	9	0	1	8	8	89%
734586	8	7	1	7	0	7	0	1	13%
737109	12	9	3	8	1	8	0	4	33%
739713	3	3	0	3	0	3	0	0	0%
742019	9	9	0	9	0	8	1	1	11%
744183	7	6	1	5	1	0	5	7	100%
744732	9	9	0	9	0	9	0	0	0%
747832	8	8	0	8	0	8	0	0	0%
753190	9	9	0	9	0	6	3	3	33%
755415	4	4	0	4	0	4	0	0	0%
756496	10	9	1	9	0	0	9	10	100%
757711	2	2	0	2	0	0	2	2	100%
758458	7	6	1	6	0	3	3	4	57%
758847	6	6	0	6	0	6	0	0	0%
760231	3	3	0	3	0	3	0	0	0%
766541	1	1	0	1	0	1	0	0	0%
768270	10	9	1	9	0	9	0	1	10%
769354	8	8	0	8	0	7	1	1	13%
770424	9	9	0	9	0	8	1	1	11%
770742	10	9	1	9	0	9	0	1	10%
772257	5	4	1	4	0	3	1	2	40%
773289	9	9	0	9	0	9	0	0	0%
774133	9	9	0	9	0	9	0	0	0%
775161	2	2	0	2	0	2	0	0	0%
775264	9	9	0	9	0	9	0	0	0%
781032	9	9	0	9	0	9	0	0	0%
781468	4	4	0	4	0	4	0	0	0%
782750	9	9	0	9	0	9	0	0	0%
783019	9	9	0	9	0	5	4	4	44%
783792	8	8	0	8	0	8	0	0	0%
787247	9	9	0	9	0	0	9	9	100%
789717	8	8	0	8	0	8	0	0	0%
793598	10	9	1	9	0	6	3	4	40%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
793628	10	9	1	9	0	9	0	1	10%
795999	10	9	1	9	0	9	0	1	10%
796277	9	9	0	9	0	8	1	1	11%
797406	9	9	0	9	0	9	0	0	0%
799650	9	8	1	8	0	8	0	1	11%
799921	8	8	0	8	0	8	0	0	0%
800651	9	8	1	8	0	7	1	2	22%
802904	9	9	0	9	0	9	0	0	0%
803572	5	5	0	5	0	4	1	1	20%
808476	1	1	0	1	0	1	0	0	0%
809028	4	4	0	4	0	4	0	0	0%
809189	9	9	0	9	0	8	1	1	11%
811059	5	5	0	5	0	4	1	1	20%
811339	9	8	1	8	0	7	1	2	22%
813795	3	3	0	3	0	3	0	0	0%
815397	9	9	0	9	0	9	0	0	0%
818062	8	7	1	7	0	7	0	1	13%
821201	8	8	0	8	0	7	1	1	13%
823119	2	2	0	2	0	2	0	0	0%
831746	5	4	1	4	0	3	1	2	40%
836494	9	9	0	9	0	9	0	0	0%
836818	9	9	0	9	0	9	0	0	0%
837719	12	9	3	9	0	9	0	3	25%
838579	9	9	0	9	0	9	0	0	0%
840686	2	2	0	2	0	2	0	0	0%
844512	9	9	0	9	0	8	1	1	11%
844519	8	7	1	7	0	7	0	1	13%
844926	9	9	0	9	0	8	1	1	11%
845029	7	6	1	6	0	6	0	1	14%
847095	8	8	0	8	0	7	1	1	13%
848379	3	3	0	3	0	3	0	0	0%
850567	10	9	1	9	0	6	3	4	40%
853607	9	9	0	9	0	9	0	0	0%
854051	8	8	0	8	0	7	1	1	13%
858711	5	5	0	5	0	2	3	3	60%
858998	11	9	2	9	0	9	0	2	18%
862198	6	5	1	5	0	4	1	2	33%
865261	8	7	1	7	0	6	1	2	25%
867371	9	9	0	9	0	7	2	2	22%

1	2	3	4	5	6	7	8	9	10
Screening Stage		Stage 1		Stage 2		Stage 3			
Particip't ID	Total Records	Unique Cases	Loss	Agreeing Cases	Incr. Loss	Further Screen	Incr. Loss	Total Loss N	%
869364	9	8	1	8	0	3	5	6	67%
870057	5	5	0	5	0	4	1	1	20%
873493	7	7	0	7	0	6	1	1	14%
874161	10	9	1	9	0	0	9	10	100%
874599	8	7	1	7	0	7	0	1	13%
875638	9	9	0	9	0	8	1	1	11%
876331	8	8	0	8	0	8	0	0	0%
876826	9	9	0	9	0	9	0	0	0%
877408	10	9	1	9	0	7	2	3	30%
878941	5	3	2	3	0	2	1	3	60%
879927	3	3	0	3	0	2	1	1	33%
882307	9	9	0	9	0	9	0	0	0%
884555	6	6	0	6	0	6	0	0	0%
884621	5	5	0	5	0	5	0	0	0%
888380	3	3	0	3	0	1	2	2	67%
889110	1	1	0	1	0	1	0	0	0%
891304	3	2	1	2	0	1	1	2	67%
891806	8	8	0	8	0	8	0	0	0%
892937	7	7	0	7	0	7	0	0	0%
892998	9	9	0	9	0	9	0	0	0%

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-08-2020	Contractor Report	

4. TITLE AND SUBTITLE
Independent Analyses of Galveston QSF18 Social Survey

5a. CONTRACT NUMBER
80LARC17C0003

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Sanford Fidell, Richard Horonjeff, Barbara Tabachnick, and Steven Clark

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
110076.02.07.06.12

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, Virginia 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)
NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NASA-CR-20205005471

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified
Subject Category 71
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES
Langley Technical Monitor: Jonathan Rathsam

14. ABSTRACT
A set of independent analyses was performed of the findings of NASA's QSF18 data collection exercise in Galveston, TX. These analyses included development of regression-based and other dosage-response relationships between the prevalence of annoyance with low-amplitude sonic booms and Perceived Level measures of single event and cumulative exposure. They also included detailed analyses of the effects of panelist attrition; of participation reminder messages sent to panelists; the ability to determine panelist locations at the time of exposure; and the effects of rattle and vibration on annoyance judgments, among other topics. The report summarizes the implications of its findings for the design of studies of community response to carpet booms created during future deployments of NASA's X-59 aircraft, and recommends several design measures to improve the efficiency of data collection.

15. SUBJECT TERMS
Galveston; community survey; dose-response relationship; low boom; sonic boom; sonic thump

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
U	U	U	UU	222	STI Help Desk(email help@sti.nasa.gov)
					19b. TELEPHONE NUMBER *(Include area code)* (757) 864-9658