NASA/CR–2019-220257

Field Evaluation of Sampling, Interviewing, and Flight Tracking of NASA's Low Boom Flight Demonstrator Aircraft

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

Richard D. Horonjeff
Consultant, Boxborough, Massachusetts

March 2019

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–2019-220257

Field Evaluations of Sampling, Interviewing, and Flight Tracking of NASA's Low Boom Flight Demonstrator Aircraft

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

Richard D. Horonjeff
Consultant, Boxborough, Massachusetts

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NNL15AA05C

March 2019

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

FINAL PROJECT REPORT
(PHASE II, MILESTONE 15)

Sanford Fidell and Richard Horonjeff

PREPARED FOR:

CONTRACT NNL15AA05C (MOD 12)
National Aeronautics and Space Administration
NASA Langley Research Center
Hampton, VA 23681

PREPARED BY:

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

December 20, 2018

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Tables

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

1 INTRODUCTION

1.1 Purpose of this Document

This document is the work product of Task 3.2.8 of Modification 12 of NASA Contract NNL15AA05C, summarizing three years' effort. The overall intent of the contract was 1) to evaluate risks to field testing of community response to low amplitude sonic booms produced by an experimental supersonic aircraft (NASA's Low Boom Flight Demonstrator, or LBFD aircraft); and 2) to identify measures for minimizing such risks.

1.2 Organization of Report

Chapter 2 summarizes the substantive effort completed in Phase I of Contract NNL15AA05C. The summary paraphrases and condenses the text of the prior Phase I report (Fidell *et al*., 2016), but retains its major figures and tables. The summary omits a mock data analysis that confirmed the workability of a suggested analysis of findings, detailed examples of the recommended geo-information system approach to route planning for LBFD flight missions[1], and preliminary (conceptual) study design recommendations. Readers interested in fuller particulars of the first year's effort may wish to consult the original report.

Chapter 3 summarizes a field demonstration of the Internet-enabled sonic boom measurement system conducted during the second year of Contract NNL15AA05C (Year 1 of Phase II.) Readers interested in further information about non-technical aspects of the second year's work (including, for example, applications to the Office of Management and Budget and NASA's Institutional Review Board) may find it in the original (Fidell *et al*., 2017) report.

Chapter 4 presents the findings of an investigation of interview completion rates in a cross-sectional social survey of a nationally representative sample conducted during the third year of Contract NNL15AA05C (Year 2 of Phase II.) It also discusses the technical rationale for developing information of interest to aircraft noise regulatory agencies; a detailed study design for variant panel sample social surveys (developed after interview completion rates in the cross-sectional study proved impractically low); and observations about the utility of a cross-sectional social survey for assessing delayed reactions to cumulative sonic boom exposure.

Chapter 5 summarizes the substantive findings of each year's efforts, and their implications for additional effort intended to further minimize risks of field testing of community reaction to LBFD flight missions.

---

[1] A modification to Contract NNL15AA05C deleted a requirement for subsequent detailed route planning.

THIS PAGE INTENTIONALLY LEFT BLANK

2 SUMMARY OF FIRST YEAR'S EFFORT (PHASE I)

The initial year's effort under Contract NNL15AA05C was reported by Fidell *et al*. (2016). Their report described the goal of a field study as the development of a highly credible dosage-response relationship between exposure to low amplitude carpet booms created by the Low-Boom Flight Demonstrator (LBFD) aircraft and the prevalence of high annoyance with such booms in overflown residential populations. Development of such a dosage-response relationship requires coordinated collection of information about both the magnitude of community exposure to low amplitude sonic booms, and about population proportions highly annoyed by them.

Sampling and prompt interviewing about short term (single event) responses to sonic boom exposure were identified as the primary risks to construction of a credible dosage-response relationship linking LBFD test flights to the prevalence of a consequential degree of annoyance with sonic boom exposure in overflown communities. Sampling risks are related to the representativeness of solicited opinions. They are linked to the accuracy and completeness of the enumeration of the population eligible for interview, and to the nature of the samples sought (independent, or cross-sectional, vs. panel). Interviewing risks are associated with interview completion rates, and with study design choices about whether to alert respondents to impending exposure to individual booms. The interviewing risks are much greater for assessments of prompt reactions to individual booms than for assessments of delayed responses to cumulative sonic boom exposure over days or weeks.

2.1 LBFD Route Planning

Fidell *et al.* (2016) also addressed the planning and selection of routes for LBFD missions. It considered the sorts of routes and communities to be overflown supersonically; addressed constraints on route selection imposed to avoid unintended exposures to sonic booms; and evaluated three illustrative flight paths and the residential populations within the carpet boom corridors underlying them. The report also described methods for planning LBFD routes, but did not identify or recommend specific routes.

A geographic information system (GIS) approach was recommended as a cost-effective tool for planning LBFD routes. Figure 1 illustrates some of the thematic layers that a GIS application would operate on to devise LBFD routes. Three nominal LBFD routes were created as examples of LBFD deployment routes, as summarized in Table 1. Basing requirements for the LBFD were not found to constrain LBFD field deployments. NASA's near-coastal facilities can easily support a range of LBFD mission routes, while the Glenn Research Center (GRC) in Cleveland provides similar mid-country options. Alternatively, mid-country routes are feasible from large numbers of military and civil airports.

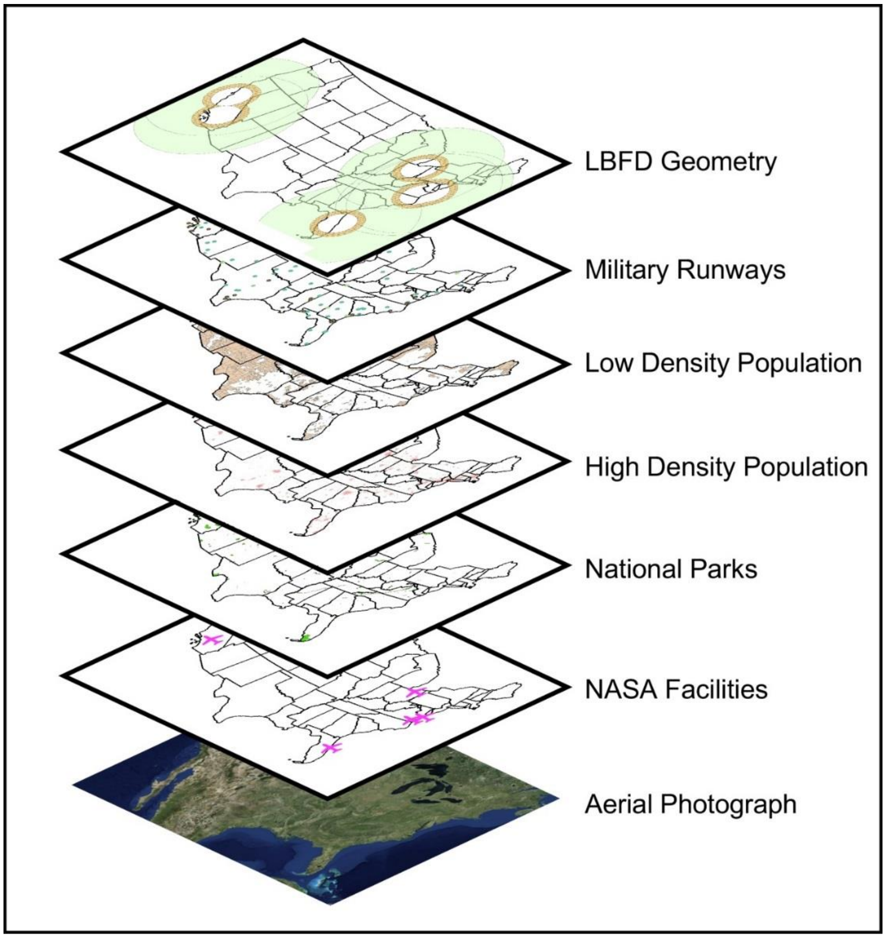


Figure 1. Schematic diagram showing superposition of thematic layers useful for LBFD route planning.

Table 1. Populations overflown on nominal LBFD routes, based on postal ZIP code statistics.

Exemplar of LBFD Route	Population within Boom Corridor	Number of Households within Corridor	Square Miles within Corridor	Average Population density (pop. per sq. mile)	Households per sq. mile
East Coast	42,186,156	15,863,033	27,629	1,527	574
West Coast	17,897,663	5,816,479	29,443	608	197
Midwestern	5,933,983	2,368,099	32,838	181	72

Figure 2 and Figure 3 illustrate two such potential routes, for West Coast and mid-country deployments, respectively. The nominal west coast LBFD route (Figure 2) overflies major U.S. population centers, including Los Angeles, as well as Bakersfield, Fresno, Merced, and Modesto. The household density in the overflown area is approximately 575 households per square mile.

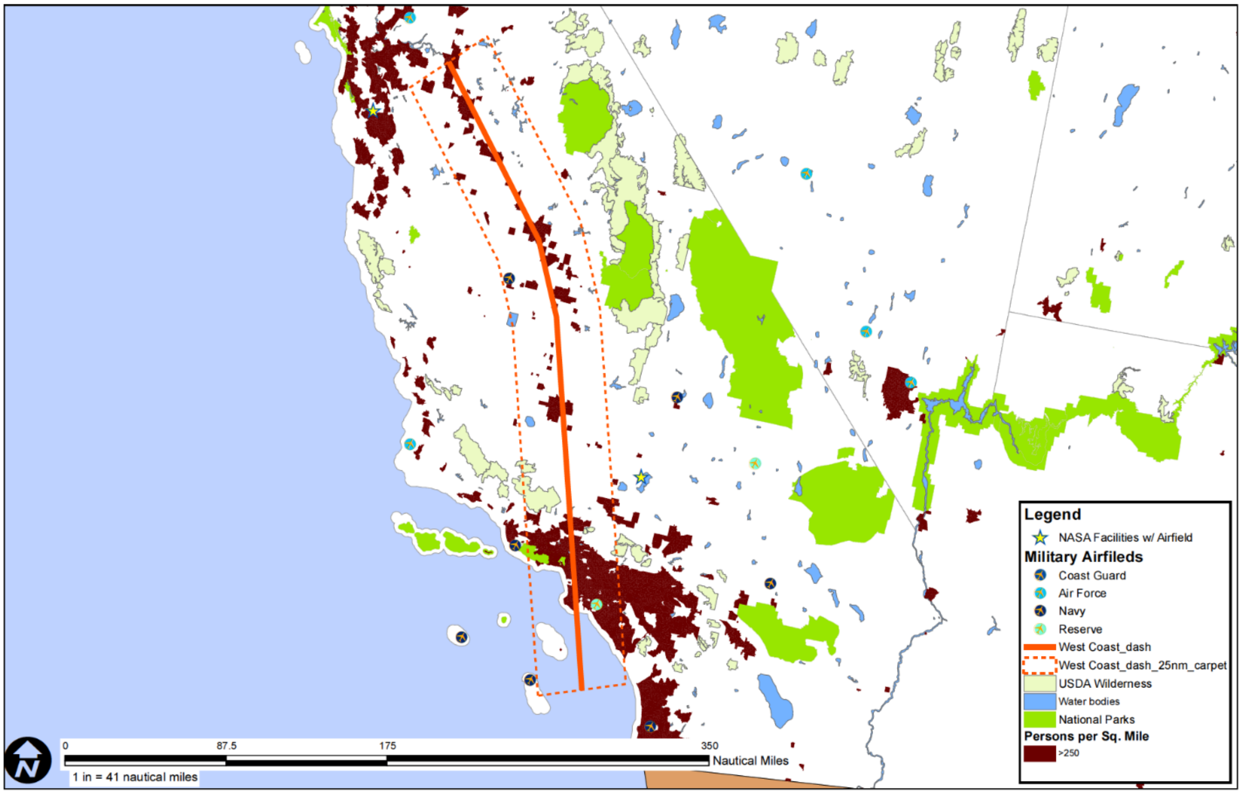


Figure 2. Illustrative example of a west coast LBFD mission route.

Figure 3 illustrates a potential mid-country LBFD route. The body of water in the upper left is Lake Michigan and the water body in the upper center is Lake Erie. The route would start and end at NASA Glenn Research Center (GRC). The route takes the LBFD to the west of GRC to expose a mid-western population to low-amplitude sonic booms, rather than overflying the east coast. The initial boom in this example could be placed in Lake Erie if the route were rotated clockwise to start its supersonic segment near Detroit (population center to left of lower tip of Lake Erie).

2.2 Suggested Sonic boom Exposure Levels for Field Trials

Figure 4 and Figure 5, adapted from Fidell *et al*. (2012), suggest the ranges of population proportions who may notice and report annoyance in varying degrees during LBFD supersonic overflights as a function of boom overpressure. The 87 data points were generated by 49 boom-accustomed residents of Edwards Air Force Base who used smartphones to self-report their reactions to a set of sonic booms and to answer a short series of questions.

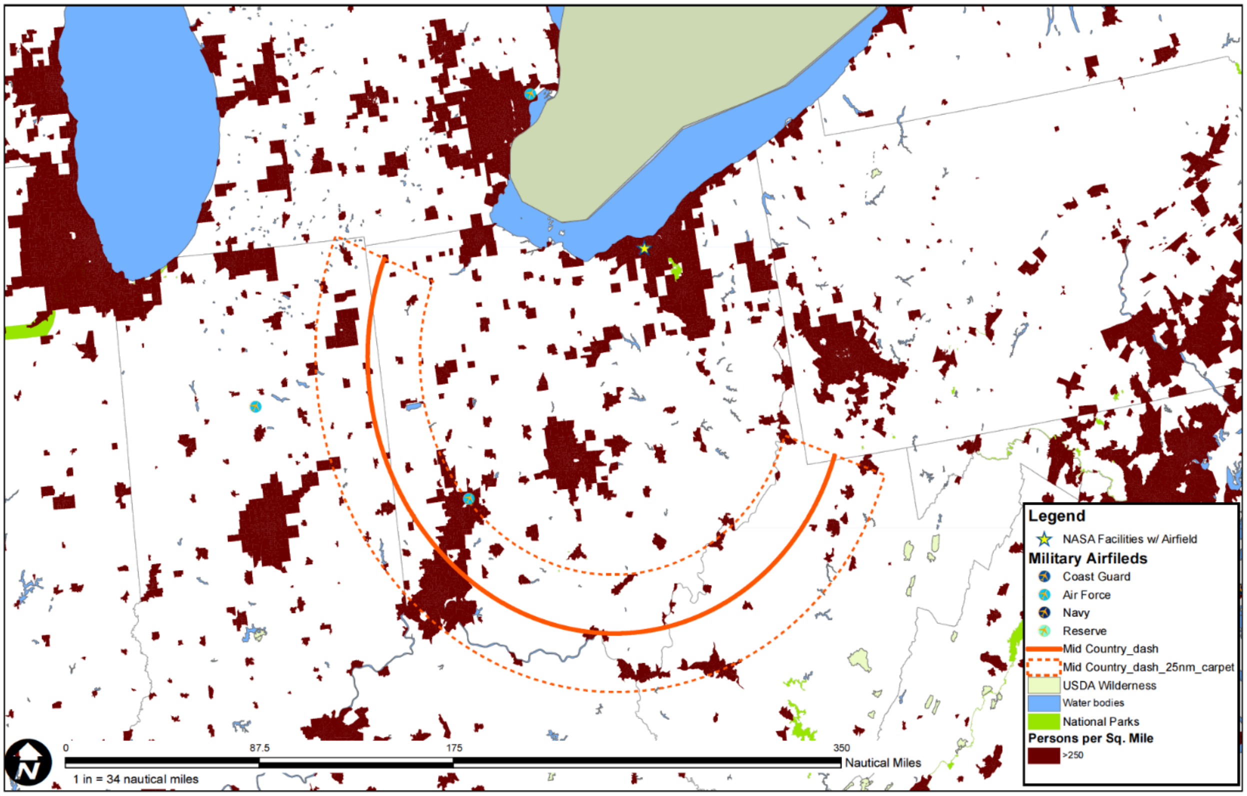


Figure 3. Illustrative example of a mid-country LBFD mission route.

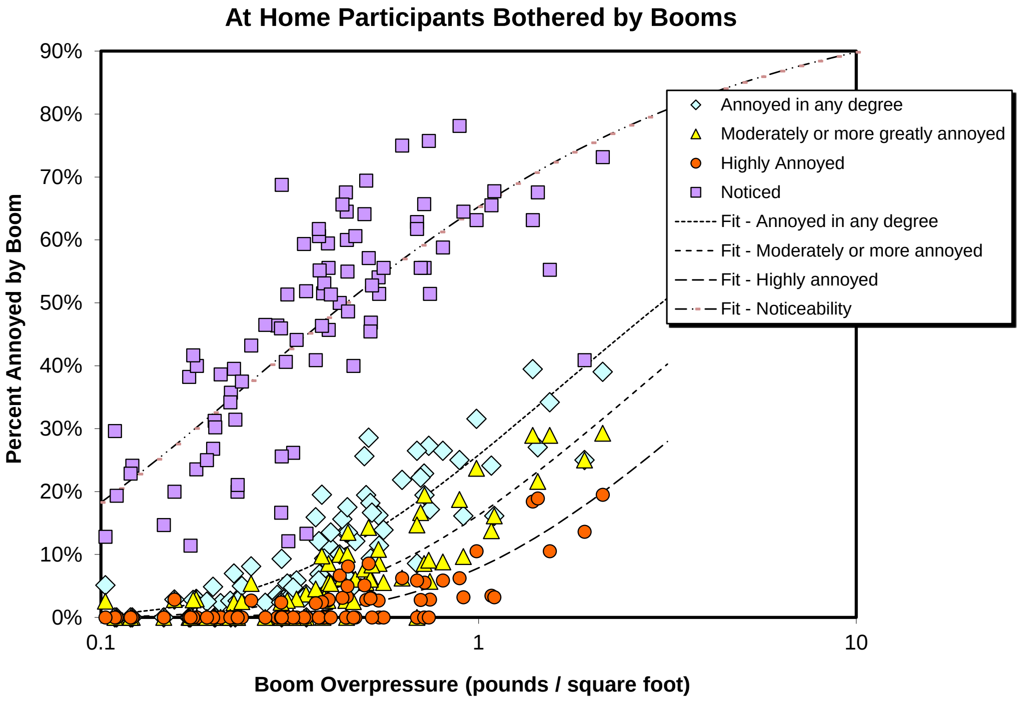


Figure 4. Illustration of interpretability of reports of notice and annoyance in immediate (single event) self-reports *versus* boom overpressure.

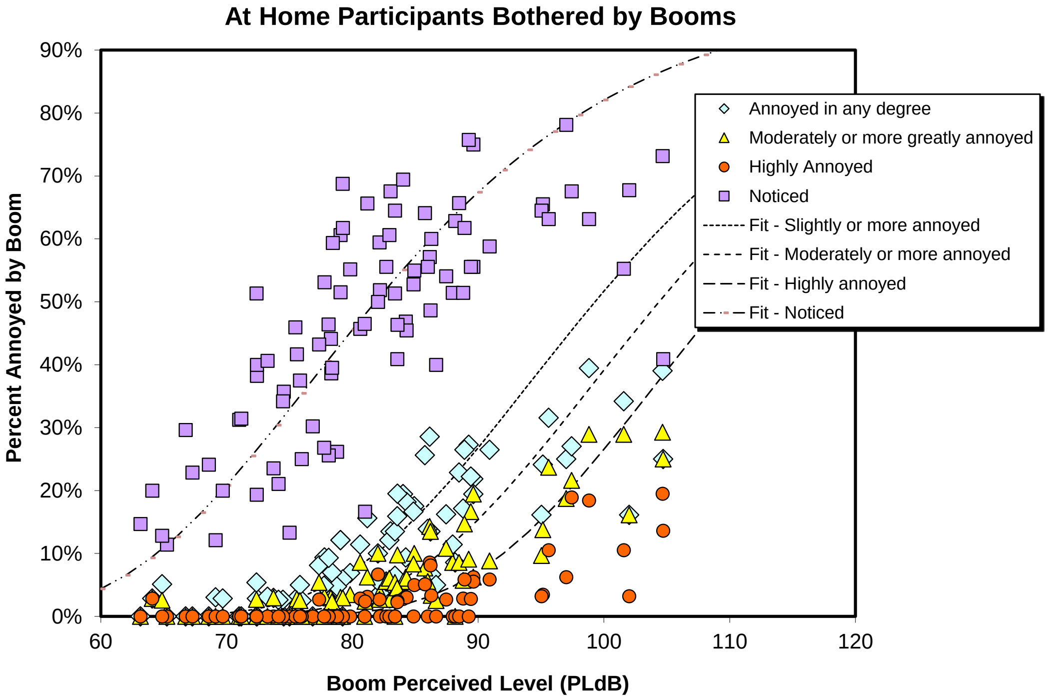


Figure 5. Illustration of interpretability of reports of notice and annoyance in immediate (single event) self-reports *versus* boom perceived level.

The data points are orderly despite the vertical scatter within the noticeability category and each of the three annoyance scale categories. Even though few respondents reported high annoyance at the highest boom intensity tested, growth in annoyance with exposure level still occurs at lower degrees of annoyance. The dashed lines in the figure are maximum likelihood parametric fits to the noticeability and three sets of annoyance category judgments. (Appendix B provides additional details of the curve fitting procedure, including a discussion of why the higher sound level data points fall below the fitting curve.) The functional forms and annoyance growth rates of all four fitting functions are identical; they are simply shifted horizontally to fit the individual data sets. All curves are asymptotic to the abscissa. The displacement between the adjacent annoyance fitting functions is approximately 4 to 5 decibels[2] for 20 Log [overpressure]. The displacement between the adjacent annoyance fitting functions for Perceived Level is approximately 5 decibels[3].

The two figures also suggest that small numbers of novel, very low amplitude sonic booms occurring at random times of day are likely to escape notice by the majority of survey respondents in boom-naïve communities. For example, at a boom overpressure of 0.3 pounds per square foot (~75 PLdB), only 20 to 50 percent of the at-home respondents noticed the boom (purple, square

---

[2] See Table 18 on page 122.

[3] See Table 20 on page 128.

plotting symbols). Since a small percentage of the participants even noticed the boom, it is hardly surprising that very small percentages of the participants were annoyed in any degree by them. Appendix B contains further discussion of this point.

2.3 Control of Cumulative Exposure Levels by Multiple within-day Flights

Cumulative levels of exposure to carpet booms are most directly controlled by numbers of daily LBFD overflights, or by combinations of numbers and levels of individual booms. For example, Figure 6 shows the relationship between the C-weighted day-night average sound level and the C-weighted sound exposure level for selected numbers of operations per day (family of lines). The graphic illustrates Equation 2 from the derivation below.

$$CDNL = CSEL + 10 * \text{Log}_{10}(N) - 10 * \text{Log}_{10}(86400) \qquad \text{Equation 1}$$

or

$$CDNL = CSEL + 10 * \text{Log}_{10}(N) - 49.37 \qquad \text{Equation 2}$$

where:

CDNL = C-weighted day-night average sound level (DNL), in decibels

CSEL = C-weighed sound exposure level (SEL), in decibels

N = number of daytime events at the given sound level

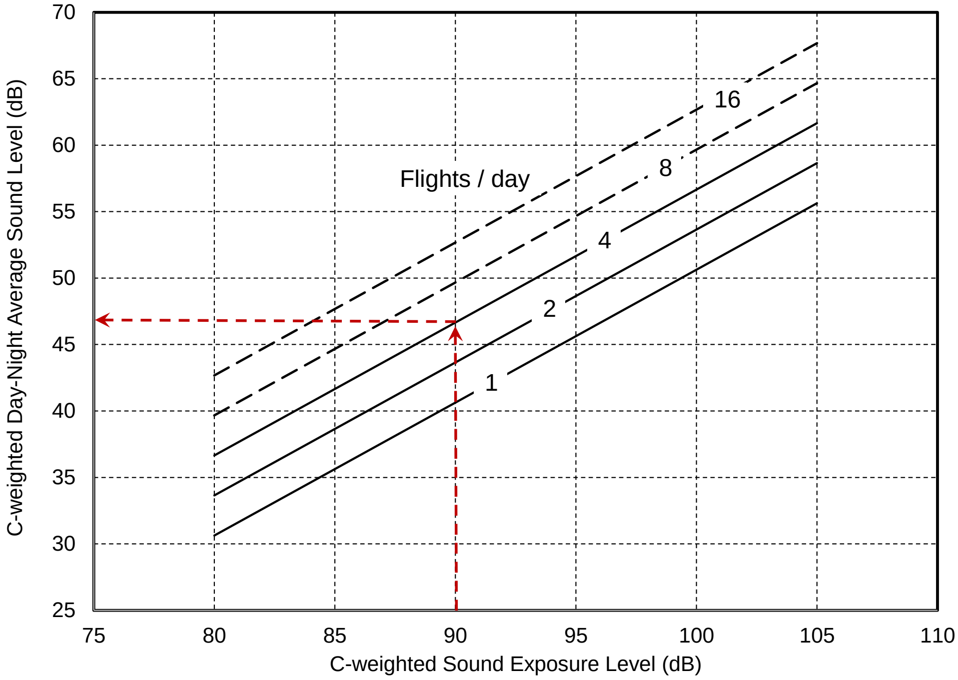


Figure 6. Relationship of C-weighted day-night average sound level to C-weighted sound exposure level and number of flights per day.

Thus, four flights per day each at a C-weighted sound exposure level of 90 dB would result in a C-weighted day-night average sound level of approximately 47 dB. (Note that more than four LBFD missions per day probably cannot be flown by a single aircraft.) This limits the range of exposure levels attainable solely by varying numbers of daily operations to 6 dB. However, if eight exposures could be accomplished (for example, by operating four out-and-back flights, as on a figure 8 flight route) then a 9 dB range could be achieved, yielding an approximate 50 dB C-weighted DNL.

Fidell *et al.* (2016), *q.v.*, transformed C-weighted SEL values into perceived levels and peak boom overpressures using empirical relationships derived from measurements made by Page *et al.* (2014) of low-amplitude sonic booms. Figure 7 plots 110 observed pairs of Perceived Level and C-weighted SEL. The slope of the linear regression is greater than unity (1.37) over the 30 dB range of measured CSELs. This is consistent with greater high frequency energy at higher boom strengths than at lower ones, to which PL would be sensitive, but C-level would not.

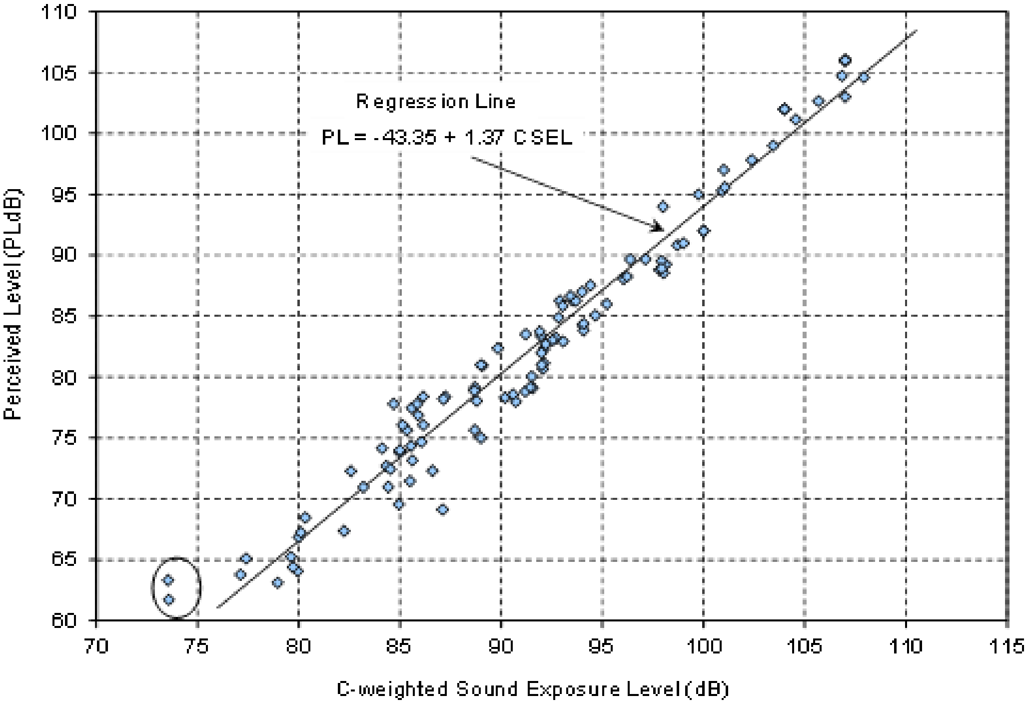


Figure 7. Observed relationship between sonic boom Perceived Level and C-weighted sound exposure level at Edwards AFB. (The two encircled data points are not included in the regression line calculation).

Figure 8 transforms the CSEL axis of Figure 6 to Perceived Level in units of PLdB via the equation shown in Figure 7:

$$CDNL = (PL + 43.35) / 1.37 + 10 * \mathrm{Log}_{10} (N) - 49.37 \qquad \text{Equation 3}$$

where:

CDNL = C-weighted day-night average sound level (DNL), in decibels
PL = perceived level, in perceived noise decibels
N = number of daytime events at the given sound level

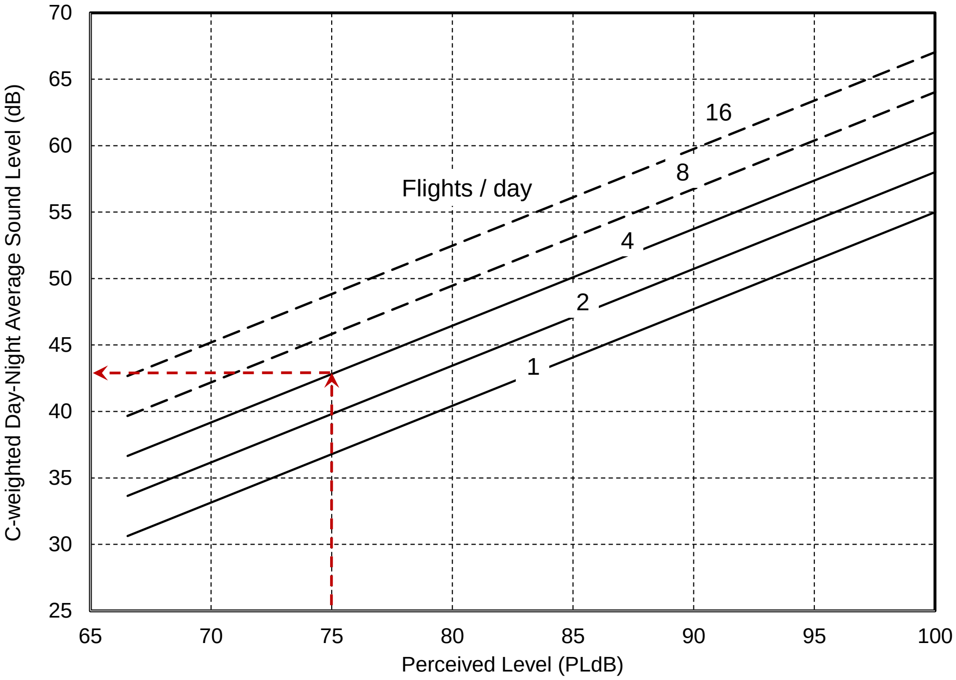


Figure 8. Approximate relationship of C-weighted day-night average sound level to perceived level and number of flights per day.

The transformation permits prediction of CDNL from a boom amplitude in perceived level and number of events. For example, four flights per day at a perceived level of 75 PLdB yields a CDNL of approximately 43 dB.

2.4 Continuum of Prompt Response Study Designs

Figure 9 shows a gamut of methods for extracting prompt opinions about the annoyance of exposure to sonic booms from exposed populations. The four designs on the left side of the figure are similar to those commonly employed in conventional, airport noise surveys. The two on the right side of the figure resemble field experimentation more than traditional social survey methods. For reasons discussed in §4.1, aircraft regulatory agencies are likely to find evidence of community reaction to low-amplitude sonic booms of the sort produced by study designs on the left side of the figure more persuasive than that produced by study designs on the other end of the continuum.

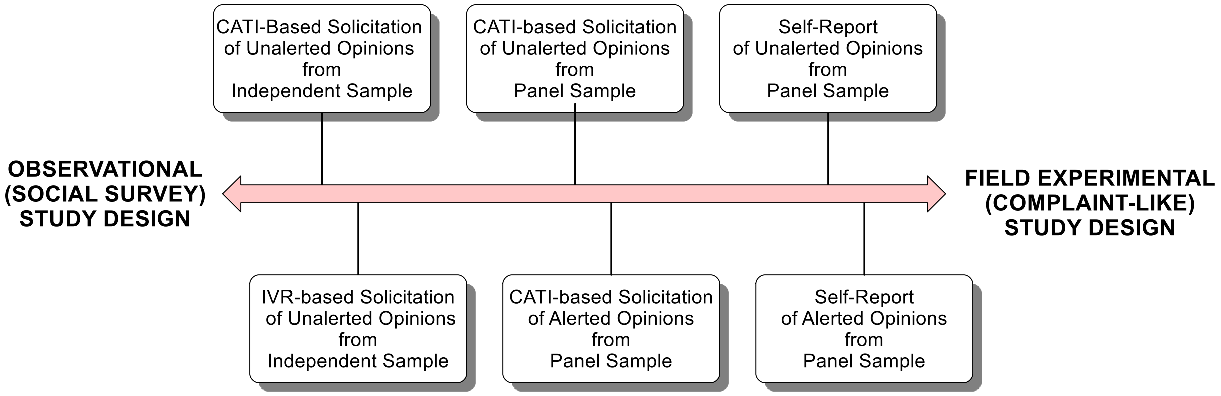


Figure 9. Continuum of study design options, ranging from conventional (purely observational) attitudinal surveys through more controlled (complaint-like) designs.

2.5 Construction and Limitations of a Telephone RDD[4] Sample Frame

EPSEM[5] RDD samples for present purposes should be generated from a sampling frame of telephone working blocks (NPA-NXX-XXYY) which fall within a carpet boom corridor. In other words, the universe of telephone working blocks which intersect the carpet boundary must be known so that the RDD sample can be drawn from it. Figure 10 shows the anatomy of a telephone number. Since working block coverage areas may bridge carpet boom corridor boundaries, some randomly selected numbers from working blocks that are only partially within boom corridors may include numbers ineligible for interview.[6]

EPSEM samples are *not* drawn from "listed samples"; that is, lists of known telephone numbers. Exact geo-locations, names, and addresses cannot therefore be assigned to individual RDD sample records *a priori*. (Such information can, however, be added to a sampling frame after a sample is drawn.) Representative sampling of all telephone subscribing households within the carpet boom boundary may be accomplished by assigning equal probabilities of selection to each telephone working block. Since the number of possible phone numbers (and working blocks) is proportional to population concentrations within the carpet boom corridor, the resulting sample is self-weighting and fully representative.

---

[4] RDD is an abbreviation of Random Digit Dialing.

[5] EPSEM is an abbreviation of Equal Probability of Selection of Elements (sampling) Method.

[6] Working block borders can span several zip codes in some cases, so sampling only those working blocks which lie predominantly or wholly within a carpet boom corridor is best accomplished with reference to directories of listed databases within boom corridors.

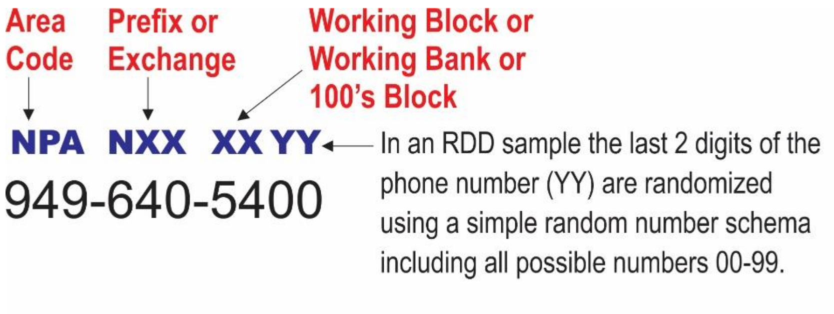


Figure 10. RDD EPSEM sample is based upon the working block ("100s block") as the major organizational unit from which the sample is drawn[7].

2.6 Interviewing Methods

The necessity for assessing prompt responses to *en route* supersonic aircraft noise poses a variety of challenges for interviewing. Standard methods of assessing aviation noise impacts are optimized for airport-centric settings, where the primary concern is assessment of long-term reactions of residents living near runways to cumulative noise exposure created by lengthy, repetitive sequences of familiar, high level, relatively long duration noise events. The same interviewing methods, however, are ill-suited for rapid assessment of immediate reactions to intermittent exposure due to rare, novel, unfamiliar (if not unrecognized), short duration, and relatively low-level sounds, over areas of thousands of square miles.

2.6.1 Prompt response

All interviewing methods potentially useful for interviewing people about their *prompt* responses to sonic boom exposure are compromises among factors such as representativeness, numbers of respondents reachable in a short time period, interview completion rates, potential sensitization of respondents to sonic booms, and cost. The required scale (rate) of interviewing contact attempts is a fundamental difficulty in assessing prompt responses to sonic boom exposure. Three[8] potential methods of interviewing residents of supersonically overflown areas about their prompt reactions include:

- Outbound interactive voice response (IVR) interviewing within independent samples;
- Conventional, computer-assisted telephone interviewing (CATI) or text messaging contacts with either independent or panel samples; and
- Smartphone-based self-report.

---

[7] NPA is an abbreviation for Numbering Plan Area.

[8] Two other methods of gauging prompt reactions to sonic booms, focus groups and content analysis of social media postings, can yield indirect supplementary information, but do not support dosage-response analyses.

The first two of the above alternatives are the most likely to support construction of credible dosage-response relationships. At least in principle, the efficiencies of scale of automated (outbound IVR) interviewing would permit collection of information about community response to low-amplitude sonic booms without calling attention in advance to the occurrence of such booms. The risk associated with automated interviewing methods is a low interview completion rate. If the interview completion rate is as low as 1%, for example, 10,000 automated, high speed contact attempts would yield only 100 completed interviews.

2.6.2 Delayed response

With little modification, interviewing methods optimized for this setting also suffice to assess *delayed* response to cumulative (weekly or longer-term) exposure to occasional sonic booms.

2.7 Questionnaire Items

Two questionnaires are required for the intended survey: one to assess prompt reactions to individual sonic booms, and one to assess delayed (*e.g.*, end-of-week) response to cumulative exposure. Both should start with a screening question to confirm that the potential respondent is 18 years of age or older. The interview attempt should be abandoned if the respondent is a minor.[9]

2.7.1 Substantive items for prompt response questionnaire

To be suitable for administration by high speed outbound IVR interviewing methods, the prompt response questionnaire should contain only closed response category responses, and must be as short as possible. This limits the number of information requests to as few items as are needed to determine:

- whether the respondent was at or near home and noticed a sonic boom within the last fifteen minutes;
- whether the respondent was bothered or annoyed by the boom, and if so, to what degree;
- whether the respondent was startled by the boom;
- whether the respondent was indoors or outdoors at the time of notice of a boom; and
- whether the respondent noticed rattling sounds, and if so, the degree of annoyance associated with hearing the rattling sounds.

---

[9] An interview conducted by a live agent could request an interview with an adult household member, but interview attempts conducted by automated means cannot readily do so.

The wording of these items should resemble that of the questionnaire items of the smartphone interviews described by Fidell *et al.* (2012). The set of questionnaire items should be roughly as follows:[10]

Item 1: "Have you been at home or nearby for the last fifteen minutes?"
*If **Yes**, ask Item 2 next; if **No,** terminate interview*

Item 2: "Did you notice a sonic boom within the last fifteen minutes?" (Yes/No)
*If **Yes**, ask Item 3 next; if **No,** terminate interview*

Item 3: "Were you bothered or annoyed by the sonic boom?" (Yes/No)

Item 4: "Press 1 if you were slightly annoyed by the sonic boom that you noticed; Press 2 if you were moderately annoyed by the sonic boom; Press 3 if you were very annoyed by the sonic boom", and push 4 if you were extremely annoyed by the sonic boom."

Item 5: "Were you startled by the sonic boom that you heard?" (Yes/No)

Item 6: "Where were you when you noticed the sonic boom? Press 1 if you were indoors at home; Press 2 if you were indoors somewhere else; Press 3 if you were outdoors."
If indoors at home or elsewhere, ask Item 7 next. If outdoors, terminate interview.

Item 7: "Did you notice any rattling sounds in addition to the sonic boom?" (Yes/No)
*If **Yes**, ask Item 7 next; if **No**, terminate interview.*

Item 8: "Press 1 if you were slightly annoyed by the rattling sounds. Press 2 if you were moderately annoyed by the rattling sounds; Press 3 if you were very annoyed by the rattling sounds or push 4 if you were extremely annoyed by the rattling sounds."

If no more cost-effective or less intrusive costly method of geo-locating respondents is identified, respondents can also be asked for their home ZIP codes at the end of the interview to further confirm their eligibility for interview.

2.7.2 Delayed response questionnaire

In principle, a questionnaire intended to document responses to cumulative exposure to sonic booms could reference periods of time longer than a day. However, interpretation of findings are most straightforward if they are confined to reactions to a single day's worth of individual exposures. (Interpretations of reactions over extended exposure periods raise complications about recall of the days and times of day when respondents were at home.)

---

[10] Respondents in automated interviewing are expected to "barge" (respond as soon as they hear a closed response category alternative of interest), and control software for automated interviewing generally repeats questions after a response timeout interval. An explicit "Press … to hear these choices again" alternative is therefore gratuitous. It is also possible to mix easily discriminated voice responses (*e.g.*, "yes" and "no") with touch tone dialed responses.

Since an end-of-day questionnaire need not be administered by high speed, high volume interviewing techniques, it need not be as succinct as the prompt response questionnaire, nor need it be limited to closed response category items. The core items for the delayed response questionnaire should include the following:

- Have you been at home or nearby (*for specific hours or most of*) today?
- How many sonic booms do you remember noticing today?
- [If any booms are recalled] Considering all of the sonic booms that you noticed today, how annoyed were you by all of them taken together?
- [If more than one boom was recalled] Was any single sonic boom that you noticed today much more annoying than any of the others?
- About what time of day did you notice the boom that was much more annoying than the rest of the booms?

2.8 Centralized Coordination of Data Acquisition

Overall coordination of a mission requires continuous, real time, centralized communication with 1) geographically-distributed aircraft tracking equipment, 2) acoustic and meteorological data collection equipment, 3) social survey call centers, and 4) NASA personnel responsible for flight coordination.[11] A centralized system can also serve data displays to other authorized parties.

Figure 11 presents a schematic overview of the recommended data collection system. To minimize the risk of a single point of failure, the central server shown in the middle of the figure should have redundant processors that shadow one another and automatically rollover in the event of a server failure. The server would be in constant communication with all data acquisition devices and services, and on an on-demand basis with project personnel. The NASA Flight Director shown in the lower right of the diagram, is the only person in real time, two-way communication with the test aircraft ("LBFD Mission.") The latter communication link has no direct connection with the server.

All clients of the system are shown in the figure surrounding the server and connected by double-ended arrows (denoting two-way communication). Since the widespread geographic dispersion of the peripheral equipment precludes hard-wired connections to the central server, all internal system communication should be Internet-based. The double-headed arrows in Figure 11

---

[11] Aircraft tracking and noise monitoring functions of a specially-designed data acquisition system (described in §2.8 and §2.9) are separately packaged and entirely independent of one another; have separate links via the Internet to the central server; and are unlikely to be co-located in field use.

thus represent (packet-switched) Internet connections, while outbound IVR calls are delivered over the public (circuit-) switched telephone network.

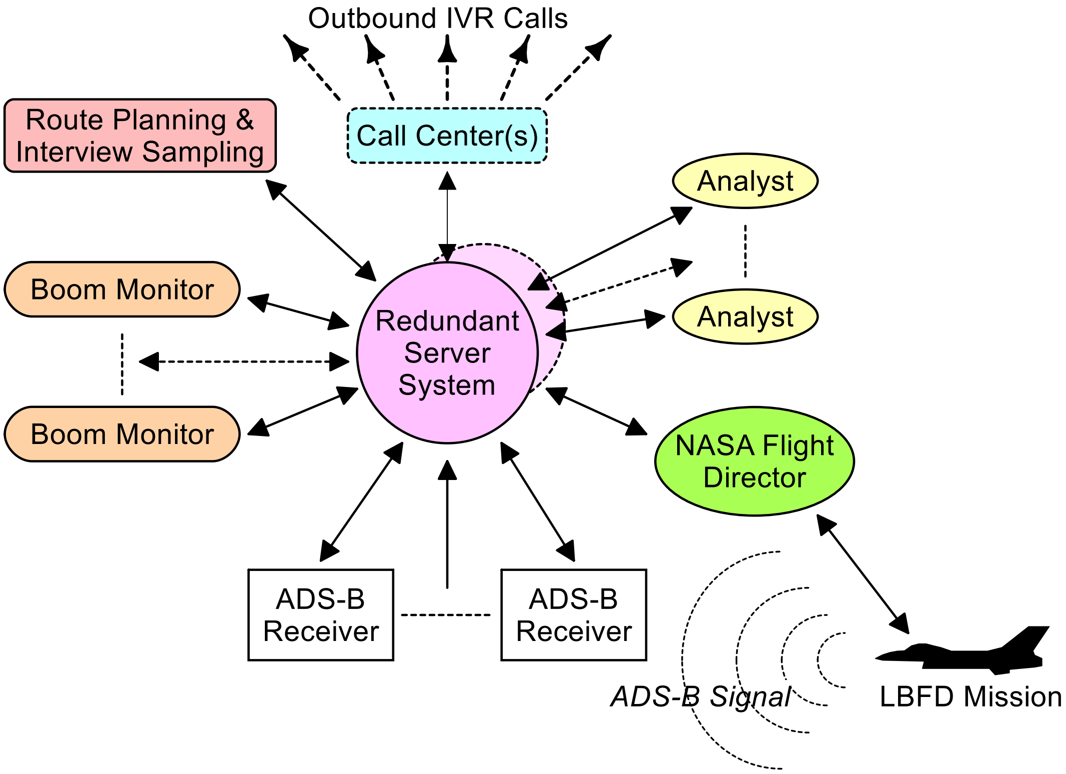


Figure 11. Schematic representation of conceptual design of data acquisition system.

2.9 Wide-Area Acoustic and Meteorological Data Collection (Boom Monitor)

Sonic boom measurements must be made near interviewing areas situated miles to either side of a ground track at least 50 miles long. Staffing all measurement sites prior to and during flight missions to verify suitable weather conditions and equipment readiness for LBFD launch is unaffordable.

The main risk mitigation considerations for acoustic data collection include:

- Ability to capture low-level sonic boom waveforms;
- ability to acquire local atmospheric conditions;
- use of proven, off-the-shelf equipment;
- redundancy;
- buffering of several days' worth of all data at the instrument site;

- Internet-accessible, real time, two-way communication with the instrumentation;
- capture of both acoustic and meteorological conditions;
- time synchronization between acoustic and social survey data collection; and
- centralized management of the data collection task.

The total number of sonic boom monitors available for field deployments constrains mission planning. A minimum of two monitors is needed to establish an empirical exposure gradient orthogonal to the route of travel of an LBFD mission. Larger numbers of monitors per community could increase the precision of boom measurement by increasing the density of measurement points, but the utility of extreme precision of measurement is open to question (see §4.1.3 and §4.1.4).

The element labeled "Boom Monitor" in Figure 11 is expanded in Figure 12 to show greater detail. This schematic diagram shows the recommended components for collecting three types of data: (1) meteorological information, (2) sound levels and sound waveforms and (3) GPS position and time-of-day information. The system would collect and store data locally, and transmit it to the server as required. The individual components are shown as gray-filled boxes. Analog signal lines between the components are shown in blue; digital signal lines are shown in red, and radio frequency lines are shown in green. The following subsections describe the system requirements and the elements of the boom monitor system that provide them. Fidell *et al.* (2017) provide additional specifics of the capabilities of this instrumentation.

2.9.1 Internet connectivity

The Boom Monitor connects to the server via a wireless cellular modem capable of 3G and 4G data connection depending on cell service available by the local provider. The cell modem and its antenna are shown in the upper right of the figure and the modem is connected to the audio recorder and embedded computer by the router shown at the lower right.

The recommended commercial off-the-shelf (COTS) cell modem shown in the figure contains its own GPS receiver, and requires only an antenna for optimal GPS signal reception. The GPS component is discussed below. The cell modem allows direct communication through the router between the server and either the embedded computer or the digital audio recorder.

2.9.2 Field instrumentation time base and current location

All components of the data collection system (server and field units) must contain internal clocks continuously synchronized to the same common time base as the server. The primary time base for the acoustic monitor station should be an Internet-based NTP (Network Time Protocol) server. The GPS receiver would provide the latitude/longitude coordinates of the monitor's position on a continuous basis. This component and its antenna are shown in the upper right of Figure 12.

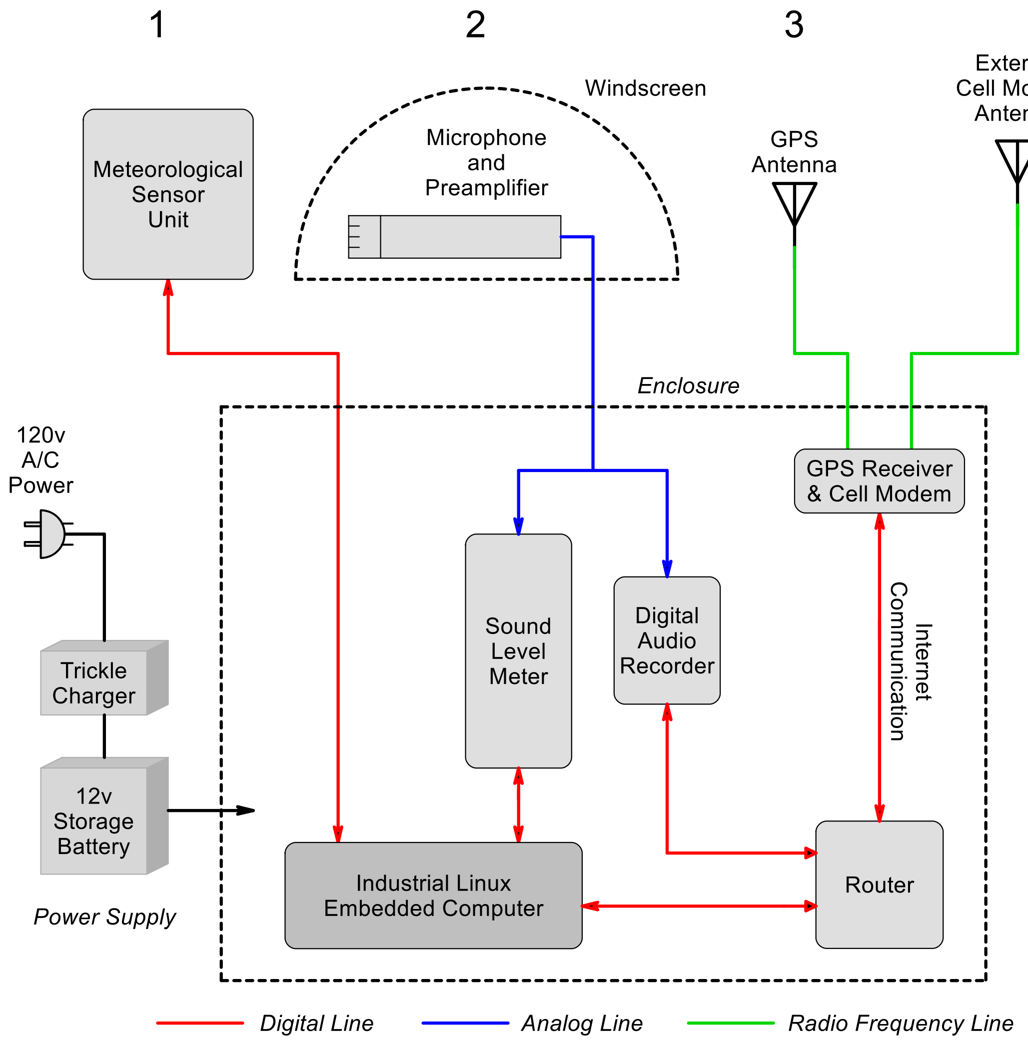


Figure 12. Schematic drawing of boom monitoring instrumentation.

2.9.3 Meteorological information

Each monitor should be equipped with its own sensor package capable of measuring air temperature, relative humidity, barometric pressure, precipitation, wind speed, and wind direction. Wind speed near the microphone is important because atmospheric turbulence can adversely affect the fidelity of recorded boom waveforms, especially low-amplitude ones.

2.9.4 Sound pressure levels and acoustic waveforms

Figure 12 shows the sound measurement components in the center of the drawing. They begin with a single 1/2-inch electret microphone and power supply. The output is shared by 1) a sound level meter and 2) a digital audio recorder. The sound level meter is connected to the embedded computer, which passes the collected data to the server in real time through the router and cell modem. The digital audio recorder contains two 16-bit A/D converter channels which with staggered gains can create a usable dynamic range well in excess of 110 dB[12]. Details of the individual data acquisition components are provided below.

Microphone and windscreen

Acoustic data is acquired by a ground plane microphone (G.R.A.S. Model 40AZ[13]) and microphone preamplifier (G.R.A.S. Model 26CG). This configuration eliminates ground reflection artifacts and reduces wind turbulence at the microphone diaphragm. The A-weighted noise floor of the recommended microphone/preamplifier system is 17 dB, adequate for measurements of expected boom levels of the LBFD aircraft. The microphone would be protected against dust and debris by a two-stage hemispherical windscreen.

Wind-induced noise at the microphone diaphragm can be a major impediment to capturing clean sonic boom waveforms. The ground plane placement of the microphone helps reduce wind-induced noise at the diaphragm, but additional turbulence reduction is also required. A 30-inch diameter, hemispherical wind screen with a commercial internal 3.5” foam windscreen mounted on the microphone itself is recommended as protection against contamination of measurements from wind noise artifact.

Sound level meter

The primary function of the sound level meter is to provide real time “slow” ambient sound levels to mission planners and analysts. The data would be transmitted to the server at one-second intervals. A-weighted, C-weighted and Z-weighted (or unweighted) sound level would be acquired continuously for this purpose. The Larson-Davis Model 831 digital sound level meter is recommended for this purpose.

Meteorological, sound level, and system health data would be transmitted in a single data block once per second. The data block would include the following parameters:

- Monitor number
- Date/Time stamp
- LeqA
- LeqC
- LeqZ
- Wind Speed
- Wind Direction

---

[12] COTS 24-bit analog-to-digital (A/D) converters with automatic archiving capability are presently not available. A COTS 16-bit converter with archiving capability, however, is available and is proposed here to provide the required functionality. Equivalent 24-bit converters are expected to be available in a few years’ time.

[13] Frequency response: 1-10,000 Hz, ±1 dB.

- Temperature
- Precipitation
- Barometric Pressure
- Battery Level
- Boom Monitor Operational Status
- GPS coordinates
- Audio Recorder Operational Status
- Audio Recorder Number of Files Stored

Analysts who monitor this information will be able to observe average values as well as extrema (and how frequently they occur) of both the sound levels as well as atmospheric conditions at the monitor. This information is essential for allowing decision makers to determine whether conditions at all sites are favorable to begin an LBFD mission.

Audio recorder

Acquisition of the boom waveform would be performed within the box labeled "Audio Recorder" in Figure 12. This box would contain a microprocessor, A/D converter, and a large amount of digital memory. The Rpi Cirrus Model 2–CR44 is proposed for this application. Two 16-bit analog-to-digital (A/D) converters operating at sampling rate of 44.1 KHz would yield a usable frequency range of 0.2 Hz to 20 KHz after passing through an internal anti-aliasing filter. Measurements will be collected *continuously* by the recorder throughout a flight test, and archived in its local memory as a sequence of time stamped 1-minute audio .wav files. No loss of data would occur between files, thus enabling later concatenation for analytic purposes. Specific 1-minute files can be retrieved from the remote monitors on command from the server via the cell modem and router.

All audio file data will remain stored in the recorder until a command from the server directs files to be deleted. Thus, files can be retrieved (or re-retrieved) at any time after the time of occurrence of a boom and a temporary loss of communication with the server poses no risk of data loss.

The server would calculate a 3-minute long time window for the shockwave arrival time. This calculation would determine which 1-minute interval most likely contained the boom signature. An additional 1-minute interval on either side would then define the files needed from the recorder. The server would request each file from the recorder. Once all three of these site-, date- and time-stamped files have been received, they would be combined into a single file and stored for subsequent processing.

2.10 LBFD Speed and Position Data Acquisition (ADS-B Receivers)

Continuous knowledge of the LBFD's speed and position is essential, since interpretation of the boom monitor data as well as management of the interview process depends on such knowledge.

2.10.1 ADS-B technology

The LBFD is expected to be equipped with an ADS-B (Automatic Dependent Surveillance-Broadcast) unit. ADS-B systems transmit information in short data blocks, once per second. Each data block consists of the aircraft's airframe-specific, 6-byte, ADS-B squawk code (*not* the secondary surveillance radar 4-digit squawk code) as well as latitude, longitude and airspeed.

Since the system is completely passive and requires no radio interrogation or triangulation technology, receivers may be located anywhere on the ground within a few tens of miles of the intended flight track. The maximum unobstructed (high altitude) air-to-ground broadcast range of onboard ADS-B equipment is approximately 180 nautical miles. A working range for receipt of ADS-B signals from the LBFD during cruise of as little as 90 nautical miles is adequate for present purposes.

2.10.2 Receiving station hardware

Field hardware would consist of an antenna, a COTS ADS-B receiver/decoder unit, a minicomputer, and a wireless cellular communications modem. The left side of Figure 13 shows these elements. The receiver/decoder turns the radio frequency signal into a digital bit stream. The minicomputer reformats the bit stream for transmission to the server.

Data transmission would be accomplished via a cellular data service to move packets from the receiver station to a cellular data service provider. Each data packet would include the IP address of the server. When received by the cellular service, data packets would be immediately forwarded to the central server via the Internet. These data packets would also be stored in local memory at the receiving station for backup purposes. Sending received data packets for all aircraft maintains maximum flexibility for displaying information about both the target and other aircraft. Analysts would be able to confirm normal operation of ADS-B receivers by observing movements of any aircraft in the area.

2.10.3 Placement of receivers

Ground receiving stations should be spaced such that at least two ADS-B receiving stations are able to simultaneously receive the LBFD ADS-B signal at all times, from takeoff to landing. Such placement provides redundancy in both the receiver hardware as well as its communication link with the data collection server shown in Figure 11. Receipt of data blocks and the assigned ADS-B squawk code by the system prior to takeoff will confirm that the aircraft's transmitter has been enabled, the target is being acquired by nearby ADS-B receivers, and that the system will be able to track the aircraft.

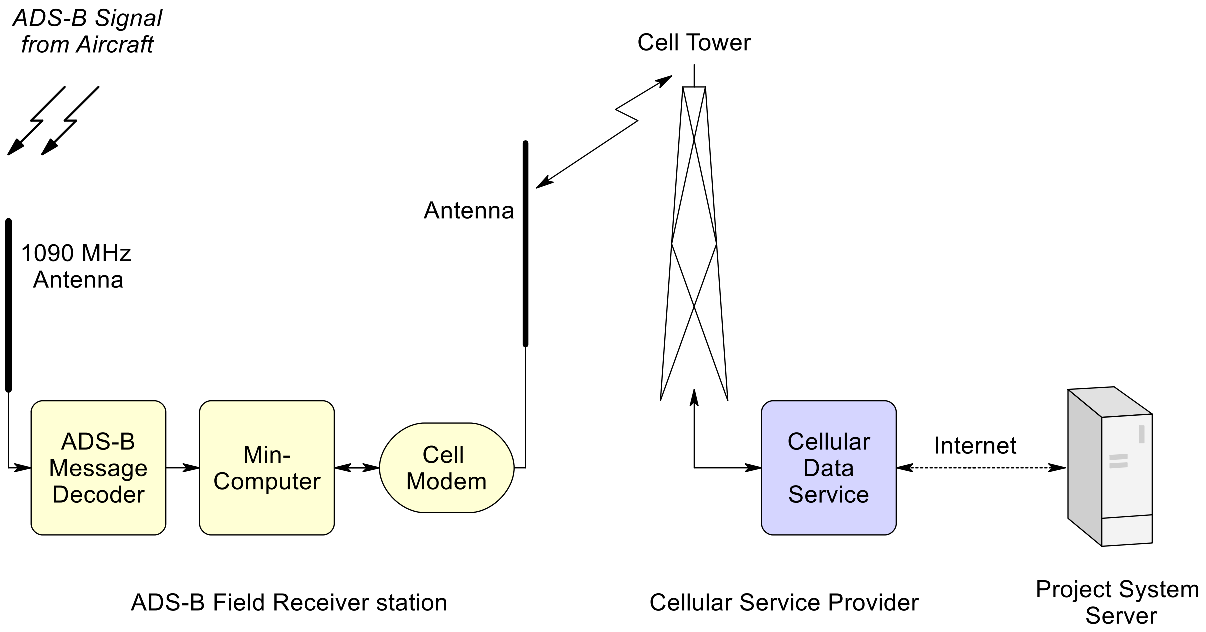


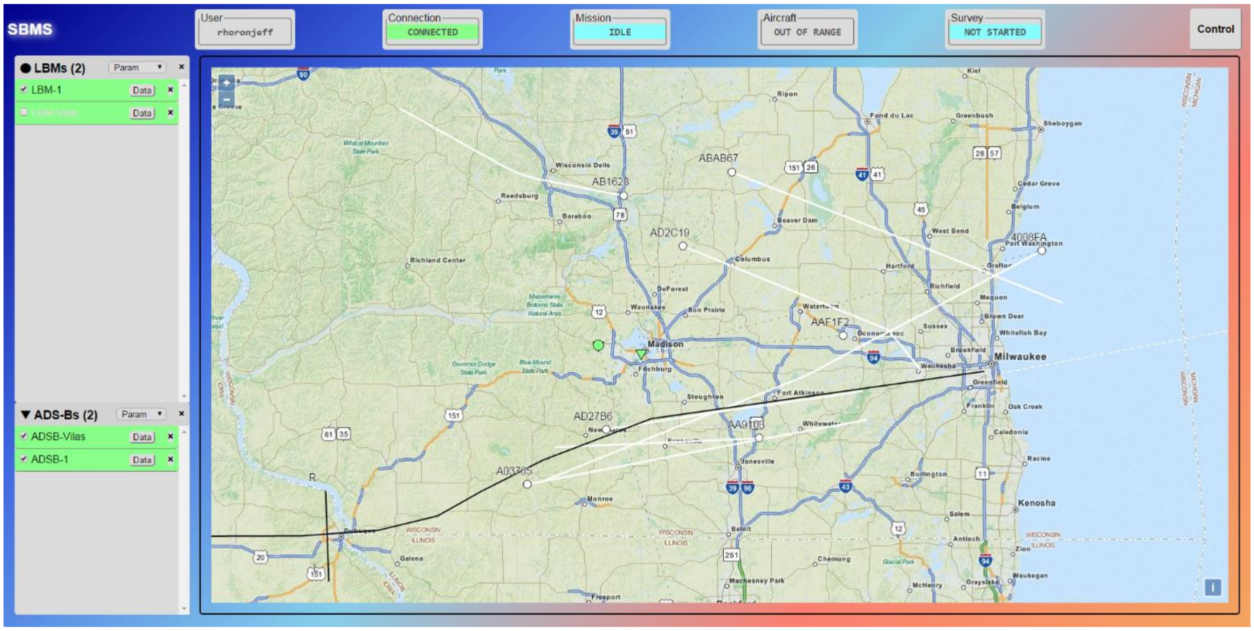
Figure 13. ADS-B receiving station block diagram.

2.11 Access to Real Time Field Data by Mission Planners and Analysts

The boom monitors and the ADS-B receivers would operate continuously throughout their deployment period, whether an LBFD mission was in progress or not. Thus, mission planners and coordinators would have continuous, real time access to sound level and atmospheric information at each noise monitor site at any time. The sound level, weather data and status of each Boom Monitor would be available in real time by means of a Web-based graphical interface available over the Internet. The Web application would show the locations of the Boom Monitors (as well as the location of the LBFD aircraft) on a map updated in real time. Detailed data for each of the boom monitors would be viewable by clicking on separate tabs that would show time-synchronized data for all of the boom monitors and ADS-B receivers. Ten boom monitors and 3 ADS-B receivers would suffice for most LBFD missions. Users would have the option of displaying all parameters for a single monitor, or a single parameter for all monitors.

Figure 14 shows a sample of the recommended Web interface display. The map would show the location of boom monitors, ADS-B receivers and LBFD aircraft. The map display would be zoomable as desired by individual users. The vertical panel on the left would show all active field units. Clicking on any one of them would display a panel of real time data and status information as shown in

Figure 15 for a boom monitor.


SBMS

User rhoronjeff

Connection CONNECTED

Mission IDLE

Aircraft OUT OF RANGE

Survey NOT STARTED

Figure 14. Recommended Web interface display.

2.12 Potential community opposition to test flights

Communities selected for supersonic test flights by the LBFD aircraft may oppose such overflights either before or after they have started.[14] Some of the likely sources of potential objections to LBFD overflights include:

- Surprise about the onset of a supersonic overflight test program;
- Apprehension about personal safety or property damage due to shock waves;
- A belief that communities were not consulted prior to the start of supersonic boom exposure; and
- Technically inaccurate or otherwise adverse popular press, talk radio, and social media accounts of test missions.

Some of these concerns may be allayed before they arise through various forms of publicity (including social media[15]), while others may have to be addressed after the fact.

---

[14] Oklahoma City initially welcomed FAA's 1964 sonic boom test program overflights, which started in February, 1964 and exposed city residents to as many as eight booms a day, at overpressures as high as 2 psf. Opposition to the test eventually forced premature termination of the project on 30 July, 1964, however.

[15] NASA already maintains a presence in social media, including Instagram (www.instagram.com/NASA), Twitter (www.twitter.com/NASA), and Facebook (www.facebook.com/NASA). Instagram appeals in large part to teens and young adults; Twitter to young through middle-aged adults; and Facebook to young through older adults.

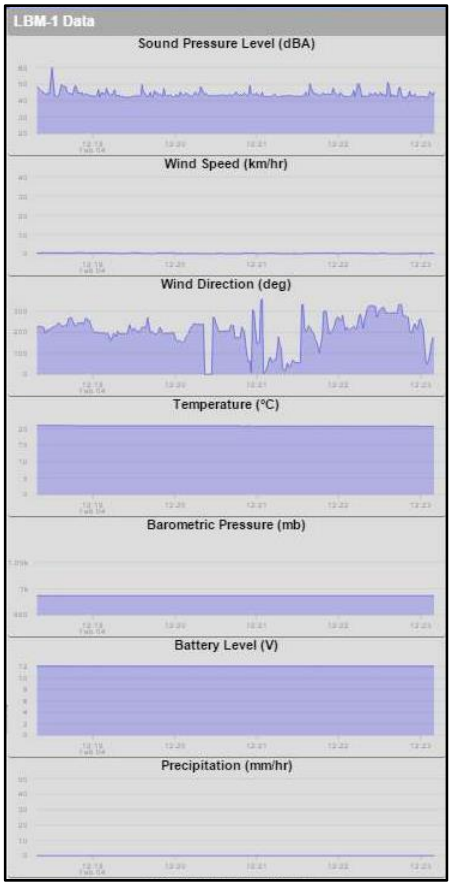


Figure 15. Boom monitor status display.

3 SUMMARY OF SECOND YEAR'S EFFORT (PHASE II – YEAR 1)

The initial year of Phase II effort under Contract NNL15AA05C is documented in Fidell *et al*. (2017). The two major substantive technical tasks accomplished in Year 1 of Phase II were 1) finalization of a detailed study plan based on the conceptual plan developed in Phase I, and 2) a field demonstration of the ability of the Internet-enabled aircraft tracking and boom-monitoring system (designed in Phase I).[16]

The detailed study plan was completed to guide the conduct of a study of interview completion rates in a cross-sectional, independent sample telephone survey without callbacks. Almost 13,000 telephone interview contact attempts were made by both automated and live agent means to estimate interview completion rates. Interview completion rates were found to be so low that cross-sectional (each respondent interviewed once, without pre-arrangement) study designs are infeasible for assessment of prompt reactions to sonic booms.

The study of the aircraft tracking instrumentation was designed to demonstrate the ability of the data collection hardware and software to:

- Continuously track an LBFD aircraft over a 100+ mile supersonic dash;
- Measure on-ground boom sound level over a wide geographic area;
- Message the attitudinal survey call center when to begin and end the survey in each interview neighborhood in near-real time; and
- Perform the above actions under centralized control with access by authorized personnel anywhere.

Each of the bulleted capabilities was successfully demonstrated for a prototype system in real-world conditions.

3.1 Finalization of Detailed Study Plan

A detailed study plan, based on the conceptual design of Phase I and a modification[17] to the original statement of work of Contract NNL15AA05C, was completed in January of 2017. The major provisions of the plan are summarized in the following subsections.

3.1.1 Desired properties of respondent sample

The representativeness of sampled opinion is of paramount importance because the purpose for collecting information about community response to low-amplitude sonic booms is to inform

---

[16] Documentation was also prepared (NASA LaRC, 2017) for the OMB and IRB review and approval processes.

[17] The original Statement of Work of Contract NNL15AA05C anticipated a demonstration of the ability to collect acoustic and community reaction information along a potential LBFD flight route. This requirement was modified to estimation of interview completion rates in a nationally representative sample of households.

nationwide regulatory policy. Most fundamentally, this means that the method of selecting potential respondents must be as free from bias as possible, so that interview findings about the annoyance and startle of low-amplitude sonic booms may eventually be freely generalized to the U.S. population. These goals are most readily achieved 1) when the opinions of each and every member of the supersonically overflown population has an equal probability of representation in responses to questionnaire items; and 2) when respondents who grant interviews express opinions about their reactions to supersonic test flights that are as close as possible to the circumstances under which they will experience them in practice.

For telephone-based interviewing, the generalizability goal is most irreproachably attained in independent, random digit dialed (RDD) telephone samples. A large enough national RDD sample is self-stratifying, in the sense that all possible classes of respondents (whether defined geographically, socio-economically, ethnically, by population density, or in any other terms) are represented in proportion to their national incidence.

3.1.2 Preference for a cross-sectional design

Ideally, a single interview would be conducted with residents of supersonically overflown communities about their prompt reactions to individual low-amplitude sonic booms.[18] This approach is preferred because it is the least susceptible to criticism on the grounds that respondents' opinions might be affected by repeated interviews, or by prior notice of the occurrence of sonic booms. The generic arguments for such potential bias are that successive interviews could call respondents' attention to low-amplitude sonic booms in ways that differ from those of a naïve public, and that respondents' opinions about sonic booms could either sensitize or habituate to multiple, repetitive interviews.

The above advantages of a cross-sectional (independent sample) study design must be balanced against the likelihood of a low interview completion rate, which can greatly increase data collection costs. Low interview completion rates can also reduce the credibility of study findings by inviting doubt about the representativeness of sampled opinions. If interview completion rates are low enough, they can render a cross-sectional study infeasible.[19]

If telephone interview completion rates are not too low to be impractical (given the time constraints imposed on interviewing by the requirement for assessment of *prompt* reactions,) it is possible to cost-effectively increase the number of interview contact attempts by means of

---

[18] The size of the population exposed to carpet booms on individual LBFD missions is expected to be large enough that there is little likelihood that multiple independent samples would exhaust the population of respondents eligible for interview within carpet boom corridors adjacent to LBFD flight tracks.

[19] This is particularly the case when the constraint of interviewing respondents within a few minutes of exposure precludes the possibility of multiple callbacks to respondents who do not answer calls or decline to grant interviews.

automated calling that relies on outgoing interactive voice response technology. It was therefore recommended that interview completion rates be assessed by both automated ("IVR", and outgoing interactive voice response), and live agent ("CATI", or computer-assisted telephone interviewing) methods.

3.1.3 Interviewing

IVR interviewing can be highly cost-effective in periodic administration of brief, simple questionnaires, composed of a small number of closed response category items, in large, geographically dispersed panel samples (*cf.* Albert *et al.*, 2016, who reported a 68% interview completion rate for monthly IVR administrations of interviews.) Pre-contacting respondents, and offering small monetary incentives, is known to improve response rates (Cantor and Williams, 2013). Under such favorable conditions, IVR methods can quickly and relatively inexpensively produce large numbers of completed interviews.

Under other conditions, however, IVR methods can yield lower response rates than live agent interviewing. These conditions include non-incentivized, one-time interview attempts, which may be indistinguishable by potential respondents from telemarketing calls. With live interviewers, CATI methods can administer lengthier and more complex questionnaires, and can accommodate more complex responses. However, large-scale administration of live agent interviews is among the less cost-effective methods for assessing prompt responses to sonic booms.

3.2 Field Test of Prototype Aircraft Tracking and Acoustic Data Collection System

Figure 16 shows the system elements and communications links used. They consist of the central System Server (located in Middleton, Wisconsin); the Boom Monitor[20] (located in Costa Mesa, CA); two ADS-B receivers[21] (located in Irvine, CA and Costa Mesa, CA); a PC computer serving as a surrogate for the call center (located in Irvine, CA); and an analyst (located in Peterborough, NH). The element labeled "Flight Director" was the engineer controlling the experiment in an office in Irvine, CA. This configuration tests all of the system elements and their communication links with the server. The links between all components were via Internet.

The field demonstration tested whether the Internet-enabled data collection system could continuously track a target aircraft with no loss of position information, contact a call center in a timely fashion to control interview start and stop times in remote neighborhoods, and capture the sound pressure waveform of a target aircraft as it overflew a boom monitor station.

---

[20] Developed under Phase I of this contract.

[21] One was developed under Phase I of this project and the second under Phase II.

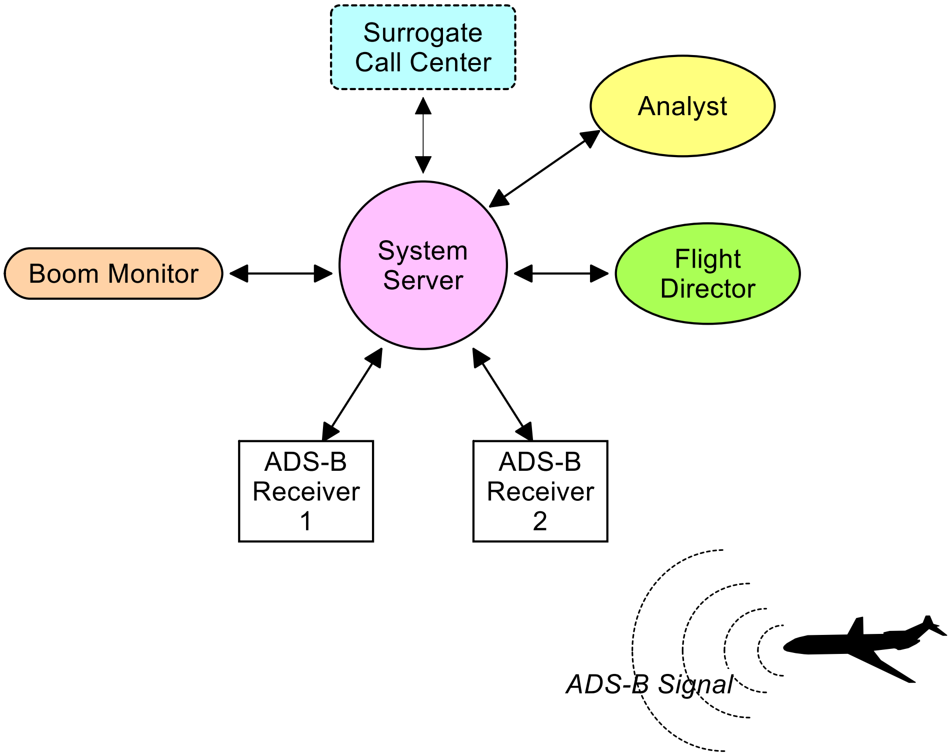


Figure 16. Portion of field measurement system tested in second year of contract.

3.2.1 Field test location

Commercial aircraft flights departing John Wayne Airport (SNA) in Orange County, CA served as LBFD surrogates. SNA uses a single runway for all commercial flight operations. The airport also has tightly controlled departure routings for commercial aircraft extending more than 50 nautical miles from the start of takeoff roll.

3.2.2 Airport and aircraft departure routes

Nearly all departures from SNA use runway 20R (196° magnetic). Figure 17 shows a collection of observed departure (blue) and arrival (green) tracks. The figure shows a large number of individual air carrier departure flight tracks from runway 20R in blue. (The tracks in green are arrivals on the downwind leg approach to Runway 20R.) After the first segment climb, departing commercial aircraft turn to a heading of 222° to follow Newport Bay (under the blue tracks) until passing the coastline. After passing the coastline, departing aircraft follow one of two corridors: 1) a right turn to the northwest, and 2) a left turn to the east. Departures are about equally split on both routes.

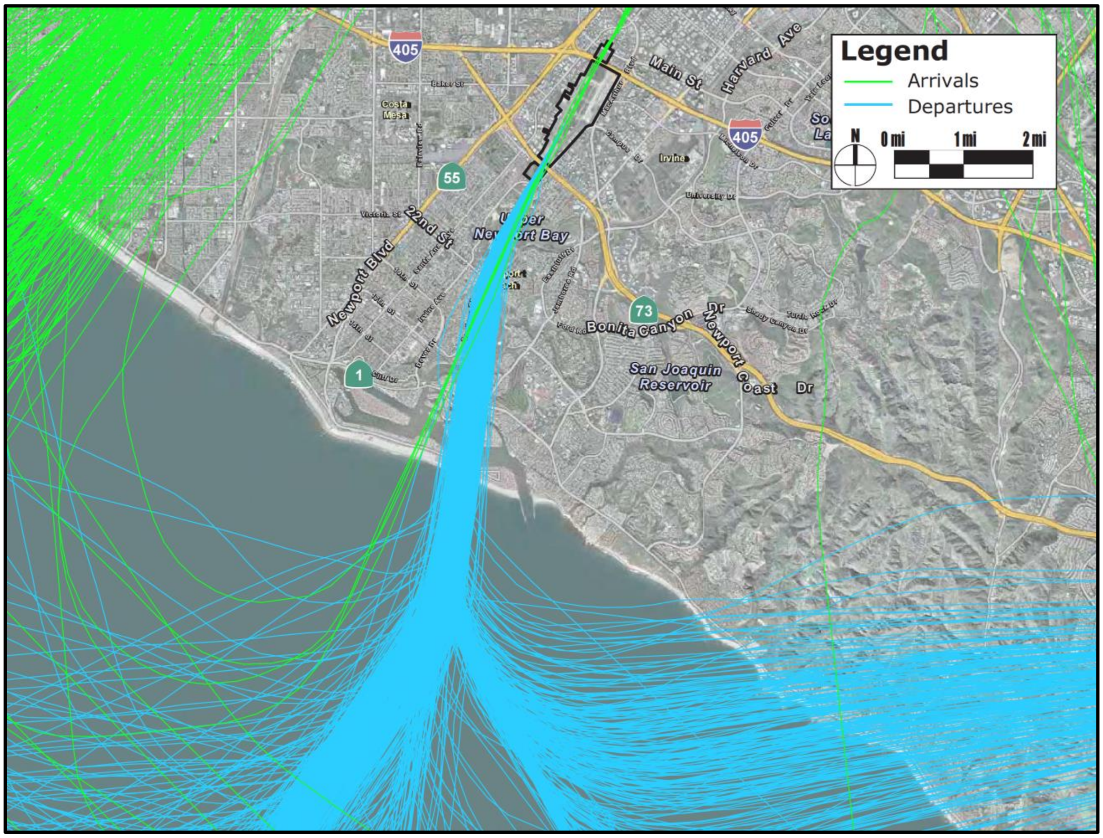


Figure 17. Typical departure flight tracks from Orange County Airport (SNA).

3.2.3 Tracking of aircraft flights

The goal of this portion of the field study was to demonstrate the reliability of a field-redundant aircraft tracking approach (two independent ADS-B receivers, each with separate Internet paths linking them to the central server). This subsection describes the field test of the tracking system.

ADS-B receiver locations

Both ADS-B receivers, each with a nominal unoccluded 200-mile reception range, were able to acquire the ADS-B messages transmitted by commercial flights selected for monitoring. One receiver was located on airport property so that an uncompromised line of sight could be maintained to aircraft as they prepared for departure, taxied for takeoff, and departed the airport. The second was located on the east side of Newport Bay 3.54 nautical miles to the south southwest nearer the coast. The locations are shown by the yellow pins in Figure 18. The nominal departure route from the airport to the coastline is shown in red.

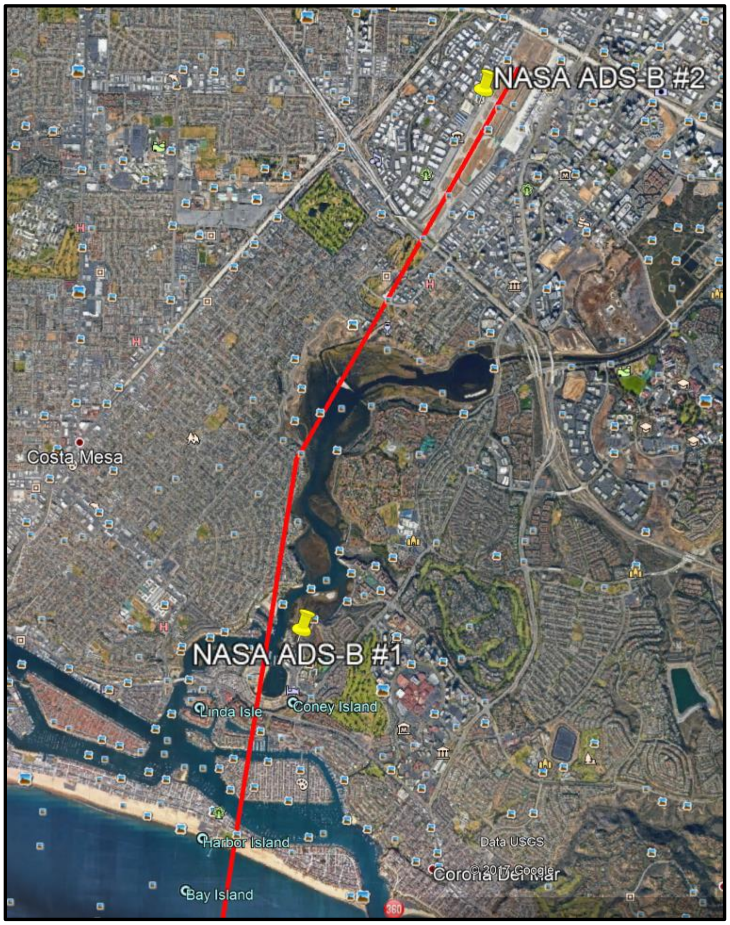


Figure 18. Locations of ADS-B receivers.

ADS-B receiver function

ADS-B equipped aircraft broadcast time, position (latitude and longitude), speed and altitude (along with other parameters) twice per second using an on-board GPS receiver for position determination. Unlike radar, the system is completely passive and any commercially available ADS-B receiver can decode the transmitted data packets. This information (along with the Mach number from the LBFD aircraft) will enable the server to continuously determine the geometric relationship between the aircraft Mach cone and any point on the ground. Hence, the server can also determine when to message the call center and provide a time-of-day to begin interviewing a particular neighborhood. Discussion of the messaging tests may be found in section 3.2.5 on page 38.

Figure 19 shows the ADS-B field receiver station components (yellow) and the data transmission path to the server. It also shows the call center (pink) and its data transmission path to and from the server.

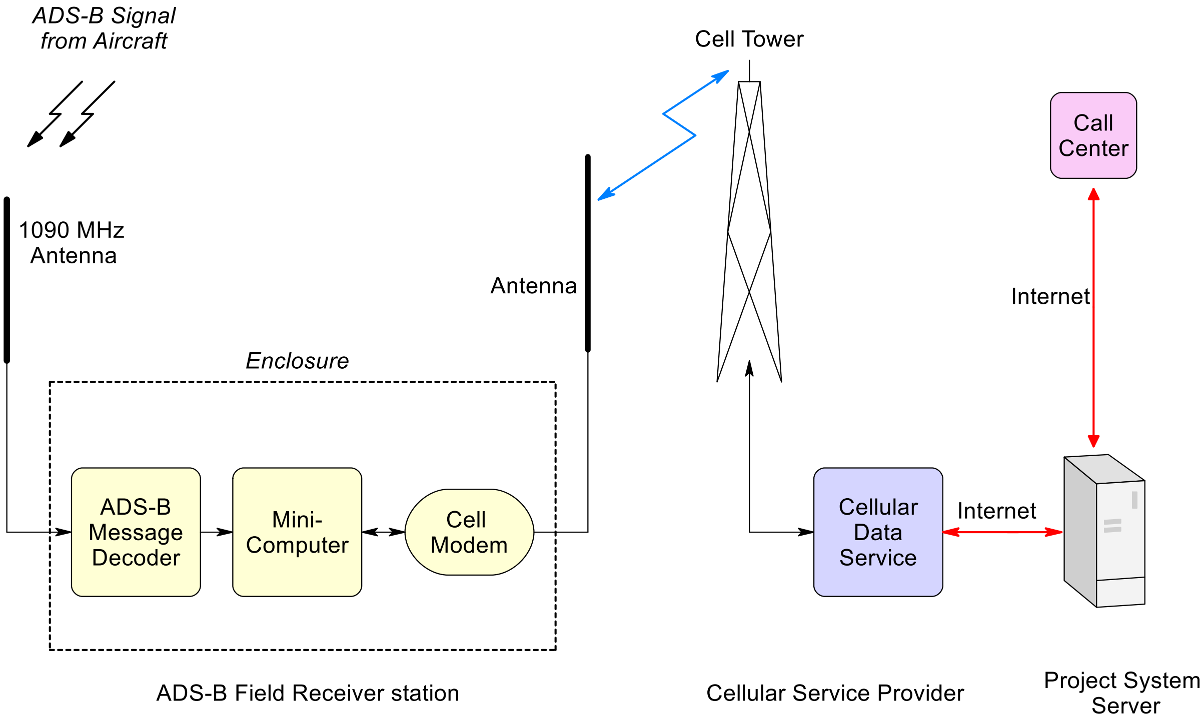


Figure 19. Data transmission diagram for interview triggering.

The ADS-B message decoder (leftmost yellow box in Figure 19) decodes all aircraft ADS-B transmissions within its reception range. Each decoded transmission is read by a mini-computer (middle yellow box) which in turn sends a corresponding data block to the server via a cellular modem (rightmost yellow box). Each data block contains the unique ADS-B code assigned to the aircraft along with parameter data. The data block is first transmitted to a nearby cell tower via a cellular modem as a series of data packets (the arrowed blue line indicates the link between the receiver station and cell tower). Since the data packets from the mini-computer contain the IP address of the server the cellular service provider can put those packets on the Internet (red line) for delivery to the server.

By receiving transmissions from all aircraft within range of the receivers the server is able to display all air traffic in the area to the user. It is also able to select just those transmissions from the target aircraft being tracked (using the user-specified ADS-B aircraft code) to monitor the progress of that aircraft over the ground.

Both the cellular data service and the Internet are well-developed and reliable technologies. However, temporary local cellular outages and data transmission latencies through the cellular data service and the Internet have the potential for interrupting the timely tracking of an LBFD aircraft. Redundant coverage of an airspace by two receivers minimizes this risk. Prior to this field test the data block message delivery latency was largely unknown, as was the availability of every half-second transmission from the aircraft. The next subsections discuss the portion of the field study that assessed the transmission latency and the temporal spacing of received data messages from the target aircraft.

Tracking an aircraft

The "mission director" (for this test a contractor engineer) watched for departing aircraft by zooming the tracking system display in to the airport terminal, taxiways and runway. Such a display of ground activity is shown in Figure 20. In this figure, the aircraft broadcasting the ADS-B code "A11AD4" has been pushed back from the gate and is taxiing into position along the northerly-most taxiway for departure on runway 20R. A second aircraft "AD6C3B" is being pushed back from its gate in the upper right. Yet a third aircraft, "A63881", is taxing north from the bottom of the figure.[22] The location of ADS-B receiver #2 may be seem as the inverted green triangle in the lower left of the figure.

Figure 20. Mission director screen view of airport activity.

At this time the mission director entered the ADS-B code into the server system as the aircraft to be tracked. When the aircraft commenced its takeoff roll down the runway the mission

---

[22] The green triangle on the left is/ the "on airport" ADS-B receiver. The dashed lines emanating from the receiver link to each aircraft being seen being seen by the receiver.

director clicked the “Start” button on his display screen to begin the tracking, data acquisition and messaging process. At that point all of the data acquisition was under server control.

As the aircraft progressed down the runway, the mission director zoomed out on the map display to observe the aircraft passing through the predetermined gates. A sample system display of a right-turning aircraft is shown in Figure 21. An example of a left-turning aircraft is shown in Figure 22.

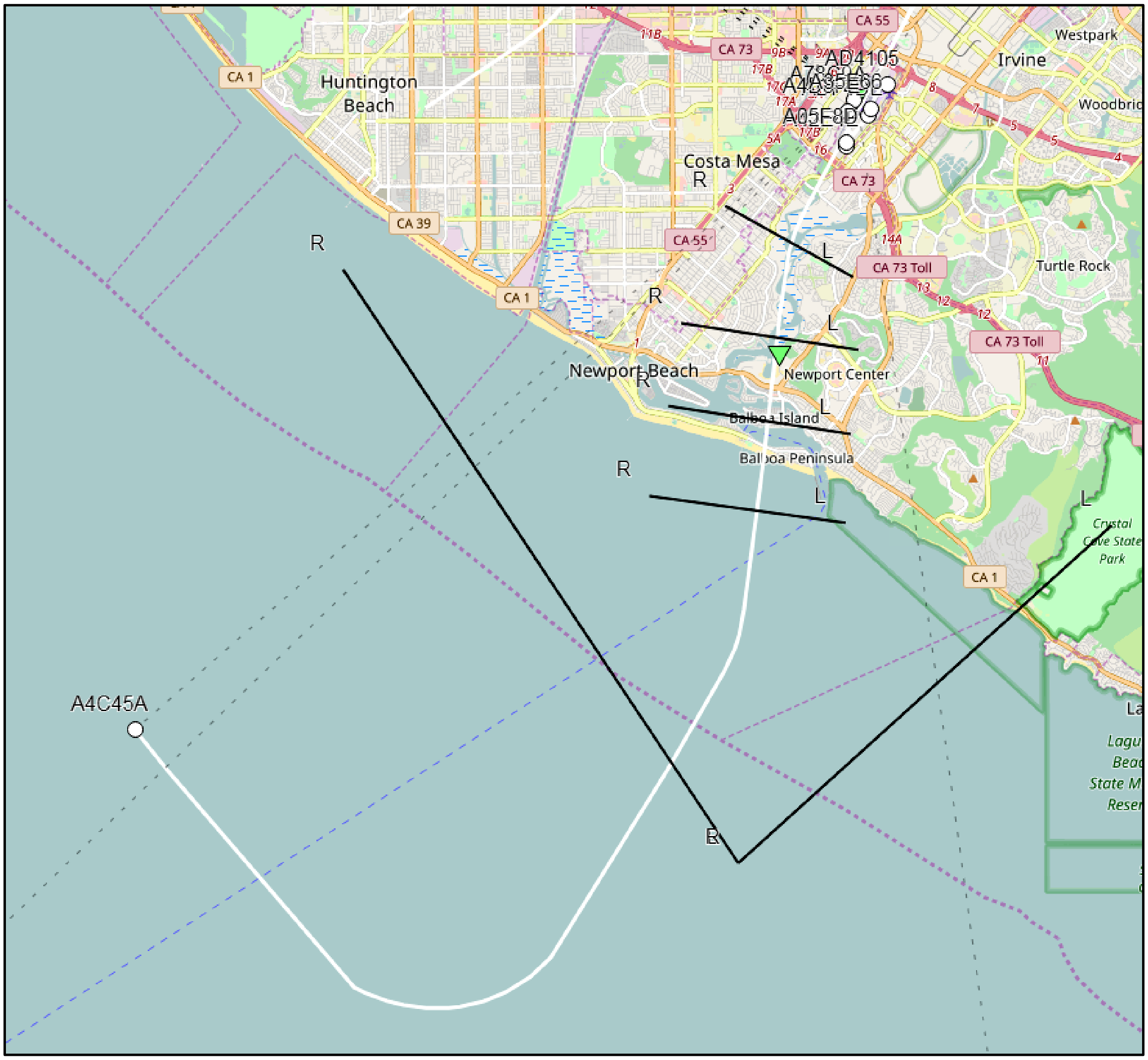


Figure 21. Screen capture of northbound aircraft departure from SNA.

The displays show the target being tracked (white circle or aircraft outline) and the path it has followed (white line). It also shows a series of black line segments. These are “gates” the aircraft crossed that were used by the system server in this test to trigger a message to a simulated call center at the times of each crossing. This messaging is discussed further in section 3.2.5. In this capacity a gate crossing served as a system-determined occurrence to trigger a message (in lieu of a system-determined occurrence of Mach cone crossing a community).

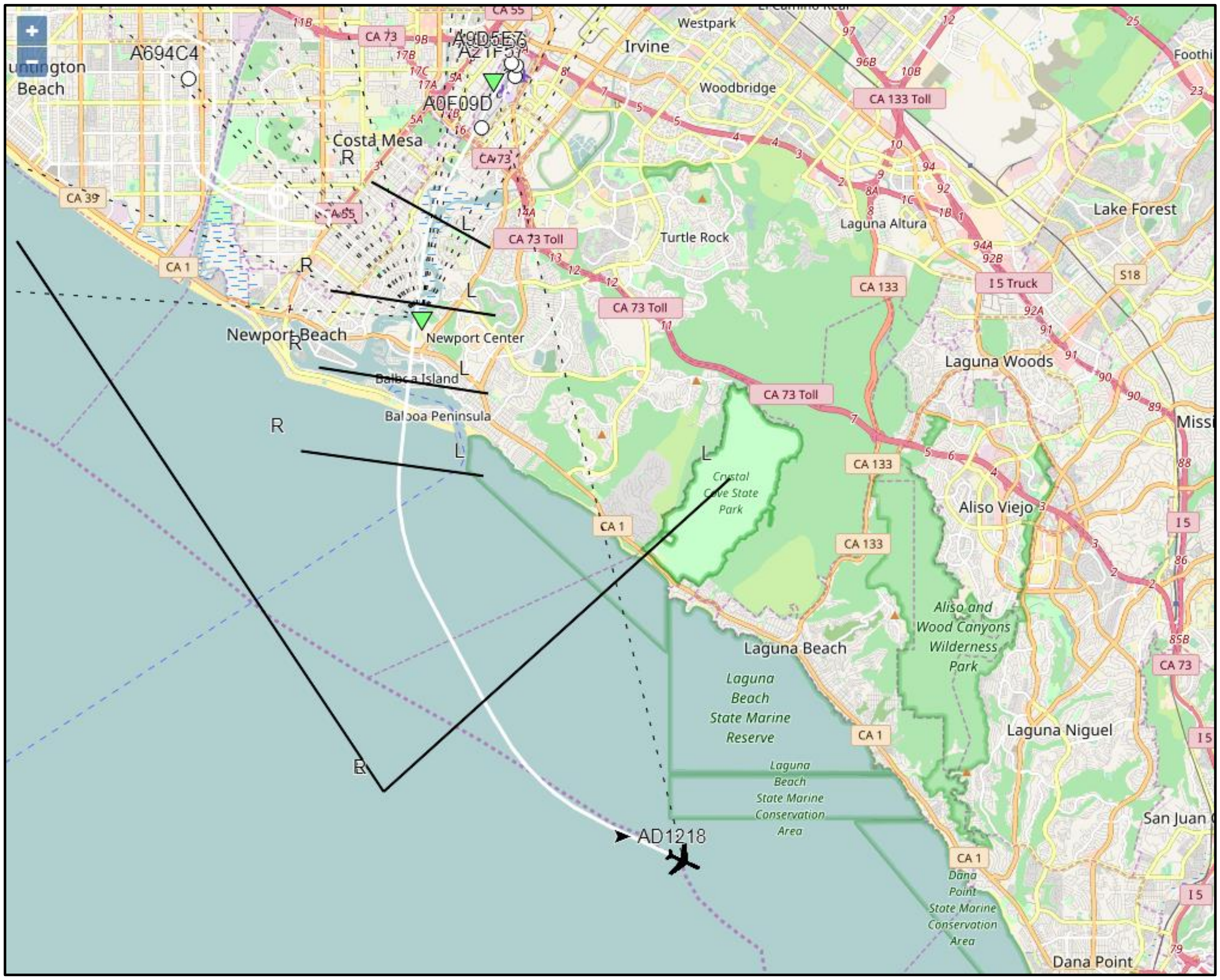


Figure 22. Screen capture of departing aircraft bound for an easterly or southerly destination.

All aircraft passed through the first four gates (conforming with the SNA noise abatement departure procedure). Gates 5 and 6 form a "V" pattern so that a departing aircraft must pass through only one or the other. The vertex of the "V" was established based on the tracks observed in Figure 17 on page 29. After the aircraft passed the last gate, the mission director clicked the "Stop" button on the system display to end the tracking process.

At this point the server placed all of the data files related to the mission in a mission-specific folder ready for upload. At that time the data set was available to anyone connected to the system. In this case the analyst in Peterborough, NH was able to upload the folder to his personal computer. However, the folder remained on the server until intentionally purged by the mission director at the end of the exercise. During the four days of testing the aforementioned analyst monitored progress of the field tests and downloaded the data as it became available. This allowed for ongoing monitoring of system performance (even down to the level of field unit battery voltages and ambient sound levels at the boom monitor) as well as performing spot checks to ensure all components were performing properly. The remaining tables and graphics presented in this report section were created by the analyst from the unloaded data.

3.2.4 Post-processing of ADS-B tracking data

Departing aircraft all passed through Gates 1 through 4 and then through either Gate 5 or 6, depending on whether they initiated a right or left turn, respectively. Figure 23 shows an example of a right turning aircraft. The plotting circles indicate individual ADS-B position reports captured by the monitoring system. The runway is shown in red in the top center of the figure. The gates are shown as blue line segments with the gate number at one end. This track consists of 362 position reports over a distance of 25.0 nautical miles from start of takeoff roll.

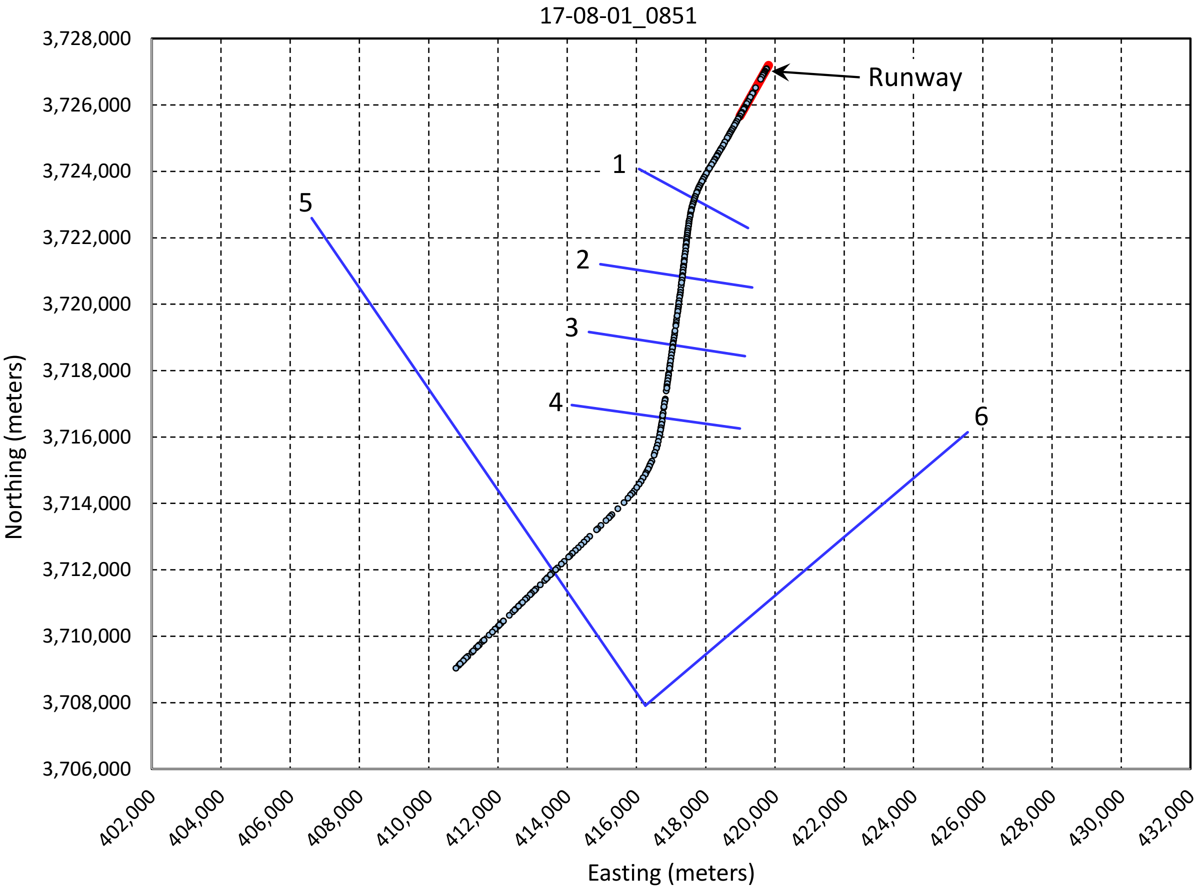


Figure 23. Example of flight track and gates (right turn passing through Gate 5).

Similarly, Figure 24 shows an example of a left-turning aircraft. This track consists of 245 position reports over a distance of 25.7 nautical miles.

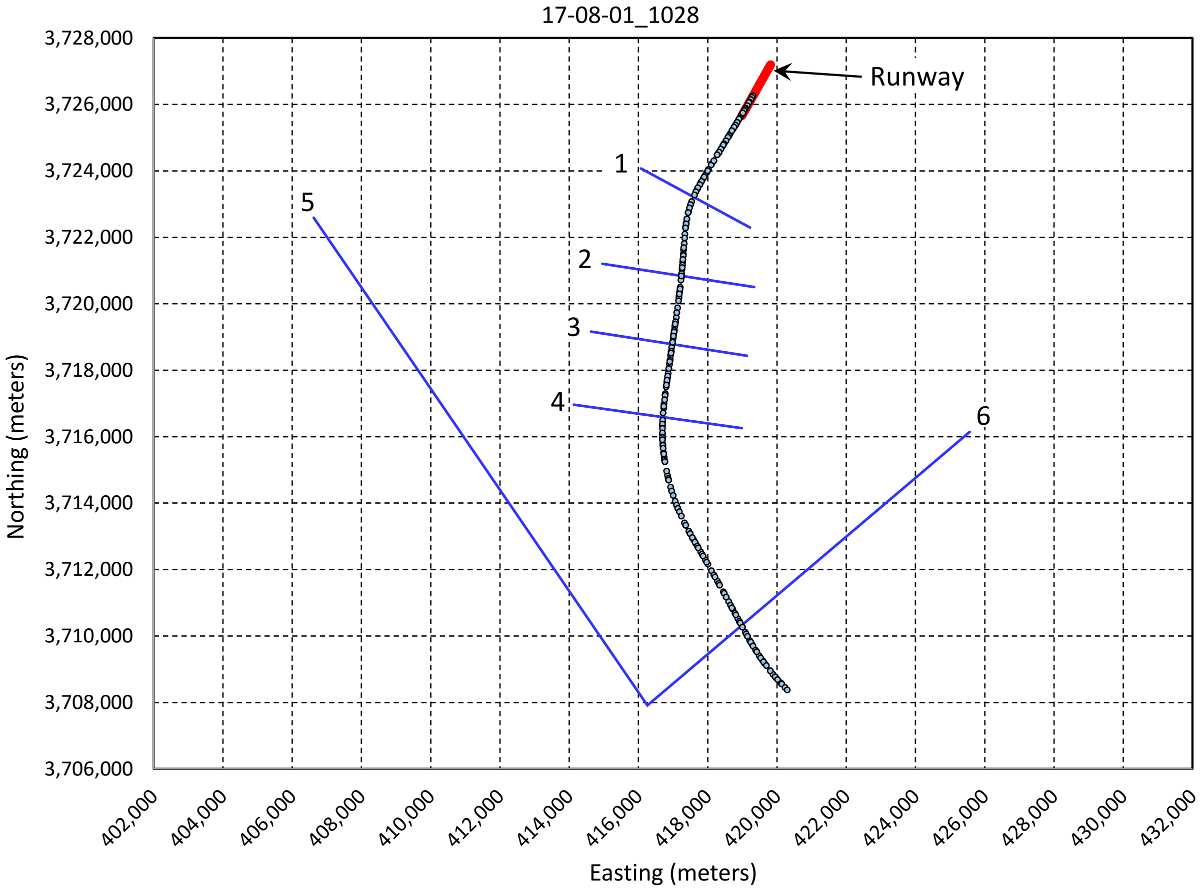


Figure 24: Example of flight track and gates (left turn passing through Gate 6).

Each ADS-B message also contained altitude and ground speed information. Graphs containing this information can be readily produced and an example is shown in Figure 25. The two parameters are plotted as a function of track of time of day. In this figure it may be seen that the rate of climb is reduced when the aircraft reaches 1000 feet MSL in accordance with the mandated power cutback at this altitude. Figure 26 shows the same parameters as a function of track distance from the start of takeoff roll. The track distance was determined by converting the latitude and longitude of each position report to Universal Transverse Mercator (UTM) coordinates in meters and then computing the distance between successive points. The latitude and longitude of the start of takeoff roll on Runway 20 were also converted so that the distance between that point and the first position report from the aircraft could also be determined.

These graphics show the level of detail available for tracking the LBFD aircraft and determining the geometric relationship between the aircraft and any point in an interviewing community. Such graphs could readily be provided for an entire LBFD mission with waypoints plotted for added interpretability.

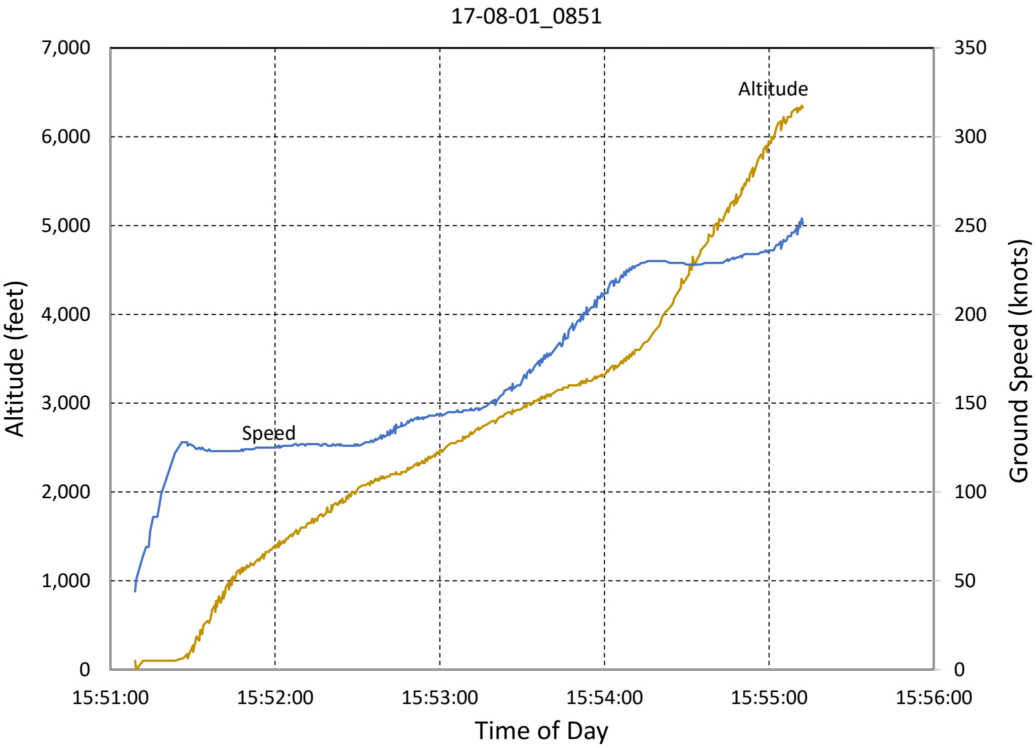


Figure 25. Example of speed and altitude profiles as function of time.

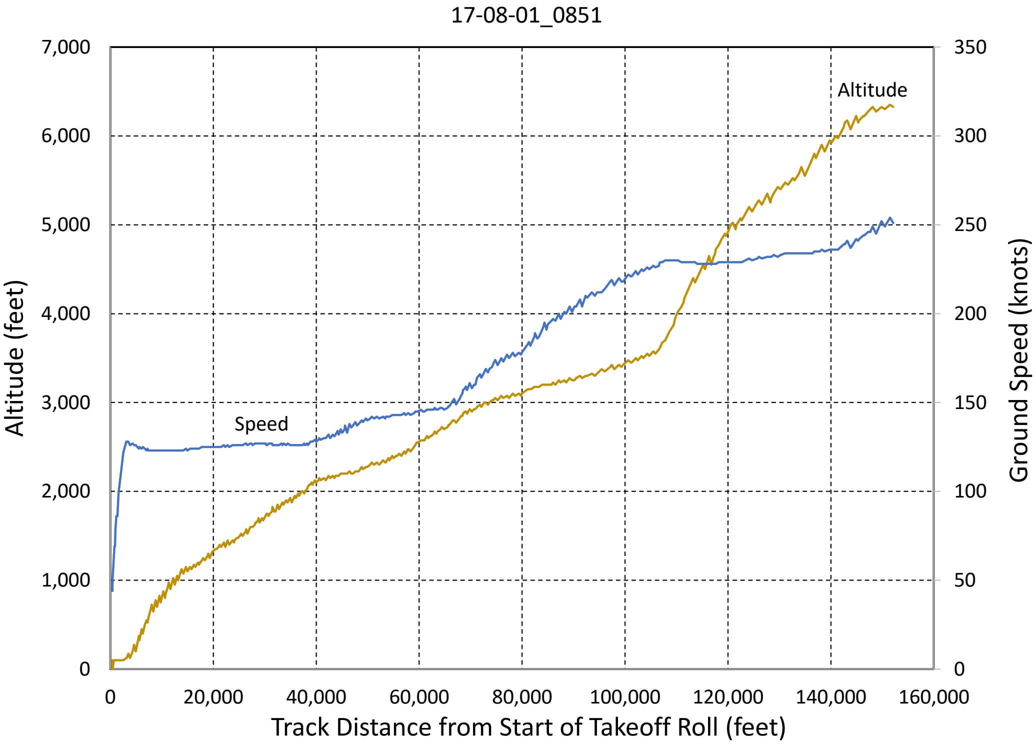


Figure 26. Example of speed and altitude profiles as function of track distance.

3.2.5 Messaging the (simulated) call center

The server monitors the progress of each tracked aircraft to decide when to send a message to the call center indicating the time of day at which to begin interviewing a particular neighborhood. During an actual LBFD flight, the server would use the aircraft trajectory, Mach number, and altitude to determine the relationship between the Mach cone and the ground and when the shock will have traversed the entire interview neighborhood. This time-of-day would typically be calculable one to two minutes in advance of the shock actually insonifying an entire neighborhood, and the message "Begin interviewing in neighborhood X at the following time-of-day" can be sent to the call center at the time it is determined.

For the present purposes of testing the reliability of the server's messaging function, there was no reason to incorporate a sophisticated sound propagation algorithm into the server software. It sufficed instead to place "gated" at various locations across the flight corridor to represent an approach to an interview neighborhood, and to determine the time that an aircraft flew through the gate, based on the ADS-B data stream. A set of several such gates was shown in Figure 21 through Figure 24. The first gate was located at the point where aircraft made a slight left turn from runway heading to head out Newport Bay. The next three gates were located along the track leading out the bay to the coastline. These gates are about 1.2 nautical miles apart. Since aircraft turn either left or right after reaching the ocean single gates can no longer be used through which all aircraft will fly. Instead, two gates were placed across each of the two routes the fifth gate (right turn) is approximately 2.9 nautical miles down track from the fourth gate, and the sixth gate is approximately 3.7 nautical miles down track from the fourth gate. These gates have a common end point forming a "V" formation. This configuration meant that air aircraft must fly through one gate but cannot fly through both. Hence, departing aircraft flew either through gates 1, 2, 3, 4, and 5 or through gates 1, 2, 3, 4, and 6.

Messaging the call center (represented by the light blue block at the top of Figure 16 on page 28) required a direct Internet connection (red line) from the server. For the purpose of this test the call center was composed of a dedicated PC with a fixed IP address. The clock on this computer was synchronized to National Institute of Standards (NIST) time as was the clock on the system server. The NIST time base is the same as the time base reported by a GPS receiver (well within one millisecond). This provided a common time base for the ADS-B messages and measurement system components.

Each time a position report from the tracked aircraft was received the server compared this position with the previous one to determine whether a gate crossing has occurred. If so, the server sent a message to the call center containing the gate number crossed. The call center logged this message's time of arrival and sent a return a message to the server containing that time for subsequent latency determinations. These time stamps were recorded to the nearest 0.01 second and saved for post-processing analysis.

The latency between the time of the aircraft gate crossing and receipt of the message at the call center has two components. The first is the combined transit time from the ADS-B receiver through the cellular data link and Internet to the server[23]. The second is the Internet transit time from the server to the call center. The server placed this timing information along with the before- and after-gate aircraft positions in a database for each flight. The next three tables show the contents of that database for an aircraft tracked on 01 August starting at 10:28 am. This aircraft turned left after reaching the ocean and flew through gates 1, 2, 3, 4, and 6.

Table 2 shows the information stored for the last position *prior to* the gate crossing. The parameters include time of day the data were generated, latitude, longitude, speed, altitude and track distance before the gate. Table 3 shows the information stored for the first position *after* each gate crossing. The parameters saved are the same in both tables.

Table 4 shows the calculated transmission latencies. The first column shows the time the server received the first ADS-B position report after the gate crossing. The second column shows the time the message was sent to the call center. The two times are the same within 0.01 seconds so server processing time was negligible. The third column (Client Recv Time) shows the time the message was received at the call center. The fourth column of Table 4 shows the interpolated gate crossing time between the two position reports before and after the gate crossing[24]. The fifth column shows the latency between the time the after crossing ADS-B message was generated and the time it was received by the server (Internet transmission time). The last column shows the latency between the time the message was sent to the interviewing contractor message and the time when it was received. The sum of the last two columns becomes the total latency.

A total of 667 such gate crossing messages were sent during the field exercise. The simulated call center received and sent return messages for all of them. Figure 27 shows the distribution of the message latencies. The top panel of the figure shows the ADS-B receiver transmission latencies to the server (this latency also includes the time between the actual gate crossing and the time of the first position report *after* the crossing (about 0.3 seconds on average). The mean latency is about 1.5 seconds and the range is approximately 0.4 to 3.2 seconds. The middle panel shows the messaging latency from the server to the call center computer. In this panel two groupings may be seen: 0.2 to 1.2 seconds and 2.6 to 4.6 seconds.

The bottom panel of Figure 27 shows the distribution of the combined latencies of the two messages. This panel indicates that the worst-case delay is on the order of 6.4 seconds. This delay

---

[23] All aircraft ADS-B messages detected by the field receiver are sent to the server. The server distinguishes the aircraft being tracked from all others.

[24] The interpolation is based on the track distances before and after the gate crossing shown in the last columns of Table 2 and Table 3. Information for first position report after gate

is well within the allowable time for alerting the call center of when to start interviewing a specific neighborhood. During a LBFD flight the shock propagating to an interview neighborhood would have left the aircraft on the order of one to two minutes prior to its arrival. That projected propagation time, plus a small margin of safety, would become the messaged time at which interviewing should commence.

Table 2. Information for last position report before gate crossing.

Mission Name:	17-08-01_1028					
Mission Director:	Eric					
ADS-B Code:	A0F09D		GMT		PDT	
Start Time:	2017-08-01T17:28:23.249Z		08/01/17	17:28:23	08/01/17	10:28:23
Gate	Last ADS-B Time Before Gate (GMT)	Last Position Before Gate [LAT]	Last Position Before Gate [LON]	Last Velocity Before Gate (kt)	Last Altitude Before Gate (ft)	Last Distance Before Gate (mi)
1	17:29:01.06	33.6470	-117.8880	210.6	1350	-0.11
2	17:29:21.14	33.6264	-117.8920	239.7	2275	-0.16
3	17:29:37.51	33.6077	-117.8945	251.2	3100	-0.15
4	17:29:53.75	33.5886	-117.8973	257.2	3900	-0.20
5						
6	17:30:42.72	33.5322	-117.8743	272.4	6200	-0.21

Table 3. Information for first position report after gate crossing.

Mission Name:	17-08-01_1028					
Mission Director:	Eric					
ADS-B Code:	A0F09D		GMT		PDT	
Start Time:	2017-08-01T17:28:23.249Z		08/01/17	17:28:23	08/01/17	10:28:23
Gate	First ADS-B Time After Gate (GMT)	First Position After Gate [LAT]	First Position After Gate [LON]	First Velocity After Gate (kt)	First Altitude After Gate (ft)	First Distance After Gate (mi)
1	17:29:02.19	33.6443	-117.8893	215.0	1475	0.09
2	17:29:21.96	33.6237	-117.8922	243.9	2400	0.03
3	17:29:38.16	33.6047	-117.8950	252.2	3225	0.06
4	17:29:55.39	33.5846	-117.8977	259.3	4050	0.08
5						
6	17:30:43.72	33.5289	-117.8723	272.4	6400	0.04

Table 4. Computed data transmission latencies.

Mission Name:	17-08-01_1028					
Mission Director:	Eric					
ADS-B Code:	A0F09D		GMT		PDT	
Start Time:	2017-08-01T17:28:23.249Z		08/01/17	17:28:23	08/01/17	10:28:23
Gate	Server Recv Time (GMT)	Server Send Time (GMT)	Client Recv Time Max (GMT)	Interpolated Gate Crossing Time	ADS-B Received Latency (O -Z)	Transmit Call Center Message Latency (R - O)
1	17:29:02.60	17:29:02.60	17:29:06.17	17:29:01.68	00:00:00.92	00:00:03.57
2	17:29:22.72	17:29:22.72	17:29:26.17	17:29:21.83	00:00:00.89	00:00:03.45
3	17:29:38.78	17:29:38.78	17:29:42.16	17:29:37.97	00:00:00.81	00:00:03.38
4	17:29:56.04	17:29:56.04	17:29:59.17	17:29:54.92	00:00:01.12	00:00:03.13
5				---	---	---
6	17:30:44.58	17:30:44.58	17:30:48.17	17:30:43.56	00:00:01.02	00:00:03.59

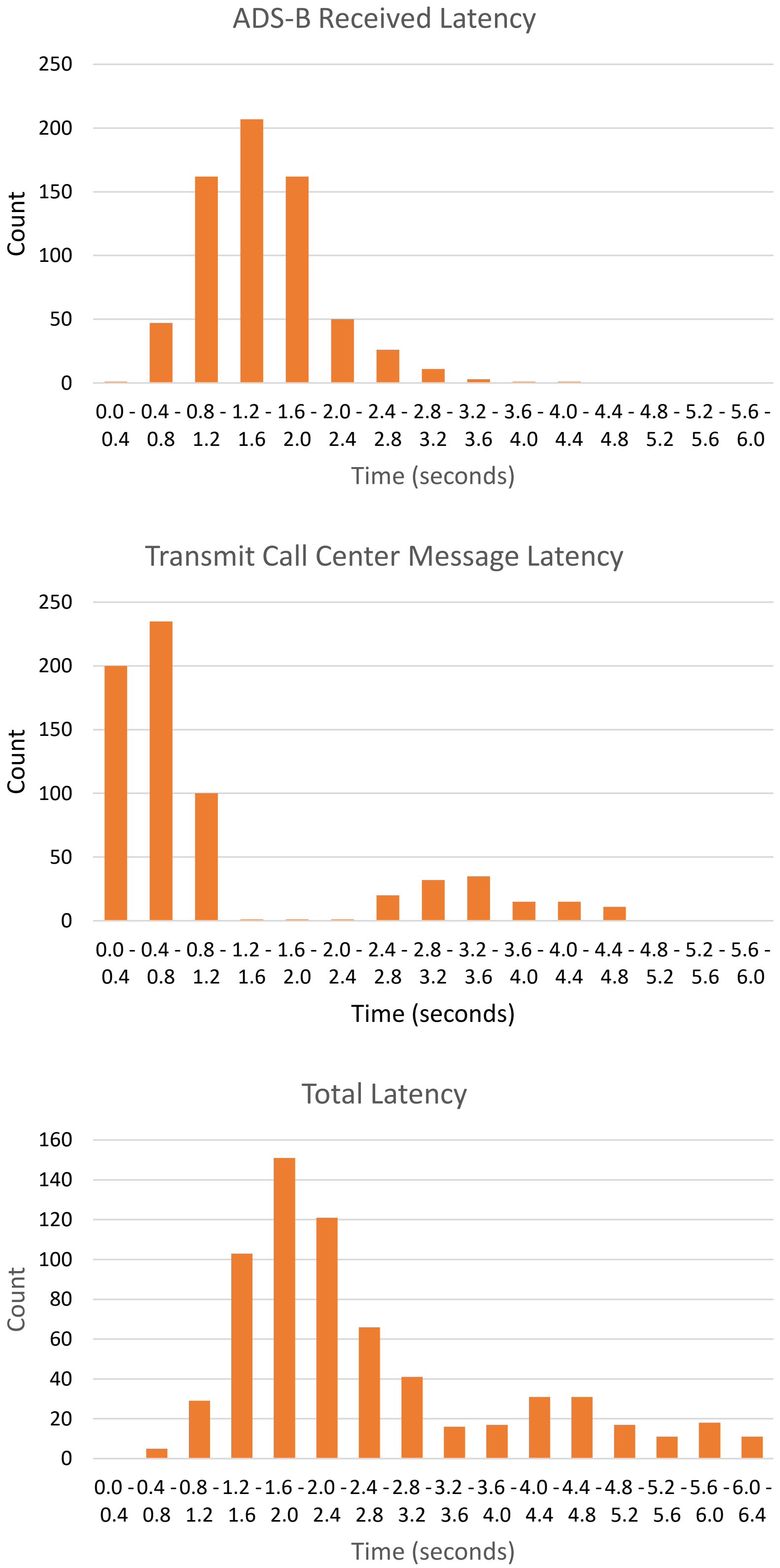


Figure 27. Message transmission latencies.

3.2.6 Capturing sound pressure waveforms

Boom monitor location

The boom monitor was located at a position collinear with Gate #1, shown as the orange pin in Figure 28. This location was in a nature reserve along Newport Bay and away from local suburban noise sources.

Figure 28. Boom monitor location.

Boom monitor operation

The boom monitor continuously recorded the sound pressure waveform from the microphone regardless of whether an aircraft was being tracked by the system. The waveform was stored within the field unit as sequential 1-minute long .wav files that began exactly on the minute (as determined by the GPS receiver within the unit). The files were named with the starting date and time.

Storing the waveform captured during missions

The time of day was noted by the system server when a tracked aircraft passed through Gate #1 The server could then determine the name of the .wav file whose time interval included the gate crossing time. The server then requested the field unit to upload this file, as well as the preceding and succeeding ones. The three files were combined during post processing into a single 3-minute file.

Figure 29 shows such a 3-minute .wav file. The passage of the aircraft is clearly visible two-thirds of the way through the record.

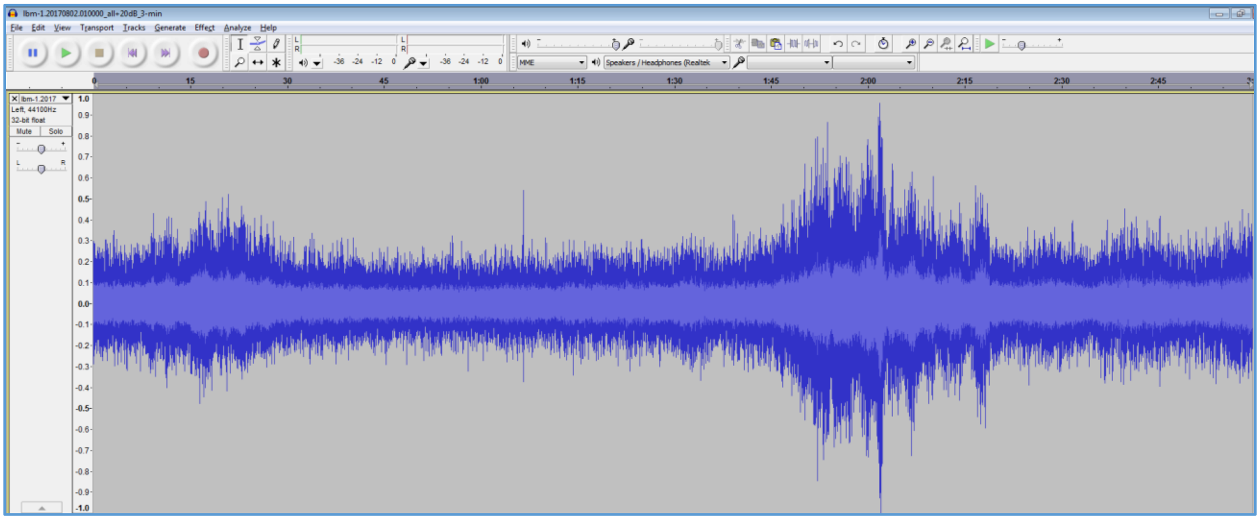


Figure 29. 3-minute waveform captured by system.

To show the aircraft passby more clearly the file was used to generate a slow-response, A-weighted sound level time history using a MATLAB-based virtual sound level meter (Muehleisen, 2016). This is shown in Figure 30. Also shown in the figure is the time at which the tracked aircraft passed through Gate #1. As expected, the maximum sound level lags the gate crossing time (in this example, by 2.3 seconds).

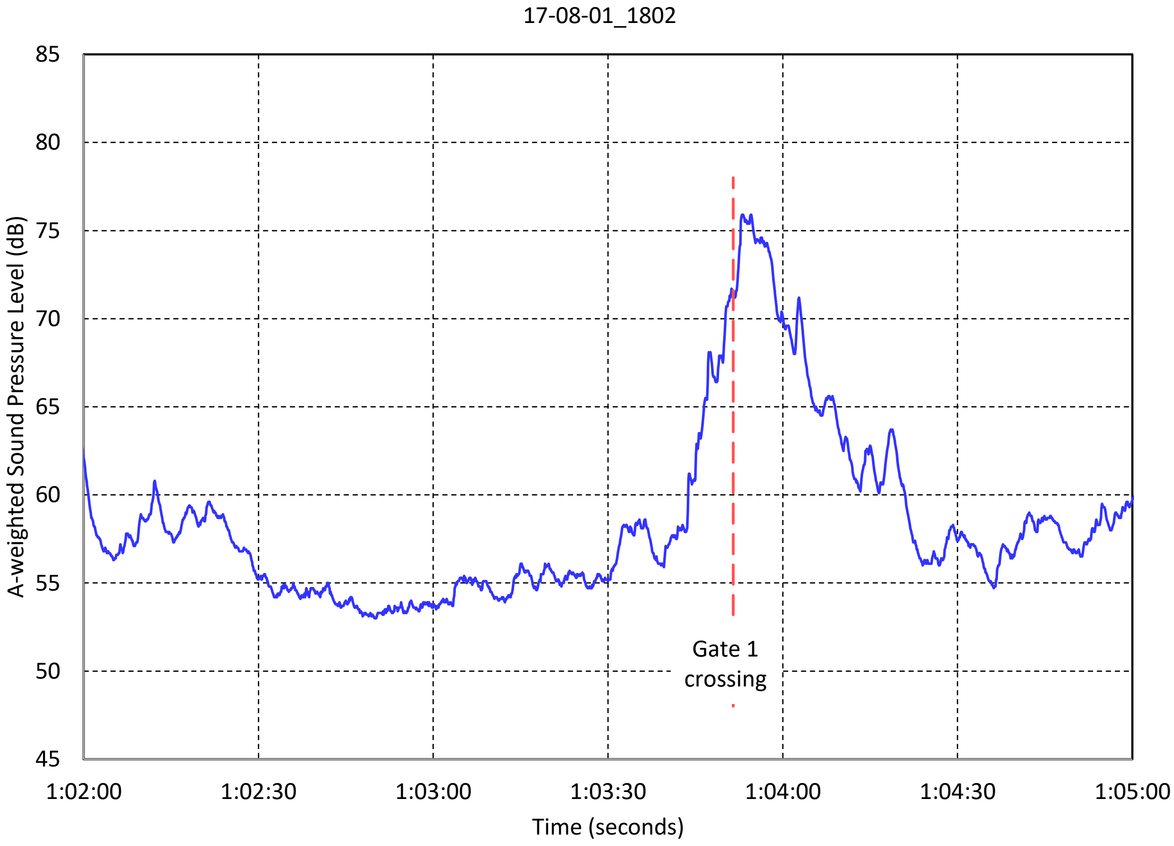


Figure 30. Example of A-weighted sound level time history derived from captured wave file.

THIS PAGE INTENTIONALLY LEFT BLANK

4 SUMMARY OF THIRD YEAR'S EFFORT (PHASE II – YEAR 2)

4.1 Community Response Technical Background

The ultimate goal of NASA's anticipated LBFD field deployments is to collect information about community reaction to low-amplitude sonic booms that may be of use to FAA and ICAO in reconsidering their current prohibitions on overland supersonic overflights (*i.e.*, CFR Title 14, Chapter I, Subchapter F, Part 91, Subpart I, Section 91.817.) Information of this sort is typically summarized in dosage-response relationships, such as those of FICON (1992) and ISO 1996-1 (2016).

For greatest face validity of findings, the annoyance and startle of exposure to low-amplitude sonic booms should be studied under conditions as similar as possible to exposure to low-amplitude booms created by a future fleet of supersonic aircraft. To be of greatest use for regulatory analyses, the findings of such studies must be fully generalizable to the national population, as well as appropriately accurate and precise. Some of the factors affecting the credibility of study findings, as well as errors of modeling and acoustic measurement, and of survey sampling, are noted in the following subsections. Study designs must control or otherwise address all of these. They must also yield information consistent with regulatory agencies' rationale for aircraft noise management policies.

4.1.1 Utility of evidence from field experiments and social surveys

Sonic boom-induced annoyance has been studied by both observational (social survey) and experimental (closely-controlled exposure) methods, in both laboratory and field settings. While laboratory studies permit greater precision of measurement, field studies offer at least the possibility of greater realism of exposure conditions and interpretability of findings for regulatory purposes. A few hybrid studies have attempted to combine the advantages of laboratory and field settings by various means. For example, some recent controlled-exposure field studies (Fidell *et al.*, 2012; Page *et al.*, 2014) restricted test participants to areas small enough to be intentionally exposed to sonic booms of credibly modeled or measured levels, and provided test participants in remote locations with convenient means for judging the annoyance of exposure in near-real time.

Historically, most field experiments of reactions to sonic booms have instructed recruited test subjects to judge the annoyance of individual booms while seated in fixed locations, either outdoors (*e.g.*, Kryter *et al.*, 1967) and/or in instrumented buildings (*e.g.*, Sullivan *et al.*, 2010). Annoyance judgments made under such circumstances still lack much of the naturalistic context in which sonic booms are experienced in real life. In real life, groups of people are not paid to listen attentively to structured series of sonic booms for extended periods of time. Further, booms heard at expected times in experimental settings are less likely to startle people than booms heard at unpredictable times, and may be less annoying because they do not interfere with ongoing life activities.

Like laboratory studies, field experiments which seek judgments of the annoyance of individual sonic booms in near-real time attempt to control sound levels of booms presented for judgment. Unlike laboratory studies, such control is not easily achieved in settings that do not confine test participants to a fixed location at the time of exposure to low-amplitude sonic booms from an LBFD flight mission. Test participants in such studies may be at home or away from home, or indoors or outdoors, when exposed to a shock wave from an LBFD overflight.

Response information inferred from individual judgments in such field experimental circumstances does not constitute evidence of actual community response to booms, but only of individual judgments by paid observers of single boom events. The credibility for regulatory purposes of a dosage-response relationship developed from a controlled, single area, fixed location field experiment may be challenged on two grounds: first, it lacks the real-life context of everyday circumstances of exposure to sonic booms; and second, it ignores potential community-specific differences. In other words, the “community” component of “community response” is missing from such a data collection exercise. Ignoring this component is equivalent to assuming that community response to sonic booms is determined exclusively by acoustic variables – a view at odds not only with common experience in airport noise controversies, but also with the latest international technical consensus standard (ISO 1991-6:2016).

Since dosage-response information of greater face validity for regulatory purposes cannot be developed from studies that constrain the normal movements and activities of test participants, means must be found 1) to estimate respondents’ levels of exposure to sonic booms under field conditions, and 2) to gauge their responses to such exposure in as prompt and non-intrusive a manner as practical.

The great bulk of information about community reaction to aircraft noise in airport communities has been collected in cross-sectional social surveys. Reactions to individual sonic booms collected in field experiments differ appreciably from reactions to habitual exposure to aircraft noise in airport communities. Conventional, airport-vicinity social surveys do not rely on self-report at times of respondents’ choosing, but rather actively solicit responses from respondents at specific times. FAA and ICAO could well object on methodological grounds to comparisons of judgments collected by disparate means of the annoyance of sonic booms and of conventional aircraft noise.

Some of the differences between naturalistic, observational field studies and controlled field experiments are unavoidable. To the extent possible, however, a field study should permit assessment of reactions to individual low-amplitude booms in as brief and effort-free (for test participants) a manner as possible. Collection of reactions to individual exposures should not require respondents to interrupt ongoing activities more than absolutely necessary, or repeatedly spending non-trivial amounts of time answering repetitive questions.

Test participants also should not be cued in advance, nor frequently reminded to listen for and attend to low-amplitude booms.[25] Repeated reminders to report reactions to low-amplitude booms also risks over-estimation of their annoyance, because such reminders could transform annoyance judgments into alerted detection reports. The distinction between mere notice of an unusual noise event and a general adverse reaction (annoyance) to the noise event could well be blurred in the minds of conscientious test subjects who have been paid to listen for particular sounds in a field experiment.

Precautions must also be taken in free-response study designs to determine whether the absence of self-reports is attributable 1) to a lack of notice of low signal-to-noise ratio sounds, 2) to panelists' absence from a boom target area at the time of exposure, or 3) to an inability to respond at a given moment. In principle, personal location information can be automatically extracted from smart phone self-reports, but only at the expense of privacy-invading geo-tracking of panelists.

4.1.2 Self-report versus solicited responses to individual sonic booms

Recent research (*e.g.*, Fidell *et al.*, 2012; Page *et al.*, 2014) on residential reaction to low-amplitude sonic booms has been conducted in specialized and geographically constrained populations, such as communities near a military airfield whose residents are commonly exposed to higher amplitude sonic booms. These studies relied on "free-response" data collection methods, in which test participants reported their reactions to booms at times of their choosing. As pilot studies based on self-selected samples of convenience, rather than on nationally representative samples, the findings of such studies are not readily generalizable beyond the opinions of test participants, and so are unlikely to be fully persuasive to FAA and ICAO.[26]

The conditions of exposure to low-amplitude booms in recent research more closely resemble controlled field experiments than real-world conditions of exposure of the sort likely to characterize exposure to low-amplitude sonic booms created by a future fleet of supersonic aircraft. Instead of occasional exposures over prolonged periods of time, participants in controlled field studies may be exposed to multiple low-amplitude booms per day on a concentrated basis, over the course of a week or more. Test participants in such studies may also be repeatedly

---

[25] Repeatedly reminding panelists to listen for low-amplitude sonic booms invites them to commit a *de facto* form of Titchener's "stimulus error" (Boring, 1921): confusing a stimulus with the sensation (or in this case, the reaction) that it creates.

[26] The goal of these studies was to address procedural issues, not to obtain generalizable samples of opinion.

reminded of the need to listen for unusual impulsive sounds, and to spend several hours over the course of a study describing their reactions to them. Paying per-response incentives in free-response studies to improve panelist retention rates over the course of a field study should be avoided, as they may bias panelists toward reporting that they have noticed low-amplitude booms, even when they are uncertain about the source of impulsive sounds.

4.1.3 Rationale for aircraft noise regulatory policy analysis

For pragmatic reasons, aircraft noise regulatory agencies routinely base their policy positions solely on information about community response to outdoor *residential* noise exposure. Geographic associations between outdoor residential noise exposure and annoyance prevalence rates are blind to the personal aircraft noise exposure of neighborhood residents. Although residents of airport-vicinity neighborhoods may spend very different amounts of time at home, and their waking hours in noise environments very different from those of their neighborhoods, their aircraft noise exposure for regulatory policy analyses is that outside their homes.

In other words, regulatory policy is based on a geographic association between cumulative outdoor noise exposure levels of households in areas within specified noise exposure boundaries, not on the actual aircraft noise exposure of household members.[27] Geographic associations of this sort, routinely inferred from cross-sectional social surveys, are the least costly and the most abundant form of epidemiological evidence about aircraft noise effects on people, but are also the weakest form of evidence.[28]

The practice of basing regulatory policy on residential rather than personal noise exposure emphasizes that a simple geographic association between aircraft noise exposure in residential neighborhoods and the prevalence of high annoyance, as inferred from cross-sectional social surveys, suffices for regulatory decision making (Fidell, 2015). By extension, the practice also implies that the personal noise exposure of residents of areas exposed to sonic booms are of secondary regulatory interest at best. For purposes of collecting information about community response to low-amplitude sonic booms for regulatory analysis, the personal exposure of individual test participants is thus also of limited interest.

Since regulatory agencies are the primary intended audience for information about community response to LBFD overflights, it follows that precise estimates of the boom exposure of participants while they are away from home in controlled exposure field experiments are not essential. If test participants report that they are not far home at the time of an LBFD overflight,

---

[27] Regulatory analyses of aircraft noise effects also fail to distinguish between indoor and outdoor exposure of household residents, even though the two forms of exposure may differ by ~20 dB.

[28] Stronger forms of epidemiological inference are based on individual-level information, as in case-control and longitudinal designs. Without information about sonic boom exposure based on personal dosimetry, however, no such evidence is available to support regulatory analyses.

and if the centerline of the boom carpet is within a few miles of its expected location, then their personal exposure will is likely to remain within a decibel or two of their nominal home exposure (*per* Figure 31 on page 52). Given regulatory agencies' customary treatment of aircraft noise exposure in 5 dB intervals, there is little reason to spend resources on excessively precise estimates of personal exposure levels.[29]

If participants in a controlled exposure study are far from home at the time of an LBFD overflight, or if atmospheric conditions shift the centerline of the boom carpet far from its expected location, then there is little justification for including their response information on such flight missions in dosage-response analyses. A pair of sonic boom monitors straddling the expected boom corridor centerline location in each community of interest can empirically verify the approximate location of the centerline with useful accuracy. Likewise, a simple questionnaire item ("Were you at or near home at the time that you noticed a low-amplitude sonic boom?") can verify that annoyance and startle reports may be usefully analyzed. A participant's response to the suggested questionnaire item can also help to distinguish a failure to notice a particularly low amplitude boom from an absence of boom exposure.

4.1.4 Methods for estimating respondent exposure to sonic booms

Sonic booms generated by supersonic aircraft flying ten miles or more above communities propagate to the ground through an inhomogeneous and unstable atmosphere, along paths that are perturbed by temporally and spatially varying temperature and wind gradients, as well as by turbulence. Sound levels created at ground level by LBFD overflights may be either modeled by long range ray tracing software, or directly measured acoustically. The former method of estimation depends heavily on assumptions about propagation conditions, and/or on labor-intensive and timely measurements of atmospheric conditions. The latter method of estimation depends on site-specific installations of specialized acoustic instrumentation.

The costs of boom level estimation by modeling and by measurement vary with the numbers and geographic dispersion of points for which estimates are required, as well as the budgets and pre-flight time available for atmospheric measurements and software predictions. Given that agencies such as FAA and ICAO routinely base their regulatory policies on estimates of place, rather than personal, noise exposure there is little need for high precision estimates of boom levels at the locations of respondents when they are *not* at home.

The credibility of boom level estimates derived by modeling inevitably depends on at least a modicum of confirmatory acoustic measurement for calibrating and verifying model predictions. In the current context, modeling of residential boom exposure levels is most useful when all social

---

[29] This 5 dB regulatory interval should not be confused with noise exposure binning for growth of annoyance or CTL estimation. Coarse binning for this purpose reduces dose certainty and could affect an empirically-derived function summarizing the growth of annoyance with exposure level.

survey respondents are located within a small area at a site in a single community, and atmospheric conditions affecting boom propagation from a known flight path to specific locations on the ground can be measured in a timely and cost-effective manner. Modeling may be less practical in study designs that address reactions to sonic booms in multiple, geographically dispersed communities, at various times of day.[30]

Fortuitously, sound levels beneath supersonic flight paths vary only slowly with distance from centerlines (Morgenstern, 2016). Figure 31 illustrates the anticipated variation in loudness (perceived level, in PLdB) as a function of distance from an LBFD track centerline (zero crosswind condition).

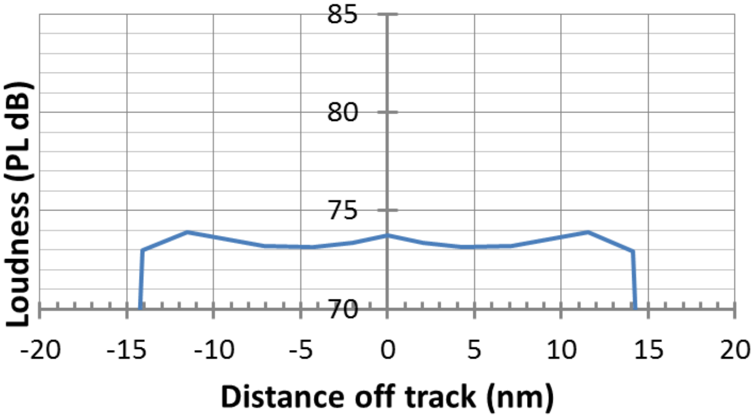


Figure 31. Cross-track variation at ground level in perceived level (PLdB) of shaped sonic boom produced by LBFD aircraft during high altitude supersonic cruise.

The figure plots the Perceived Level on the ordinate and the lateral distance in nautical miles to either side of the centerline on the abscissa. The aircraft is assumed to be flying at 55,000 feet MSL. The plot emphasizes the minor (less than about ± 1 dB) variation in perceived level from the aircraft ground track centerline to the carpet edge. Even if the centerline of a boom corridor were to shift by several miles, it would still be possible to estimate a neighborhood boom exposure with better precision than FAA typically accepts for dosage-response functions for conventional aircraft noise.

Figure 31 was developed through modeling, and does not include effects of atmospheric turbulence or other localized randomness in atmospheric propagation. With boom durations of

---

[30] The Internet-enabled, multi-site, real time sonic boom measurement system described by Fidell and Horonjeff (2017) is a potential solution to this problem. The costs of constructing and deploying multiple field monitoring stations may well be lower than the costs of the labor repeatedly required to measure, model, and analyze boom levels in multiple geographically-dispersed communities. Since acoustic measurements are critical in any event, and since data acquisition costs do not scale directly with numbers of monitors, field measurements might as well be made with a sufficiently dense network of Internet-enabled boom monitoring stations to develop empirical exposure gradients in interviewing areas.

less than 200 milliseconds, atmospherically-induced, random fluctuations in boom intensity may be considerably greater than the small fluctuations shown in Figure 31. One means of dealing with this situation would be to measure at a number of locations, so that an average measurement value could be calculated to characterize an entire neighborhood's exposure. The standard error of the average dose assigned to test participants could then be determined from the entire set of such simultaneous measurements, and applied to the dose-response analysis to characterize the uncertainty of exposure level estimates on annoyance growth estimates.

Since the exposure gradient of carpet boom levels is relatively shallow within a few miles of the centerlines of supersonic flight tracks, and since measurements made at a two or more points permit estimation of an empirical cross-track sound level gradient, errors of boom level estimation should not be great.[31] Estimates of absolute boom levels derived by direct field measurement should, of course, also be checked against estimates derived from prospective modeling. If the two forms of estimate disagree greatly, estimates derived from field measurements should be preferred to those obtained by assumption and modeling. Multiple simultaneous measurements could also confirm the extent of any lateral shift of predicted boom levels. This would permit determination of whether a portion of the community may have lain outside the carpet boom corridor, so that unexposed households could be excluded from analyses. This approach reduces exclusive reliance of boom exposure predictions derived by modeling.

Judicious selection of interviewing areas (*e.g.*, far from lateral cutoff boundaries of carpet boom corridors) would help to minimize violations of near-uniform exposure assumptions. Inferring whether a portion of an interviewing site lies beyond the lateral cutoff edge of a boom corridor on a particular flight from an empirical sound level gradient is at least as credible an approach to exposure estimation as boom modeling.

Field confirmation of actual exposures created by each LBFD flight within an interviewing area can be made by installing a boom monitors at two (or more) sites perpendicular to the expected flight track. One of the monitors should be located near the point in the neighborhood nearest the ground track, while the other would be located near a boundary farthest from the track. The expected difference in measured boom levels should be small, so that an average value between the two could be used as an estimator of neighborhood exposure. However, if the more distant monitor revealed a boom level considerably lower than the nearer, then it could be assumed that the carpet boundary penetrated the neighborhood and that dose estimation may have been compromised at a subset of the surveyed residences.

---

[31] Serious doubts prior to a planned LBFD mission about boom corridor mis-placement of more than a few miles would presumably lead to a decision to cancel that flight. The Internet-enabled monitoring system provides continuous, real-time knowledge of ground-level wind speed and direction to aid such decisions.

4.2 Results of Single-Call Contact Attempt Study

A cross-sectional study of the feasibility of collecting opinions from a nationally representative, independent sample of households about reactions to sonic booms was undertaken in the second year of Phase II. The substantive questionnaire items were perfunctory in nature, because no actual supersonic overflights accompanied the interviews, and the primary concerns of the data collection exercise were 1) to estimate interview completion rates achievable with single contact attempts (no callbacks)[32], and 2) to quantify potential differences in interview completion rates as functions of day of the week, time of day, and geographic region. Interview completion rates were expected to be low due to the lack of callbacks, but it was believed that positive public attitudes toward the NASA "brand"[33] might yield a higher response rate than might otherwise be obtained.

4.2.1 Sample construction

A random digit dialed ("RDD") sample frame was composed to support the data collection exercise. The frame contained both landline and wireless telephone numbers within the 48 contiguous United States. The RDD sample frame was stratified by the ten postal ZIP code regions, and balanced in proportion to state populations to minimize excessive concentrations of respondents in small numbers of densely populated urban areas. The entire frame was pre-dialed shortly prior to the start of data collection to exclude disconnected and non-residential (business and other organizational) telephone numbers from the sampling frame. This precaution maximized the number of potential contacts with residential households, rather than with unproductive telephone numbers.

4.2.2 Dialing schedule and mechanics

Attempts to contact potential respondents were made by two interviewing methods: automated and live agent calling. The automated contact attempts utilized outbound interactive voice response ("IVR", or "robocall") technology, while the live agent contact attempts were made by conventional computer-assisted telephone interviewing ("CATI") methods. Live agent calls were conducted from 30 April through 6 May 2018 (Monday through Sunday). Automated calls were conducted from 30 April through 7 May (Monday through Monday).[34] The interview contact attempts were authorized by the Office of Management and Budget under Control Number 2700-

---

[32] This constraint is imposed by the requirement to assess "prompt" (within 15 minutes) reactions to exposure to individual low-amplitude sonic booms.

[33] The "CNAM" (Caller ID) string displayed on a potential respondent's telephone cannot be reliably controlled, since it depends on the willingness of competitive and incumbent local exchange telephone carriers ("CLECs" and "ILECs", respectively) terminating carriers to pay for a "registry dip." The terminating carrier typically contracts with only one of several CNAM database providers. Even if the originating (outbound) CNAM string of a call is "NASA", it may be variously displayed as "Out of Area", "Unknown", or otherwise by the party receiving the inbound call.

[34] A small number of IVR calls dialed on Monday, 7 May 2018 were analyzed as though they had been made on Monday, 30 April 2018.

0166. Of the 12,734 dialings, 6,275 contact attempts were made by sixteen live agents, while 6,459 were made by software-controlled IVR dialing engines.

4.2.3 Disposition of contact attempts not yielding completed interviews

Figure 32 and Figure 33 summarize the disposition of CATI and IVR contact attempts. Call disposition categories differed for the two forms of interviewing because live agents can distinguish more categories of call dispositions (for example, by detecting language difficulties). IVR call dispositions are limited to categories that can be algorithmically distinguished.

For example, the duration of the "hello" uttered by whomever (or whatever) answers a call is used in automated calling to differentiate between a live respondent and electronic answering devices. If the duration of the initial answer of an automated call is less than 1.8 seconds, the contact is classified as answered by a person. If the duration of the initial answer is greater than 1.8 seconds, the call is considered to be answered by machine.

If the response to an IVR call is classified as made by a person, the possible call dispositions are limited to “Complete” (if answers are provided to all questionnaire items) or “Refusal” (if the respondent stops answering questions before providing an answer to the last questionnaire item.) For present purposes, the greater number of call disposition categories for CATI calls were consolidated into categories similar to those used for IVR calls. The bulk of the “Additional categories” for live agent calls were disconnected and re-directed telephone numbers.[35]

Figure 34 and Figure 35 show very little variability in interview completion rates between weekday and weekend dialings, for both CATI and IVR interviewing methods. The day-of-week differences in percentages of completed interviews were too small, in both relative and absolute terms, to be of practical utility for purposes of LBFD mission scheduling.

---

[35] The live-agent disconnect screening undertaken prior to calling cannot identify all disconnected numbers in any sample. To avoid bothering potential respondents, live-agent disconnect screening is conducted with a maximum of four rings per call. Many phone numbers continue to ring more than four times before reaching a disconnect message, however. The number of rings prior to a disconnect message depends in part on the relative level of activity on telephone switching equipment and on the overall bandwidth of each exchange’s circuit switching infrastructure. Some exchanges with older equipment still rely on actual tape messages, which can extend wait times for the disconnect message to play. Ongoing changes in the activity status of telephone numbers can also affect the success of disconnect purging. Additional phone numbers will reach disconnected status between the time disconnect purging is completed and interviewing starts.

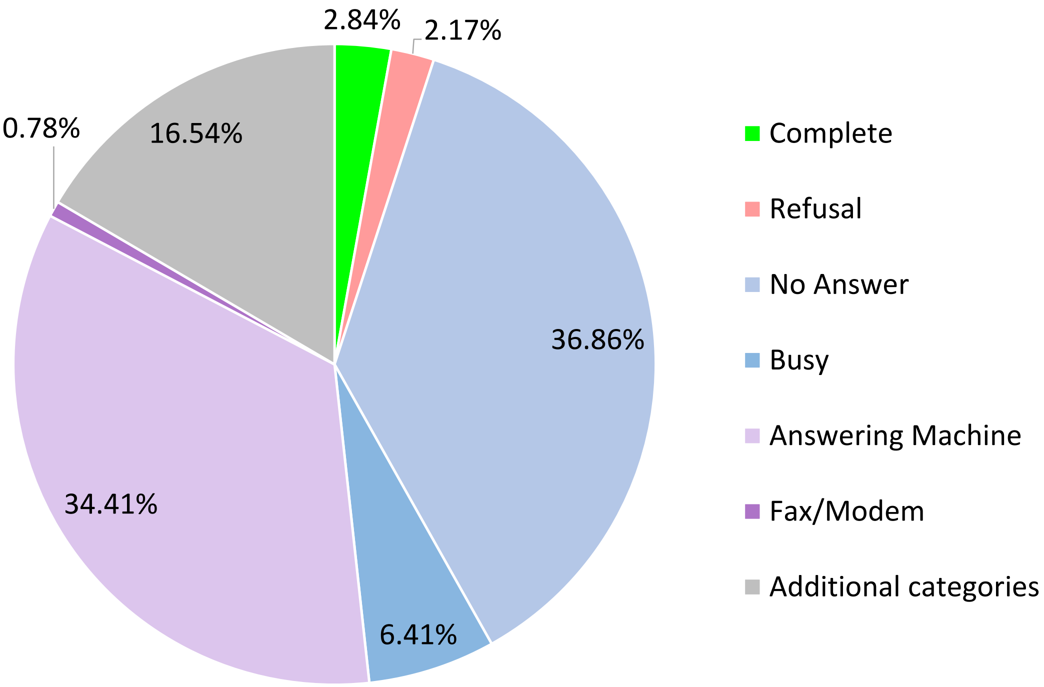


Figure 32. Live agent (CATI) contact attempt dispositions.

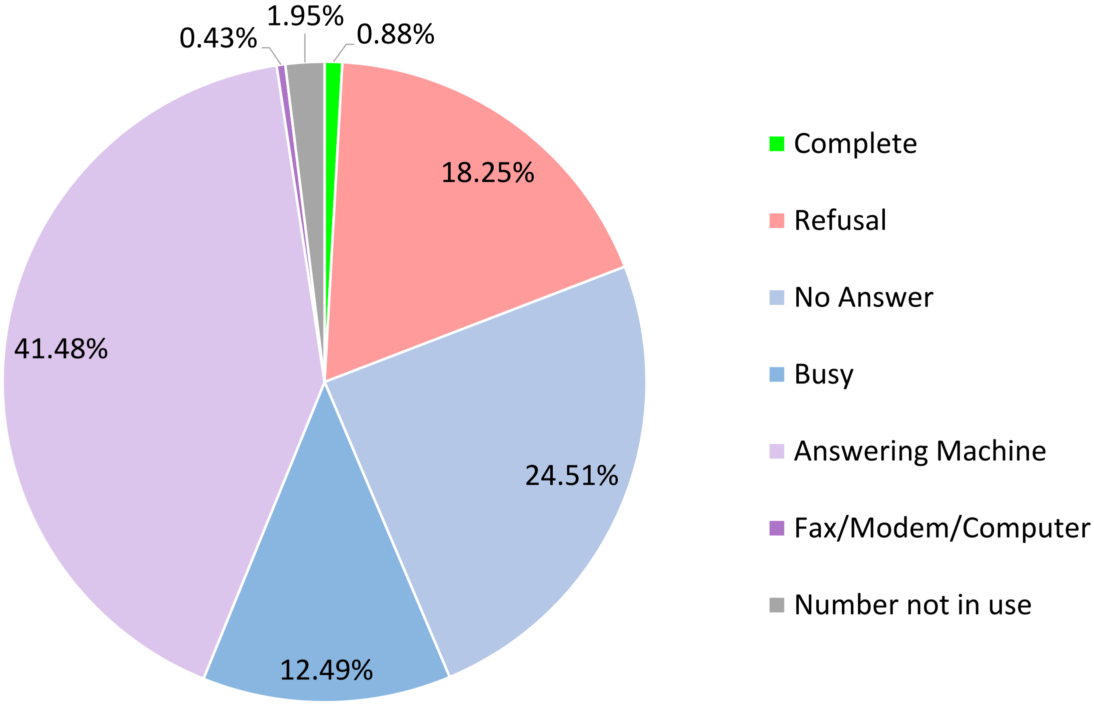


Figure 33. Automated contact (IVR) attempt dispositions.

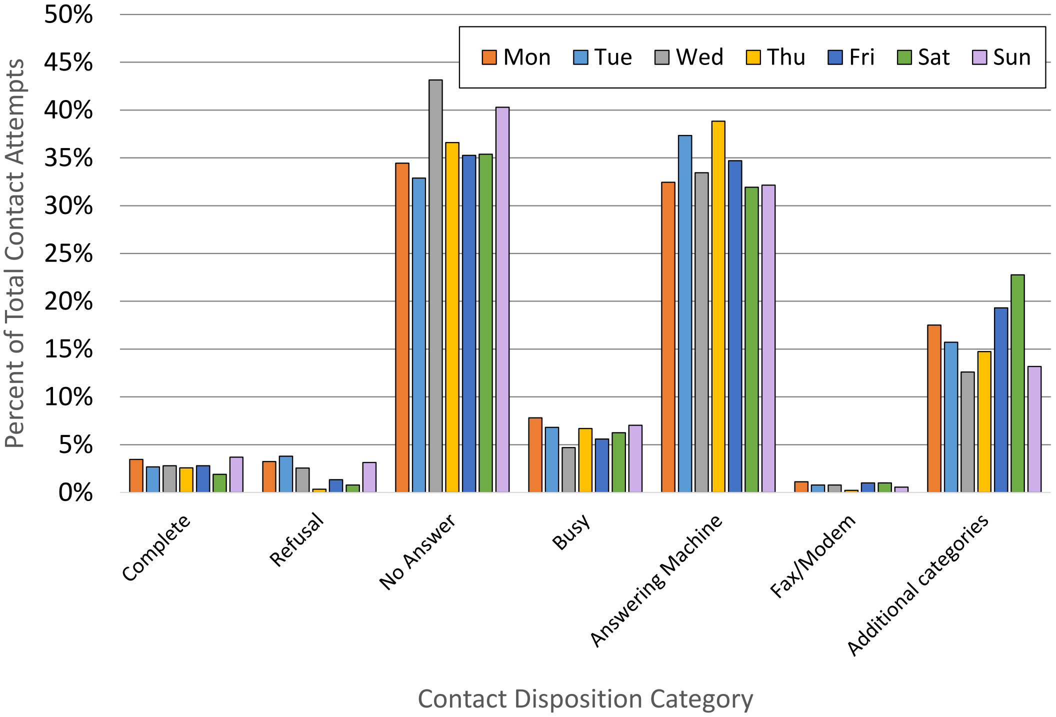


Figure 34. Live agent (CATI) contact attempt dispositions by day of week.

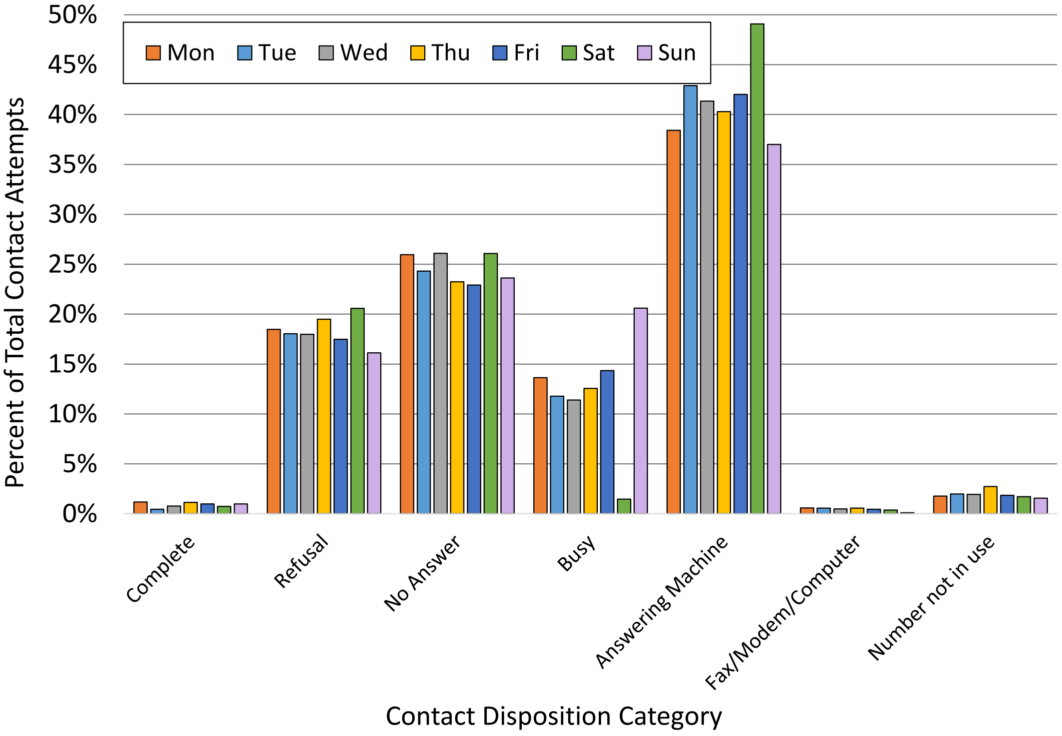


Figure 35. Automated contact (IVR) attempt dispositions by day of week.

4.2.4 *Percentages of dialings yielding completed interviews*

Of the 6,275 and 6,459 live agent and automated dialings, 178 (2.84%) and 55 (0.88%) yielded complete interviews, respectively. Figure 36 shows the automated and live agent completion rates, along with their 95% confidence interval bounds. The differences in completion rates between the two forms of interviewing, although relatively small, are nonetheless statistically significant because of the substantial numbers of contact attempts.

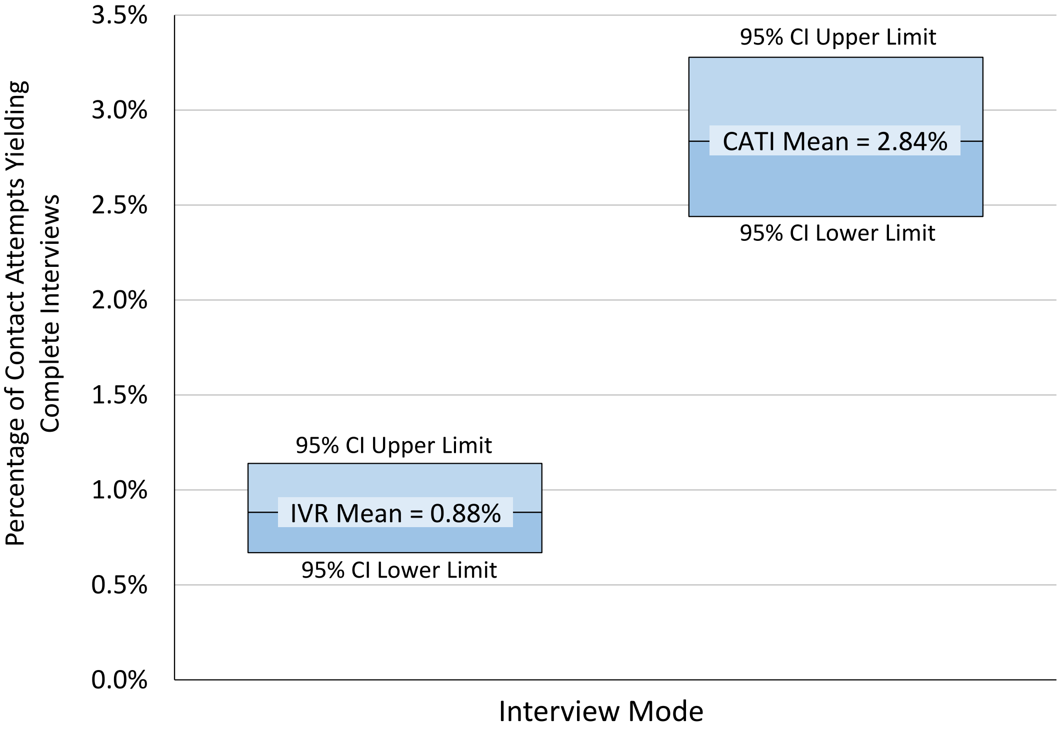


Figure 36. Interview completion rates for IVR and CATI contact attempts.

4.2.5 *Analysis of interview completion rates by hour of the day*

Table 5 shows cumulative (across interviewing days) numbers of CATI and IVR contact attempts by hour of the day. The table also displays percentages of attempts by hour. Figure 37 plots the hourly percentages seen in Table 5 for both interviewing modes. The percentages across all hours add to 100% for each interview mode.[36]

---

[36] The high percentages of contact attempts by hour are an artifact of a non-uniform distribution of dialing hours across the 11-hour period. Two (or more) dialing time blocks overlapped by 2 hours on two days because the hourly rate of CATI and IVR dialings had to be adjusted to extend the interviewing period over an entire week.

Table 5. CATI and IVR contact attempts by hour of the day.

Hour		CATI		IVR	
From	To	Attempts	% of Total	Attempts	% of Total
8:00 AM	9:00 AM	0	0.00%	10	0.15%
9:00 AM	10:00 AM	396	6.31%	604	9.35%
10:00 AM	11:00 AM	254	4.05%	457	7.08%
11:00 AM	12:00 PM	295	4.70%	538	8.33%
12:00 PM	1:00 PM	433	6.90%	620	9.60%
1:00 PM	2:00 PM	457	7.28%	599	9.27%
2:00 PM	3:00 PM	484	7.71%	815	12.62%
3:00 PM	4:00 PM	1,030	16.41%	550	8.52%
4:00 PM	5:00 PM	1,083	17.26%	488	7.56%
5:00 PM	6:00 PM	658	10.49%	589	9.12%
6:00 PM	7:00 PM	529	8.43%	630	9.75%
7:00 PM	8:00 PM	651	10.37%	540	8.36%
8:00 PM	9:00 PM	5	0.08%	19	0.29%
Totals		6,275	100.00%	6,459	100.00%

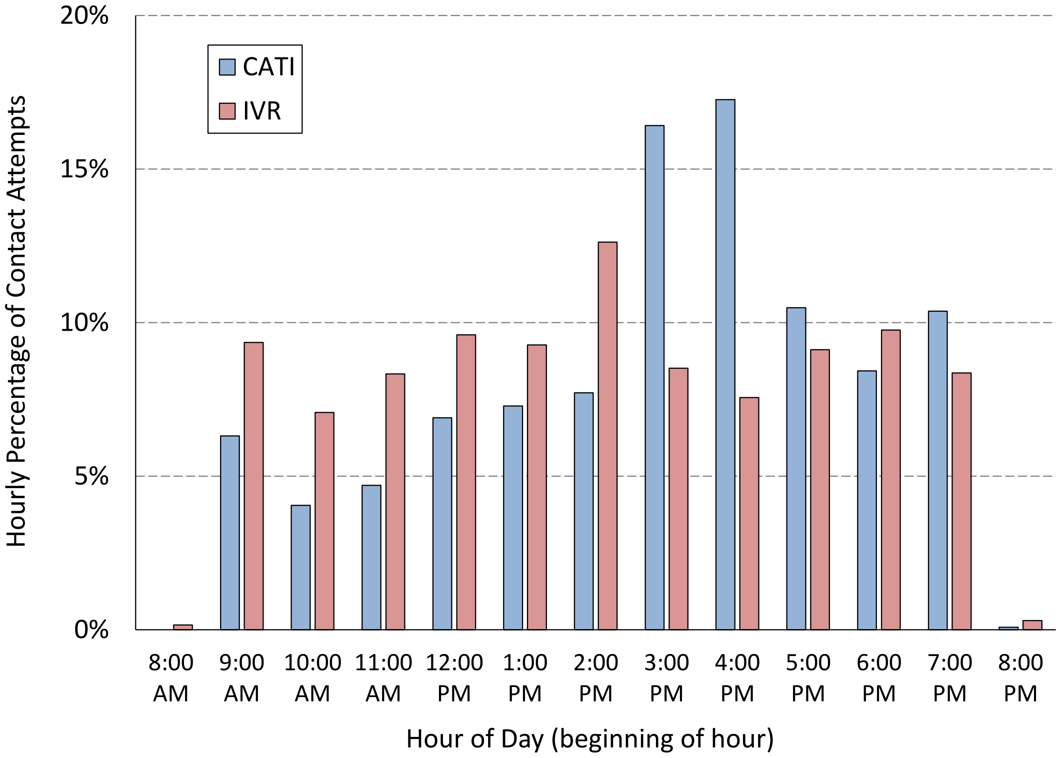


Figure 37. CATI and IVR distributions of contact attempts by hour of the day.

Table 6 divides the 11-hour interviewing period (9 am to 8 pm) into three periods of day. The table shows numbers of contact attempts, numbers of completed interviews, and the percentage of completed interviews (relative to the number of contact attempts) for each of the interviewing modes and time periods.

Table 6. CATI and IVR contact attempts and interview completion rates, by period of the day.

LOCAL TIME OF DAY		CATI			IVR		
From	To	Attempts	Completes	% Complete	Attempts	Completes	% Complete
7:30 AM	12:00 PM	945	22	2.33%	1,608	9	0.56%
12:00 PM	5:00 PM	3,487	100	2.87%	3,072	27	0.88%
5:00 PM	8:30 PM	1,843	56	3.04%	1,777	19	1.07%
Totals		6,275	178	2.84%	6,457	55	0.85%

Figure 38 plots percentages of contact attempts yielding completed interviews for each period of the day. Interview completion rates are smallest in the morning and greatest in the evening hours.

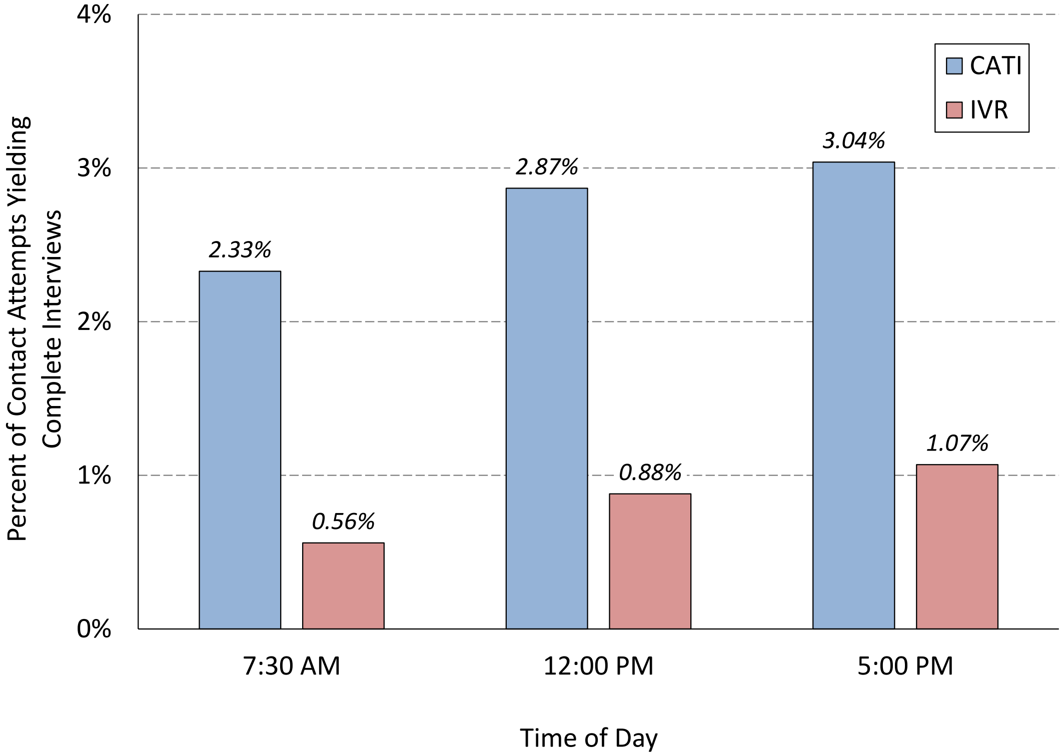


Figure 38. CATI and IVR interview completion rates by period of the day.

4.2.6 Analysis of interview completion rates by day of week

Table 7 shows numbers of contact attempts by day of the week for both interviewing modes. For each day, the table shows the number of contact attempts, the number of completed interviews, and the percentage of completed interviews relative to the number of contact attempts. These numbers are cumulative across all interview hours (9:00 AM to 8:00 PM, local time).

Table 7. CATI and IVR contact attempts and completed interviews by day of week.

Day of Week	CATI			IVR		
	Attempts	Completes	% Complete	Attempts	Completes	% Complete
Sun	896	33	3.68%	924	10	1.08%
Mon	897	31	3.46%	682	10	1.47%
Tue	897	24	2.68%	915	4	0.44%
Wed	897	25	2.79%	1,029	8	0.78%
Thu	896	23	2.57%	1,067	12	1.12%
Fri	896	25	2.79%	921	9	0.98%
Sat	896	17	1.90%	921	6	0.65%
Total	6,275	178		6,459	59	

Figure 39 plots the total number of contact attempts for each day of the week by interview mode. This figure shows that the numbers of daily contact attempts were roughly equal over the interviewing period. Figure 40 plots the percentage of contact attempts yielding completed interviews for each day of the week. The completion rates are greatest on Sunday and Monday, and least on Saturday, for both CATI and IVR contact attempts. (Note however, that none of the daily interview completion rates is great enough to be of meaningful use for purposes of scheduling LBFD flight missions.)

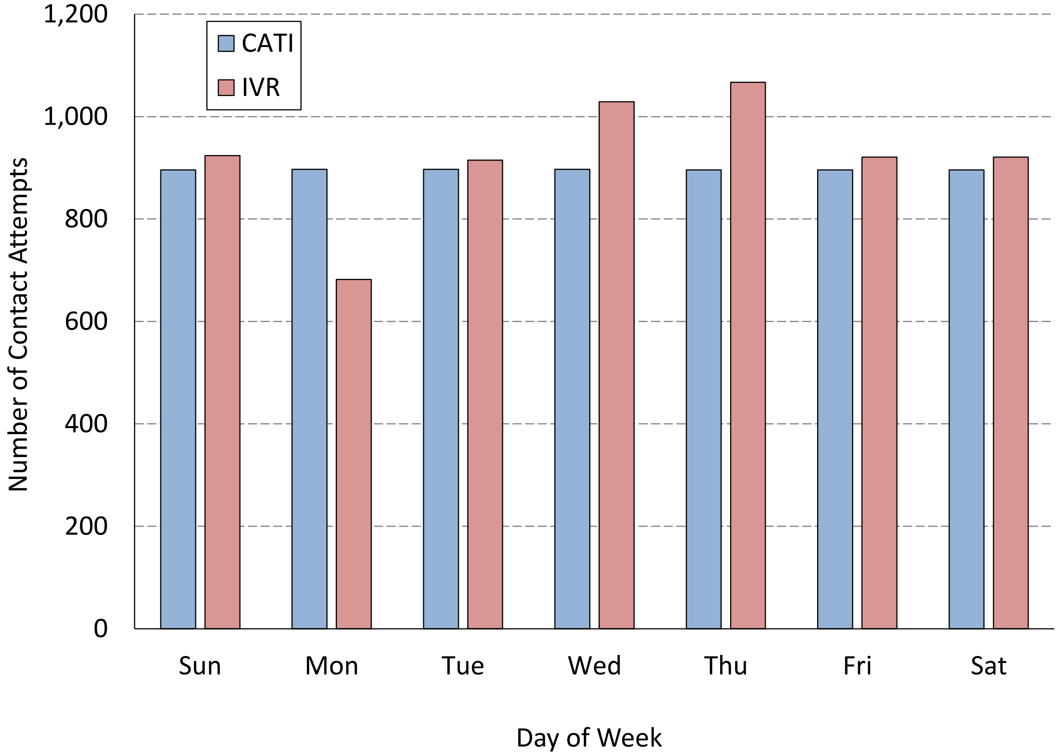


Figure 39. Daily contact attempts for live agent and automated dialing.

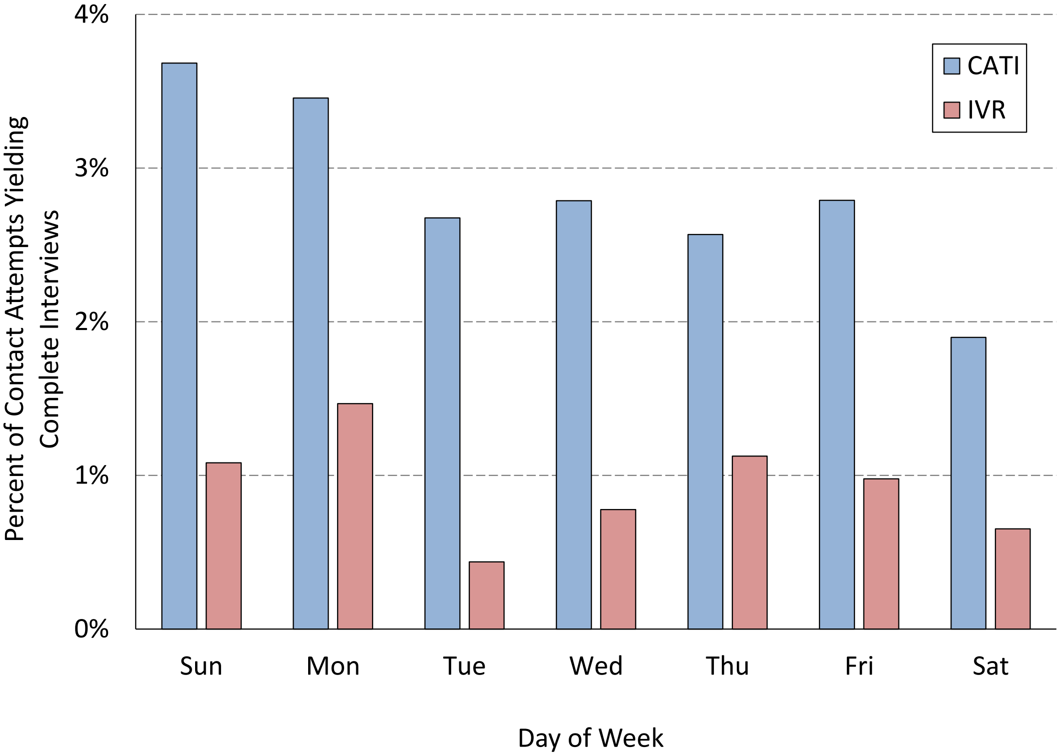


Figure 40. CATI and IVR interview completion rates by day of the week.

Figure 41 shows the locations of respondents who granted and completed interviews in the current data collection exercise. The geographic distribution of dot density predictably reflects the population density throughout the country.

Figure 41. Locations of completed interviews by CATI (blue) and IVR dialing (red) methods.

4.2.7 Analysis of interview completion rates by ZIP code region

Table 8 shows calling results by individual ZIP code region. The first column in the table indicates the ZIP region (first digit of the ZIP code). The second and third columns show the number of contact attempts for CATI and IVR interviews, respectively. The numbers in these columns are plotted in Figure 42. The fourth and fifth columns show the number of contact attempts in each region as a percentage of the total number of attempts for each contact mode.

The sixth and seventh columns show numbers of completed interviews for CATI and IVR interviews. Figure 43 plots the number of completed interviews by ZIP region. The eighth and ninth columns show numbers of completed interviews as percentages of the number of contact attempts for each region. These percentages are plotted by ZIP region in Figure 44.

Table 8. Contact attempts and completed interviews by ZIP region.

ZIP Region	Contact Attempts		Percent of Total		Completes		% Complete	
	CATI	IVR	CATI	IVR	CATI	IVR	CATI	IVR
0	525	533	8.37%	8.25%	17	2	3.24%	0.38%
1	726	790	11.57%	12.23%	18	4	2.48%	0.51%
2	674	665	10.74%	10.30%	15	2	2.23%	0.30%
3	967	989	15.41%	15.31%	28	6	2.90%	0.61%
4	590	628	9.40%	9.72%	14	4	2.37%	0.64%
5	344	356	5.48%	5.51%	9	6	2.62%	1.69%
6	453	434	7.22%	6.72%	11	7	2.43%	1.61%
7	680	716	10.84%	11.09%	18	4	2.65%	0.56%
8	396	388	6.31%	6.01%	17	8	4.29%	2.06%
9	920	960	14.66%	14.86%	31	15	3.37%	1.56%
Total	6,275	6,459	100.00%	100.00%	178	58		

Figure 42 shows that the greatest numbers of contact attempts occurred in ZIP code regions 3 and 9. Region 3 includes the southeastern states including Florida. Region 9 includes the three west coast states. The fewest contact attempts occurred in regions 5, 6 and 8. These three regions are largely rural.

Figure 44 shows the greatest interview completion rates in ZIP region 8, and the least in region 2. Regional differences between the greatest and least interview completion rates, however, are less than ± 1%, hardly enough to be of meaningful utility for purposes of planning LBFD mission routes. Further, the absolute numbers of completed interviews (particularly IVR interviews) per region are too small in many cases to warrant meaningful inferential analyses.

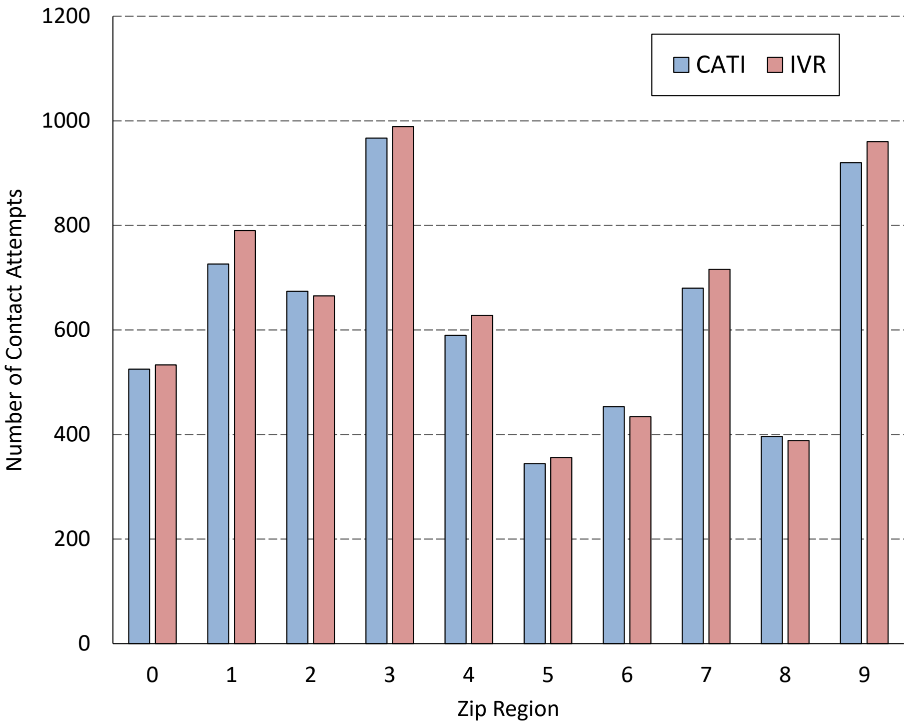


Figure 42. Numbers of contact attempts by ZIP region.

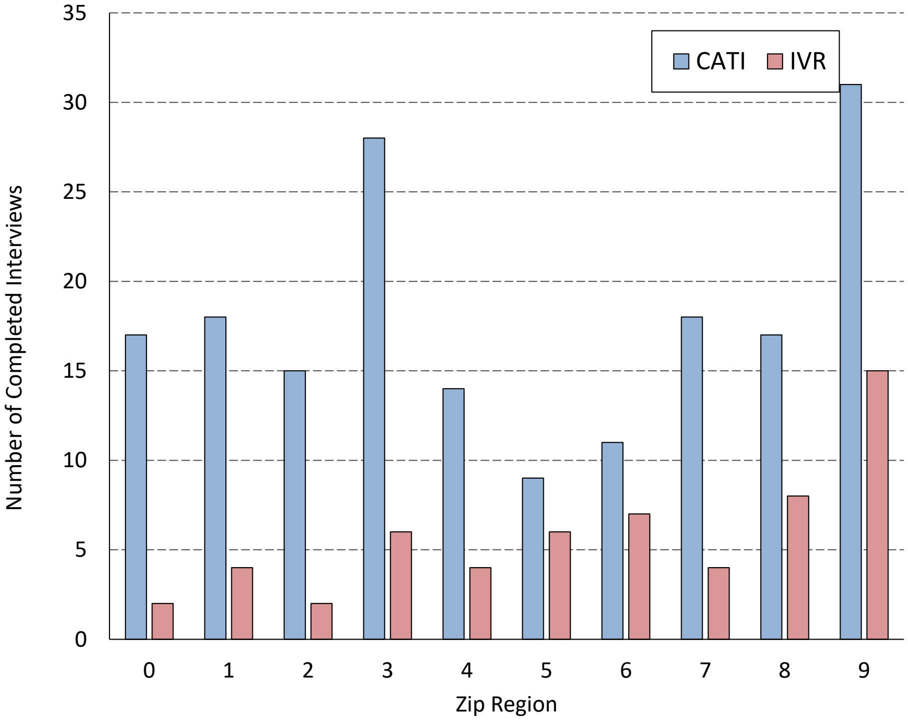


Figure 43. Numbers of completed interviews by ZIP region.

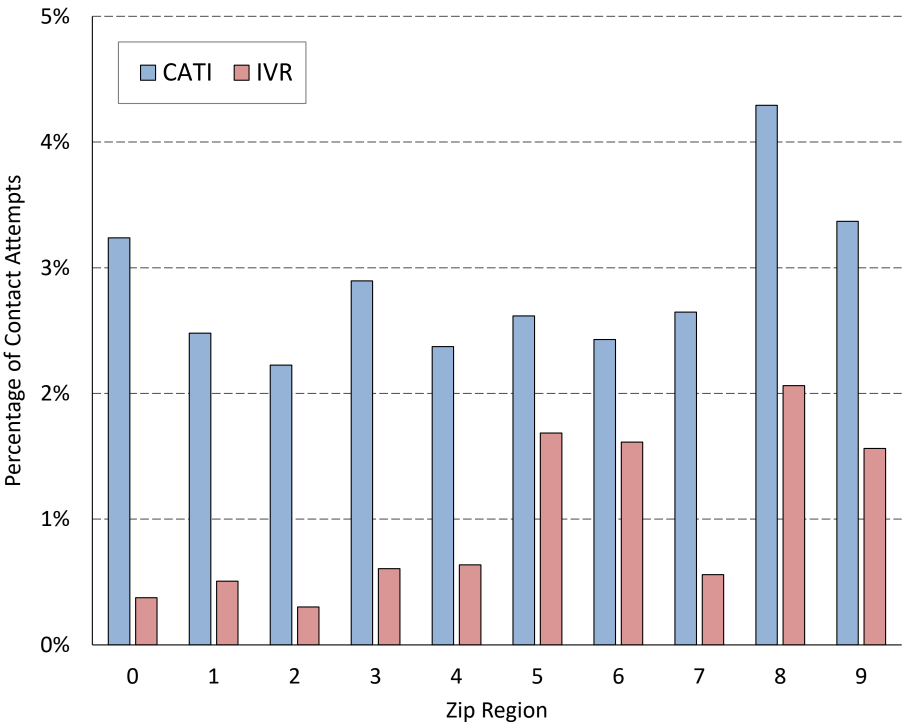


Figure 44. Percentages of contact attempts yielding completed interviews by ZIP region.

4.2.8 Analysis of interview completion rates by population density

Table 9 summarizes interview completion rates by ZIP code population density, in people per square mile. The first column shows population density intervals chosen for analysis. The remaining columns tabulate completion rates of CATI and IVR contacts. Numbers of contact attempts are shown for both CATI and IVR interviewing modes, followed by numbers of completed interviews and percentages of contact attempts resulting in completed interviews.

Only 20 and 1 contact attempts, respectively, were made in two highest population density categories (highlighted in Table 9.) These few observations are insufficient in number to specify interview completion rates to an accuracy of 1 or 2 percent. (All but one of the contact attempts in the 64,000 to 128,000 people per square mile interval were made in New York City, while the remaining one was made in San Francisco.)

Figure 45 plots the percentage of contact attempts yielding completed interviews as a function of population density as shown in Table 9. The figure reveals a reasonably consistent trend of declining interview completion rates for live agent interviews (CATI) as a function of increasing population density. No such trend is apparent for automated interviews (IVR). Even in the lowest population density areas, however, the absolute interview completion rates are so low that they are of little practical interest for current purposes.

Table 9. Completed interviews by population density.

Population Density Range (persons / sq. mi.)	CATI			IVR		
	Attempts	Completes	% Compl	Attempts	Completes	% Compl
0 - 1000	3,488	105	3.01%	3586	40	1.12%
1000 - 2000	777	25	3.22%	797	3	0.38%
2000 - 4000	932	26	2.79%	895	5	0.56%
4000 - 8000	630	14	2.22%	648	8	1.23%
8000 - 16000	269	5	1.86%	269	1	0.37%
16000 - 32000	82	1	1.22%	78	1	1.28%
32000 - 64000	76	1	1.32%	93	0	0.00%
64000 - 128000	20	1	5.00%	14	0	0.00%
128000 - 256000	1	0	0.00%	2	0	0.00%
All	6,275	178	2.84%	6,382	58	0.91%

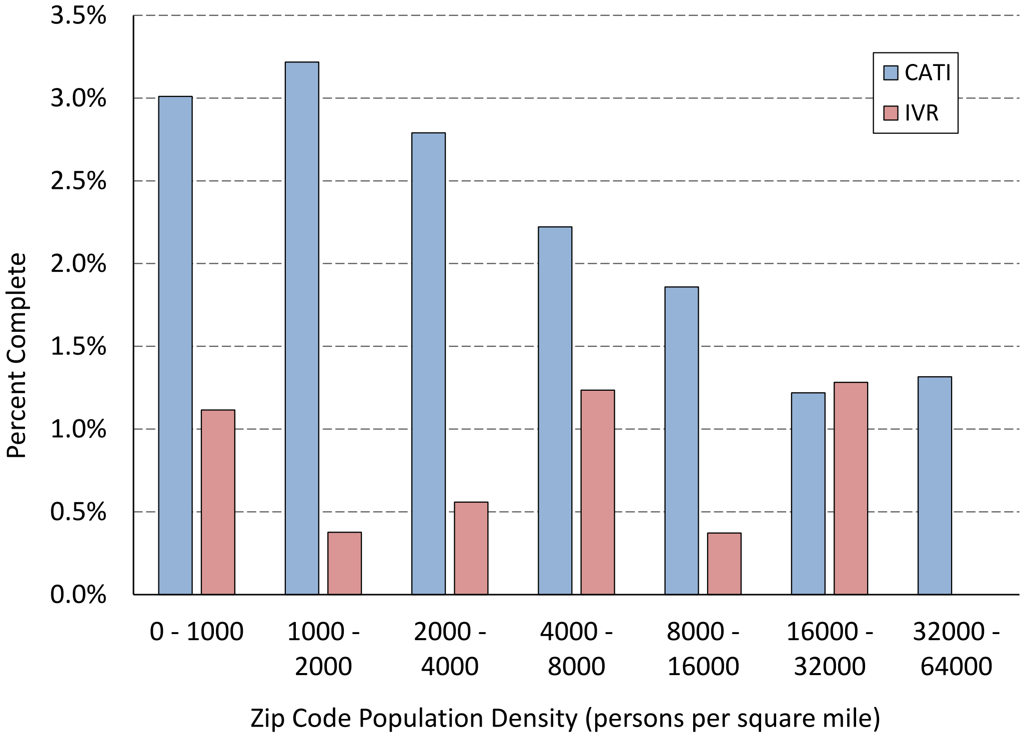


Figure 45. CATI and IVR interview completion rates by population density.

4.2.9 Discussion of interview completion rates

The findings described in this Chapter are consistent with the trend of declining telephone interview completion rates of recent decades. Figure 46 adds the rightmost data point from the current study to the Pew Research Center's (2012) findings of interview completion rates in recent years.[37]

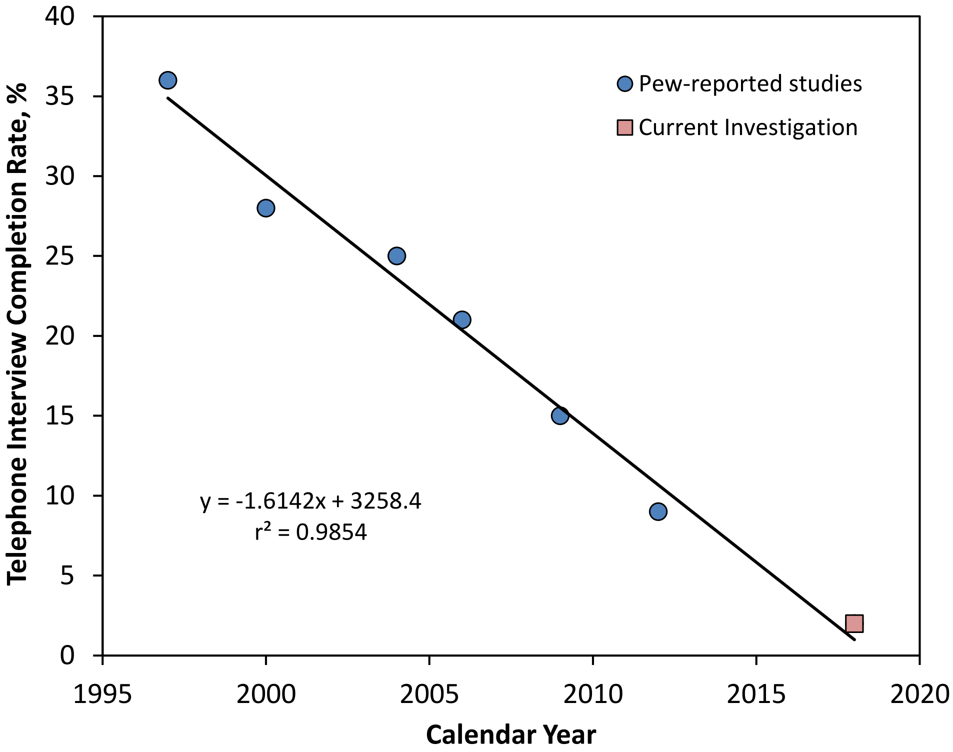


Figure 46. Declining trend in telephone interview completion rates in recent years (rightmost point represents current findings).

4.2.10 Responses to substantive questionnaire items

Notice of aircraft noise

Figure 47 shows responses to the question "Have you noticed any aircraft noise today?" for both CATI and IVR interviews. Of the 178 CATI respondents, 12.9% reported hearing aircraft noise. 15.5% of the 58 IVR respondents reported hearing aircraft noise. The difference between the CATI and IVR percentages is not significant at the 95% level of confidence.

---

[37] Numbers of callback attempts were not consistent among all of the studies, but the general temporal trend is nonetheless clear: telephone interview completion rates have declined precipitously in recent years.

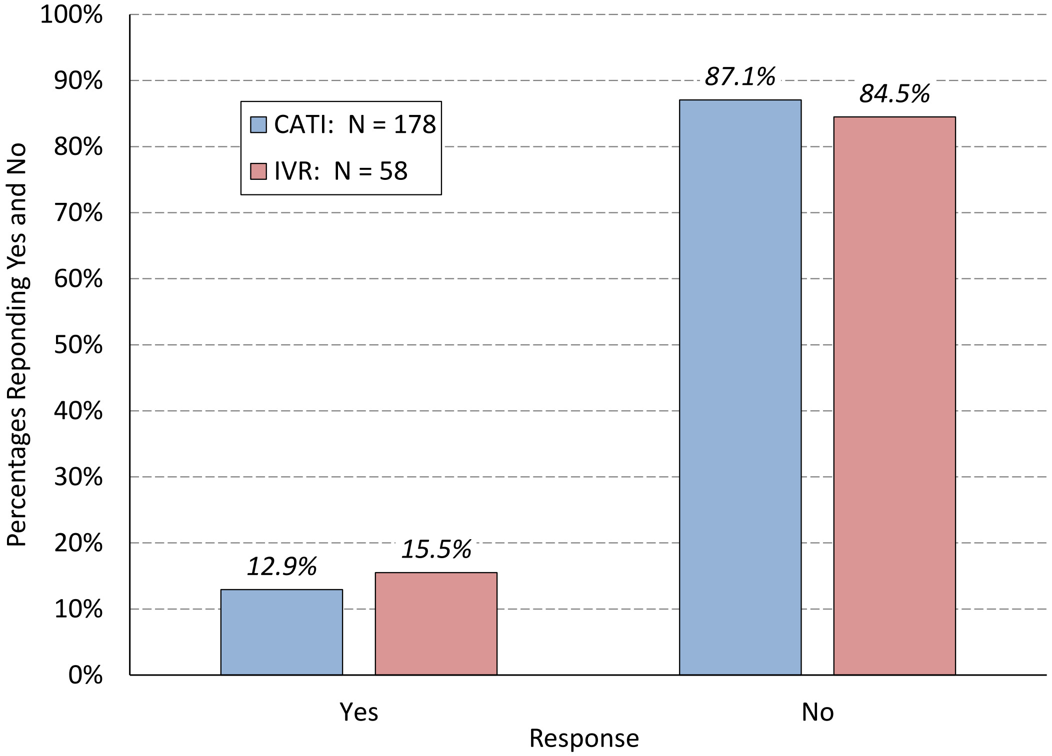


Figure 47. Responses to the question: "Have you noticed any aircraft noise today?".

Notice of sonic booms

Figure 48 shows responses to the question "Have you noticed any sonic booms today?" for CATI and IVR interviews. Of the 178 CATI respondents, 2.8% (5) reported hearing sonic booms. Of the 58 IVR respondents, none reported hearing a sonic boom.

Annoyance of sonic booms

Figure 49 shows responses to the question "Were you bothered or annoyed by the sonic boom?" for the CATI interviews only (no IVR respondents reported hearing a sonic boom.) Of the five CATI respondents who reported hearing a sonic boom, none reported that they had been bothered or annoyed by it.

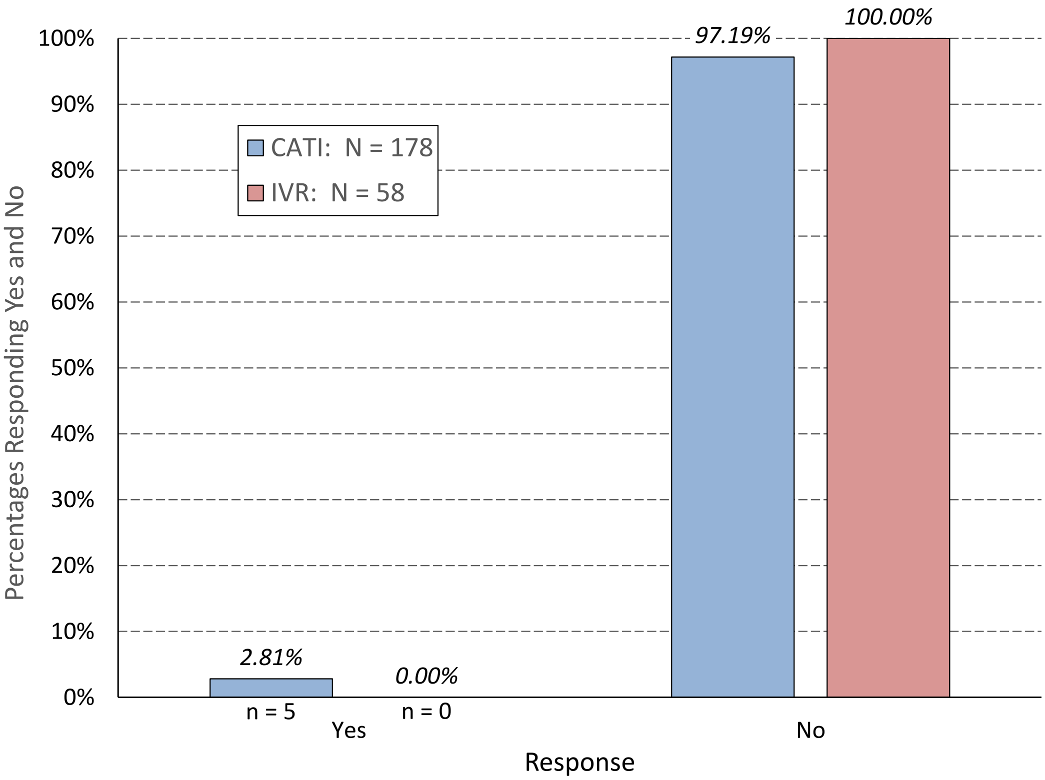


Figure 48. Responses to question: "Have you noticed any sonic booms today?".

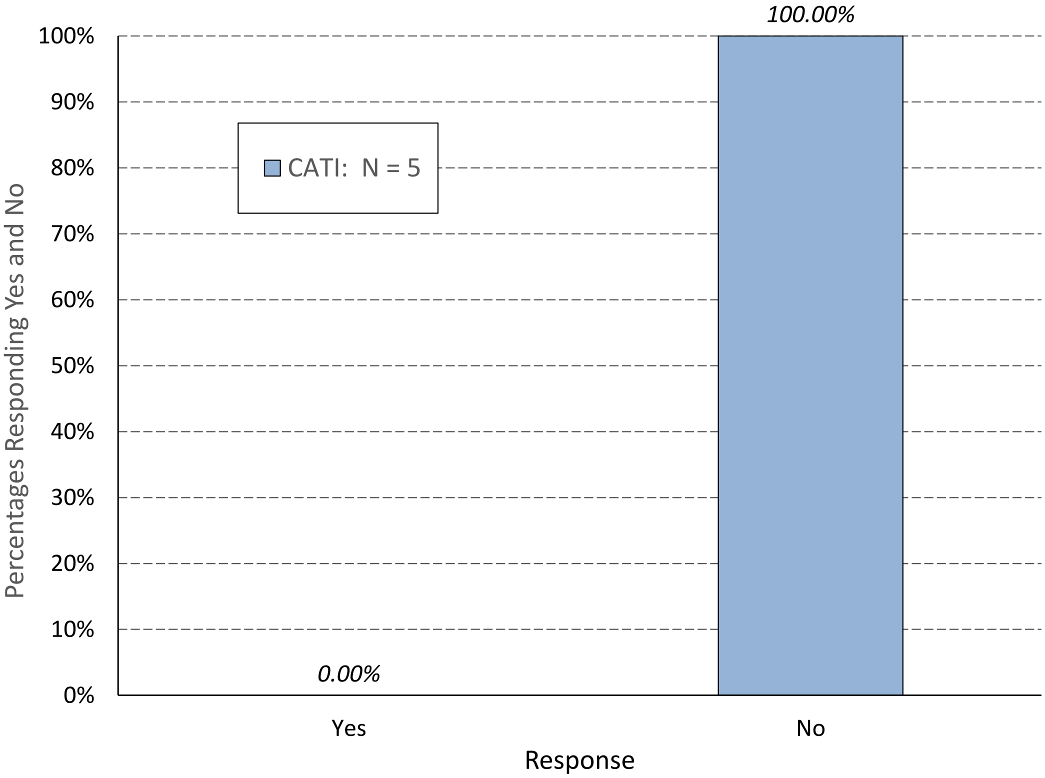


Figure 49. Responses to question: "Were you bothered or annoyed by the sonic boom?".

4.3 Alternate (Panel Sample) Study Designs

4.3.1 Overview of panel sample design variants

Supersonic LBFD flights can rapidly overfly many communities on an individual sortie. Given the substantial costs of LBFD deployments and flight hours, the most cost-effective study designs for assessing community response to low-amplitude sonic booms are those which maximize the amount of community response information collected during each LBFD sortie. Efficient study designs should therefore evaluate reactions to low amplitude sonic booms in multiple overflown communities on each flight mission.

Study designs premised on carefully planned LBFD flight routes and real-time synchronization of interviewing to shock wave arrival times in geographically dispersed communities should therefore be preferred, even (as addressed later) at the potential cost of somewhat lower precision of measurement of boom exposure.[38] Any observed differences in annoyance prevalence ratings in overflown communities at similar noise exposure levels should be quantified by CTL analysis.

Separate panels and sub-panels should be formed within circumscribed areas of several communities near the expected centerlines of anticipated boom corridors studies. As in cross-sectional study designs, panelists should be interviewed about their prompt reactions to individual low-amplitude sonic booms. Outdoor boom exposure levels within each of the defined interviewing areas should be directly measured in real time by Internet-enabled boom monitors, at a minimum of two points orthogonal to the direction of flight in each community, as shown in Figure 50. The figure suggests that the measurement points in each interviewing area be separated by about the distance corresponding to a 2 dB difference in cross-track boom levels. If an LBFD carpet boom corridor then drifts to one side or the other of the expected centerline due to atmospheric conditions, it will still be possible to estimate boom exposure levels at panelists' home addresses with useful precision.

Details of the study designs of the sort recommended below must be tailored to factors such as numbers of LBFD flight missions and flight routes, and the geographic relationships between these routes and communities targeted for interviewing. Absent such information, the level of detail of study designs described in this Chapter must be considered exemplary, rather than deployment-specific.

---

[38] Atmospheric factors may unpredictably shift the centerlines of carpet boom corridors in some interviewing areas from one overflight to the next, but extensive, near-real time meteorological measurements needed for ray tracing modeling along the entire track of a lengthy LBFD flight route are impractical. Even if atmospheric conditions immediately before a lengthy LBFD mission are predicted to excessively displace carpet boom corridors, it would probably be too late to conduct interviews in alternate interviewing areas, or to re-route or re-schedule the flight. Flight cancellations may be the best option in such cases.

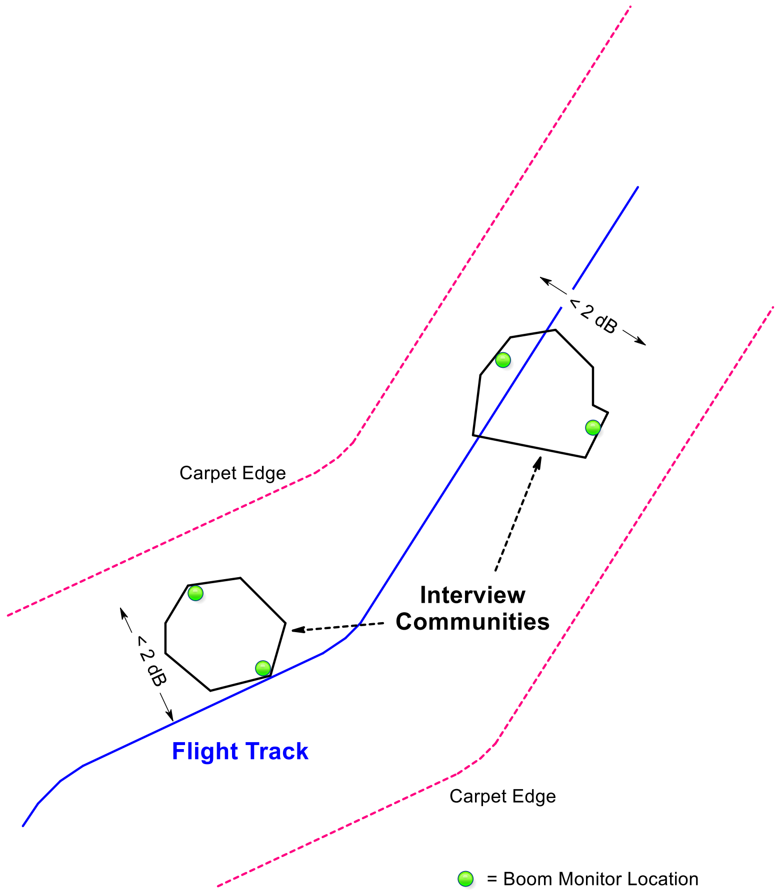


Figure 50. Recommended geographic relationship of interview communities and boom monitor locations with respect to carpet boom edges.

4.3.2 Panel sampling as a remedy for low interview completion rates

Panel sampling is the obvious remedy for the low interview completion rates of single contact attempt, cross-sectional social surveys documented in §4. As cross-sectional telephone interviewing interview completion rates have declined precipitously in recent years (Pew Research Center, 2015), panel samples have become increasingly popular, both in marketing (*e.g.*, Research Now, 2018) and non-commercial studies. Respondents in panel sample studies agree in advance to participate in multiple rounds of interviews, potentially greatly improving interview completion rates.

Panel sample study designs nonetheless have a number of potential disadvantages with respect to cross-sectional designs:

1) Respondents who agree in advance to participate in a panel sample are, in effect, self-selecting into the sample. Their motives for participating in a panel could reflect biases in

favor or against the subject of interviews which differ from those of the target population as a whole;

2) Panelists who respond multiple times to the same interview could exhibit sequential biases – either sensitization or habituation – in responding to multiple identical interviews[39];

3) Panel sampling is more costly than independent sampling, due to additional administrative costs of recruiting panelists and tracking their participation in successive rounds of interviews, as well as the costs of incentives offered to panelists to encourage their continuing participation in the study. Pew Research Center (2015) observes that the infrastructure to construct and manage an on-going panel study requires substantial up-front and continuing investments of staff time and expertise; and

4) A substantial budget is required to prepare, defend, and revise documentation required by an Institutional Review Board (IRB) and the Office of Management and Budget (OMB) for a study design that is more complex than a simple cross-sectional social survey. An IRB, for example, could object that participation incentives large enough to be effective might also be considered coercive, and that geo-tracking of panelists is an unacceptable invasion of privacy.

4.3.3 Generic problems of repeated interviewing

Repeated interviewing of panel members immediately following LBFD overflights presents a number of practical challenges.[40] Ideally, all panelists' annoyance and startle reactions to individual sonic booms would be solicited within fifteen minutes of the arrival of the shock waves produced by each LBFD overflight. The requirement for time-critical assessment of reactions to single noise events precludes written forms of data collection (including postal and Internet-based administration of questionnaires, as described by Miller *et al.*, 2014), in which an unverified household member controls the timing and pace of interview completion.

The most straightforward means for soliciting prompt reactions to low-amplitude booms is by pre-arranged agreement with panelists to answer telephone calls with a specifiable caller ID string. If live agent telephone interviewing is the preferred method of interview administration, the services of relatively large numbers of interviewers will be required for brief periods of time during each LBFD overflight. If multiple LBFD overflights are scheduled several hours apart on

---

[39] Sequential biases due to repeated interviews and to cumulative exposure can be difficult to distinguish unless the correlation between numbers of interviews and numbers of overflights can be broken by design.

[40] Coordination of interview timing with shock wave arrival times is not one of them, however. An ADS-B based, Internet-enabled flight tracking and boom monitoring system has been shown to support prompt and reliable determination of interview start times and boom levels produced in overflown communities (*cf.* § 3.2, and Fidell *et al.*, 2017).

the same day, interviewers could be idle for extended periods of time between flights. This is not the most cost-effective use of their time.

Automated (outbound interactive voice response, or IVR) interviewing can considerably reduce telephone interviewing costs, but such automated interviews require a simplified questionnaire. Given that panel members will have agreed to multiple interviews; can be familiarized with the interview protocol in advance; and receive monetary incentives to complete interviews, a simplified questionnaire may not be an excessive price to pay to reduce per-interview data collection costs. As a check on the false-positive response rate the IVR approach could be used at little incremental cost to make calls when no LBFD mission had been conducted.

Some form of either self-report (self-administered or "mixed-mode"[41] responding) could be an alternative to either exclusively live agent or exclusively automated interviewing for repeated solicitations of prompt reactions to LBFD overflights. For example, panelists with smartphones could be provided with an application program similar to the one developed for use in prior community response research at Edwards Air Force Base (Fidell *et al.*, 2012; Page *et al.*, 2014). Those with landline telephone service, or with text message-incompatible ("clamshell") wireless service, could be interviewed by live agents.

The most convenient means (for researchers) for conducting multiple interviews with panelists is via self-administration of the survey instrument, as enabled by automated digital communication of some form. One problem with doing so via web-based application software for smartphone and other mobile and fixed devices, however, is that such methods may yield only a sample of convenience, rather than a fully representative and generalizable sample.

For example, a web-based application program that permits panelists to choose between smartphone and other Internet-enabled response methods systematically excludes the opinions of a large segment of the population: people who lack text messaging wireless telephone service or other forms of web access. The most obvious population segment under-represented in such a sample is the elderly, among whom wireless phone service is far from universal.[42] Others who may be under-represented include those of limited literacy.

Self-administration of interviews also sacrifices tight control over their timing. Panelists may respond to impulsive noise events other than low-amplitude sonic booms (false alarms) prior

---

[41] "Mixed mode" in the present context implies a combination of response modalities, such as prompted and unprompted (free response), automated and non-automated, self-administered and interviewer-administered reporting.

[42] Blumberg and Luke (2015, 2017) estimate that a little less than half of U.S. households still rely primarily on landline telephone subscription. The percentage of wireless-only and wireless-mostly households continues to increase, albeit with state-by-state differences that may be related to population age and income. In Texas, for example, about a quarter of households still cannot be reached by wireless telephone.

to the actual arrival times of shock waves produced by LBFD overflights at their homes, or with latencies longer than preferred. They may also not self-report for a variety of reasons: because they were away from home at the time of an LBFD overflight; because it was inconvenient to respond promptly; or because they simply did not notice a sonic boom. Unless the locations of respondents who self-administer interviews are continuously tracked, it can be difficult to distinguish among these possibilities. If the timing of interviews is centrally-determined rather than left to the discretion of panelists, however, a simple questionnaire item (*e.g.*, of the form "Have I reached you at- or near your home?") can resolve the ambiguity.

In short, Web-based self-reporting of the prompt annoyance of low-amplitude sonic booms may be challenged on several grounds:

1) Latencies between exposure and responses are not centrally controlled in a self-administered interview, so that a small percentage of responses may be received outside of the desired 10 – 15 minute post-exposure temporal window (Fidell *et al.*, 2012). Opinions expressed in late-arriving reports of annoyance and/or startle may be less extreme than those expressed in prompt reports, and can confound the association of response with dose when it is uncertain which of two or more overflights generated a response.

2) A failure by panelists to complete a self-administered, web-based interview may indicate either a *bona fide* failure to notice a low-amplitude boom, or might simply mean that panelists were not near- or at-home, or otherwise without web access, at the time of an overflight. Uncertainty of this sort could result in underestimates of the slope of dosage-response functions relating boom levels to annoyance and startle prevalence rates.

3) The opinions of panelists who participate in web-based-only interviews may be challenged as unrepresentative of those of the general residential population. The percentage of households with web access is notably smaller than the percentage with telephone and postal service. Groves *et al.* (2004) report that those who lack internet access tend to be older and of lower income than the general population.

4) Opinions of panelists who tolerate installation of an application program on their smartphones which permits GPS tracking of their locations at all times may differ systematically from those of people unwilling to permit such invasions of privacy. People who do not agree to permit their locations to be remotely tracked might, for example, be older and/or otherwise less tolerant of noise exposure than those willing to be tracked.

5) The great bulk of the peer-reviewed literature on the annoyance of aircraft noise exposure has been collected via interviewer-administered questionnaires. Comparisons of the findings of self-administered with interviewer-administered questionnaires may be challenged as potentially influenced by systematic but unknown methodological biases.

4.3.4 Attrition in short-duration panels

The bulk of the literature on attrition rates over the duration of longitudinal studies deals with long term (years or decades-long) studies. Causes of panel attrition in long term studies can include re-location and deaths of panelists. These are negligible risks in short term studies. The only longitudinal study about the annoyance of aircraft noise (Fields and Powell, 1985) found that contained a modicum of information about participation in multiple daily interviews concluded that "it was possible to ensure the respondents' cooperation during the extended (*22 repeated interview*) study period."

Fumagalli *et al.* (2013) stress the importance of dealing with two aspects of attrition:

> "Two important elements are the success of the survey at avoiding non-response at the first wave (initial non-response) and at avoiding subsequent loss of sample members due to non-response at latter stages (sample attrition)."

In other words, panelists must be retained between the time of their acceptance to the panel and the time at which the first round of interviews is scheduled. For present purposes, this implies that the recruitment process should not start more than about a month before the start of LBFD overflights, and that the process should be compressed to a period of roughly three weeks.

Panel members must also be retained during the survey period itself. Groves *et al.* (2004, 2006) suggest that attrition becomes more likely when panel members are "insufficiently motivated or interested in the survey." Schoeni *et al.* (2013) point out the cost-effectiveness of timely incentive payments as a means for maintaining panelist engagement in a study, and stress that the administrative burden and statistical treatments intended to compensate for high attrition rates can be more costly than participation incentives. Non-trivial monetary incentives for continuing participation appear to be the best way to maintain motivation and interest in the present case. Panelists who miss more than two successive rounds of interviews would ideally be contacted to attempt to encourage them to continue their participation.

Herzog and Rodgers (1988) caution that attrition tends to increase with increasing age of the participant. If continuing participation in a short term panel study is truly age-specific, and severe enough to alter the age balance of the panel, it may be advisable to compensate for any appreciable age imbalance in later rounds of interviewing by appropriately weighting the opinions of old and young members of the panel.

4.3.5 Continuum of number of rounds of interviews

Panel sample designs appropriate for soliciting reactions to multiple LBFD flights may be thought of as arrayed along a continuum that ranges from two interviews per panelist to as many interviews per panelist as intended LBFD flight missions. At the lower extreme of this continuum, each panelist would be interviewed about the annoyance and startle of exposure to the booms produced on the first and last missions of each LBFD deployment.[43] At the opposite extreme, all panelists could be interviewed about their reactions to each and every LBFD overflight.

The benefits of study designs that interview some panelists more or less often than others are two-fold. One reason for doing so is to control the costs of recruiting solicitations and of interviewing. Between-round attrition rates in repeated waves of interviews may be as great as about 9% (Pew Research Center, 2015), particularly in long duration panel studies. The same source also reports that only about half of all respondents invited to participate in a survey panel agree to do so. Considering these two factors, if a minimum of about 600 respondents were required on the final of 12 rounds of interviews (corresponding to 12 LBFD overflights per deployment) at a 9% between-round attrition rate, about 1700 panelists would have to be recruited. To recruit 1,700 panelists for an initial round of interviews, however, a minimum of twice as many, or 3,400, would have to be solicited.

If the initial solicitation of interest in participating in a panel sample were undertaken by mail, and if a monetary incentive were included with each invitational mailing, then at least 3,400 such incentive payments would be required to recruit panelists for each LBFD field deployment. If the initial incentive offered for panel participation were $5.00, the initial mailing would require a budget of at least $17,000, exclusive of bookkeeping and other costs required for follow-up mailings.

Further, if all panelists were offered monetary incentives of comparable size to encourage their continued participation in subsequent rounds of interviews, the cumulative sum of subsequent incentive payments could amount to tens of thousands of dollars at the 9% between-round attrition rate. Confidence interval widths for the initial rounds of interviews would also be unnecessarily and wastefully narrow, and the resultant statistical power would be excessive.[44]

One way to reduce overall panel recruitment costs is to create multiple sub-panels, each to be interviewed fewer times than the number of expected LBFD flight missions for a given

---

[43] The rationale for interviewing all panelists about their reactions to the first and last flights is to permit statistical analyses that could distinguish the effects of total numbers of exposures from the effects of multiple interviews. All households in the same community would be exposed to the same numbers of LBFD overflights, but only some of them would be susceptible to sequential biases due to multiple interviews alone.

[44] Confidence interval widths depend on standard error sizes, calculated as standard deviations divided by numbers of observations. All other things being equal, precision of measurement increases as the square root of the number of observations. The square root of 2100 is more than 2.5 times greater than the square root of 300, so that confidence interval widths for the initial round of interviews will be considerably (and wastefully) smaller than necessary to distinguish among annoyance prevalence rates associated with sonic boom exposure levels.

deployment.[45] For example, suppose the 600 required panelists completing 12 rounds of interviews in the example above were split into two equal sub-panels. Each sub-panel would be interviewed for half the number of rounds as before, or 6 rounds. At the end of 6 rounds the desired number of panelists still participating would be just 300. However, two panels of 300 provides the same number of remaining participants at the end of all rounds as one panel of 600, but the attrition over 6 rounds would not be as great as over 12 rounds. Assuming the same 9% attrition rate as above, approximately 480 initial panelists would be needed for each panel's first round, or about 960 across the two panels. This is a substantial reduction over the first example of 1,700 initial panelists. Again assuming half of the households invited actually accept and participate in a panel sample, then the two panel example would require about 1,920 invitations to be mailed instead of the 3,400 in the first example, a reduction of 43 percent.

If not all sub-panels are interviewed after each flight, then complete data from certain combinations of exposure levels may not be available. Nonetheless, a dosage-response function summarizing the findings of interviews with all panelists could still be constructed by combining annoyance prevalence rates across sub-panels. (This is analogous to the way that dosage-response functions are commonly inferred in meta-analyses by combining information collected from multiple independent samples of respondents in separate studies.)

4.3.6 Sampling frames appropriate for panel sampling

An ideal – but nonexistent – sampling frame is fully exhaustive of the target population and completely current. Voter registration, motor vehicle registration databases, marketing and other specialized databases, do not fully enumerate all households. Even if the goal of a social survey is to accurately reflect the opinions of a population within a limited geographic area, rather than nationwide, the fundamental requirement for producing generalizable findings from any given sample is that each and every member of the population must have an equal probability of contributing an opinion to the study.[46]

Even the two databases that are widely viewed as the most complete are less than fully exhaustive of the U.S. population. NORC (2018) estimates that address-based samples include only about 92% of the national population. Blumberg and Luke (2018) report that about 97% of the national population subscribes to some form of telephone service. Table 10 summarizes the demographics of U.S. wireless telephone ownership.

---

[45] This tactic also permits distinctions to be drawn between sequential biases due to numbers of exposures and those due to numbers of repeated interviews.

[46] It is conceivable in low population density communities near some LBFD flight tracks that it might be possible to exhaust rather than sample all households, thereby avoiding any sampling error (but requiring a finite population correction to standard errors of estimated population proportions.).

RDD- and postal address-based sampling can each yield representative samples. Generic arguments can be made for either sampling strategy, but may in the end depend on specifics of LBFD flight routes and areas suitable as interviewing sites. Arguments of this sort concern non-coverage and uncertainties associated with the actual addresses of postal box renters, but also include multiplicity of listings and the cost of sampling logistics. The relative difficulty of composing and administering different forms of survey samples are not limited to the initial selection of eligible respondents. For RDD telephone sampling, geo-targeting and accurate geo-coding are more complex than for ABS postal sampling, because an ABS sample contains a nominal address for every record (other than RFD and post office boxes), simplifying appending latitude/longitude coordinates to each record.

Table 10. U.S. cellphone/smartphone owner demographics in early 2018.

Demographics of Cellphone/Smartphone Ownership (% of U.S. adults)			
Demographic Category	**Any cellphone**	**Smartphone**	**Cellphone, but not smartphone**
Total	95%	77%	17%
Men	95%	80%	16%
Women	94%	75%	19%
Ages 18-29	100%	94%	6%
30-49	98%	89%	9%
50-64	94%	73%	21%
65+	85%	46%	40%
White	94%	77%	17%
Black	98%	75%	23%
Hispanic	97%	77%	20%
Less than high school graduate	90%	57%	33%
High school graduate	92%	69%	24%
Some college	96%	80%	16%
College graduate	97%	91%	6%
Less than $30,000	92%	67%	25%
$30,000-$49,999	98%	82%	15%
$50,000-$74,999	98%	83%	15%
$75,000+	98%	93%	5%
Urban	96%	83%	13%
Suburban	94%	78%	16%
Rural	91%	65%	26%

A combination of RDD and postal respondent selection is also possible, if initial recruitment for a panel sample is more cost-effective, because it would provide accurate addresses for each potential panel member. Table 11 compares the advantages and disadvantages of address-based and telephone RDD sample frames. It is also possible to augment RDD landline, RDD wireless, and Enhanced-Wireless™ sample records with information from an address-based frame to reach some of the ~3% of the U.S. population that lacks telephone service. (Note, however, that it can be difficult to determine in advance the extent of non-response bias in such blended sampling schemes.)

Table 11. Summary of advantages and disadvantages of postal and telephone initial recruitment contact methods.

Address-based/Postal Sample Frame	Telephone RDD Landline + Wireless Sample Frame
92% coverage of US Population (NORC), may be increased to 97% by addition of consumer, voter, and other data sets	97% coverage of the US population overall. (Ranges from 98.6% for Delaware to 95.5% for North Dakota)
Allows for pre-mailer to be sent in an effort to increase response rates. Detailed information may be sent which participants may review at their convenience	A pre-mailer can still be sent to about 50% of households (depending upon the area) based upon matching of RDD landline and wireless numbers
Over-representation of non-Hispanic white and more highly educated households	Younger 18 to 24 year olds are traditionally difficult to reach in telephone studies, however with proper stratification and field management, this can be mitigated.
Issues of multiplicity: Millions of households identified in PO Box ZIP codes are also listed under the residential address frame within the USPS CDS/DSF database.	Many people have more than one telephone number
Some percentage of mailpieces will be returned if they do not reach selected households, but an unknowable percentage of undelivered mailpieces will NOT be returned	Non-responding and non-contacted households are easily determined as part of CATI data collection
ABS requires appending addresses based upon consumer, white page (Landline Listed), known wireless data with address (Enhanced-Wireless), and Voter Data in the area to the ABS sample frame to improve completeness. ABS samples are known in some cases to be less than comprehensive; for example, where rural or general delivery addresses are numerous	RDD frame when including both all potential landline and wireless working blocks (NPA-NXX-XX) within the carpet boom corridor are already present and complete -- with one exception: Foreign exchanges based upon wireless subscribers who have moved and carried with them a wireless phone that originated in another area, state, or city.
Difficulty knowing the actual physical location of households whose only mailing address is a PO BOX	RDD sample records are geocoded by ZIP code based upon where the telephone working block predominately falls. During panel recruitment, respondents can be geolocated accurately based upon their address.
Centralized delivery points such as guard gated communities present problems	Centralized delivery points such as guard gated communities do not affect RDD sample configuration

Address-based/Postal Sample Frame	Telephone RDD Landline + Wireless Sample Frame
Misclassification of residential addresses as business addresses can cause over-sampling and sending mail-pieces to non-qualifying addresses	Businesses within RDD samples can be screened by standard business purging. Nevertheless, some business numbers will remain in the final sample (individual non-listed desk phones, for example)
Under-coverage of rural areas	RDD samples do not typically under-represent rural areas as long as phone service is available. Extremely rural areas and remote wooded areas may lack wireless (but rarely landline) telephone service
Ability to match most addresses to a precise geographic location is a strength of address-based search. Using accurate address, latitude and longitude can be appended to each record and be helpful in restricting interviewing within the carpet boom corridor.	Addresses can be appended to RDD sample using targeted sampling frames such as Landline Listed (white page data), Enhanced-Wireless (self-reported wireless data), voter data, and other consumer databases.
Substantial mail-piece production, printing, and postage costs are inescapable	No mandatory upfront mail-piece production, printing, or postage costs
Incentive provided with *each* mail-piece to encourage to encourage participation; incentive will be wasted on non-respondents.	No incentive required for respondents who do NOT join the panel
Phone number only available for a subset of sample records, but can be obtained in mail-back	Phone numbers are available to 100% of RDD frame
Matching names to addresses prior to mailing is not recommended in most cases. When correct, names can help with the deliverability/open rate, but when incorrect, can appreciably reduce the ABS effectiveness.	Name-matching not necessary to conduct telephone based data collection.
High labor costs for manual processing of large number of returned mail pieces; relatively high mailing and/or courier service costs;	Low cost, automated means available for tracking and following-up on CATI contacts; live agent discretion and assistance

4.3.7 Methods of identifying and recruiting potential panelists

To preserve the ability to generalize interview findings beyond any particular sample of opinion, persons eligible to join a panel of respondents must be selected in a manner representative of a definable target population. The default population is assumed to be all residential households within carpet boom corridors produced by LBFD flights.

Initial postcard pre-notification

Recruitment of panelists for a panel sample should begin with an initial postcard notifying households eligible to participate in a panel study of a subsequent invitational mailing. The latter mailing is the one intended to describe and solicit interest in panel participation of household members within LBFD overflight corridors. The initial notification postcard is recommended for several reasons:

> 1) Miller *et al.* (2014) report that "households sent ... *[a]* mail survey are much more likely *[than households contacted by telephone]* to respond to the survey, after accounting for all other variables. (Note, however, that the Miller *et al.* finding pertains to a cross-sectional, rather than longitudinal study design.)[47]
>
> 2) It is likely that the response rate to a suitably prepared mailpiece soliciting interest in panel participation will be greater than for an initial telephone solicitation, even with multiple callback attempts (see §4.4). Recent aircraft noise survey experience (Cointin *et al.*, 2016) suggests an approximate 3:1 advantage of postal over telephone response rates.
>
> 3) An initial postal invitation for participation permits immediate evidence of a more credible and tangible incentive for panel participation (for example, a token monetary incentive paper-clipped to the letter of invitation) than a telephone contact can provide.
>
> 4) Follow-up mailings, sent soon after receiving consent to participation in a panel, permit a more thorough explanation of details of NASA's field tests than a live agent telephone interviewer can provide, such as a brochure that a potential panelist can study at leisure.

Both address-based and telephone RDD-based recruitment strategies should begin with a random selection of eligible households within a geographically-defined sample frame for an expected sonic boom carpet corridor. In the case of an address-based sampling frame, the cross-track width of the interviewing site will depend on population density and proximity to the boom

---

[47] Note further that Miller *et al.* (2014) also report that "households having a matching telephone number, living in a census block with a high percentage of persons aged 50 and over, or living in a census block with a low percentage of Hispanics, are more likely to respond to the survey." These findings suggest that English-fluent households with landline telephone service and more elderly household residents are likely to be over-represented in airport-vicinity social surveys. Miller *et al.* (2014) did not observe that noise exposure *per se* was meaningfully related to willingness to participate in their survey.

carpet edge (lateral cutoff distance) under worst tolerable uncertainty about the location of the carpet boom centerline. In communities of greater population density, the cross-track width of interviewing areas can be narrower, potentially providing greater homogeneity of boom exposure levels than in communities of lesser population density. In the case of telephone RDD sampling frame, the cross-track width of the interviewing area will depend on the geographic distribution of telephone working blocks fully contained within a distance of the flight track centerline corresponding to boom exposure levels of ±2.5 dB or less.

Non-residential (*i.e.*, commercial, institutional, *etc.*) addresses should be excluded from areas selected for boom exposure that are defined by address-based sampling. It will also be necessary to attempt to associate potential panelist names with street addresses. Mailed notifications sent to the fraction of initial postal solicitations of interest in panel participation that cannot be associated with a resident name will have to be addressed simply as "Resident," and the salutation of the subsequent invitational mailing will have to be "Dear Sir or Madam." Neither practice is likely to improve response rates.[48]

If the initial enumeration of households eligible for interview is made by telephone RDD sampling, then both names and street addresses will have to be associated with telephone numbers. Name and address matches to about half of eligible telephone numbers can be expected. A pre-dialing of the RDD sample can determine which of the numbers are non- residential (*i.e.*, business, fax/modem, *etc.*)

Formal invitational mailings

As described by Cointin *et al.* (2016), households that do not return a self-addressed, stamped acknowledgment of interest in participation should receive a "thank-you/reminder postcard" approximately one week after the initial survey mailing. Regardless of the source of the sample frame used to identify potential panelists, a subsequent formal invitational mailing should be sent to all recipients of the original notification postcards who responded positively. To attract notice and attention, every effort should be made to make the invitational mailing as distinctive as possible. For example, mailings should be hand-stamped and addressed, oversize mailpieces, with a NASA logo and return address.

The postal invitation should deliver an information packet briefly describing the purpose and nature of the panel interview. The packet should include a stamped, self-addressed envelope for postal return of a one-page form on which addressees potentially interested in joining a panel can provide basic information for further contact, as well as an informed consent form. The back

---

[48] Link *et al.* (2008) report that use of surnames where available in an address based sample may be counterproductive.

of the envelope should display a large type printed message to the effect "This is a request for paid assistance in a NASA flight test program...do not discard."

Efforts to contact non-respondents to the invitational mailing must then be pursued with further mailings and/or telephone callbacks. Options for doing so are discussed below.

Option 1 – Initial Address-Based Sample Frame

The sample frame is composed in this approach from an address database such as the USPS delivery sequence database. The process flow for this option is shown in

Figure 51. Beginning in the upper left of the figure, households eligible for interview are enumerated for a geographically-defined subset of addresses. Telephone numbers and names of household members are appended to the sample records as available from other databases. A pre-notification post card is sent to a random sub-set of suitable size, or if the interviewing area is sufficiently small and the population density is sufficiently low, to all households in the frame.

A week later, formal invitation letters with cash incentives are mailed to the same addressees. Some of these addressees will return an acceptance letter to NASA. After another week, non-respondents will receive a first follow-up contact attempt. For sampling frame records that have no telephone number associated with them, the first follow-up mailing would be similar to the first. For sampling frame records that *do* contain telephone numbers, live agents would attempt a telephone follow-up. If contact is made and an acceptance obtained by telephone, the panelist would be added to an acceptance list.

Some fraction of the first follow-up attempts will result in acceptances. The remainder will receive a second follow-up attempt by courier service or next-day postal delivery to draw additional attention to the solicitation. The second follow-up process is diagrammed in

Figure 52. The mailpiece would contain a slightly modified version of the first follow-up letter (stressing the importance of participation), along with another mailback form and a self-addressed, stamped envelope.

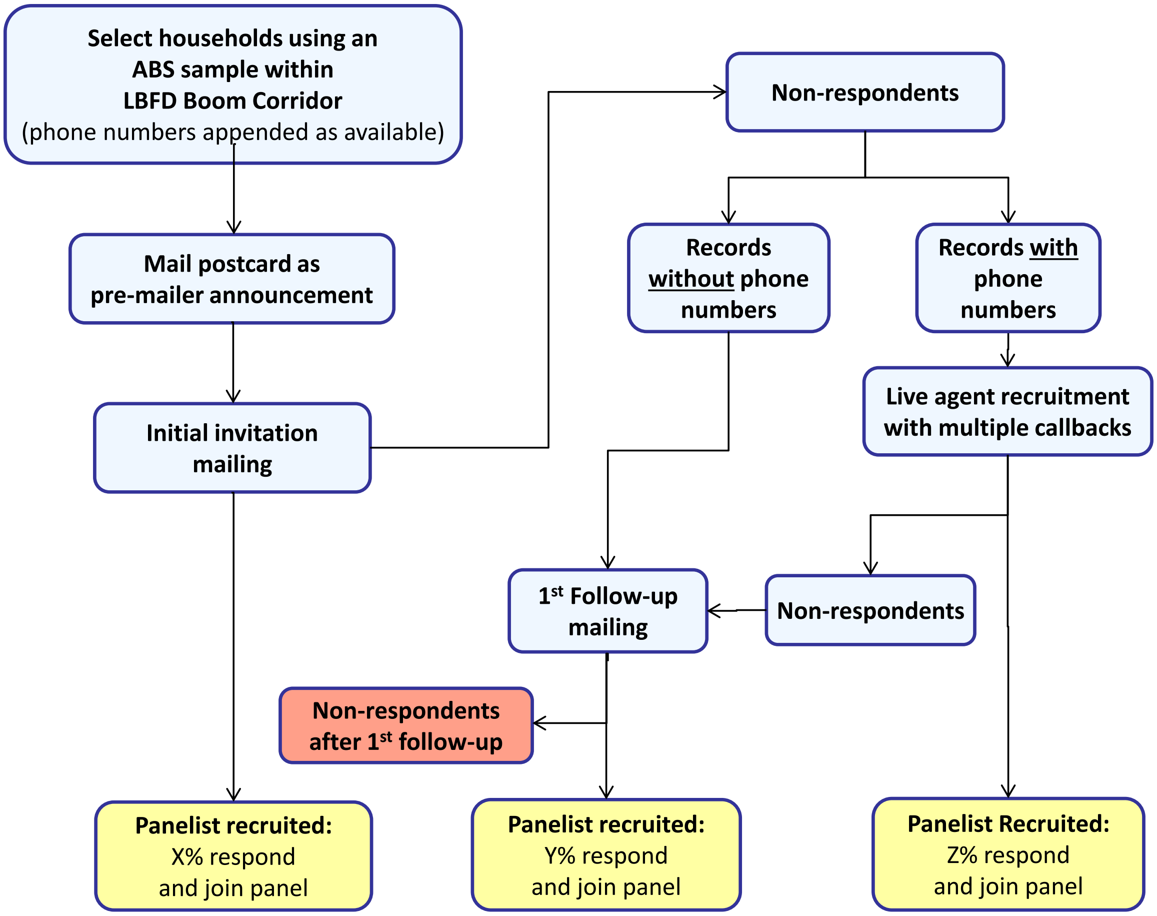


Figure 51. Flowchart for recruitment contacts from an address-based sampling frame.

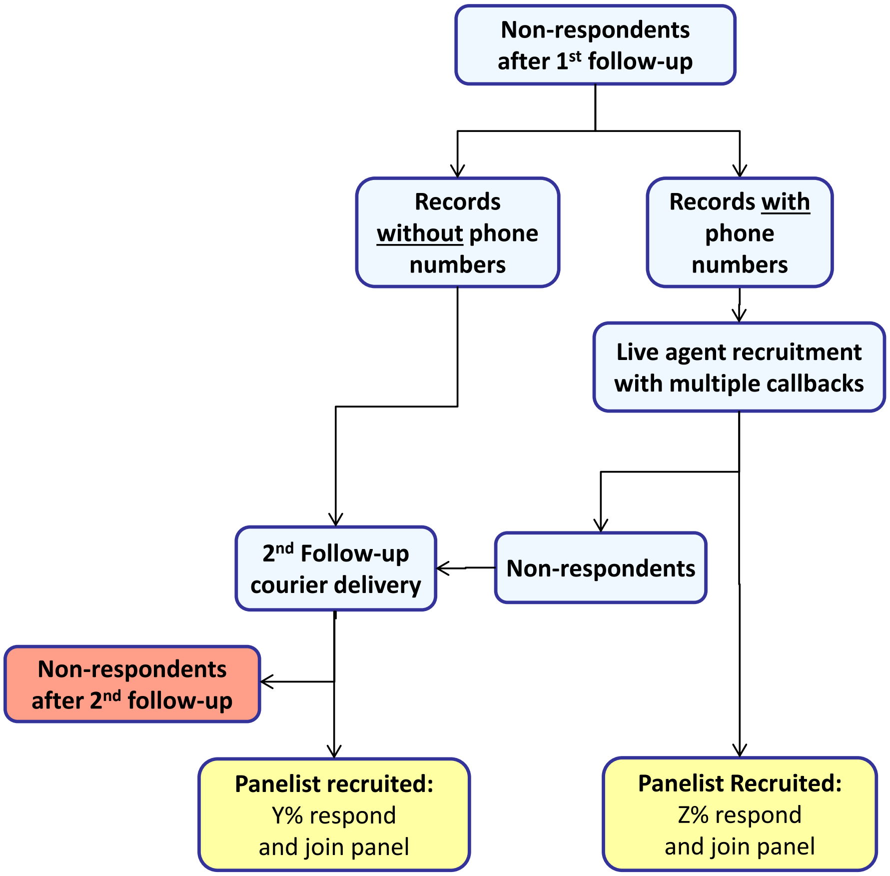


Figure 52. Flowchart for a second follow-up (courier or other overnight) based recruitment solicitation.

Option 2 – Telephone-based sample frame

The sample frame is composed from an RDD EPSEM database of telephone working blocks fully contained within the expected sonic boom corridor and the initial invitation would be conducted by telephone if possible. The process flow for this option is shown in Figure 53. Beginning in the upper left of the figure, households eligible for interview are identified. The sample frame is updated with current telephone company information to eliminate known disconnected numbers. Mailing addresses and names of household members would be appended to the sample records as available in other databases.

Sample records *with* appended addresses would be sent the initial postcard mailing as in Option 1. This mailing, however, would alert the recipient to an initial live agent telephone invitation to join the study. A week later, these records would be (repeatedly, if necessary) contacted by a live agent interviewer. Those sample records *without* an appended address would receive the live agent recruitment attempt directly. The initial telephone invitation will recruit some panelists. The remaining non-respondents would receive a first follow-up invitation by mail, and a second follow-up by courier or next-day mailing as in Option 1.

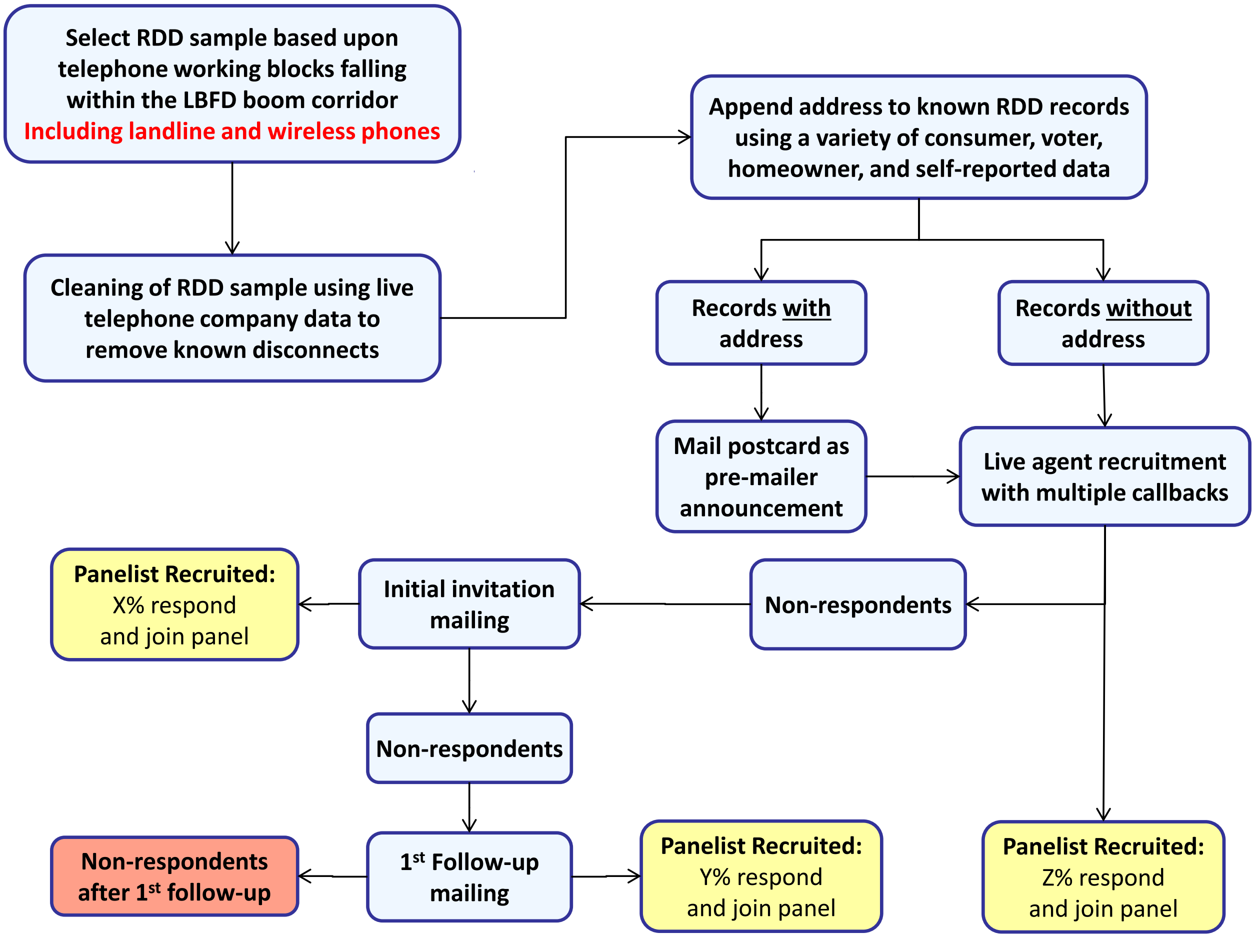


Figure 53. Flowchart for recruitment contacts from an RDD EPSEM telephone sample frame.

Information provided in panelist recruitment materials

The initial solicitation of interest in panel participation should describe the LBFD flight test program in broad terms only, to avoid preferentially attracting panelists with strong positive or negative opinions about supersonic flight and aircraft noise. Follow-up information subsequently provided to potential panelists can be more specific about the purposes of the flight test program, but must also avoid suggestive wording that could influence responses to interview questions.

Estimating required numbers of invitations and panelists

The Pew Research Center (2015) reports that slightly more than half of all individuals invited to participate in its American Trends panel study agreed to do so:

> "A 54% majority of those who were invited to join the panel agreed to do so. Of those invited *[to participate in the panel]*, 43% joined and responded to at least one wave *[of interviews]*..., 33% responded to at least five waves, and 20% responded to all *[nine]* waves."

These figures imply about a 9.1% constant attrition rate in panel participation from round to round. If a minimum of about 300 panelists are to contribute opinions to nine rounds of interviews, then about 640 respondents must be recruited for the first round of interviews. Similarly, if 12 overflights are anticipated during an LBFD deployment, and the opinions of about 300 panelists are sought in the final round of interviews, then about 850 panelists would have to be recruited initially.

Note, however, that the Pew estimates of panel attrition rates were derived from long-term studies in which substantial numbers of panelists may move (or even die) from year to year. For short-terms studies of the present sort, the Pew attrition estimates should be considered an upper bound, and lower attrition rate estimates are more likely. Table 12 shows the effects of compound attrition rates from 2% to 8% per round of interviews on the numbers of panelists remaining after the numbers of rounds of interviews shown in the leftmost column. Even at these lower attrition rates, the clear inference to be drawn from these figures is that many more than 300 panelists are required initially if 300 panelists are required in a final round of interviews after exposure to tens of LBFD flight missions.

Table 12. Remnant panel size after specified rounds of interviews and compound attrition rates for a study with 300 initial panelists.

Round	Attrition per Round			
Number	0.02	0.04	0.06	0.08
2	288	276	265	254
5	271	245	220	198
10	245	199	162	130
15	222	163	119	86
20	200	133	87	57
25	181	108	64	37
30	164	88	47	25
35	148	72	34	16
40	134	59	25	11

Advisability of reminder notifications

Reminder calls are an expedient for minimizing panel attrition over the course of a field experiment. On days when LBFD missions are planned, panelists could be called to alert them to the possibility of noticing a low-amplitude sonic boom.[49] Reminder notifications when reactions to booms are gauged by self-report may be arguably useful in minimizing panel attrition between boom exposures, but are probably superfluous, if not counter-productive, when opinions of panelists are actively solicited. Alternate remedies for expected attrition, such as simple over-sampling to recruit enough panelists to leave a minimally useful number still active in the final boom exposures, are preferable for purposes of improving the face validity of study findings.

As noted in Chapter 3, participation in a longitudinal panel sample, by its very nature, calls undesired attention to the annoyance and startle of unexpected sonic booms. Reminder notifications further exacerbate differences between experimental and real-world exposures to low amplitude sonic booms. On balance, reminder notifications are not the best way to counter potential attrition. They can be an unnecessary measure that reduces the fidelity of the circumstances of experimental exposure to sonic booms and real-world exposure to booms produced by fleets of future supersonic aircraft.

4.3.8 Estimating costs of panel sample studies

The circumstances of individual LBFD deployments (basing options, flight times to communities of appropriate size and location with respect to LBFD flight tracks, air traffic and other operational constraints, *etc*.) dictate flight patterns, numbers, and times of exposure. For example, if it is discovered in earlier deployments that the range of boom exposure levels chosen does not yield useful information about annoyance prevalence rates of interest, different exposure ranges and numbers of flight might be selected for later deployments. It is premature to estimate such costs before LBFD mission routes have been selected.

Appendix A estimates relative costs of data collection costs for small, intermediate, and large scale cases.

4.3.9 Measures for controlling panel attrition between rounds of interviews

While some panelists will undoubtedly fail to complete interviews on successive rounds of interviews, the literature provides little guidance for confident prediction of attrition rates in short term panel studies. The main defense against inter-round attrition is some form of monetary incentive to panelists to remain engaged in the study.

---

[49] If false alarm rates for self-reporting are of interest, reminder notifications could also be made on days when no boom exposures are intentionally created. Such a practice would verge on deceptive, however, and might not be tolerated by an Institutional Review Board.

A flat rate incentive (for example, a periodic payment such as $25/week) could be offered to all panelists who complete a minimum number of interviews per week. Such a salary-like payment to all qualifying panelists would simplify the administrative expense of closely tracking each panelist's daily performance, but might not provide sufficient incentive to answer every telephone call requesting an interview, particularly if opinions about multiple booms per day are solicited. A large enough flat rate incentive would almost certainly encourage continuing participation in a panel, but could unduly burden project budgets. An Institutional Review Board would probably also view an offer of a substantial payment for participation in a study as a coercive impediment to obtaining true informed consent.

Piece-rate (per interview) payment is a preferred alternative to a flat rate incentive payment.[50] (Note, however, that it could encourage self-reporting of marginally noticeable booms.) Such incentives must be large enough to be meaningful, but not so large as to be unaffordable within a project budget, or coercive for panelists. Piece-rate incentives would require close tracking of the participation of individual panelists, with its attendant administrative costs. Such tracking, however would also permit end-of-day reminder contacts that could serve as another measure for improving panelist retention.

4.3.10 Questionnaire

The questionnaire to be repeatedly presented to panelists should be as brief and direct as feasible. The substantive items should be limited to questions about recent notice of a sonic boom, whether it was startling, and its degree of annoyance. The form of the questions and closed response category alternatives should closely resemble those of the cross-sectional study described in §3. The questionnaire should then inquire whether the panelist was at or near home at the time of notice of the sonic boom, and whether the panelist was indoors or outside home when the boom was noticed.

An additional question may be posed if the panelist was not at home at the time of notice of a sonic boom, to the effect "If you had heard this boom while at home, how annoying would it have been?" The following subsections address the rationale for adding such a question.

4.3.11 Necessity for estimating boom exposure away from panelists' homes

Dosage-response functions are uncertain on both their exposure (abscissa) and opinion (ordinate) axes. Excessive uncertainty on either axis leads to an unreliable dosage-response function of little utility for regulatory interpretations. As noted in §4.1.3 and §4.1.4, policy

---

[50] If piece rate incentives are offered, care must be taken to solicit panelist opinions at a few times when no low-amplitude booms have been produced, in order to permit estimation of false alarm rates.

decisions of regulatory agencies are based on geographic associations between outdoor residential noise exposure and the prevalence of a consequential degree of annoyance. Actual personal noise exposure is irrelevant to such regulatory analyses.

In controlled exposure field experiments whose findings are not intended to be generalized to the national population, precise information about respondent locations is useful for reducing the uncertainty of estimates of predicted boom levels. In social surveys whose findings are intended to be fully generalizable, a larger sample of respondents is useful for reducing the uncertainty of estimates of annoyance prevalence rates. High precision of boom exposure levels may not be as essential in the social survey case for at least two reasons:

1) Since noise regulatory policy points are conventionally set 5 dB apart, and the resolution of the great bulk of noise exposure estimates routinely considered by regulatory agencies in formulating policy is on the order of ± 2.5 dB, very high precision of boom level estimates is unwarranted, especially if it comes at the expense of limited resources for improving the precision of measurement of annoyance and/or startle effects; and

2) Cross-track gradients of sonic boom levels are shallow near centerlines of carpet boom corridors. If communities to be overflown on LBFD missions are judiciously chosen, and if exposure levels are estimated directly by real-time monitoring rather than by modeling, errors of estimate of exposure levels may be small enough to be negligible for regulatory purposes.

The most expedient way to deal with the issue of opinions expressed by panelists who are not at home at the time of boom exposure may be at the data analysis stage. Dosage-response relationships should be constructed only for panelists who describe themselves as at- or near-home at the time of interview. (Note however, that such a dosage-response relationship will almost certainly contain a nationally disproportionate number of responses from the frequently-at home demographic – typically, the retired, elderly, and ill.)

4.4 Sampling Opinions about Delayed Reactions to Booms

As described in §4.2, independent (single, non-pre-arranged telephone interview per respondent) sampling is infeasible for purposes of assessing prompt community reaction to individual low-amplitude sonic booms. Although independent sampling of opinion about the annoyance and startle of individual sonic boom exposures is ill-suited to time-critical data collection, it may be useful for assessment of reactions to cumulative sonic boom exposure. For example, with enough callbacks, single interview per respondent study designs may be useful for assessing community reaction to the total impulsive noise exposure of an entire LBFD deployment's worth of sonic booms.

Cumulative carpet boom exposures should be estimated both by residential addresses of individual respondents, and under the assumption of homogeneity of exposure for all respondents.

The former estimates should be derived from exposure gradients derived from as dense a network of measurement points as is affordable, while the latter estimates should be based on modeling. The precision of boom exposure estimates need not differ greatly from the exposure bin widths adopted for dosage-response analyses.

4.4.1 Some recent experience with multiple callbacks

Callback attempt information was analyzed for eight recent telephone surveys: three helicopter noise annoyance studies conducted for ACRP (Mestre *et al.*, 2017) and five surveys conducted by California State University, Fullerton. Table 13 displays the surveys, numbers of callbacks, and sample sizes.

Table 13. Multiple callback social surveys.

Venue	Study	Year Completed	Number of Callbacks	Total Sample Size
California	Post-college employment	2017	16+	25,628
Northern California	Public transportation	2018	9+	1,201
Dominguez Hills, CA	Public opinion (various topics)	2017	6+	384
San Francisco County	Water conservation	2018	6+	402
National	Small Business Microloan	2017	20+	195
Las Vegas	ACRP 02-48 Helicopter Noise	2015	10	741
Long Beach	ACRP 02-48 Helicopter Noise	2015	9	1089
Washington, DC	ACRP 02-48 Helicopter Noise	2016	10	442

Table 14 shows numbers of completed interviews by call attempt for each study. The absolute numbers are of little interest, but the rates at which these numbers decay with successive call attempts have implications for the design of a cross-sectional design of community reaction to cumulative sonic boom exposure.

An exponential decay rate provides a good approximation to the interview completion rates shown in Table 14. The natural logarithms of the numbers of completed interviews in Table 14 are plotted in Figure 54 against call attempt number. The near-linearity of the data points in most of the charts confirms the exponential decay. Linear least square fits to the data sets yielded coefficients of determination (r^2 values) greater than 0.9 for all but two studies.

The exponential decay exponents derived from the analyses presented in Figure 54 are shown in the second column of Table 15. The third column of Table 15 shows attrition rates, expressed as a percentage of the previous call attempt's yield in completed interviews (determined by the relationship: Attrition = $100 \bullet [1 - e^{\text{EXPONENT}}]$). The slope derived from each study was then used to predict a *normalized* number of completed interviews per successive contact attempt.

Table 14. Completed telephone interviews by call attempts.

Attempt	Study							
	CA CC	Nor CA	CSU DH	SF County	Small Business	Las Veg	Long Beach	Wash DC
1	4362	406	151	82	22	306	287	83
2	4060	263	90	104	25	137	158	48
3	2982	184	51	100	16	89	126	41
4	2451	142	40	75	25	88	96	46
5	1887	92	26	26	20	51	70	31
6	1712	71	15	15	10	33	66	27
7	1364	16			12	22	45	20
8	1170	13			10	10	39	19
9	1043	10			7	5	46	8

Table 15. Exponential decay rates derived from eight recent telephone surveys.

Study	Exponent	Attrition per call attempt
California Community College - 2017	-0.190	17.3%
Northern California - 2018	-0.490	38.8%
CSU Dominguez Hills - 2017	-0.443	35.8%
San Francisco County - 2018	-0.370	30.9%
Small Business Microloan - 2017	-0.147	13.7%
Las Vegas - 2015	-0.468	37.4%
Long Beach - 2015	-0.233	20.8%
Washington DC - 2016	-0.235	21.0%
Average	-0.322	26.9%
Standard Deviation	0.1364	9.9%
N	8	8

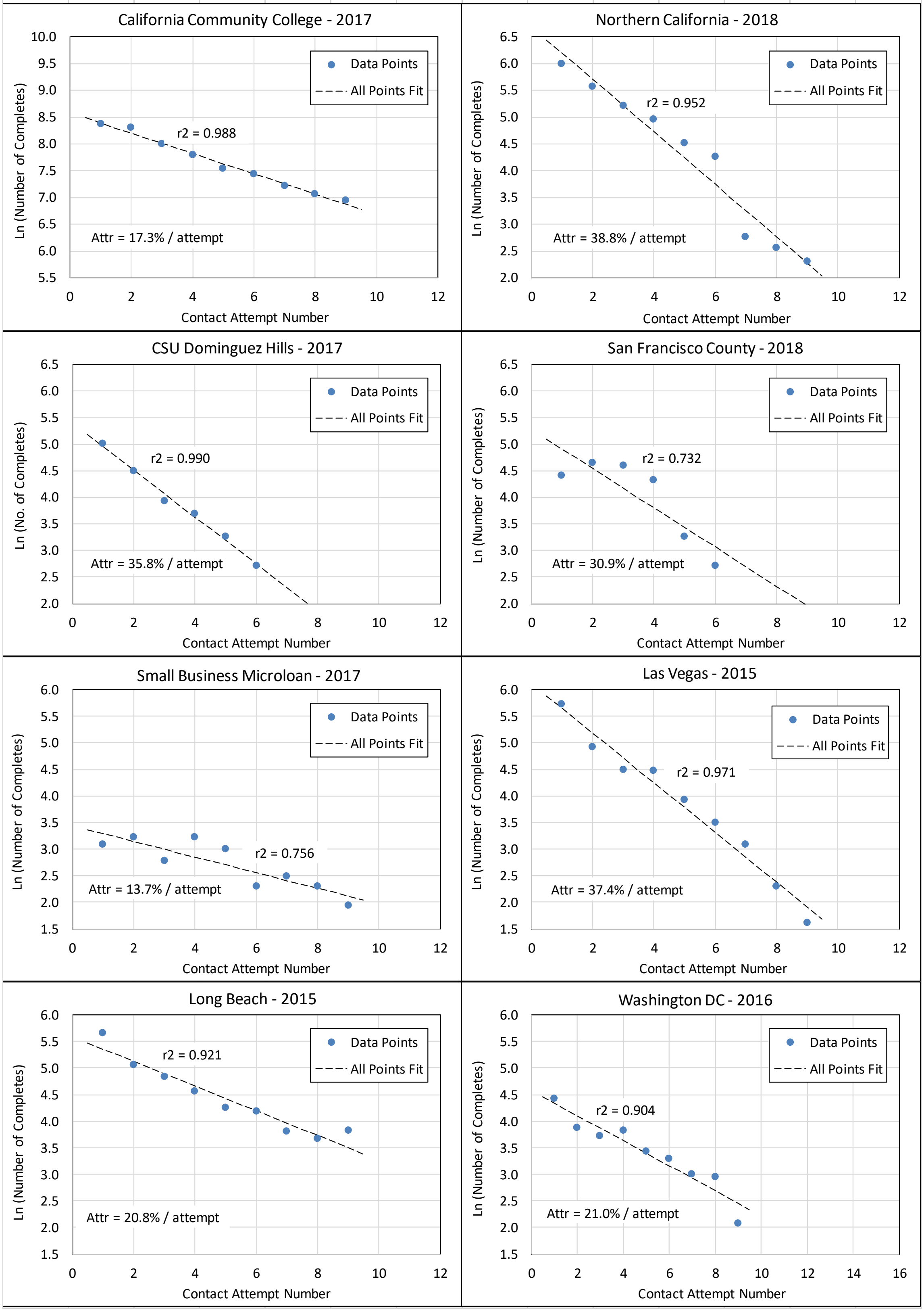


Figure 54. Exponential decays in numbers of completed interviews with increasing numbers of contact attempts.

The normalized number of interviews is the fraction of first contact attempt completed interviews obtained with each additional call attempt. Figure 55 plots these normalized numbers *versus* contact attempt. The form of the curves is:

$$F = \exp(m * [N - 1]) \qquad \text{Equation 4}$$

where: F = fraction of first-attempt completed interviews
m = slope derived from fit in Figure 54
N = contact attempt number

The curves in Figure 55 appear to cluster in two groupings, albeit for no obvious methodological reasons.[51] The range in exponent values is -0.147 to -0.490, or a factor of 3.3. The red curve in Figure 54 is a prediction based on the average decay slope shown at the bottom of Table 15.

Figure 56 shows the normalized data of Figure 55 plotted in cumulative form. The ordinate in this figure plots the normalized cumulative number of completed interviews acquired after N rounds of calls. For example, at Las Vegas (orange diamonds), after nine call attempts the total number of completed interviews was 2.6 times greater than that obtained from just the first round. The red curve is the prediction using the average decay slope. The range of completed interview multiples after nine call attempts is 2.5 to 5.4.

4.4.2 Summary of analyses of multiple callback findings

The major implications of the above analyses are 1) that five callback attempts are likely to improve interview completion rates by a factor of at least two, and 2) that nine contact attempts provide no more than about a factor of five improvement in the interview completion rate of a single contact attempt. Thus, if the nominal interview completion rate on the first contact attempt is as great as 5%, then five additional contact attempts are likely to provide interview completion rates between 10% and 20%. Yet more callbacks could yield only a marginal improvement in interview completion rates, or at best, an interview completion rate of about 25%.

It is readily apparent that the utility of multiple telephone contact attempts for assessing community reaction to cumulative sonic boom exposure over periods of less than weeks is limited. Interview completion rates on the order of 20 – 25% may be similar to those obtained in recent years by postal surveys, but are not great enough to allay all potential objections to the generalizability of study findings.

---

[51] The subject matter of the social survey is probably not a primary reason for the different exponential decay rates, since few of the early failed contact attempts were refusals. Callback intervals may have differed somewhat among the non-time critical studies, however. A minimum of three hours is commonly adopted for callback scheduling, but the duration of the callback interval is not standardized in opinion research.

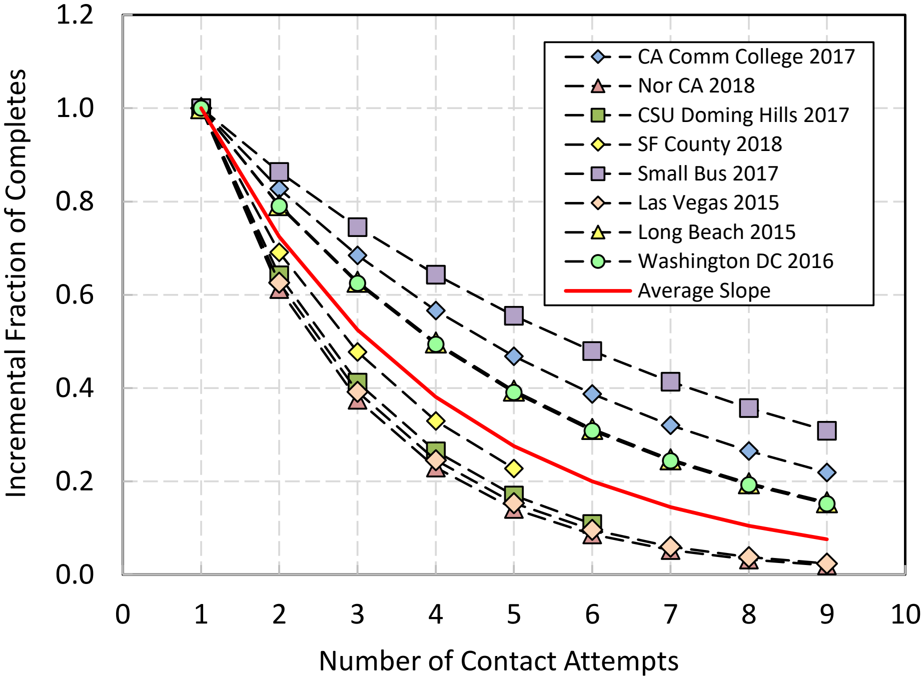


Figure 55. Example of effect of number of contact attempts on number of completed interviews per attempt.

Further, if calling is limited to the time period from the last LBFD flight of the day to 8:00 PM or 9:00 PM local time, then as little as three to four hours per day may be available for callbacks. Even if callbacks are scheduled as often as every two hours, calling hours would not permit more than two callback attempts on the same day. If interviewing about reactions to cumulative exposure were to begin on the day after the last flight, then as many as five callbacks might be possible within a day of the last exposure episode, and as many as ten callbacks might be possible within two days after the final exposure. Interviews about delayed reactions to cumulative boom exposure should also avoid weekday, daytime periods. Particularly in higher population density areas, response rates are considerably greater during evening and weekend time periods, when more people are home from work.

The greatest utility for repeated contact attempts in a cross-sectional study design is thus for gauging community response for a week or more's exposure to sonic booms. Little improvement in interview completion rates can be expected from scheduling callbacks often enough to complete two or more contacts following the last of a single day's LBFD overflights. (Multiple callbacks could also be attempted if necessary to improve interview completion rates in a longitudinal (panel) sample study of opinions about long-term, cumulative exposure.)

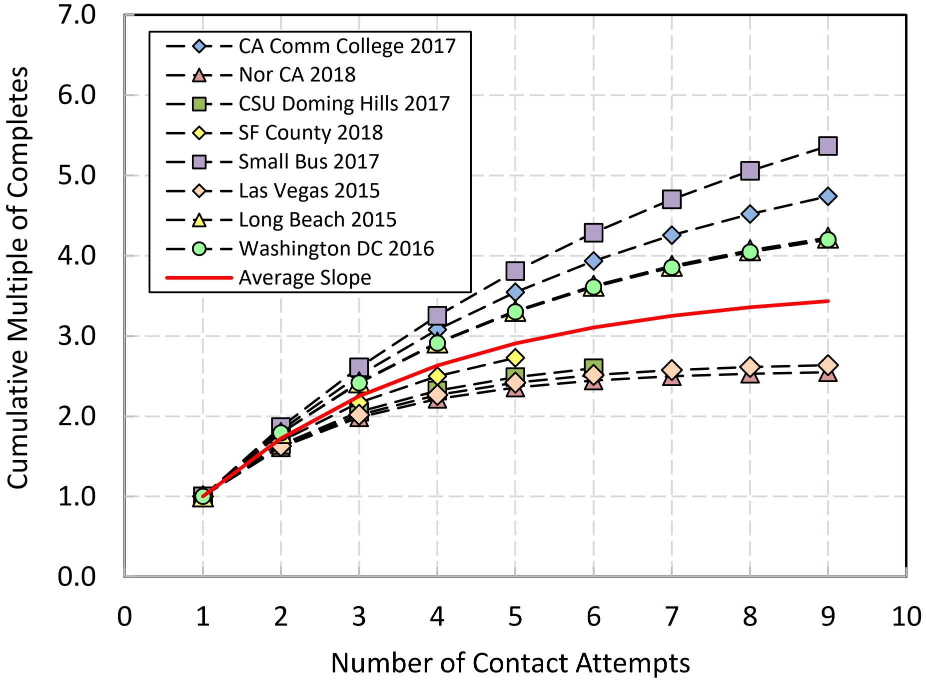


Figure 56. Examples of effect of repeated contact attempts on cumulative numbers of completed interviews.

THIS PAGE INTENTIONALLY LEFT BLANK

5 SUMMARY AND CONCLUSIONS

5.1 Summary of Major Findings

5.1.1 Summary of first year's (Phase I) results

The first year's effort identified sampling and interviewing as the principal risks to assessment of prompt reactions to overflights producing low-amplitude sonic booms. It also 1) established the utility of geo-information system-based route planning for LBFD flight missions, and 2) developed and demonstrated a prototype of a geographically-distributed, Internet-enabled instrumentation system capable of wide-area tracking of LBFD aircraft in near-real time. The latter system permits synchronizing the conduct of interviews in multiple overflown communities with arrival times of shock waves at interviewing sites; and of measuring, archiving, and processing their acoustic signatures.

Means were also recommended for constructing representative, telephone-based samples of eligible respondents living in households within carpet boom corridors adjacent to LBFD flight tracks, and for conducting interviews with cross-sectional (independent) samples of such respondents about their prompt reactions to exposure to low-amplitude sonic booms.

5.1.2 Summary of second year's (Phase II) effort

A detailed study design was prepared and accepted by NASA for a set of single-contact attempt telephone interviews with a nationally representative sample of households. The study design focused on testing automated and live agent interview completion rates obtainable without callbacks.

A minimal (two monitoring station) version of the aircraft tracking system was built and installed near a civil airport in a successful demonstration of the system's ability to detect and track aircraft movements. The field exercise also demonstrated the ability of the system to capture the acoustic emissions of departing aircraft, and to serve aircraft position and sound level information to remote, geographically-distributed analysts in near-real time.

5.1.3 Summary of third year's (Phase II) effort

Upon approval of OMB and IRB of the detailed study plan, a stratified, nationally representative sample of landline and wireless telephone-subscribing households was constructed. A total of 12,734 telephone interview contact attempts of the sort required by a straightforward cross-sectional study design were then made. These contact attempts demonstrated the impracticality of conducting a time-critical, cross-sectional study of prompt community response to low-amplitude sonic booms by means of "independent" (single contact attempt per respondent for each LBFD flight mission) telephone samples of respondents. The observed interview completion rates for these single telephone contact attempts were so low (~ 1% to 3% for automated and live agent interviews, respectively) that:

> 1) the representativeness of collected opinions would be susceptible to intuitive challenge as inadequate, even absent conclusive evidence of non-representativeness. Refuting challenges to representativeness would have to demonstrate that the composition of the actual sample did not differ from that of the target population, a task that is tantamount to proving a negative;
>
> 2) the information required to refute allegations of non-representativeness would require a questionnaire considerably lengthier than that required simply to determine the prevalence of boom-induced startle and annoyance. Such a questionnaire would have to inquire about potentially sensitive and intrusive matters, including respondents' age, gender, education, employment, home ownership, income, ethnicity, family size, and other demographic factors; and
>
> 3) unreasonable numbers of attempts would be required to re-contact households with unsuccessful initial contact attempts, given the limited time available for doing so. For example, if about 500 completed interviews were desired in a supersonically overflown community, approximately 50,000 automated interview attempts would have to be made within ten to fifteen minutes of each LBFD overflight. Such large numbers of contact attempts could well exceed the numbers of households available for interview in areas of similar boom exposure levels in some communities near LBFD flight tracks. Such large numbers of interviews could be cost-effectively undertaken only by means of automated (*i.e.*, outgoing interactive voice response) interviewing, a data collection method ill-suited for complex and sensitive questionnaire items.

The infeasibility of independent sampling for evaluating prompt responses to LBFD overflights in a cross-sectional study is due in large part to simple non-response: that is, potential respondents – particularly those contacted on wireless telephones – refusing to answer calls with unfamiliar caller IDs. It is also due in part, however, to 1) the lack of time to attempt to contact the same respondent more than once within a few minutes after the arrival of a shock wave at the respondent's location; and 2) the need to place calls during weekday/daytime hours, when response rates are notably lower than during evenings and weekends.

Despite the poor interview completion rates achieved under the above constraints, cross-sectional assessments of delayed reactions to LBFD overflights could still be feasible, if multiple attempts could be made to contact respondents during evening and weekend time periods, over extended time periods.

Detailed plans for a longitudinal (panel) sample were developed as an alternative to a cross-sectional sample design.

5.2 Conclusions and Recommendations

The findings summarized above support the following conclusions and recommendations:

1) Route planning for LBFD missions should be conducted by means of customized geo-information system software. The most cost-effective missions for data collection purposes are those that permit collection of community response information from multiple communities underlying carpet boom corridors on individual flights.

2) The information most directly useful to FAA and ICAO in reconsidering their prohibitions on overland supersonic flight is that produced in studies which resemble social surveys more closely than controlled exposure field experiments. Differences in reactions to low-amplitude sonic booms from one community to the next are likely to be comparable or greater in magnitude to differences associated with exposure to booms varying in magnitude by a few decibels. The former differences are not cost-effectively measured in a series of single-community studies.

3) Respondents in a cross-sectional social survey of community response to low-amplitude sonic booms would ideally complete a single, brief telephone interview inquiring – within minutes of the time that shock waves produced by LBFD overflights reach their communities – about their annoyance and startle due to single sonic booms.

4) The requirement for rapid collection of prompt reactions to overflights renders such straightforward solicitations of opinions infeasible in LBFD field testing. Daytime, week day telephone interview completion rates without callbacks are prohibitively low, and the additional callbacks needed to improve interview completion rates are inconsistent with the requirement to assess prompt reactions to boom exposure.

5) A simpler cross-sectional study design is still potentially useful for quantifying week-or-longer delayed reactions to the cumulative boom exposure of LBFD overflights.

6) Sequential biases potentially introduced by a longitudinal (panel) study designs can be detected and quantified if multiple sub-panels are formed, of which only sub-sets are interviewed following each LBFD overflight.

7) A range of alternate (panel sample) study designs appears feasible, but need to be adapted to deployment-specific factors such as anticipated numbers of LBFD flight missions and of communities to be overflown on each mission. Such study designs should be pre-tested prior to implementation to further minimize data collection risks by optimizing details of panelist recruitment and determining the effectiveness of retention measures.

8) Only small percentages of study participants are likely to notice (and judge as highly annoying) low amplitude booms with PLdB values on the order of 75 to 85 PLdB. As discussed in Appendix B, this creates analytic difficulties in constructing credible dosage-

response relationships. A form of single-event CTL (Fidell *et al.*, 2011) analysis applied to lower annoyance levels can resolve such difficulties.

6 ACKNOWLEDGEMENTS

The authors thank Ms. Laura Gil-Trejo, Director of the California State University Fullerton Social Science Research Center, for supplying information about recent telephone interview completion rates as a function of number of callback attempts. Thanks are also extended to Dr. Barbara Tabachnick and Mr. Steven Clark for contributions to Chapter 4 of this report. Drs. Jonathan Rathsam and Kevin Shepherd of NASA Langley Research Center participated in helpful discussions of several technical issues.

THIS PAGE INTENTIONALLY LEFT BLANK

7 REFERENCES

Albert, S., King, J., and Keene, R. (**2016**) “Assessment of an interactive voice response system for identifying falls in a statewide sample of older adults,” Preventive Medicine, Volume 71, pp. 31-36.

Blumberg, S., and Luke, J. (**2015**) “Wireless substitution: early release of estimates from the National Health Interview Survey, July-December 2014,” National Center for Health Statistics, Centers for Disease Control and Prevention, Atlanta, GA.

Blumberg, S., Luke, J., (**2017**) “Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2016,” National Center for Health Statistics, Centers for Disease Control and Prevention, Atlanta, GA.

Blumberg S. and Luke, J. (**2018**) “Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2017,” National Center for Health Statistics, Centers for Disease Control and Prevention, Atlanta, GA.

Boring, E. G. (**1921**) “The Stimulus-error,” The American Journal of Psychology, Vol. 32, No. 4, pp. 449-471.

Cantor, D., and Williams, D. (**2013**) “Assessing Interactive Voice Response for the National Crime Victimization Survey,” Final Report, Westat, Rockville, MD.

Cliatt II, L., Haering Jr., E., Jones, T., Waggoner, E., Flattery, A. and Wiley, S. (**2014**) “A Flight Research Overview of WSPR, A Pilot Project for Sonic Boom Community Response,” 32nd AIAA Applied Aerodynamics Conference. 10.2514/6.2014-2268.

Cointin, R., Sizov, N., and Hileman, J. (**2016**) “U.S. Civil Aircraft Noise Annoyance Survey Design,” Internoise (Hamburg).

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

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. (**2015**) “A Review of U.S. Aircraft Noise Regulatory Policy,” Fidell, S., Acoustics Today, 11 (4), 26 – 34.

Fidell, S., (**2018**) “A Modern Standardized Method for Predicting Community Response to Aircraft Noise”, Civil Engineering and Architecture 6(2): 71-77, 2018 http://www.hrpub.org DOI: 10.13189/cea.2018.060204.

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., Horonjeff, R., Mestre, V., Clark, S., Fitzpatrick, J., and Sanchez, G. (**2016**) “Final Report: Test Plan for Phase II of Low-Boom Flight Demonstrator Risk Reduction Study,” NASA Contract NNL15AA05C.

Fidell, S. and Horonjeff, R. (**2017**) “Phase II, Year 1 Interim Report”, Contract NNL15AA05C.

Fields, J., and Powell, C. (**1985**) “A community survey of helicopter noise annoyance conducted under controlled noise exposure conditions”, NASA Technical Memorandum 86400.

Fumagalli, L., Laurie, H. and Lynn, P. (**2013**) “Experiments with methods to reduce attrition in longitudinal surveys,” J. R. Statist. Soc. A., 176, Part 2, pp. 499–519.

Groves, R. (**2006**) “Non-response rates and non-response bias in household surveys,” Public Opinion Quarterly, 70(5), Special Issue 2006, 646–675.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (**2004**) “Survey Methodology,” Wiley, Hoboken, NJ.

Herzog, A. & Rodgers, W. (**1988**) “Age and Response Rates to Interview Sample Surveys,” Journal of Gerontology, 43(6), November 1988, S200–S205.

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

Kryter, K., Johnson, P., and Young, J. (**1967**) “Psychological experiments on sonic booms conducted at Edwards Air Force Base,” SRI Project ETU-6065, Air Force Contract AF49 (638(-1758, National Sonic Boom Evaluation Office, Arlington, VA.

Link, M., Battaglia, M., Frankel, M., Osborn, L. and Mokdad, A (**2008**) “A Comparison of Address-Based Sampling (ABS) *Versus* Random-Digit Dialing (RDD) For General Population Surveys,” Public Opinion Quarterly, Vol. 72, No. 1.

Mestre, V., Fidell, S., Horonjeff, R., Schomer, P., Hastings, A., Tabachnick, B., and Schmitz, F. (**2017**) “Assessing Community Annoyance of Helicopter Noise,” Report 181, Airport Cooperative Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine.

Miller, N., Cantor, D., Lohr, S., Jodts, E., Boene, P., Williams, D., Fields, J., Gettys, M., Basner, M., and Hume, K. (**2014**) “Research Methods for Understanding Aircraft Noise Annoyances and Sleep Disturbance,” Final Report, ACRP 02-35, Transportation Research Board, National Academies of Science, Washington, D.C.

Morgenstern, J. (**2016**) Personal communication.

Muehleisen, R. (**2016**) "Virtual Sound Level Meter for analyzing calibrated sound files," https://muehleisen.github.io/VSLM/, visited 10/26/2018.

NASA LaRC (**2017**) Submissions of applications and supplementary information to Office of Management and Budget and NASA’s Institutional Review Board for review and approval of field studies under Contract NNL15AA05C.

National Opinion Research Center (NORC) (**2018**), University of Chicago, “Amerispeak Panel Sample,” https://amerispeak.norc.org/, accessed on 28 August 2018.

Page, J., Hogdon, V., Krecker, P., Cowart, R., Hobbs, C., Wilmer, C., Koening, C., Holmes, T., Gaugler, T., Shumway, J., Rosenberger, J. and Philips, D. (**2014**) "Waveforms and Sonic Boom Perception and Response (WSPR): Low-Boom Community Response Program Pilot Test Design, Execution, and Analysis," NASA CR-2014-218180.

Pew Research Center (**2012**) “Assessing the Representativeness of Public Opinion Surveys,” Washington, D.C.

Pew Research Center (**2015**) “Building Pew Research Center’s American Trends Panel,” Washington, D.C.

Research Now (**2018**) https://www.researchnow.com/, accessed on 18 August 2018; see also https://en.wikipedia.org/wiki/Research_Now, accessed on 12 December 2018.

Schoeni, F., Stafford, F., McGonagle, K. and Andreski, P. (**2013**) “Response Rates in National Panel Surveys,” Ann Am Acad Pol Soc Sci., 645(1), 60–87.

THIS PAGE INTENTIONALLY LEFT BLANK

8 APPENDIX A. COST ESTIMATES FOR PANEL DATA COLLECTION

This Appendix contains illustrative examples of estimates of data collection costs of variant panel sample study designs. Each example assumes that the total recruited sample will be divided into sub-panels to permit the distinction of effects of sequential bias from those of cumulative exposure. The spreadsheet that generated the cost estimates permits users to vary the following seven parameters:

1) the number of households receiving postal solicitations of interest in participating in repeated interviews about reactions to low-amplitude sonic booms;

2) the fraction of households responding positively to the invitation to join a panel;

3) the fraction of participants actually granting interviews in each round of interviews;

4) the total number of LBFD flight missions per deployment[52];

5) the number of sub-panels into which the respondents are divided (to permit analyses that can distinguish between potential bias due to numbers of repeated interviews and potential bias due to cumulative exposure);

6) the monetary incentive provided to individual panelists for completing an interview; and

7) the cost of an individual telephone interview attempt.

The first three parameters concern panel recruitment. The remaining parameters concern interview administration during an LBFD field deployment. The spreadsheet combines user-selected values of these parameters to estimate a total data collection cost, the total number of completed interviews, and the cost per interview. These estimates are shown in **red** in the following tabulations.

Note that the cost estimates so-calculated pertain only to interview-related costs for a single panelist per household; exclude costs associated with other aspects of the conduct of field studies of community reaction to low-amplitude sonic booms; and are hence *not* intended as estimates of total study costs. They are, however, useful for comparative purposes, since they provide study designers with relative cost information for the subjective data collection component of alternate panel sample studies.

The values of the above parameters in three case study examples are summarized in Table 16.

---

[52] For simplicity of cost estimation, each flight is assumed to be accompanied by a round of interviews. More elaborate study designs may interview only sub-panels of respondents during a subset of LBFD flights.

Table 16. Values of cost estimation parameters assumed for purposes of illustration in small, intermediate, and large scale LBFD field tests.

Parameter	Case1	Case 2	Case 3
Desired number of households participating in repeated interviews about reactions to low-amplitude sonic booms	2,000	5,000	10,000
Fraction of households responding positively to the invitation to join a panel in initial and follow-up contact attempts	0.25	0.25	0.25
Fraction of participants actually granting interviews in each round of interviews	0.85	0.85	0.85
Total number of LBFD flight missions per deployment	18	36	72
Number of sub-panels into which the respondents are divided	6	6	6
Monetary incentive provided to individual panelists for completing an interview	\$3.00	\$3.00	\$3.00
Cost of an individual telephone interview attempt	\$1.00	\$1.00	\$1.00

The total estimated data collection costs for the three cases summarized in Figure 58, Figure 60 and Figure 62 are \$95,672, \$273,491, and \$713,799 for the small, intermediate, and large scale cases, respectively. The corresponding costs per interview are \$19.83, \$12.33 and \$9.18, for the three cases. A summary breakdown is shown in Table 17.

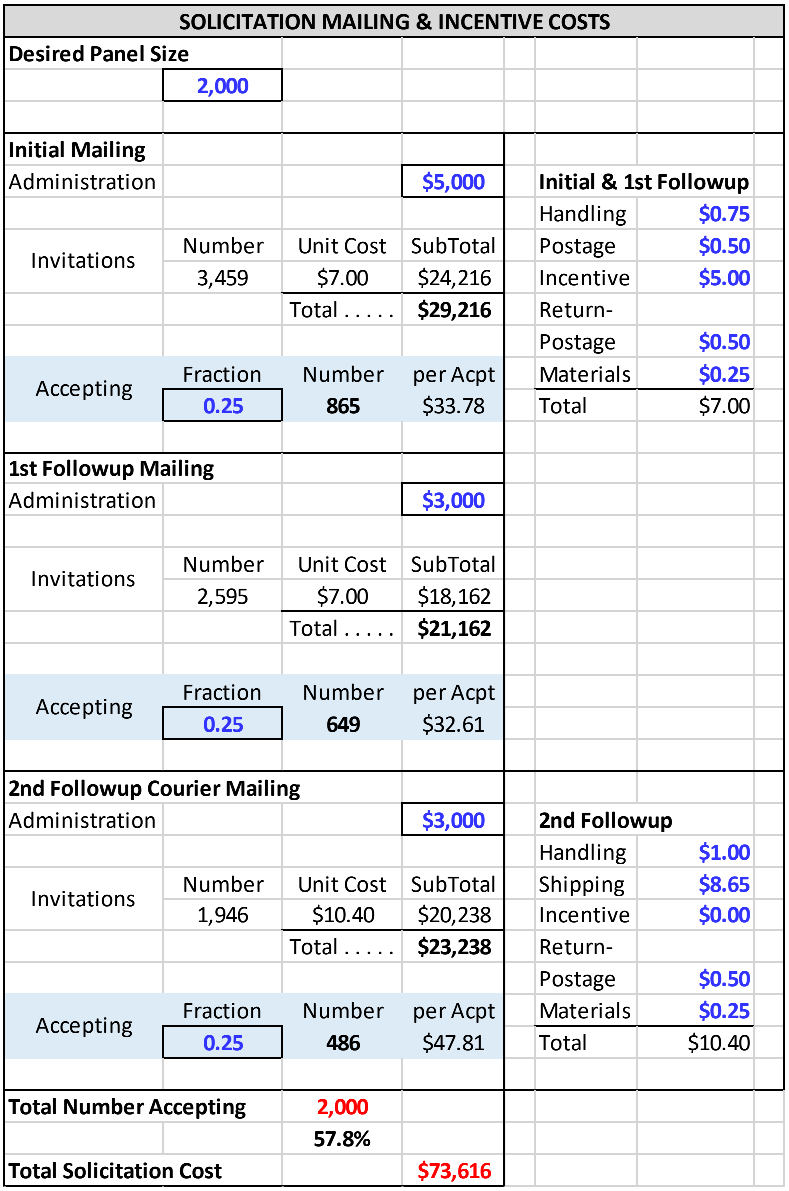
Table 17. Summary of data collection cost estimates for three cases.

Case	Study Scale	Desired Panel	Invitations Required	Completed Interviews	Cost Component			Cost per Interview
					Invitation	Interview	Total	
1	Small	2,000	3,459	4,824	\$73,616	\$22,056	\$95,672	\$19.83
2	Intermed.	5,000	8,649	22,176	\$167,541	\$105,870	\$273,411	\$12.33
3	Large	10,000	17,297	77,778	\$324,081	\$389,718	\$713,799	\$9.18

The three above cases are hardly the only ones of potential interest, nor are they necessarily recommended for further consideration. They are considered here merely to roughly approximate a plausible range of data collection parameters. Deployment-specific factors (*e.g.*, geographic distribution of communities in proximity to operating bases and flight tracks, availability and shape of airspace for initial acceleration to supersonic flight, flight time and basing budgets, and the like) must be specified before meaningful inferences can be drawn about data collection costs.

8.1 Example of an Interviewing Cost Estimate for a Small Scale Study

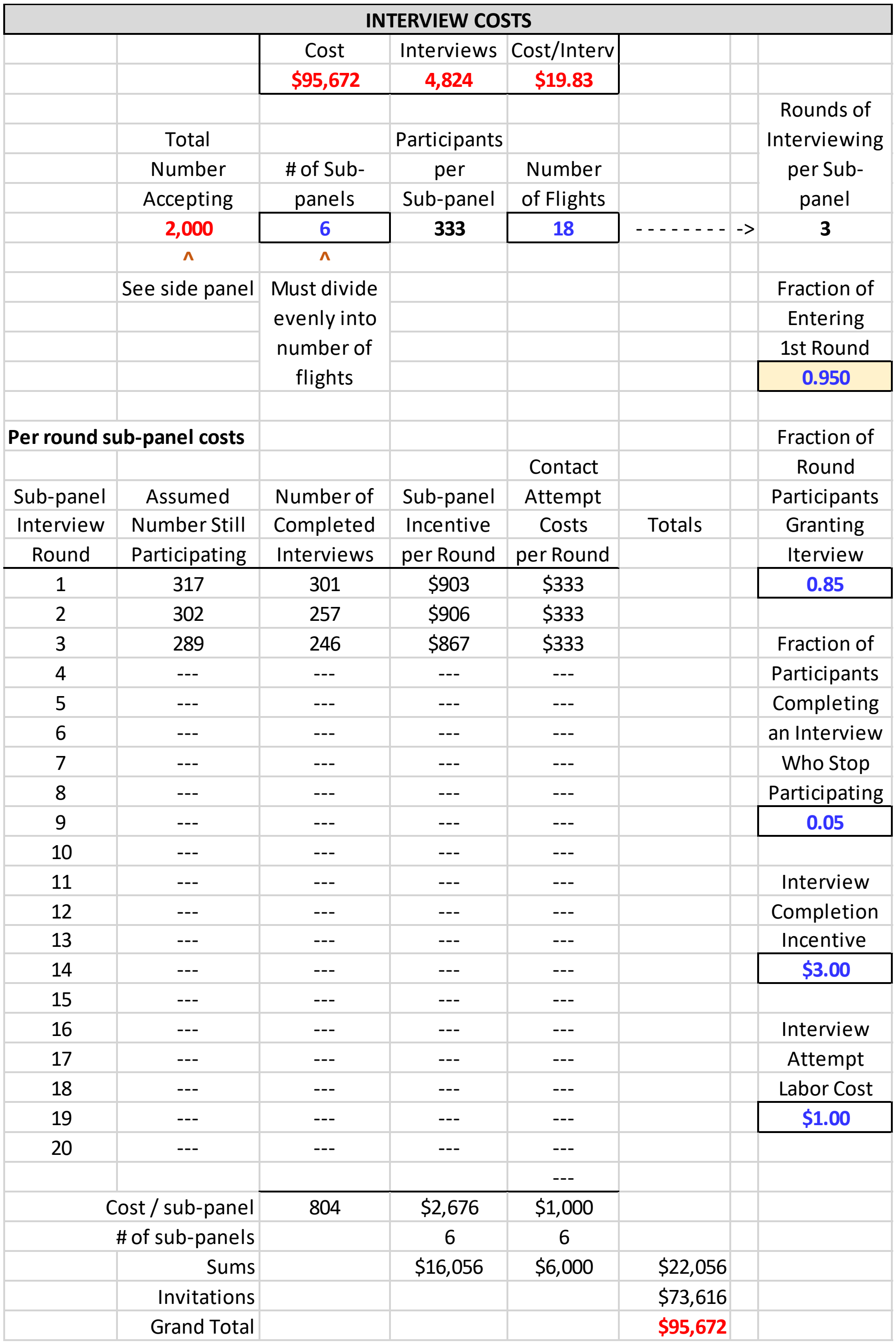
The first example resembles a small scale, controlled field experiment-like study of opinions solicited for a modest number of LBFD flight missions. Two thousand (2,000) panelists are desired in this study, per Figure 57.



SOLICITATION MAILING & INCENTIVE COSTS					
Desired Panel Size					
	2,000				
Initial Mailing					
Administration			$5,000	**Initial & 1st Followup**	
				Handling	$0.75
Invitations	Number	Unit Cost	SubTotal	Postage	$0.50
	3,459	$7.00	$24,216	Incentive	$5.00
		Total	**$29,216**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**865**	$33.78	Total	$7.00
1st Followup Mailing					
Administration			$3,000		
Invitations	Number	Unit Cost	SubTotal		
	2,595	$7.00	$18,162		
		Total	**$21,162**		
Accepting	Fraction	Number	per Acpt		
	0.25	**649**	$32.61		
2nd Followup Courier Mailing					
Administration			$3,000	**2nd Followup**	
				Handling	$1.00
Invitations	Number	Unit Cost	SubTotal	Shipping	$8.65
	1,946	$10.40	$20,238	Incentive	$0.00
		Total	**$23,238**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**486**	$47.81	Total	$10.40
Total Number Accepting		**2,000**			
		57.8%			
Total Solicitation Cost			**$73,616**		

Figure 57. Estimated costs associated with panel formation for a small scale study.

Given the assumptions of the spreadsheet, 3,036 households should be sent one or more solicitation letters seeking their participation in a study involving five rounds of interviews during the course of a limited-duration study. Each panelist is expected to grant three interviews during a set of 18 LBFD overflights, as shown in Figure 58. Interview costs of $22,056 are summed with the panel invitation costs of $73,616 to yield a cost estimate of $95,672 for panel formation and interviewing.



INTERVIEW COSTS

		Cost	Interviews	Cost/Interv			
		$95,672	4,824	$19.83			
							Rounds of
	Total		Participants				Interviewing
	Number	# of Sub-	per	Number			per Sub-
	Accepting	panels	Sub-panel	of Flights			panel
	2,000	6	333	18	- - - - - - - -	->	3
	^	^					
	See side panel	Must divide					Fraction of
		evenly into					Entering
		number of					1st Round
		flights					0.950
Per round sub-panel costs							Fraction of
				Contact			Round
Sub-panel	Assumed	Number of	Sub-panel	Attempt			Participants
Interview	Number Still	Completed	Incentive	Costs	Totals		Granting
Round	Participating	Interviews	per Round	per Round			Iterview
1	317	301	$903	$333			0.85
2	302	257	$906	$333			
3	289	246	$867	$333			Fraction of
4	---	---	---	---			Participants
5	---	---	---	---			Completing
6	---	---	---	---			an Interview
7	---	---	---	---			Who Stop
8	---	---	---	---			Participating
9	---	---	---	---			0.05
10	---	---	---	---			
11	---	---	---	---			Interview
12	---	---	---	---			Completion
13	---	---	---	---			Incentive
14	---	---	---	---			$3.00
15	---	---	---	---			
16	---	---	---	---			Interview
17	---	---	---	---			Attempt
18	---	---	---	---			Labor Cost
19	---	---	---	---			$1.00
20	---	---	---	---			

	Cost / sub-panel	804	$2,676	$1,000			
	# of sub-panels		6	6			
	Sums		$16,056	$6,000	$22,056		
	Invitations				$73,616		
	Grand Total				$95,672		

Figure 58. Costs associated with repeated interviews in small scale study.

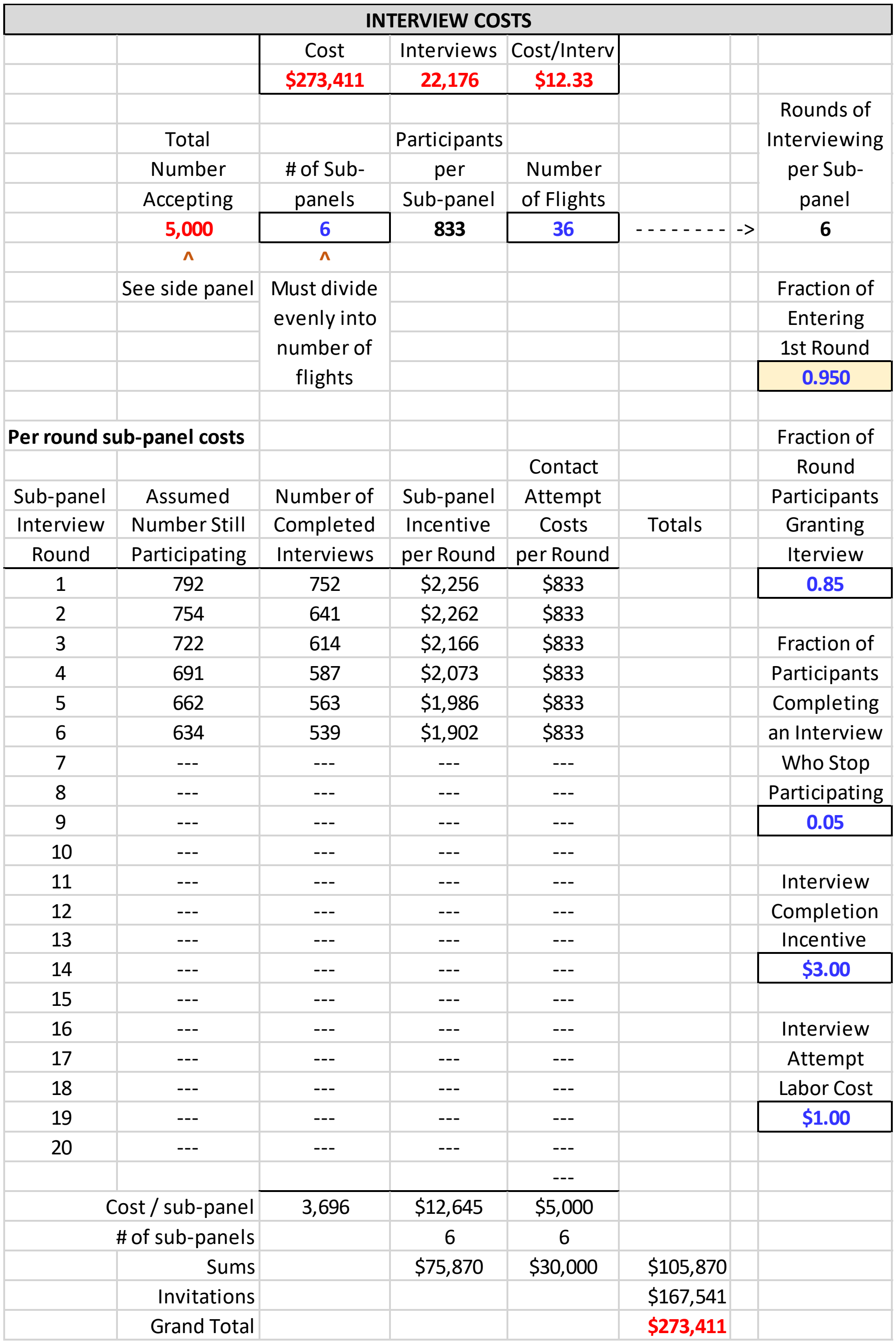
8.2 Example of an Interviewing Cost Estimate for an Intermediate Scale Panel Study

The second example is an intermediate scale study of opinions solicited to a modest number of LBFD flight missions. Five thousand (5,000) panelists are desired in this study, per Figure 59.

SOLICITATION MAILING & INCENTIVE COSTS					
Desired Panel Size					
	5,000				
Initial Mailing					
Administration			$5,000	**Initial & 1st Followup**	
				Handling	$0.75
Invitations	Number	Unit Cost	SubTotal	Postage	$0.50
	8,649	$7.00	$60,541	Incentive	$5.00
		Total	**$65,541**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**2,162**	$30.31	Total	$7.00
1st Followup Mailing					
Administration			$3,000		
Invitations	Number	Unit Cost	SubTotal		
	6,486	$7.00	$45,405		
		Total	**$48,405**		
Accepting	Fraction	Number	per Acpt		
	0.25	**1,622**	$29.84		
2nd Followup Courier Mailing					
Administration			$3,000	**2nd Followup**	
				Handling	$1.00
Invitations	Number	Unit Cost	SubTotal	Shipping	$8.65
	4,865	$10.40	$50,595	Incentive	$0.00
		Total	**$53,595**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**1,216**	$44.07	Total	$10.40
Total Number Accepting		**5,000**			
		57.8%			
Total Solicitation Cost			**$167,541**		

Figure 59. Estimated costs associated with panel formation for an intermediate scale study.

Given the assumptions of the spreadsheet, 8,649 households should be sent one or more solicitation letters seeking their participation in a study involving five rounds of interviews during the course of a limited-duration study. Each panelist is expected to grant six interviews during a set of 36 LBFD overflights, as shown in Figure 60. Interview costs of $105,870 are summed with the panel invitation costs of $167,541 for a total of $273,411 in data collection costs.



INTERVIEW COSTS

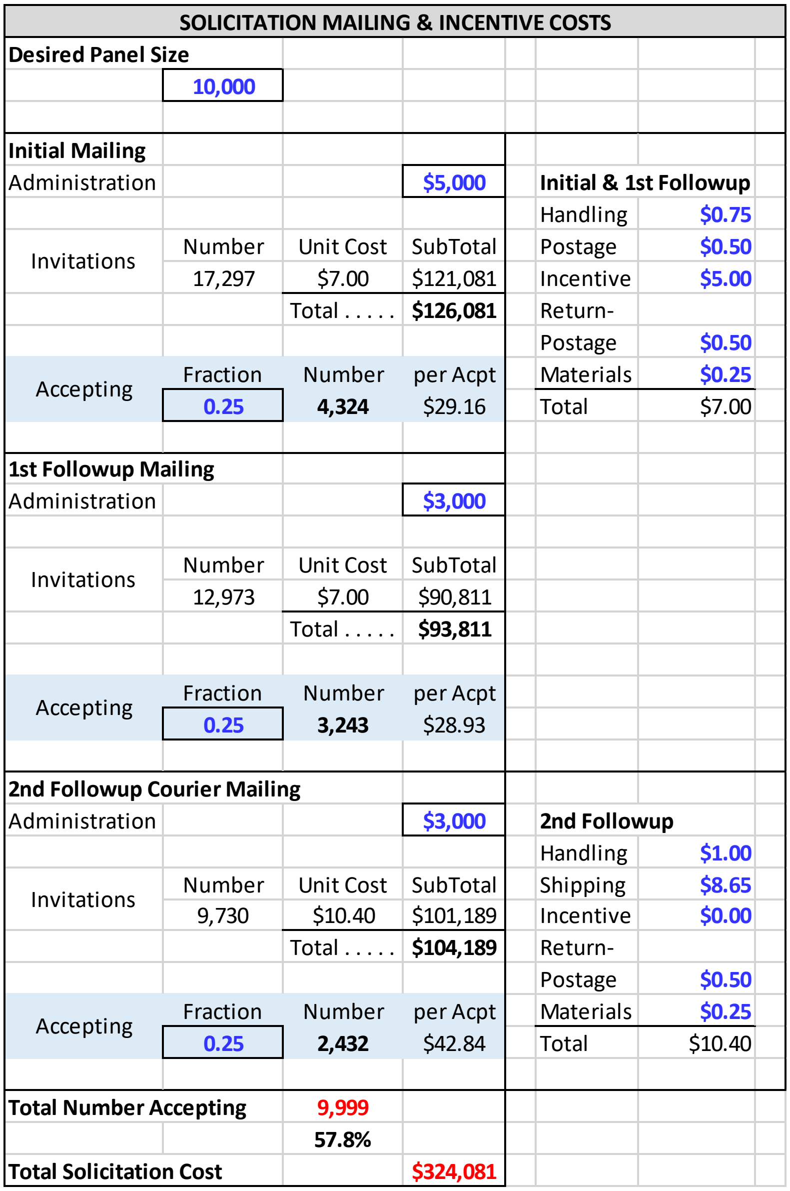
		Cost	Interviews	Cost/Interv			
		$273,411	22,176	$12.33			
	Total Number Accepting	# of Sub-panels	Participants per Sub-panel	Number of Flights			Rounds of Interviewing per Sub-panel
	5,000	6	833	36	- - - - - - - -	->	6
	^	^					
	See side panel	Must divide evenly into number of flights					Fraction of Entering 1st Round
							0.950
Per round sub-panel costs							Fraction of Round Participants Granting Iterview
Sub-panel Interview Round	Assumed Number Still Participating	Number of Completed Interviews	Sub-panel Incentive per Round	Contact Attempt Costs per Round	Totals		
1	792	752	$2,256	$833			0.85
2	754	641	$2,262	$833			
3	722	614	$2,166	$833			Fraction of Participants Completing an Interview Who Stop Participating
4	691	587	$2,073	$833			
5	662	563	$1,986	$833			
6	634	539	$1,902	$833			
7	---	---	---	---			
8	---	---	---	---			
9	---	---	---	---			0.05
10	---	---	---	---			
11	---	---	---	---			Interview Completion Incentive
12	---	---	---	---			
13	---	---	---	---			
14	---	---	---	---			$3.00
15	---	---	---	---			
16	---	---	---	---			Interview Attempt Labor Cost
17	---	---	---	---			
18	---	---	---	---			
19	---	---	---	---			$1.00
20	---	---	---	---			

	Cost / sub-panel	3,696	$12,645	$5,000			
	# of sub-panels		6	6			
	Sums		$75,870	$30,000	$105,870		
	Invitations				$167,541		
	Grand Total				$273,411		

Figure 60. Estimated costs associated with repeated interviews in an intermediate scale study.

8.3 Example of an Interviewing Cost Estimate for a Large Scale Panel Study

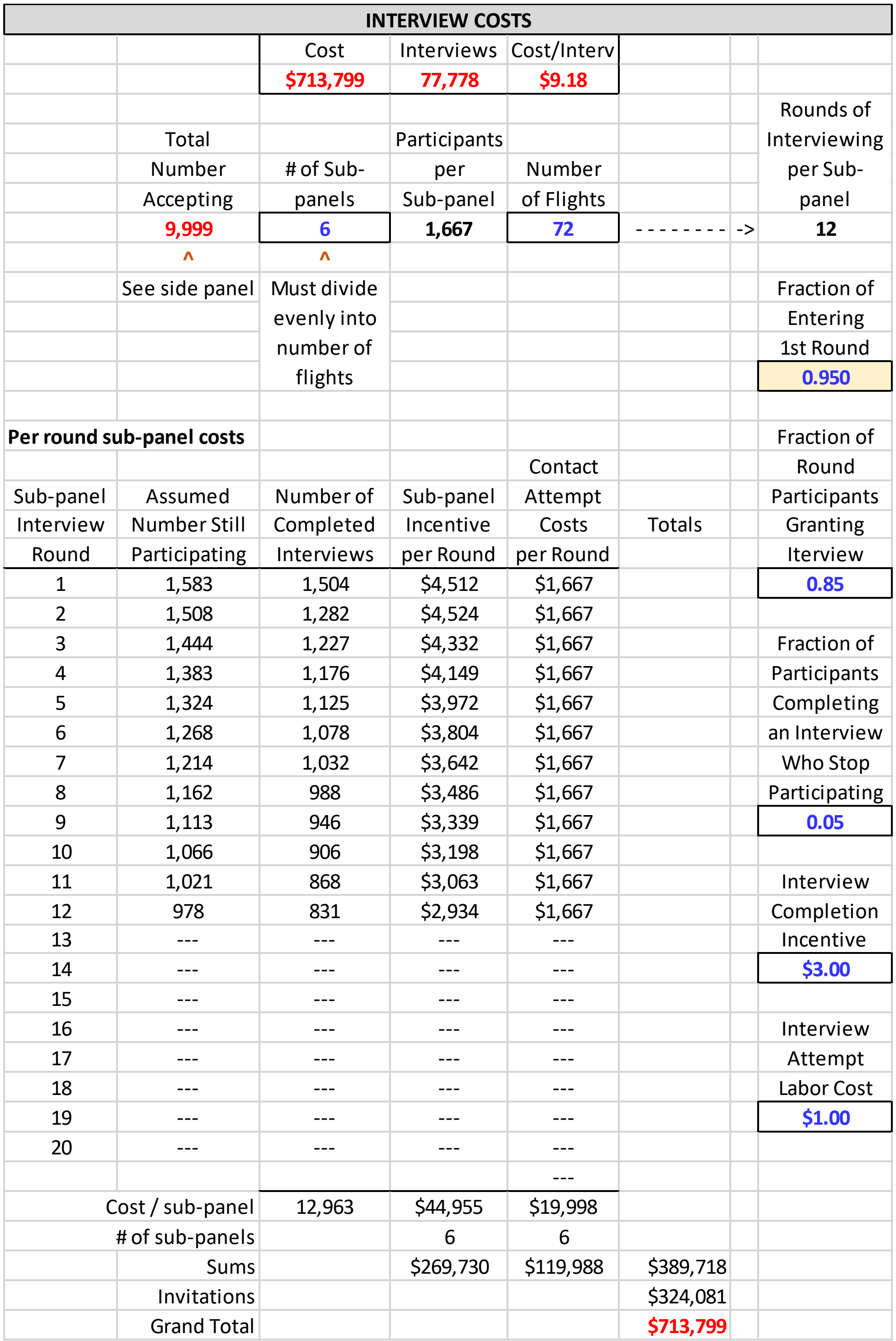
The third example is a large scale study of opinions solicited to a modest number of LBFD flight missions. Ten thousand (10,000) panelists are desired in this study, per Figure 61.



SOLICITATION MAILING & INCENTIVE COSTS					
Desired Panel Size					
	10,000				
Initial Mailing					
Administration			$5,000	**Initial & 1st Followup**	
				Handling	$0.75
Invitations	Number	Unit Cost	SubTotal	Postage	$0.50
	17,297	$7.00	$121,081	Incentive	$5.00
		Total	**$126,081**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**4,324**	$29.16	Total	$7.00
1st Followup Mailing					
Administration			$3,000		
Invitations	Number	Unit Cost	SubTotal		
	12,973	$7.00	$90,811		
		Total	**$93,811**		
Accepting	Fraction	Number	per Acpt		
	0.25	**3,243**	$28.93		
2nd Followup Courier Mailing					
Administration			$3,000	**2nd Followup**	
				Handling	$1.00
Invitations	Number	Unit Cost	SubTotal	Shipping	$8.65
	9,730	$10.40	$101,189	Incentive	$0.00
		Total	**$104,189**	Return-	
				Postage	$0.50
Accepting	Fraction	Number	per Acpt	Materials	$0.25
	0.25	**2,432**	$42.84	Total	$10.40
Total Number Accepting		9,999			
		57.8%			
Total Solicitation Cost			$324,081		

Figure 61. Estimated costs associated with forming a panel for a large scale study.

Given the assumptions of the spreadsheet, 17,297 households should be sent one or more solicitation letters seeking their participation in a study involving five rounds of interviews during the course of a limited-duration study. Each panelist is expected to grant six interviews during a set of 72 LBFD overflights, as shown in Figure 62. Interview costs of $389,718 are summed with the panel invitation costs of $324,081 for a total of $713,799 in data collection costs.



INTERVIEW COSTS

		Cost	Interviews	Cost/Interv			
		$713,799	77,778	$9.18			
							Rounds of
	Total		Participants				Interviewing
	Number	# of Sub-	per	Number			per Sub-
	Accepting	panels	Sub-panel	of Flights			panel
	9,999	6	**1,667**	72	- - - - - - - -	->	**12**
	^	^					
	See side panel	Must divide					Fraction of
		evenly into					Entering
		number of					1st Round
		flights					0.950
Per round sub-panel costs							Fraction of
				Contact			Round
Sub-panel	Assumed	Number of	Sub-panel	Attempt			Participants
Interview	Number Still	Completed	Incentive	Costs	Totals		Granting
Round	Participating	Interviews	per Round	per Round			Itervew
1	1,583	1,504	$4,512	$1,667			0.85
2	1,508	1,282	$4,524	$1,667			
3	1,444	1,227	$4,332	$1,667			Fraction of
4	1,383	1,176	$4,149	$1,667			Participants
5	1,324	1,125	$3,972	$1,667			Completing
6	1,268	1,078	$3,804	$1,667			an Interview
7	1,214	1,032	$3,642	$1,667			Who Stop
8	1,162	988	$3,486	$1,667			Participating
9	1,113	946	$3,339	$1,667			0.05
10	1,066	906	$3,198	$1,667			
11	1,021	868	$3,063	$1,667			Interview
12	978	831	$2,934	$1,667			Completion
13	---	---	---	---			Incentive
14	---	---	---	---			$3.00
15	---	---	---	---			
16	---	---	---	---			Interview
17	---	---	---	---			Attempt
18	---	---	---	---			Labor Cost
19	---	---	---	---			$1.00
20	---	---	---	---			

	Cost / sub-panel	12,963	$44,955	$19,998			
	# of sub-panels		6	6			
	Sums		$269,730	$119,988	$389,718		
	Invitations				$324,081		
	Grand Total				$713,799		

Figure 62. Estimated costs associated with repeated interviews in large scale study.

9 APPENDIX B. EXAMPLES OF SINGLE-EVENT ANALYSES OF PROMPT RESPONSE DATA

This Appendix demonstrates several CTL-like (Fidell *et al.*, 2011) analyses of low-boom, prompt dose-response information. Data from an Edwards Air Force Base study (Fidell *et al.*, 2012) are used to demonstrate how future data sets can support parsimonious single-event dose-response analyses. Whereas regression analysis requires estimates of two parameters – slope and intercept – of a fitting function, CTL analyses assume a fixed slope (that of the growth of loudness with sound level), so that only the position of the fitting function on the abscissa requires estimation.

CTL analysis was developed to characterize community response as a function of DNL, a 24-hour, cumulative noise exposure metric. The current application requires assessment of community response to individual, impulsive noise events, however. A metric of overpressure or impulsive Perceived Level of individual low-amplitude booms is therefore substituted for cumulative, A-weighted sound level (DNL) for which the analysis procedure was originally developed. For purposes of this Appendix, the single event tolerance metric is represented symbolically as either $L_{\text{ct (single event OP)}}$ or $L_{\text{ct (single event PL)}}$ for use with overpressure of Perceived Level as the response predictor, respectively.

The resulting single event tolerance levels cannot be directly compared with those from prior studies using Day-Night Average Sound Level (DNL) to characterize cumulative noise exposure, since they differ by at least 49.4 dB (10^*Log_{10} [86,400 seconds per day].

9.1 Single-event CTL Analysis in Units of Overpressure

Figure 63 shows the high annoyance dose-response data from 87 booms to which test participants were exposed over a two-week period in the Fidell *et al.* (2012) study[53]. (The full data set is tabulated in Table 22 at the end of this appendix.) The abscissa plots the maximum overpressure (20 Log [maximum OP]) of the boom, while the ordinate plots the percentage of at-home respondents who described themselves as highly annoyed by them. The solid black line is the maximum likelihood fit (the value of *A*) to the data points based on the following equations:

$$\%HA = 100 \bullet e^{(-A/m)} \qquad \text{Equation 5}$$

where:

$\%HA$ = percentage of respondents highly annoyed

A = community sensitivity factor

$$m = 10^{(20 \text{ Log } [OP] / 10) \bullet 0.3} \qquad \text{Equation 6}$$

or,

---

[53] Residents were exposed to a total of 106 boom events, of which only 87 were analyzable.

$$m = OP^{2 \cdot (0.3)} = OP^{0.6} \qquad \text{Equation 7}$$

The dotted blue lines show the 95% confidence interval on the derived $L_{ct\ (\text{single event OP})}$ value. That is, they show the uncertainty in the *lateral* position of the curve. Since the slope of the function in $\%HA = e^{(-A/m)}$ is fixed by the assumption of the exponent factor of 0.3 in Equation 6, the maximum likelihood is determined by a single parameter, A. $L_{ct\ (\text{single event OP})}$ is related to A by:

$$L_{ct\ (\text{single event OP})} = 33.3333 \cdot \text{Log}\,[A] + 5.3\ \ (\text{dB}) \qquad \text{Equation 8}$$

$L_{ct\ (\text{single event OP})}$ anchors the CTL curve to the abscissa at the point at which it crosses the 50 percent highly annoyed point on the ordinate. This definition, which affects only a constant, allows community tolerances for sonic booms to be compared among interviewing areas across exposure levels. Such *differences* would be the same for any other chosen percent highly annoyed anchor point (*e.g.*, 25 percent, or 12 percent). For the data shown in Figure 63, $L_{ct\ (\text{single event OP})}$ = 18.8 dB, with a 95% confidence interval of -1.01 to +1.04 dB.

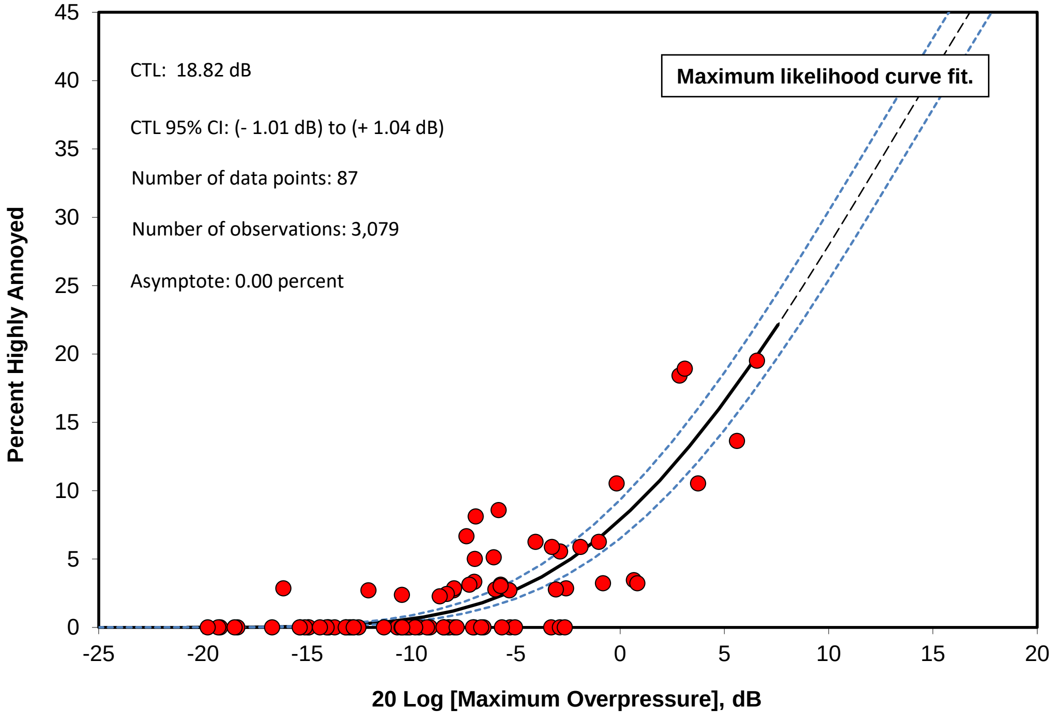


Figure 63. Prevalence of high annoyance vs overpressure for single event data of Fidell *et al.* (2012).

The same analysis permits similar, CTL-like estimates for lesser degrees of annoyance as well. For example, the analysis can be applied to percentages of respondents who describe themselves as moderately or more (moderately, very, or extremely) annoyed. Such an analysis is shown in Figure 64. The data points in this figure have the same abscissa values as those in Figure 63, but the prevalence of annoyance values are greater. Hence, the fitting curve is shifted to the left of that shown in Figure 63. For this characterization of the data, $L_{\text{ct (single event OP)}}$ = 13.9 dB, and the 95% confidence is -0.81 to +0.81 dB.

Likewise, the analysis can also be applied to percentages of respondents who describe themselves as slightly or more (slightly, moderately, very, or extremely) annoyed by individual sonic booms. Such an analysis is shown in Figure 65. Once again, the data points in this figure have the same abscissa values as those in Figure 63, but the annoyance prevalence rates are greater, resulting in a further left-shifting of the fitting curve from that in the prior figure. For this interpretation of the data, $L_{\text{ct (single event OP)}}$ = 9.7 dB, and the 95% confidence is -0.71 to +0.71 dB.

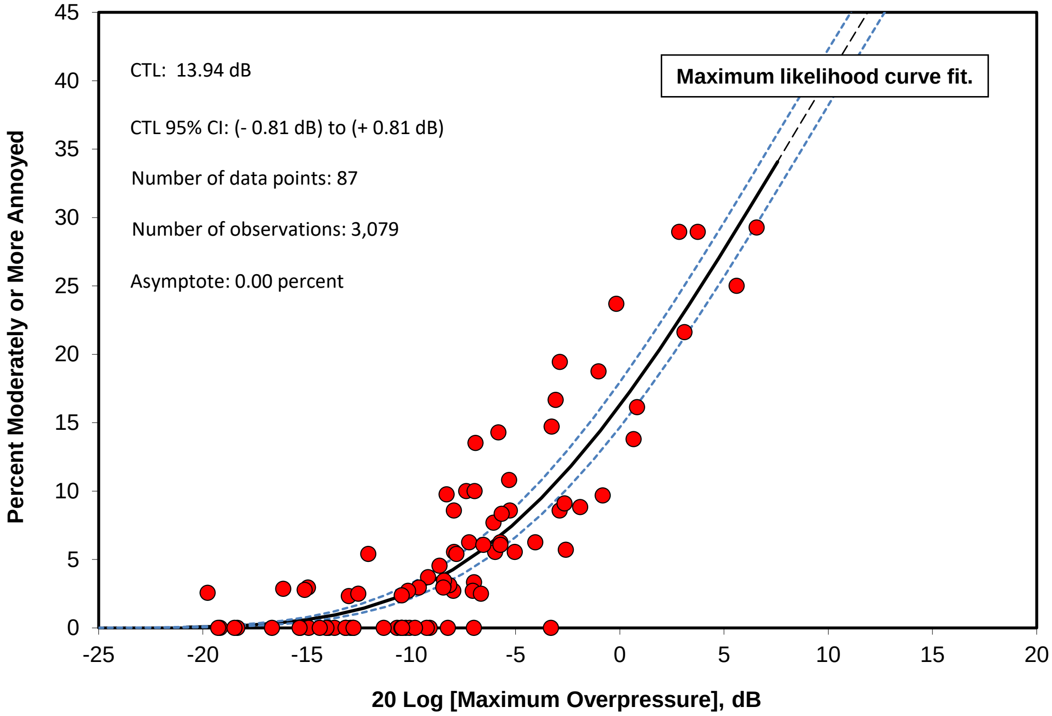


Figure 64. Prevalence of moderate or greater annoyance vs overpressure for single event data of Fidell *et al.* (2012).

The confidence intervals in these three figures become smaller when greater numbers of data points lie on the steeper portions of the curve. This occurs because greater numbers of data points influence the lateral position of the curve. (Points lying along the abscissa have provide *no* information about the lateral position of the fitting function).

The number of observations associated with each data point influence the most likely value of the curve fit parameter(s) in a maximum likelihood analysis. For completeness, the number of individuals responding to each boom is shown in Figure 66. In this case, the numbers of respondents are fairly consistent between exposures. However, those data points with greater numbers of respondents have more influence on the curve position than do those with lesser numbers. If only a few participants had responded to a single boom, then the data point associated with that boom would have had much less influence on the fitting function than any of the others.

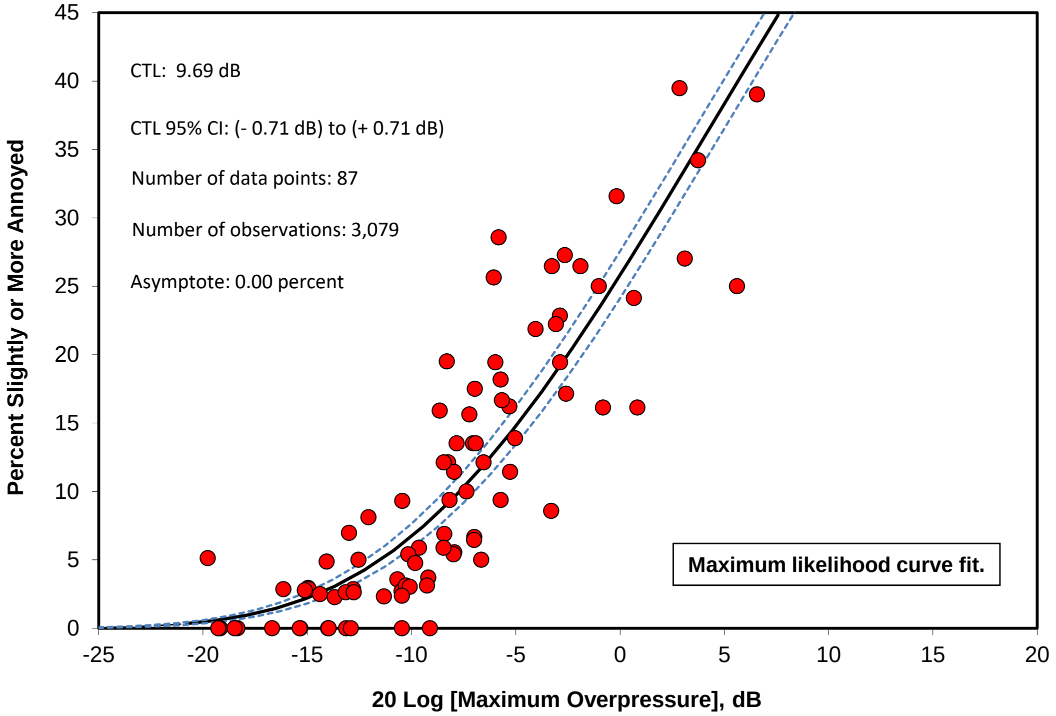


Figure 65. Prevalence of slight or greater annoyance vs overpressure for single event data of Fidell *et al.* (2012).

Table 18 summarizes the parameters associated with each of the above three curve fits. The first column shows the cumulative annoyance category. The second column tabulates the single event tolerance level (in dB). The third column shows the 95 percent confidence interval on the derived single event tolerance level. The fourth column shows the difference in tolerance level from that of the annoyance category directly above. This column shows that the lateral difference between annoyance categories is about 4 to 5 dB. The last column shows the approximate confidence intervals on the *differences* shown in the fourth column.

Table 18. Comparison of community sensitivity values for three annoyance categories in units of overpressure.

Level of Annoyance	L_{CT} (dB)	95% CI (dB)	Δ CTL (dB)	Δ 95% CI (dB)
Slightly+	18.82	0.71		
Mod+	13.94	0.81	4.88	1.09
Highly	9.69	1.03	4.25	1.32

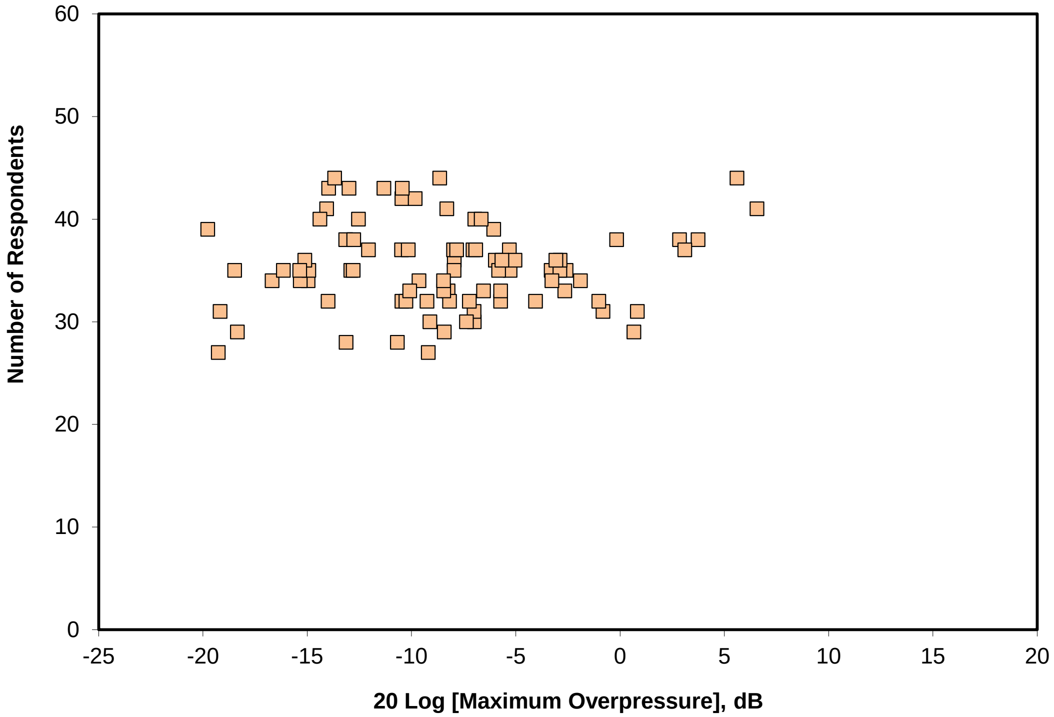


Figure 66. Numbers of respondents vs overpressure for single event data of Fidell *et al.* (2012).

9.2 Single-event CTL Analysis in Units of Perceived Level

The method described in the prior subsection was also undertaken using perceived level as a predictor of annoyance prevalence rates. Figure 67 shows the high annoyance dose-response data from 87 boom events to which local residents were exposed over a two-week period. The abscissa plots the Perceived Level of the boom, while the ordinate plots the percentage of at-home respondents who rated the boom as highly annoying. The solid black line shows the maximum likelihood fit to the data points using $\%HA = e^{(-A/m)}$ as a fitting function.

$$L_{ct\ (\text{single event PL})} = 33.3333 \cdot \text{Log}\ [A] + 5.3\ \ (\text{dB})$$

Equation 9

where:

$$m = 10^{(PL/10\ \cdot\ 0.3)}$$

Equation 10

As in Figure 67, Figure 68 plots the percentages of respondents who describe themselves as moderately or more (moderately, very, or extremely) annoyed, and Figure 69 plots percentages of respondents who described themselves as slightly or more (slightly, moderately, very, or extremely) annoyed.

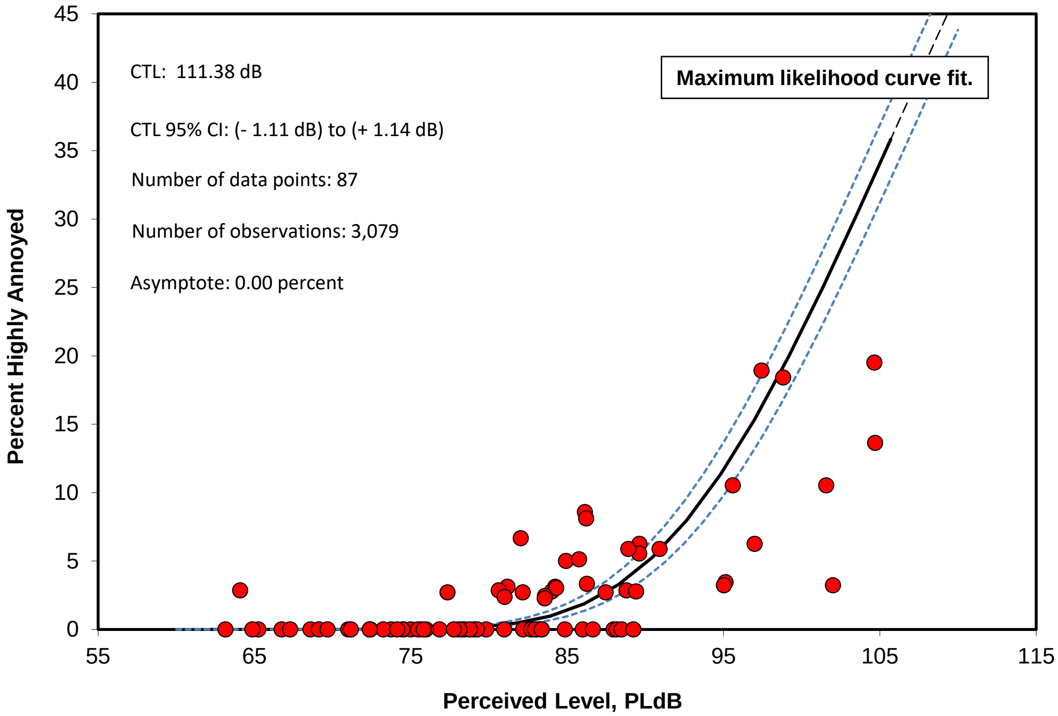


Figure 67. Prevalence of high annoyance vs. Perceived Level for single event data of Fidell *et al.* (2012).

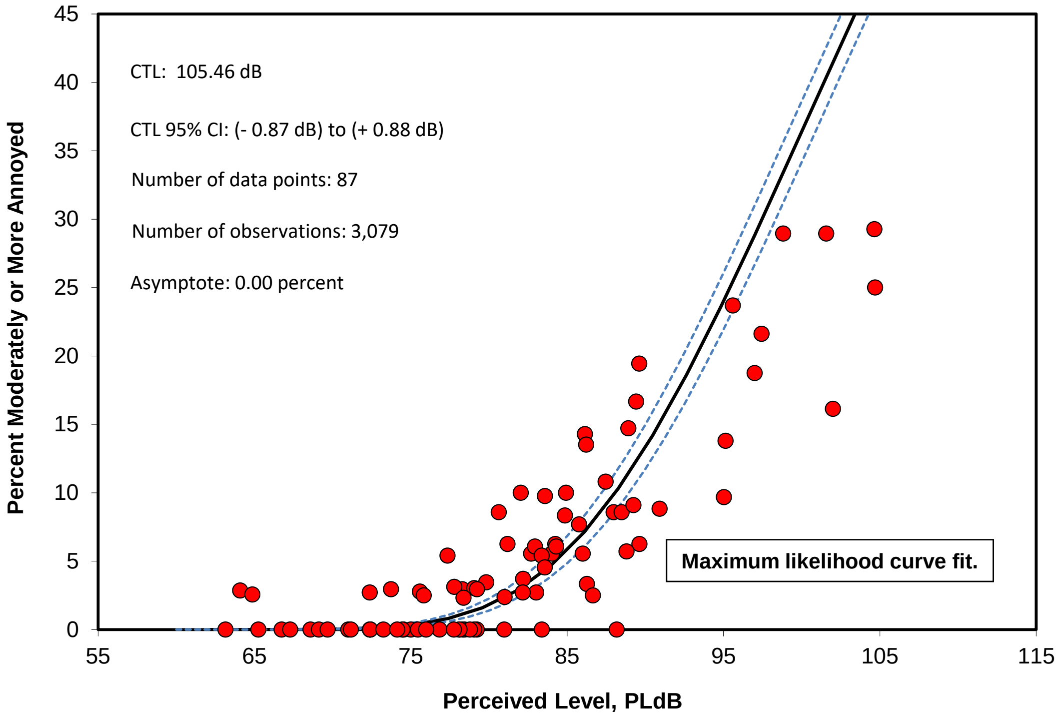


Figure 68. Prevalence of moderate or greater annoyance vs. Perceived Level for single event data of Fidell *et al.* (2012).

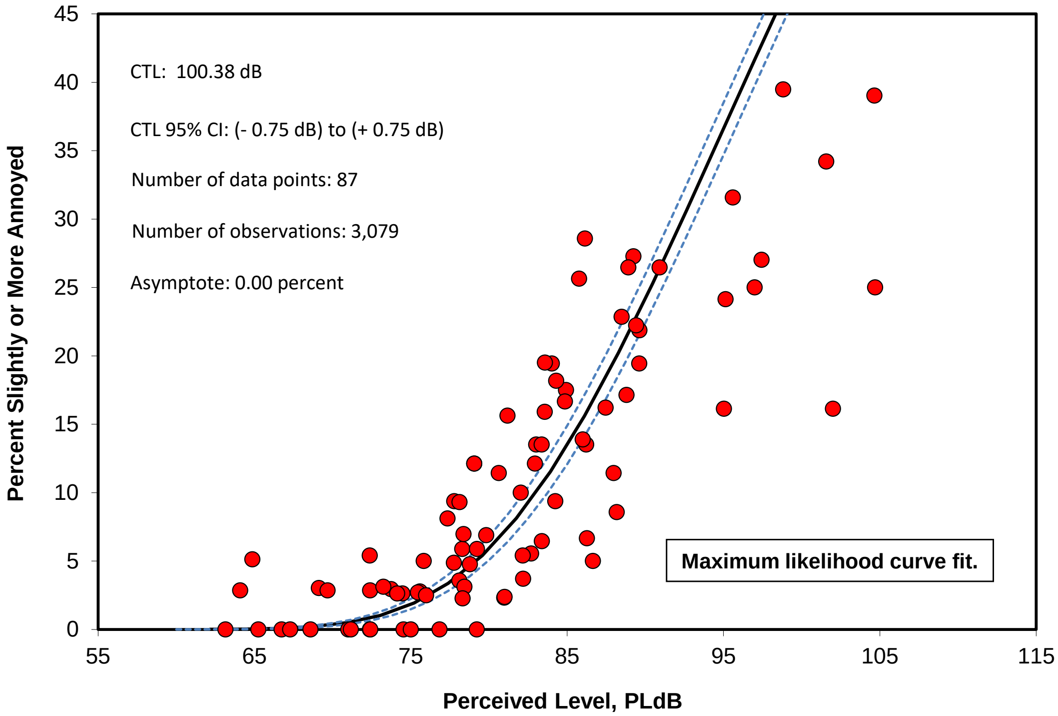


Figure 69. Prevalence of slight or greater annoyance vs. Perceived Level for single event data of Fidell *et al.* (2012).

Evident from Figure 67, Figure 68 and Figure 69 is the observation that the higher exposure level data points in the Perceived Level figures do not align as well with the curve fit as they did with the overpressure metric shown in Figure 63 through Figure 65. (It may be that the spectral content of the high-exposure booms affected the PL differently from the overpressure.) However, the data points with levels of less than 91 PLdB *do* follow the shape of the curve, suggesting that any data collected as part of the LBFD program might be amenable to analyses in which annoyance growth rates are assumed to be fixed (the exponent of 0.3) across all conditions of noticeability through high annoyance. This point is discussed further in §9.4 below.

Figure 70 shows the number of respondents associated with each data point in the preceding three figures. The Perceived Level is plotted on the abscissa and the number of at-home respondents is plotted on the ordinate.

Table 19 tabulates the parameters associated with each of the above three curve fits. The first column shows the cumulative annoyance category. The second column tabulates the single event tolerance level (in PLdB). The third column shows the 95 percent confidence interval on the derived single event tolerance level. The fourth column shows the difference in tolerance level from that of the annoyance category directly above. This column shows that the difference between annoyance categories is about 5 to 6 PLdB. The last column shows the approximate 95 percent confidence intervals on the differences shown in the fourth column.

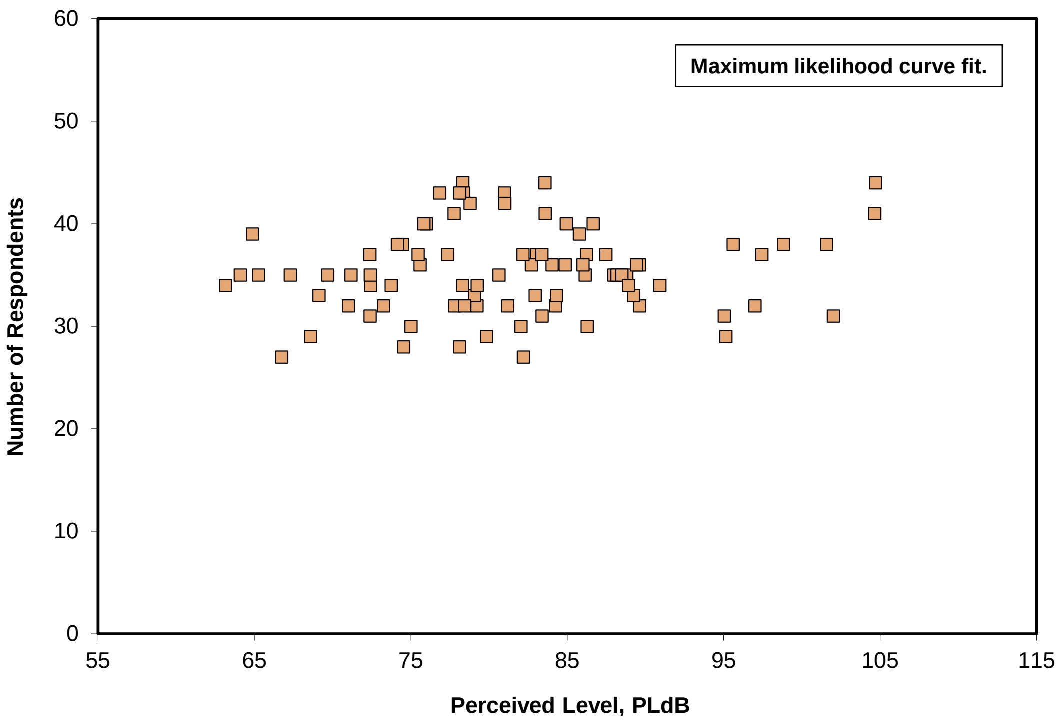


Figure 70. Numbers of respondents vs Perceived Level for single event data of Fidell *et al.* (2012).

Table 19. Comparison of community sensitivity values for three annoyance categories – Perceived Level.

Level of Annoyance	L_{ct} (dB)	95% CI (dB)	Δ CTL (dB)	Δ 95% CI (dB)
Slightly+	100.38	0.75		
Mod+	105.46	0.88	5.08	1.16
Highly	111.38	1.12	5.92	1.42

9.3 Relationship between Perceived Level and 20 Log [Overpressure]

Figure 71 plots the Perceived Level and 20 Log [Maximum Overpressure] pairs for each of the 87 data points used in the analyses as an aid to interpreting the relationships between the reported data sets. The slope is nearly 10 PLdB in Perceived Level for a 5 dB change in 20 Log [Overpressure].

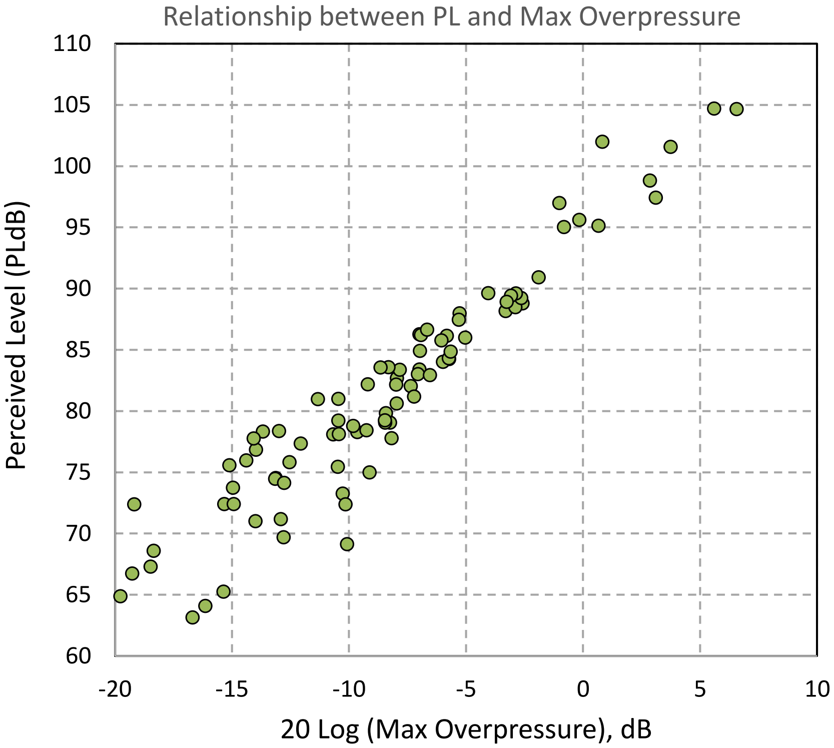


Figure 71. Observed relationship between Perceived Level and 20 Log [Maximum Overpressure].

9.4 Additional Perceived Level Analyses

The observation in §9.2 regarding the CTL curve fit to the higher Perceived Level data points (91 PLdB and greater) is further elaborated in this subsection. In this analysis, data points with Perceived Levels of 91 PLdB and above were removed from the data set to illustrate the type of curve fit more likely to be observed from LBFD boom levels. Figure 72 shows the "high annoyance" dose-response data from 77 boom events of less than 91 PLdB to which local residents were exposed over a two-week period. The abscissa plots Perceived Levels of booms, while the ordinate plots percentages of at-home respondents who rated the boom as highly annoying. The solid black line shows the maximum likelihood fit to the data points using $\%HA = e^{(-A/m)}$. The fitting function plotted in Figure 72 provides a good representation of the central tendency of the data.

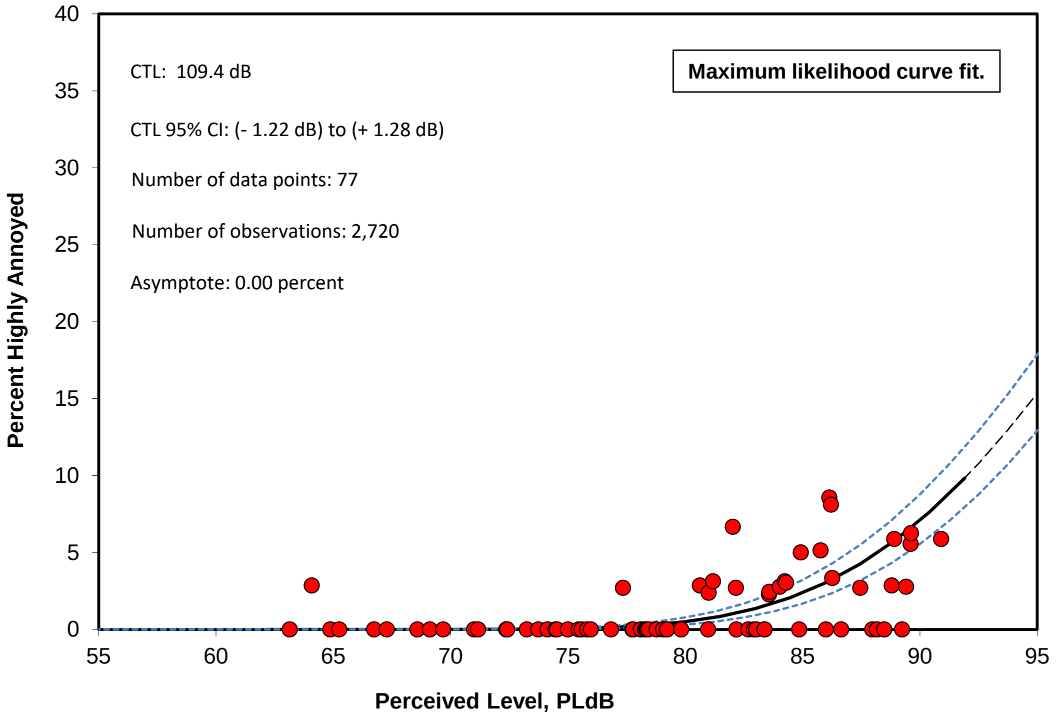


Figure 72. Prevalence of high annoyance vs. Perceived Level (<91 PLdB) for single event data of Fidell *et al.* (2012).

Likewise, the CTL fitting functions for lower levels of annoyance using the truncated data set (Figure 73 and Figure 74) also provide good representations of the data. Confidence intervals on the lateral positions of the curves in Figure 72 through Figure 74 are -1.22 to +1.28, -0.94 to +0.95 and -0.80 to +0.79 PLdB, respectively.

The numbers of at-home respondents for each of the data points are shown in Figure 75. The curve fit parameters for the three annoyance levels are provided in Table 20. The lateral distance between the three annoyance level curves is on the order of 5.0 to 5.1 PLdB.

Table 20 Comparison of community tolerance values for three annoyance categories – Perceived Level <91 PLdB.

Level of Annoyance	CTL_{SE} (dB)	95% CI (dB)	Δ CTL (dB)	Δ 95% CI (dB)
Slightly+	99.29	0.80	--	--
Mod+	104.29	0.94	-5.00	1.23
Highly	109.40	1.25	-5.11	1.56

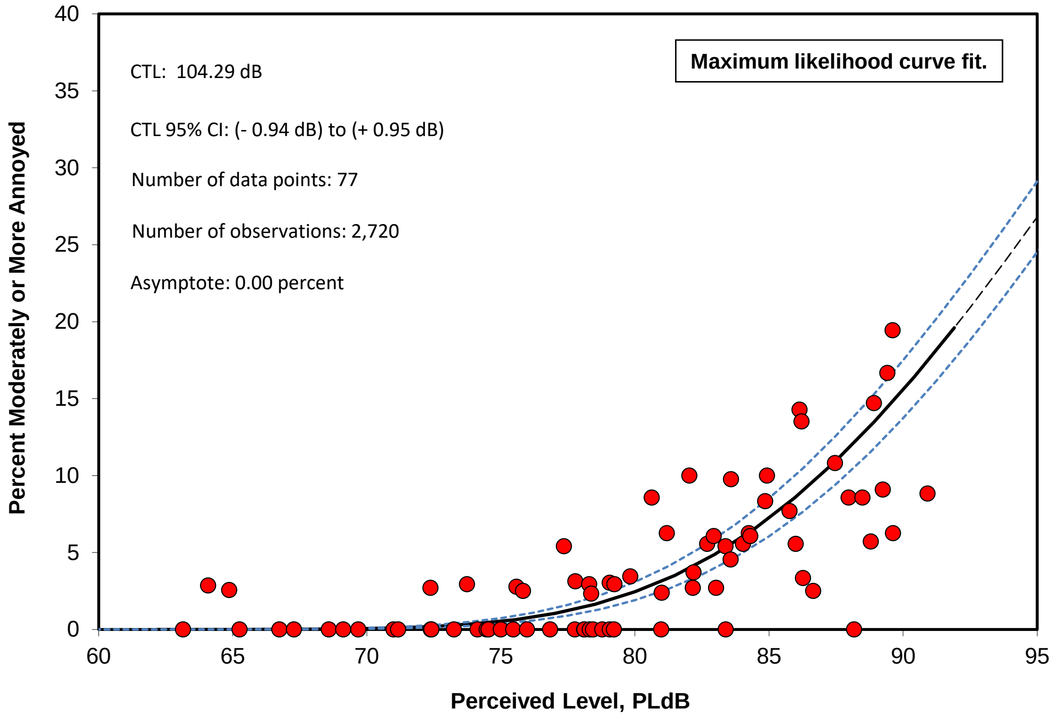


Figure 73. Prevalence of moderate or greater annoyance vs. Perceived Level (<91 PLdB) for single event data of Fidell *et al.* (2012).

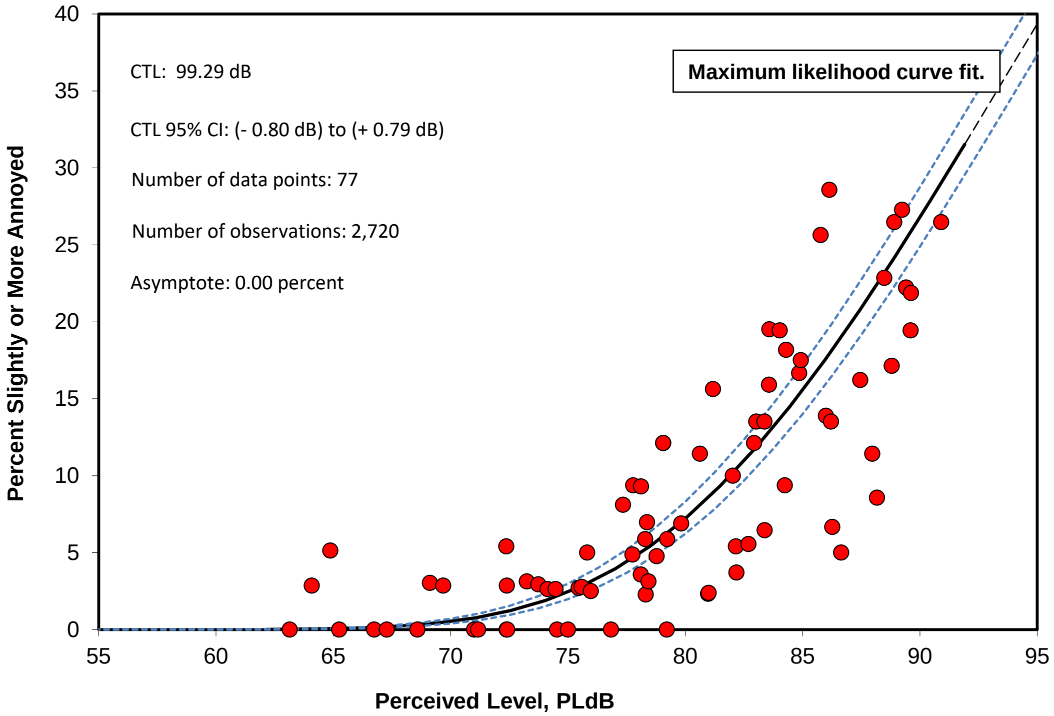


Figure 74. Prevalence of slight or greater annoyance vs. Perceived Level (<91 PLdB) for single event data of Fidell *et al.* (2012).

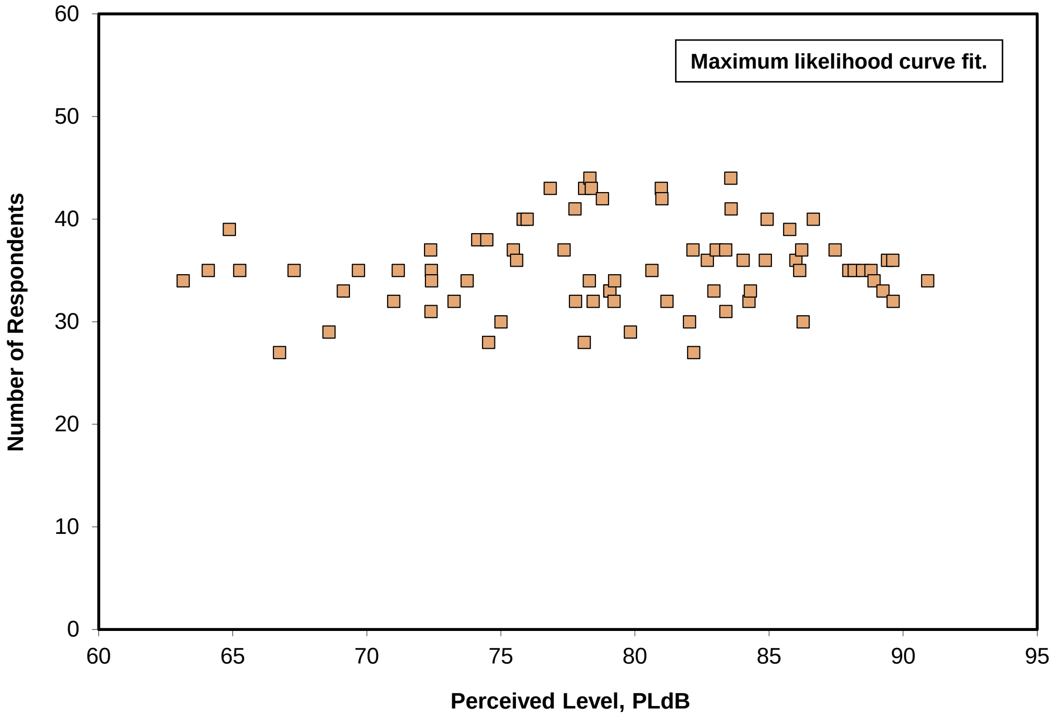


Figure 75. Numbers of respondents vs Perceived Level (<91 PLdB) for single event data of Fidell *et al.* (2012).

9.5 Relationship between Noticeability and Annoyance

A further question that arises is how well the CTL functional form might fit the noticeability dose-response data set. Specifically, can the CTL function be used to estimate the difference in sound level between that required for noticeability and that required for any of the annoyance categories? Figure 76 shows the maximum likelihood fit to the functional forms of Equation 5 and Equation 7 superimposed on the noticeability *versus* 20 Log (Maximum Overpressure) data set obtained at Edwards AFB. Similarly, Figure 77 plots noticeability data against Perceived Level. The PL transform for use in $\%HA = e^{(-A/m)}$ is shown in Equation 5 and Equation 9. The fits in both figures appear quite good for exposure levels below 1 pound per square foot (0 dB on the abscissa in Figure 76, and 91 PLdB in Figure 77), similar to the findings in §9.2.

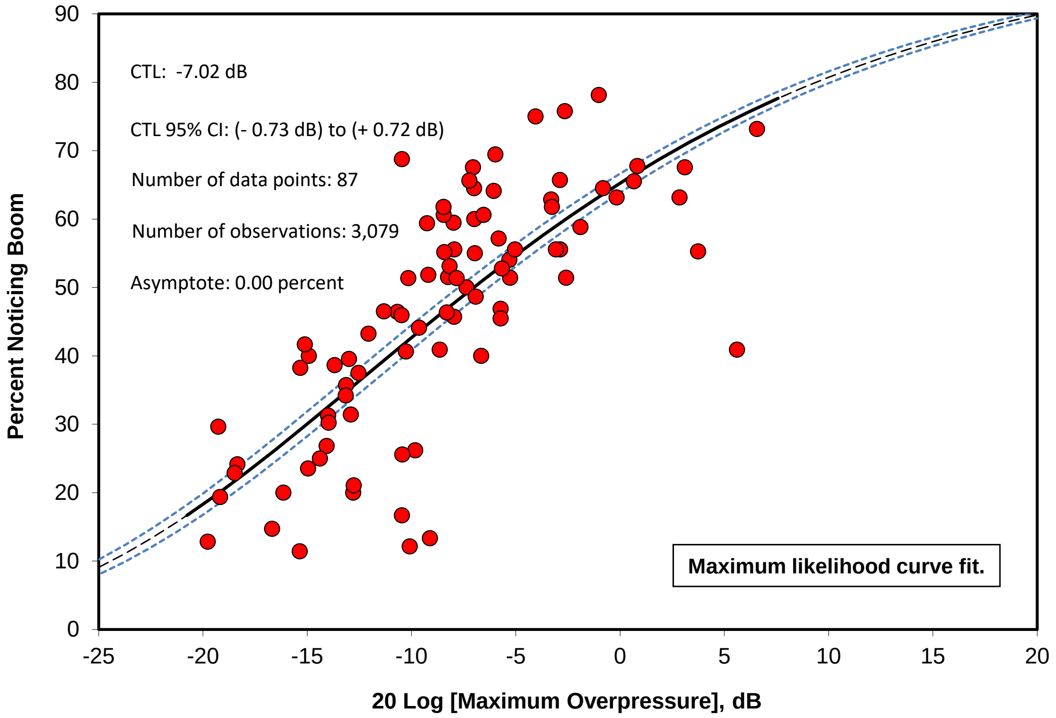


Figure 76. Noticeability *versus* 20 Log (Maximum Overpressure) for single event data of Fidell *et al.* (2012).

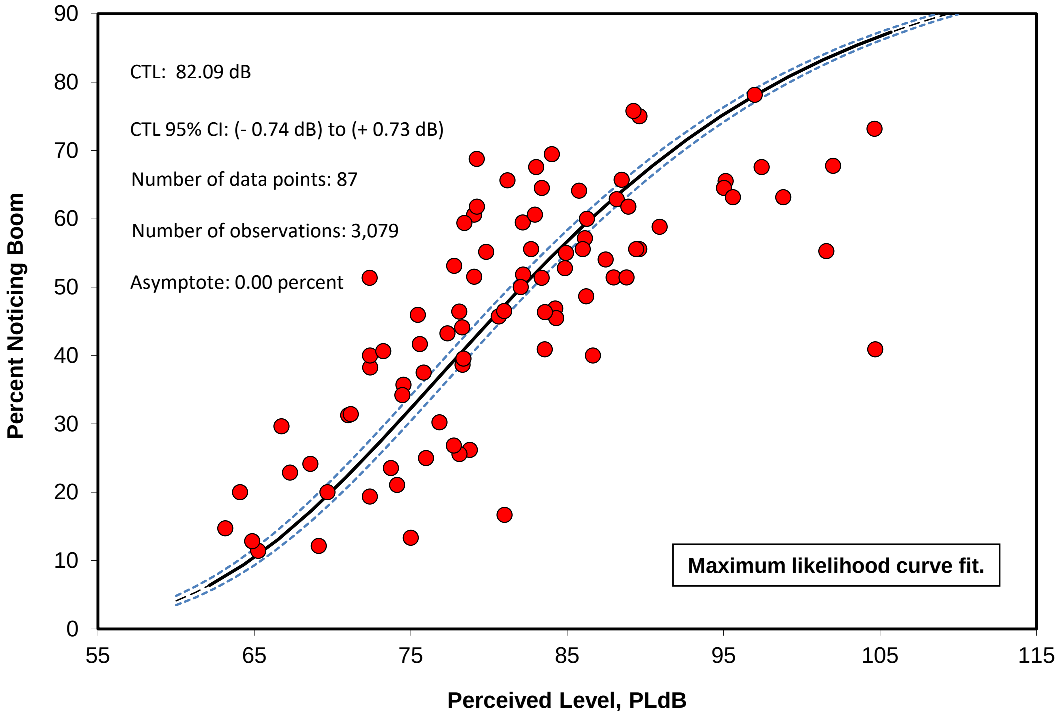


Figure 77. Noticeability *versus* Perceived Level for single event data of Fidell *et al.* (2012).

These higher level data points were removed from the Perceived Level data set as described in §9.4 to form a subset from which the curve fit was recalculated. The revised graphic for that subset is shown in Figure 78. The CTL functional form for this subset provides an excellent fit to the entire data subset which more than covers the Perceived Level range of interest for the LBFD exercises.

Considerable scatter about the fitting function is apparent in Figure 78. Nonetheless, the 95% confidence interval on the single event tolerance level is acceptably small for current purposes, due to the relatively large number of data points, their position well above the abscissa, a range in dose of almost 30 PLdB, and the large number of respondents across all data points (2,720).

It is likely that boom levels in future LBFD field testing will range from approximately 75 to 85 PLdB. An estimate of the effect on prediction of community response to booms of such a limited exposure range may be obtained by analyzing just those data points from the full Edwards data set within this range. Figure 79 shows the outcome of this analysis. For this truncated data subset, the width of the 95% confidence interval increases to approximately ± 1.04 PLdB (as opposed to the ± 0.77 PLdB confidence interval of the larger data set in Figure 78.) However, the estimated single event tolerance level changes by only 0.24 (81.62 – 81.38) PLdB, less than half the confidence interval. Thus, these two estimates do not differ meaningfully.

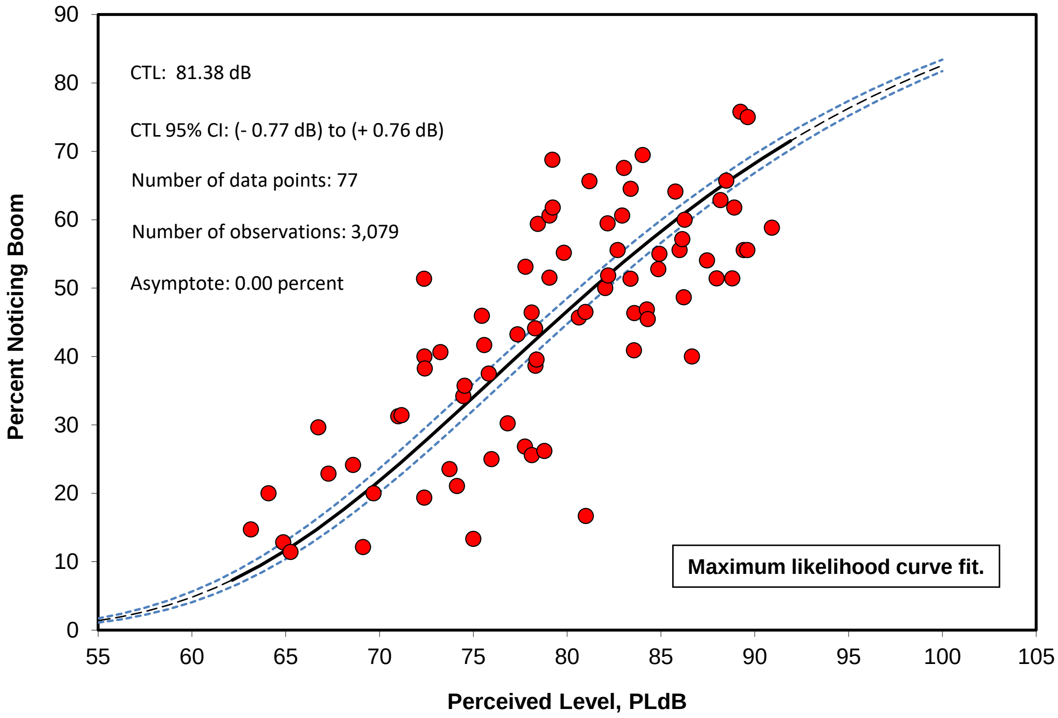


Figure 78. Noticeability *versus* Perceived Level (<91 PLdB) for single event data of Fidell *et al.* (2012).

Table 21 shows the statistics of the noticeability category as well as the three annoyance categories. The first column indicates the category, starting with noticeability and then progressing to higher levels of annoyance. The second column shows the derived value of the single event tolerance level, CTL_{SE}. The third column shows the confidence interval on the CTL_{SE} estimate.

The fourth column shows the difference in CTL_{SE} between successive categories. For example, the CTL_{SE} for noticeability was 81.4 PLdB, and that for slight or greater annoyance was 99.3 dB, a difference of 17.5 PLdB. The incremental differences between the additional annoyance categories were 5.1 and 5.0 PLdB.

Column 5 shows the cumulative difference between noticeability and any of the annoyance categories. For example, the difference between noticeability and slight or greater annoyance is 17.5 PLdB. Between noticeability and high annoyance, the difference is 27.6 PLdB. These large differences should be kept in mind during the planning of future LBFD community response field tests.

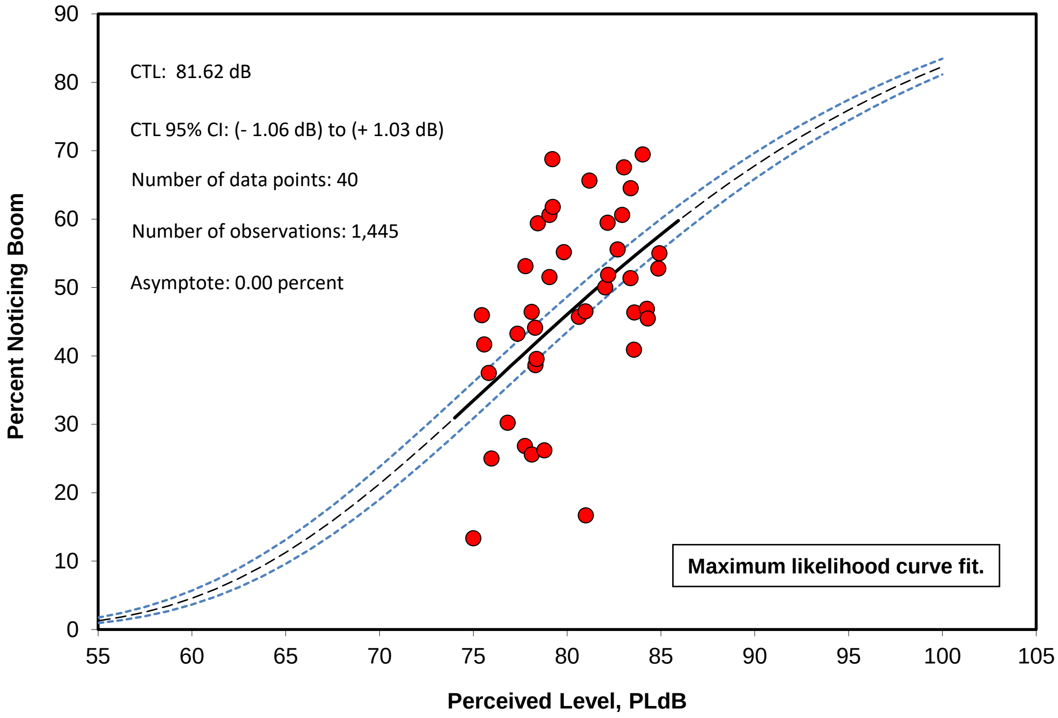


Figure 79. Noticeability *versus* Perceived Level (75 to 85 PLdB) for single event data of Fidell *et al.* (2012).

Table 21. Comparison of community tolerance values for noticeability and three annoyance categories for booms of Perceived Level <91 PLdB.

Category	CTL_{SE} (PLdB)	95% CI (PLdB)	Δ CTL (PLdB)	Cum Δ CTL (PLdB)	Δ 95% CI (PLdB)
Notice	81.83	0.77	--	0	--
Slightly+	99.29	0.80	-17.46	17.46	1.11
Mod+	104.29	0.94	-5.00	22.46	1.23
Highly	109.40	1.25	-5.11	27.57	1.56

9.6 Data Used in Constructing the Tables and Figures in this Appendix.

Table 22 provides the complete data set from which the plotted data points in this appendix were derived. The data was acquired as part of the dose-response investigation in Fidell *et al*. (2012).

The first column shows the boom sequence number. The next three columns provide the dose metrics for each boom. The numbers in these columns represent the average of the estimated boom intensities across all residences. The fifth column shows the number of at-home participants at the time of the boom. The remaining four columns provide the percentages of at-home respondents who (a) noticed the boom, (b) were slightly or more annoyed by the boom, (c) were moderately or more annoyed, and (d) were very or more (highly) annoyed by the boom.

Table 22. Data from Fidell *et al*. (2012).

Boom Number	Partic Avg max OP	Participant Avg 20 Log[max OP]	Partic Avg PL	Number of At-Home Participants	Percentages of At-Home Respondents Noticing or Annoyed: Notice	Slightly+	Mod+	Highly
1	1.081	0.67	95.14	29	65.5	24.1	13.8	3.4
2	0.349	-9.20	82.19	27	51.9	3.7	3.7	0.0
3	0.448	-6.98	86.27	30	60.0	6.7	3.3	3.3
4	0.911	-0.82	95.03	31	64.5	16.1	9.7	3.2
5	0.110	-19.18	72.39	31	19.4	0.0	0.0	0.0
6	0.121	-18.34	68.59	29	24.1	0.0	0.0	0.0
7	0.221	-13.13	74.54	28	35.7	0.0	0.0	0.0
8	0.109	-19.26	66.74	27	29.6	0.0	0.0	0.0
9	1.100	0.83	102.00	31	67.7	16.1	16.1	3.2
10	0.293	-10.68	78.11	28	46.4	3.6	0.0	0.0
11	0.301	-10.45	79.22	32	68.8	0.0	0.0	0.0
12	0.447	-6.99	83.39	31	64.5	6.5	0.0	0.0
13	0.445	-7.05	83.03	37	67.6	13.5	2.7	0.0
14	0.179	-14.95	73.74	34	23.5	2.9	2.9	0.0

Boom Number	Partic Avg max OP	Participant Avg 20 Log[max OP]	Partic Avg PL	Number of At-Home Participants	Percentages of At-Home Respondents Noticing or Annoyed			
					Notice	Slightly+	Mod+	Highly
15	0.171	-15.33	72.41	34	38.2	0.0	0.0	0.0
16	0.545	-5.27	87.97	35	51.4	11.4	8.6	0.0
17	0.401	-7.95	82.69	36	55.6	5.6	5.6	0.0
18	0.179	-14.92	72.40	35	40.0	2.9	0.0	0.0
19	0.176	-15.12	75.58	36	41.7	2.8	2.8	0.0
20	0.399	-7.98	82.16	37	59.5	5.4	2.7	2.7
21	0.627	-4.05	89.63	32	75.0	21.9	6.3	6.3
22	0.350	-9.12	75.00	30	13.3	0.0	0.0	0.0
23	0.429	-7.36	82.03	30	50.0	10.0	10.0	6.7
24	0.518	-5.72	84.25	32	46.9	9.4	6.3	3.1
25	0.436	-7.22	81.19	32	65.6	15.6	6.3	3.1
26	0.890	-1.01	97.00	32	78.1	25.0	18.8	6.3
27	0.804	-1.90	90.91	34	58.8	26.5	8.8	5.9
28	0.504	-5.98	84.03	36	69.4	19.4	5.6	2.8
29	0.472	-6.54	82.94	33	60.6	12.1	6.1	0.0
30	0.687	-3.30	88.17	35	62.9	8.6	0.0	0.0
31	0.743	-2.59	88.80	35	51.4	17.1	5.7	2.9
32	0.387	-8.25	79.06	33	51.5	12.1	0.0	0.0
33	0.330	-9.64	78.29	34	44.1	5.9	2.9	0.0
34	0.401	-7.96	80.63	35	45.7	11.4	8.6	2.9
35	0.717	-2.88	88.49	35	65.7	22.9	8.6	0.0
36	0.738	-2.64	89.24	33	75.8	27.3	9.1	0.0
37	0.390	-8.18	77.78	32	53.1	9.4	3.1	0.0
38	0.378	-8.46	79.06	33	60.6	12.1	3.0	0.0
39	0.379	-8.42	79.83	29	55.2	6.9	3.4	0.0
40	0.378	-8.46	79.24	34	61.8	5.9	2.9	0.0
41	0.300	-10.47	75.46	37	45.9	2.7	0.0	0.0
42	0.345	-9.25	78.44	32	59.4	3.1	0.0	0.0
43	0.200	-13.99	71.00	32	31.3	0.0	0.0	0.0
44	0.308	-10.26	73.25	32	40.6	3.1	0.0	0.0
45	0.312	-10.15	72.38	37	51.4	5.4	2.7	0.0
46	0.227	-12.91	71.17	35	31.4	0.0	0.0	0.0
47	0.120	-18.48	67.29	35	22.9	0.0	0.0	0.0
48	2.131	6.57	104.66	41	73.2	39.0	29.3	19.5
49	0.250	-12.06	77.35	37	43.2	8.1	5.4	2.7
50	0.512	-5.82	86.14	35	57.1	28.6	14.3	8.6
51	0.450	-6.96	84.93	40	55.0	17.5	10.0	5.0
52	0.499	-6.05	85.77	39	64.1	25.6	7.7	5.1
53	0.543	-5.30	87.46	37	54.1	16.2	10.8	2.7
54	0.451	-6.92	86.22	37	48.6	13.5	13.5	8.1

Boom Number	Partic Avg max OP	Participant Avg 20 Log[max OP]	Partic Avg PL	Number of At-Home Participants	Percentages of At-Home Respondents Noticing or Annoyed			
					Notice	Slightly+	Mod+	Highly
55	0.719	-2.87	89.61	36	55.6	19.4	19.4	5.6
56	0.406	-7.83	83.38	37	51.4	13.5	5.4	0.0
57	0.313	-10.09	69.12	33	12.1	3.0	0.0	0.0
58	0.173	-15.36	65.26	35	11.4	0.0	0.0	0.0
59	0.230	-12.79	69.69	35	20.0	2.9	0.0	0.0
60	0.147	-16.68	63.15	34	14.7	0.0	0.0	0.0
61	0.157	-16.14	64.09	35	20.0	2.9	2.9	2.9
62	0.702	-3.07	89.42	36	55.6	22.2	16.7	2.8
63	0.687	-3.26	88.91	34	61.8	26.5	14.7	5.9
64	0.560	-5.03	86.00	36	55.6	13.9	5.6	0.0
65	0.517	-5.73	84.30	33	45.5	18.2	6.1	3.0
66	0.522	-5.66	84.86	36	52.8	16.7	8.3	0.0
67	0.987	-0.16	95.61	38	63.2	31.6	23.7	10.5
68	1.542	3.75	101.58	38	55.3	34.2	28.9	10.5
69	1.391	2.86	98.82	38	63.2	39.5	28.9	18.4
70	1.434	3.11	97.43	37	67.6	27.0	21.6	18.9
71	0.466	-6.66	86.65	40	40.0	5.0	2.5	0.0
72	0.272	-11.32	80.98	43	46.5	2.3	0.0	0.0
73	0.201	-13.97	76.84	43	30.2	0.0	0.0	0.0
74	0.208	-13.68	78.32	44	38.6	2.3	0.0	0.0
75	0.225	-12.99	78.37	43	39.5	7.0	2.3	0.0
76	1.908	5.61	104.70	44	40.9	25.0	25.0	13.6
80	0.300	-10.46	81.00	42	16.7	2.4	2.4	2.4
90	0.323	-9.82	78.79	42	26.2	4.8	0.0	0.0
91	0.384	-8.31	83.59	41	46.3	19.5	9.8	2.4
96	0.301	-10.44	78.12	43	25.6	9.3	0.0	0.0
97	0.370	-8.64	83.57	44	40.9	15.9	4.5	2.3
98	0.198	-14.07	77.76	41	26.8	4.9	0.0	0.0
99	0.191	-14.39	75.98	40	25.0	2.5	0.0	0.0
100	0.236	-12.54	75.83	40	37.5	5.0	2.5	0.0
101	0.220	-13.15	74.47	38	34.2	2.6	0.0	0.0
102	0.230	-12.77	74.13	38	21.1	2.6	0.0	0.0
103	0.103	-19.77	64.87	39	12.8	5.1	2.6	0.0

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD-MM-YYYY)* 01-03-2019	**2. REPORT TYPE** Contractor Report	**3. DATES COVERED** *(From - To)*

4. TITLE AND SUBTITLE Field Evaluations of Sampling, Interviewing, and Flight Tracking of NASA's Low Boom Flight Demonstrator Aircraft	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER NNX15AA05C
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Fidell, Sanford; Horonjeff, Richard D.	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER 110076.02.07.02.03
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-2019-220257

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified-Unlimited
Subject Category 71
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

Langley Technical Monitor: Jonathan Rathsam

Final Report Phase II, Milestone 15

14. ABSTRACT

This report summarizes three years' effort, conducted in two phases. The overall intent of the project was twofold: 1) to evaluate risks to field testing of community response to low amplitude sonic booms produced by an experimental supersonic aircraft (NASA's Low Boom Flight Demonstrator, or LBFD aircraft); and 2) to identify measures for minimizing such risks. Chapter 2 summarizes the planning effort of Phase I. Chapter 3 summarizes a field demonstration of an Internet-enabled sonic boom measurement system. Chapter 4 presents the findings of an investigation of interview completion rates in a time-constrained, cross-sectional telephonic social survey of prompt reactions to individual sonic booms, as observed in a nationally representative sample of households. It also discusses the rationale for developing information of interest to aircraft noise regulatory agencies; a detailed study design for variant panel sample social surveys (developed after interview completion rates in the cross-sectional study proved impractically low); and observations about the feasibility of a cross-sectional social survey for assessing delayed reactions to cumulative sonic boom exposure. Chapter 5 summarizes the substantive findings of each year's efforts, and their implications for additional effort intended to further minimize risks of field testing of community reaction to LBFD flight missions.

15. SUBJECT TERMS

Community response; Interview; Sonic boom; Survey

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			STI Help Desk (email: `help@sti.nasa.gov`)
U	U	U	UU	150	**19b. TELEPHONE NUMBER** ***(Include area code)*** (757) 864-9658