Partnership for AiR Transportation
Noise and Emissions Reduction
An FAA/NASA/Transport Canada-
sponsored Center of Excellence

Development of a Model of Startle Resulting from Exposure to Sonic Booms

A PARTNER Projects 8 and 24 Report

prepared by
Andrew J. Marshall, Patricia Davies

December 2013

REPORT NO. PARTNER-COE-2013-003

Development of a Model of Startle Resulting from Exposure to Sonic Booms

A PARTNER Projects 8 and 24 Report

Andrew J. Marshall, Patricia Davies

PARTNER-COE-2013-003

December 2013

This work was funded by the US Federal Aviation Administration Office of Environment and Energy under FAA Award No. 07-C-NE-PU, Amendment Nos. 006, 007, and 022, and Award No. 07-C-NE-PU, Amendment Nos. 013 and 028, and Award No. 09-C-NE-PU, Amendment Nos. 001, 008, 009, 014, and 019. These projects were managed by Laurette Fisher and Mehmet Marsan of the FAA.

The Partnership for AiR Transportation Noise and Emissions Reduction — PARTNER — is a cooperative aviation research organization, and an FAA/NASA/Transport Canada-sponsored Center of Excellence. PARTNER fosters breakthrough technological, operational, policy, and workforce advances for the betterment of mobility, economy, national security, and the environment. The organization's operational headquarters is at the Massachusetts Institute of Technology.

The Partnership for AiR Transportation Noise and Emissions Reduction
Massachusetts Institute of Technology, 77 Massachusetts Avenue, 33-240
Cambridge, MA 02139 USA
http://partner.mit.edu

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	x
ABSTRACT	xviii
1 INTRODUCTION	1
1.1 Problem Statement	4
1.2 Basic Sonic Boom Generation and Propagation	6
1.3 Sonic Boom Reproduction and Low Frequency Perception	9
1.3.1 Tactile Responses Evoked by Low Frequency Sound	10
1.4 Human Response to Sonic Booms and Similar Sounds	12
1.5 Startle	15
1.5.1 Basic Properties of Startle and its Anatomy	16
1.5.2 Startle Modification	23
1.5.2.1 Pre-pulse Inhibition Models	25
1.5.3 Sonic Boom Evoked Startle	33
1.5.4 Thresholds of Startle: Booms and Other Signals	38
1.6 Methods	39
1.6.1 Skin Conductance	42
1.6.2 Pulse-rate	43
1.6.3 Electromyography	48
1.7 Hypothesis	48
1.8 Organization and Summary	51
2 COMPARISON OF RESPONSES BETWEEN TWO SEMANTIC DIFFERENTIAL EXPERIMENTS WITH TWO METHODS OF PLAYBACK	52
2.1 Methods	52
2.1.1 Word Scales	53
2.1.2 Apparatus	53
2.1.3 Signals	54
2.1.4 Procedure	55
2.1.5 Subjects	56
2.1.6 Metric Calculation	57
2.2 Results	58
2.2.1 Comparison of Mean Ratings	59
2.2.2 Metric Analysis	63
2.2.2.1 Single Number Metrics	65

			Page
		2.2.2.2 Models of two or more metrics	69
	2.3	Discussion	71
	2.4	Summary	75
3	A COMPARISON OF SUBJECTIVE RATINGS AND PHYSIOLOGICAL MEASURES OF STARTLE		77
	3.1	Experimental Methodology	78
		3.1.1 Task	79
		3.1.2 Semantic Differential Component	81
		3.1.3 Signals	81
	3.2	Procedure	82
		3.2.1 Subjects	83
	3.3	Results	83
		3.3.1 Post-processing of physiological measures	85
		3.3.2 Analysis of Physiological Measures	86
		3.3.3 Latency Analysis	91
		3.3.4 Subjective Ratings versus Physiological Responses	95
	3.4	Summary	97
4	REPEATABILITY OF PHYSIOLOGICAL MEASURES EVOKED BY LOW AMPLITUDE SONIC BOOMS		99
	4.1	Methods	99
		4.1.1 Signals	100
		4.1.2 Procedure	102
		4.1.3 Subjects	105
	4.2	Results	105
		4.2.1 Skin Conductance	106
		4.2.2 Heart-rate	112
		4.2.3 Electromyography	116
		4.2.4 Paired-comparison Analysis	119
		4.2.5 Comparison of Physiological Measures and Subjective Ratings	124
	4.3	Summary and Conclusions	125
5	DEVELOPMENT OF A STARTLE MODEL		128
	5.1	Methods	129
		5.1.1 Signals	129
		5.1.2 Procedure	133
		5.1.3 Subjects	133
	5.2	Results	134
		5.2.1 Metrics Analysis: Single Metric Models	137
		5.2.2 Metrics analysis: models of two or more metrics	139
	5.3	Summary and Conclusions	142

Page

6 GENERALIZATION OF STARTLE MODEL WITH DIVERSE IMPULSIVE STIMULI 143
6.1 Methods 143
6.1.1 Signals 143
6.1.2 Procedure 144
6.1.3 Participants 146
6.2 Results and Discussion 146
6.2.1 Analysis of Single Metric Models 147
6.2.2 Analysis of Candidate Models 153
6.3 Summary and Conclusions 161

7 CONCLUDING COMMENTS AND RECOMMENDED FUTURE WORK . . 163
7.1 Summary of Work Completed 163
7.2 Contributions of this Work 166
7.3 Recommendations for Future Work 167

LIST OF REFERENCES 169

APPENDICES
Appendix A: Semantic Differential Experiments 175
Appendix B: Initial Startle Experiment 190
Appendix C: Startle Repeatability Study 192
Appendix D: Startle Modeling Experiment 193
Appendix E: Verification Experiment 198
Appendix F: Miscellaneous Computer Programs 203

VITA 238

LIST OF TABLES

Table Page

1.1 Maximum of Moore and Glasberg's time-varying loudness of low-level signals that have been found to evoke startle responses compared to range used in author's previous study. 39

1.2 Summary of typical response latencies for physiological measures used in this research. Values tabulated from (Brown et al., 1991); (Turpin et al., 1999) and (Dawson et al., 2000) 41

2.1 Semantic Differential Word Pairs. $*$ and $+$ denote scales used only in the earphone and simulator tests respectively. 53

2.2 Description of signals used in each semantic differential test. $*$ are signals used in the earphone experiment. 55

2.3 Coefficients of determination (R^2) for single number metrics predicting average loudness, annoyance and startle ratings. Coefficient of determination for Earphone experiment are in parenthesis. *denotes not significant at $\alpha = 0.05$. Mm is maximum of Moore loudness, Zm is maximum of Zwicker loudness. 66

2.4 Coefficients of determination (R^2) for linear models predicting subject-rated annoyance and startle utilizing Zwicker's time-varying loudness in simulator experiment. Coefficients of determination for the models in the earphone experiment are in parentheses. * denotes models with partial F-test p-values of >0.05 i.e., not significant at the 5% level. Zm denotes maximum loudness, ZRT denotes loudness rise time, ZdL denotes loudness derivative and S denotes sharpness. 71

2.5 Coefficients of determination (R^2) for linear models predicting subject-rated annoyance and startle utilizing Moore's time-varying loudness for both experiments. Coefficients of determination for the models in the earphone experiment are in parentheses. * denotes models with partial F-test p-values of >0.05, i.e. not significant at the 5% level. Mm denote the maximum loudness, MRT denotes loudness rise time, MdL denotes loudness derivative and S denotes sharpness. 74

3.1 Stimuli used in both parts of the experiment. In the task part of the test signals 1-12 were used and in the semantic differential test signals 1-9 were used. The maximum of Moore and Glasberg's time-varying loudness for each signal is given as a reference. 82

Table		Page
3.2	Coefficient of determination between probability of response and mean startle ratings. * denote p-values >0.05, i.e., not significant at the 5% level. SD = Semantic differential test.	95
4.1	Coefficient of determination (R^2) between probability of physiological responses and BTL startle scores. p-values are displayed in parenthesis. *'s denote negative correlations. HRA - heartrate increase, HRD - heartrate decrease, SCM - sternocliedomastiod, AC - anterior scalene, TRP - trapezius and SC - skin conductance.	125
4.2	Coefficient of determination (R^2) between probability of co-occuring physiological responses and BTL startle scores. p-values are displayed in parenthesis. *'s denote negative correlations. HRA - heartrate increase, HRD - heartrate decrease, SCM - sternocliedomastiod, AC - anterior scalene, TRP - trapezius and SC - skin conductance.	126
5.1	Correlation coefficient (R) between metrics of stimulus set. LN_{max} is maximum Moore's long-term loudness, Dur is duration.	132
5.2	Description of block content by signal type.	132
5.3	Coefficient of determination for selected models. N is N_{max}, dL is dL_{max}, S is S_{max}, Dr is Dur and LN is LN_{max}.	140
5.4	Estimated parameters for selected models. N is N_{max}, dL is dL_{max}, S is S_{max}, Dr is Dur and LN is LN_{max}.	141
6.1	Correlation coefficient (R) between metrics of stimulus set. LN_{max} is maximum Moore's long-term loudness, Dur is duration, Z_{max} is maximum Zwicker's time-varying loudness and PL is Steven's Mk7 perceived level.	144
6.2	Coefficient of determination (R^2) between metrics of stimulus set and average subject responses. LN_{max} is maximum Moore's long-term loudness, D is duration, Z_{max} is maximum Zwicker's time-varying loudness and PL is Steven's Mk7 perceived level. Results from Chapter 5 experiment are in parenthesis.	150
6.3	Coefficient of determination (R^2) between metrics of stimulus set and average subject responses. LN is LN_{max}, N is N_{max}, S is S_{max}, dL is dL_{max} and Dr is duration. M_8 is model 8 and M_6 is model 6. Coefficients of determination (R^2) for the set of signals also used in the test described in Chapter 5 are in parenthesis.	156
6.4	Variables and their coefficients in Models 12 and 21. The third-term coefficient is zero if $Dur < 500$ ms for Model 21.	161

Appendix Table | Page

A.1 Statistics of subject responses in the earphone semantic differential experiment. Signal numbers correspond to Table 2.2 and Scale numbers correspond to Table 2.1. Scales ranged from -16 to +16. 176

A.2 Statistics of subject responses in the simulator semantic differential experiment. Signal numbers correspond to Table 2.2 and Scale numbers correspond to Table 2.1. Scales ranged from -16 to +16. 182

A.3 Metric values for signals using in Earphone (1st number) and Simulator (2nd number) semantic differential experiments. 188

A.4 Metrics used in models of two or more metrics in earphone (1st number) and simulator (2nd number) semantic differential experiments 189

B.1 Letters correct in arithmetic and memory task for each subject. 190

B.2 Duration of each letter block for each subject. Signal playbacks are display in Figure 3.3. 191

C.1 Probabilities of first sound (row) being judged more startling than second sound (column) on day 1. 192

C.2 Probabilities of first sound (row) being judged more startling than second sound (column) on day 2. 192

D.1 Mean ratings, standard deviation of the estimated mean and metric values for the startle modeling experiment stimuli. N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones) and Dur is duration (ms). 193

E.1 Mean ratings, standard deviation of the estimated mean and metric values for the startle ratings in Chapter 6. N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones), Dur is duration (ms), PL is Steven's Mk7 PL (dB) and Z_{max} is maximum Zwicker loudness (sones). 198

E.2 Mean annoyance ratings, standard deviation of the estimated mean, acceptability fractions, confidence intervals and associated metric values for the sounds in Chapter 6. Ann is Annoyance, Acc is acceptability, N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones), Dur is duration (ms), PL is Steven's Mk7 PL (dB) and Z_{max} is maximum Zwicker loudness (sones). 201

LIST OF FIGURES

Figure Page

1.1 a) A traditional sonic boom with labeled signal properties and b) the corresponding energy spectral density of traditional sonic boom. 2

1.2 a) The time history of a simulated low boom and b) the corresponding energy spectral density. 3

1.3 a) The time-history and b) corresponding energy spectral density of non-linearly steepened 20 ms, 50 Hz sine wave at propagation distances of 0 (black), $2c$, $4c$, $6c$, $8c$ and $10c$ (lightest gray). c is the linear sound speed (343 m/s). 8

1.4 Diagram of startle neuro-circuitry. Startle modification circuitry is labeled in box with dashed outline. 19

1.5 Signal flow of prepulse inhibition model of Schmajuk and Larrauri (2006). . 29

1.6 Output prepulse model (short dash, black), instantaneous loudness (long dash, black), short-term loudness (solid black) and loudness derivative (gray solid) for a 50 ms, 80 dB white noise pulse. 31

1.7 Maximum startle output of Schmajuk's model versus maximum short-term loudness derivative before first loudness peak ($R^2 = 0.83$) for boom-like white noise signals. 32

1.8 a) Schematic of sound presentation scenario for simulation of a 90 dB pulse with recorded background noise. b) Maximum startle due to 80 db pulse with recording noise present normalized by maximum startle when there is no recording noise, plotted against duration of recording noise prior to pulse event ($Tdur$). Results predicted using Schmajuk and Larrauri's (2006) model for recording levels of 45 (black), 50, 55 and 65 dB (lightest gray). 34

1.9 Illustration of onset and peak latency definitions for an arbitrary response to an event (gray vertical line). 41

1.10 A typical skin conductance response to a sonic boom which started at t = 0 seconds. Relevant response properties are labeled. 44

1.11 a) A typical signal produce by an optical heart-rate sensor in response to sonic boom. b) Band-pass filtered pulse signal. c) Associated heart-rate calculated via instantaneous frequency (black, solid), zero-crossing (dark gray, long dashes) and peak-picking (light Gray, short dashes). Boom onset denoted by vertical black line. Note presence of movement artifacts. 47

Figure Page

1.12 EMG measurement of a trapezius muscle response when exposured to a loud sonic boom. Vertical black line indicates onset of sonic boom. 49

2.1 a) Zwicker loudness time history, b) Loudness derivative time history and c) Zwicker Sharpness time history of signal 8. 58

2.2 Factor loading of a) Simulator experiment and b) Earphone experiment. Black is loudness factor, dark gray is the temporal factor, white is the spectral factor and light gray is industrial factor (only present in earphone experiment). . . 61

2.3 a) Annoyance versus loudness ratings (R^2 = 0.80), b) startle versus loudness ratings (R^2 = 0.75), and c) annoyance versus startle ratings (R^2 = 0.93) for simulator experiment. d) Annoyance versus loudness ratings (R^2 = 0.78), e) startle versus loudness ratings (R^2 = 0.83), and f) annoyance versus startle ratings (R^2 = 0.87) for earphone experiment. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder. 62

2.4 Earphone v. Simulator ratings of a) and d) loudness (R^2 = 0.55), b) and e) annoyance (R^2 = 0.75), and c) and f) startle (R^2 = 0.59). For d)-f) marker size is based on average ratings of the "No Feel-Feel" scale. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, filtered booms are stars and diamonds are thunder. . . 64

2.5 Subject-rated a) loudness (R^2 = 0.83), b) annoyance (R^2 = 0.63) and c) startle (R^2 = 0.64) versus Maximum Moore's loudness for Simulator Study results. Subject-rated d) loudness (R^2 = 0.88), e) annoyance (R^2 = 0.69) and f) startle (R^2 = 0.75) versus Maximum Moore's loudness for Earphone Study results. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder. 67

2.6 Subject-rated a) loudness (R^2 = 0.89), b) annoyance (R^2 = 0.61) and c) startle (R^2 = 0.59) versus Maximum Zwickers's loudness for Simulator Study results. Subject-rated d) loudness (R^2 = 0.51), e) annoyance (R^2 = 0.36) and f) startle (R^2 = 0.28) versus Maximum Zwicker's loudness for Earphone Study results. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder. 68

2.7 Maximum Zwicker's time-varying loudness versus maximum Moore and Glasberg's time-varying loudness for a) simulator experiment and b) earphone experiment. 69

Figure Page

2.8 Annoyance and startle ratings versus predictions of linear model of Mm and 1/MRT in a) & c) the earphone and b) & d) the simulator experiments, respectively. Annoyance and startle ratings versus predictions of linear model of Mm and MdLmax in e) & g) the earphone and f) & h) the simulator experiments, respectively. Refer to Table 2.5 for R^2 values. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder. 72

2.9 Annoyance and startle ratings versus predictions of linear model of Zm and ZdLmax in a) & c) the earphone and b) & d) the simulator experiments, respectively. Annoyance and startle ratings versus predictions of linear model of Zm and Smean in e) & g) the earphone and f) & h) the simulator experiment, respectively. Refer to Table 2.4 for R^2 values. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder. 73

3.1 Diagram of EMG locations: a) sternocliedomastoid (SCM), b) anterior scalene, and c) upper trapezius. For anatomical reference, d) is the collarbone. . . . 78

3.2 Letter recall screen under perfect recall of letter sequence: H,J,Q,Y,F, and T. 80

3.3 a) Letter correct for each subject through time. Shades of gray indicate number of letters correctly. White indicates all correct. Letter blocks with sound exposures are marked with X's. b) Mean letters correct through time for sounds (black) and no sounds (white). Errorbars are standard deviation of the estimated mean. 84

3.4 a) Time taken by each subject (in seconds, 0-140 s) for each letter block. Letter task blocks with sound exposures are marked with X's. b) Mean time per letter block over the course of the experiment for blocks with sounds (black) and without sounds (white). Errorbars are standard deviation of the estimated mean. 85

3.5 Profile of Subject 7's task experiment. Triangles correspond to letter recall screens. (6 correct = black, 5 correct = gray, 4 correct = light grey). a) Earphone voltage: dashed circles denote signals of interest, b) EMG: SCM is black, scalene is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute. 88

3.6 A closer examination of the sound exposure of subject 7's task around 270 seconds. a) Earphone voltage, b) EMG: SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute. 89

Figure Page

3.7 A closer examination of the sound exposure of subject 7's task around 532 seconds. a) Earphone voltage, b) EMG , SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute. Note that due to the high frequency content of this signal, the earphone voltage levels are low. 90

3.8 A profile of Subject 7's semantic differential physiological measures. a) Earphone voltage, b) EMG , SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μ- Siemens), d) filtered skin conductance (μ-Siemens), and e) heart rate in beats-per-minute. For the EMG plot, SCM is black, trapezius is gray, and scalene is light grey. Both Skin conductance measures are in μSiemens and the heart rates are in beats per minute. 92

3.9 Fraction of EMG responses, heart-rate responses and skin conductance responses observed for each sound exposure for a), c) & e) the task portion and b), d) % e) the semantic differential portion, respectively. For EMG, black is SCM, gray is scalene and white is trapezius. For heart-rate: black is short latency increase; dark gray is decrease, then increase; gray is medium latency increase; light gray is long latency increase and white is medium latency decrease. 94

3.10 Fraction of EMG responses, heart-rate responses and skin conductance responses observed for each signal for a), c) & e) the task portion and b), d) % e) the semantic differential portion, respectively. For EMG, black is SCM, gray is scalene and white is trapezius. For heart-rate: black is short latency increase; dark gray is decrease, then increase; gray is medium latency increase; light gray is long latency increase and white is medium latency decrease. . . 96

4.1 a) Maximum loudness versus maximum loudness derivative of first loudness peak for stimuli. Levels of stimuli used in Thackray et. al. (1974) are stars. High amplitude stimulus is a square. The five stimuli in the main study are circles. Loudness time-histories of each of the signals: b) signal 1, c) signal 2, d) signal 3, e) signal 4 and f) signal 5 (black) and signal 6 (gray). 103

4.2 Fraction of skin conductance responses obtained from a) 5 sound stimuli, b) high amplitude stimuli and c) pictures on both days. d) Fraction of responses for each subject on both days. for a) skin conductance, b) heart-rate, c) SCM, d) scalene and e) trapezius. Vertical line separates Day 1 and Day 2 responses. 107

4.3 Fraction of heart-rate responses obtained from a) 5 sound stimuli, b) high amplitude stimuli and c) pictures on both days. d) Fraction of responses for each subject on both days. Short-latency heart-rate increases, decrease then increases, medium latency increases, long latency increases and decreases only are denoted by black though white respectively. Vertical line separates Day 1 and Day 2 responses. 108

Figure Page

4.4 Fraction of a) SCM, c) scalene and e) trapezius responses obtained for 5 main stimulis. Fraction of b) SCM, d) scalene and f) trapezius responses for high amplitude stimuli. For Black denotes responses with a magnitude of greater than 10% of head circle, gray denotes all responses from the sound start, white denotes responses from the end of the sound. Vertical line separates Day 1 and Day 2 responses. 109

4.5 Fraction of a) SCM, c) scalene and e) trapezius responses obtained for pictures. Fraction of b) SCM, d) scalene and f) trapezius responses for each subjects on both days. For Black denotes responses with a magnitude of greater than 10% of head circle, gray denotes all responses from the sound start, white denotes responses from the end of the sound. Vertical line separates Day 1 and Day 2 responses. 110

4.6 Filtered skin conductance time-histories for subject a) 2, b) 4 and c) 6. Black is the first day and gray is the second day. 111

4.7 Average skin conductance response magnitude for both high amplitude stimuli (squares) and 5 original stimuli (circles). Black is day 1 and gray is day 2. . 113

4.8 Heart-rate profiles of a) Day 1 main stimuli, b) Day 2 main stimuli, c) Day 1 high stimulus and d) Day 2 high stimulus. Black through light gray filled circles correspond to 1st through 5th sound exposures. Black through light gray triangles correspond to 6th through 10th sound exposures. For a) and b) exposure order was signal 5, 2, 1, 4, 3 for the 1st-5th and 6th-10th exposures respectively. Error bars are standard deviation of the estimated mean. 115

4.9 Average magnitude of EMG responses (in % of maximum obtained in calibration procedure) for both high amplitude stimuli (squares) and 5 original stimuli (circles) measured from the a) SCM, b) scalene and c) trapezius. Black is day 1 and gray is day 2. 118

4.10 BTL startle scores for a) first day, b) second day and c) both days. Black are scores calculated using all subjects' data, gray are scores calculated for all but one subject's data. Dark gray diamond are scores from after removing subject 9. 120

4.11 BTL startle scores calculated for 1st session (light gray), 2nd and 3rd session (dark gray) and all sounds (black) for a) first day, b) second day and c) both days. 121

4.12 BTL startle scores versus a) N_{max} ($R^2 = 0.61$) and b) dL_{max} ($R^2 = 0.58$). . . 122

4.13 BTL startle scores versus estimates of the linear model of maximum loudness and maximum loudness derivative ($R^2 = 0.88$). 123

Figure Page

5.1 a) N_{max} versus dL_{max}, b) S_{mean} versus N_{max} and c) S_{mean} versus dL_{max} for test stimuli in the presence of background noise. Synthesized and industry simulated boom signals are dark gray. Boom-like noise signals are light gray. 131

5.2 a) Average startle ratings in signal presentation order (Gray circles). Black dots are average N_{max} for each signal presentation. Errorbars are standard deviation of the estimated mean. Solid vertical lines correspond to boundaries of blocks with 30 signals. Dashed vertical lines correspond to boundaries of blocks with 28 signals. 135

5.3 a) Average startle ratings of booms presented in blocks 1-4 and booms presented in block 5 . b) Block averaged startle ratings for booms in 1-4, boom-like noises and booms in block 5 with equivalent N_{max}. Booms in blocks 1-4 are white, Booms in block 5 are black and boom-like noises are gray. Error bars are standard deviation of the estimated mean. 136

5.4 Average startle ratings versus a) N_{max} (R^2all = 0.45, R^2booms = 0.69), b) dL_{max} (R^2all = 0.31, R^2booms = 0.31), c) S_{mean} (R^2all = 0.34, R^2booms = 0.41), d) S_{max} (R^2all = 0.56, R^2booms = 0.08), e) LN_{max} (R^2all = 0.46, R^2booms = 0.77) and f) Dur (R^2all = 0.32, R^2booms = 0.73). Black circles are booms, white circles are boom-like noises. Error bars are standard deviation of the estimated mean. 138

6.1 a) N_{max} versus dL_{max}, b) S_{max} versus N_{max}, c) S_{max} versus dL_{max} and d) LN_{max} versus N_{max}. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 145

6.2 Signal presentation order-averaged a) startle ratings, b) annoyance ratings and c) acceptability fractions. 148

6.3 a) Average annoyance ratings versus average startle ratings (R^2 = 0.93). b) Acceptability fraction versus average startle ratings (R^2 = 0.87). c) Acceptability fraction versus average annoyance ratings (R^2 = 0.90). Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. . 149

Figure Page

6.4 a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus maximum of short-term loudness (N_{max}). d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus maximum long-term loudness LN_{max}. R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 151

6.5 a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus perceived level (PL). d) Average startle ratings, e) Average annoyance ratings and f) Acceptability fractions versus maximum of Zwicker's time-varying loudness (Z_{max}). R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 152

6.6 a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus dL_{max}. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus S_{max}. g) Average startle ratings, h) average annoyance ratings and i) acceptability fractions versus Dur. R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 154

6.7 Group averaged startle residuals (startle ratings without N_{max}) versus a) dL_{max}, b) S_{max} and c) Dur. 155

6.8 Average startle ratings versus a) linear model of LN_{max} and S_{max} and b) model of $LN_{max}S_{max}$ and $dL^2_{max}S_{max}$. R^2 values are in table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 157

6.9 a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus model 8+Dur. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus model 6+Dur. R^2 values are in Table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 159

Figure Page

6.10 a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus model 8+$LN_{max}Dur$. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus model 6+$LN_{max}Dur$. R^2 values are in Table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections. 160

Appendix Figure

F.1 a) Time-history and b) loudness time-history of a sonic boom. Solid line is instantaneous loudness, long dashes are short-term loudness and short dashes are long-term loudness. 204

F.2 a) Pressure time-history of original signal, b) pressure time-history of boom-like noise and c) loudness time-history of boom-like noise (dash) and original signal (solid). The loudness time-history of the boom-like noise is time-shifted to be more noticeable. 233

ABSTRACT

Marshall, Andrew J. Ph.D., Purdue University, May 2012. Development of a Model of Startle Resulting from Exposure to Sonic Booms. Major Professor: Dr. Patricia Davies, School of Mechanical Engineering.

Aircraft manufacturers believe that it is possible to create supersonic business jets that would have quieter sonic booms than those that lead to the current ban on overland commercial supersonic flight over the US. In order to assess if the impact of these "low booms" is acceptable to the public, new human subject testing must occur. In recent studies, it was found that subjects' judgments of annoyance were highly correlated to judgments of startle and were unable to be fully explained by loudness judgments alone. However, this experiment utilized earphones for playback, which was unable to reproduce low frequencies (< 25 Hz) well. Building upon this study, an additional semantic differential experiment was conducted using a sonic boom simulator for playback which could reproduce these frequency components. Results of both experiments were similar and again it was found that average startle and annoyance ratings were highly correlated and that statistics of time-varying loudness were highly correlated with subjects' responses. However, it was unclear if subjects' judgments of startle corresponded to physiological responses associated with startle. To examine if physiological responses associated with startle were evoked by the low booms, two studies were conducted; a pilot study and a repeatability study. While physiological responses associated with startle were evoked by low booms, startle responses were found to have occurred infrequently. However, subjects' judgments of startle were found to be correlated with physiological responses and to have less day-to-day and subject-to-subject variance. Candidate startle models were estimated from data obtained from an experiment where subjects' judged the startle evoked by a series of low amplitude sonic booms and boom-like noises. These candidate startle models were then tested in an additional study which used a more diverse set of stimuli. It was found that

a linear model consisting of the maximum long-term Moore and Glasberg's time-varying loudness, maximum sharpness and duration (as calculated from time-varying loudness) was the best simple model that predicted subject responses in both modeling studies well. Details of the physiological responses commonly evoked by low booms and investigations into the use of loudness-rate metrics to predict startle will also be discussed.

1. INTRODUCTION

One of the main legislative barriers to the operation of commercial supersonic aircraft is the ban on overland supersonic flight that was established in 1973. The primary reason for that ban was the impact of the noise produced by these aircraft. This noise, called a sonic boom, is a result of the craft moving at a velocity greater than the speed of sound, which creates shock waves which propagate to the ground. Traditional sonic booms, the cause of the original ban, are characterized by a large peak pressure (>100 Pa), termed the overpressure, and a fast (<4 ms) rise time (Figure 1.1). Classically, these booms have a shape similar to a capital "N", the origin of the common moniker "N-wave". In terms of frequency content, traditional N-waves are dominated by low frequency components, but, due to the fast rise time, also contain some high frequency components (100-1000 Hz) that are audible. Recently, there has been interest in the production of supersonic business jets that would have sonic booms that are modified or "shaped" so that the sound may be acceptable to the public. The concept of manipulating the shape of an aircraft to create a quieter sonic boom is not new. The original theoretical work on this was pioneered by Seebass (1972). This recent interest was ignited by advances in computational models of sonic boom generation as well as the experimental verification of the practicality of boom shaping completed in the shaped sonic boom demonstrator project (Plotkin, Graham, Coen, Haering, Maglieri, Pilon, Salamone, Page, Pawlowski, Schein, McCurdy, Murraym, Ehernberger, and Bobbit, 2004). In general, these shaped booms are designed to have reduced overpressures (<50 Pa) and more gradual rise times, resulting in a quieter boom. A simulated low boom time-history and the corresponding spectrum is in Figure 1.2. The shape of the low boom is smoother and longer in duration than the traditional boom and there are relatively fewer high frequency components in low booms than in classic booms.

However, before any low boom craft can be deployed, the reason for the ban on overland commercial supersonic flight need to be addressed: the impact of these sounds on people.

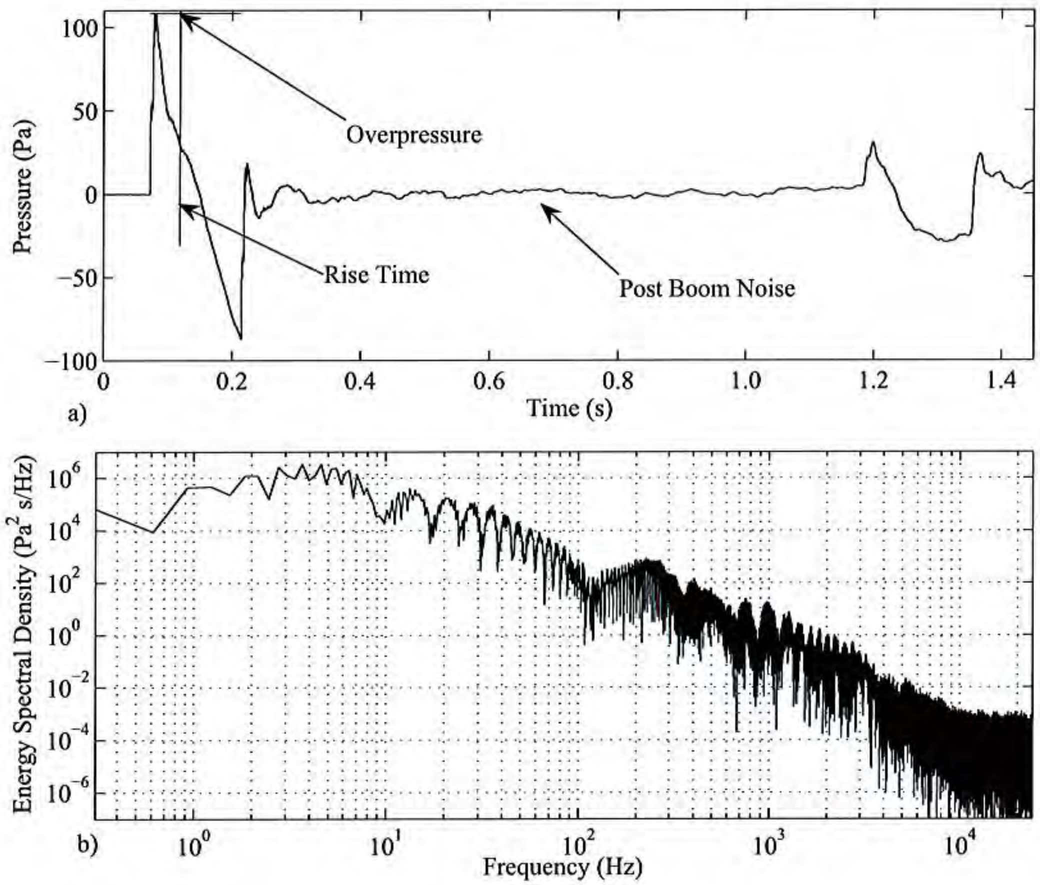

Figure 1.1. a) A traditional sonic boom with labeled signal properties and b) the corresponding energy spectral density of traditional sonic boom.

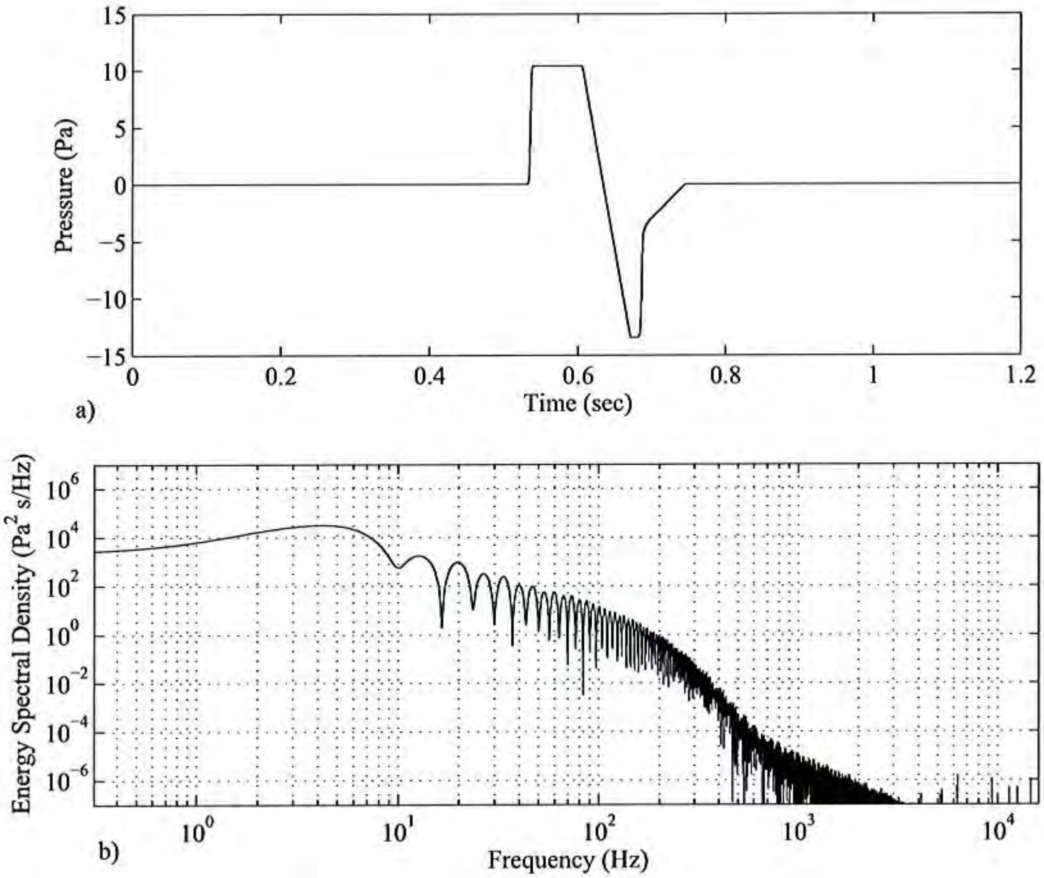

Figure 1.2. a) The time history of a simulated low boom and b) the corresponding energy spectral density.

The historical origins of this ban stem from larger political movements directed at aircraft noise. Throughout the 1960s, there was increasing public discontent towards the noise generated by aircraft, mostly caused by commercial jet airline traffic and military jet noise. This concern and some federal court cases (notably American Airlines v. Hempstead, US Supreme Court, Vol 411 U.S. 624,1973) elevated the regulation of aircraft noise to a national political issue. Evidence of the importance of this issue to voters can be seen in the "War on Aircraft Noise" which was lead by Rep. Herbert Tenzer of New York (Widdle, 2005). In 1973, congress passed an amendment that gave the Federal Aviation Administration (FAA) the authority to regulate aircraft noise. With this authority, the FAA banned commercial overland supersonic flight.

This ban was a blanket prohibition of commercial supersonic flight; no criterion for what an "acceptable" boom might be was addressed. If any assessment is to be made regarding the impact of low boom craft, more research is needed to evaluate what aspects of a low boom are objectionable to the public and to be able to quantize the acceptability (or lack thereof) of potential low boom waveforms. In previously conducted research, which will be discussed below, sound characteristics that people respond to have been identified and preliminary models that predict the annoyance or loudness of these sounds have been developed. However, more research is needed to fully develop these models so they can be used to accurately predict human response for a variety of boom signatures and other impulsive sounds heard in a similar context. The goal would be to provide a model that could be used to predict the potential impact of noise generated from hypothetical supersonic aircraft on a community.

1.1 Problem Statement

The main objective of this research is to develop models that predict annoyance experienced when subjects are exposed to transient sounds similar to and including low level sonic booms. In the author's previous research it was found that subjects' ratings of how startling the sounds were are highly correlated to their ratings of annoyance and thus to develop an annoyance model it is necessary to understand acoustic startle reactions and how sound characteristics affect them.

For people to relate to the numbers produced by the annoyance model, it should be possible to use the model to predict annoyance for transient sounds with which they are familiar. This is in contrast to the sonic boom metrics arising from previous research which were developed to compare likely responses to various types of sonic booms only. Also, because the actual noise signatures that will be heard on the ground from the various supersonic aircraft designs being proposed are still unknown, it is necessary to probe more deeply into how various transient sound characteristics would impact annoyance, so that deviations from predicted signatures would not lead to a need to develop a new annoyance model. When sonic booms are heard indoors the sounds are more complex than when heard outdoors because of the reverberations, absorption, spectral shaping by windows and walls, and rattle. The acoustic startle model developed in this research should be applicable to these more complex sounds. However, the annoyance due to these complex sounds heard indoors will be affected by the context and also by the additional sound characteristics. In this research, the focus will be on annoyance and startle in response to sounds heard outdoors. When subjects evaluate the loudness of transients like sonic booms, it is likely correlated to the maximum loudness, as predicted by various loudness algorithms. For slower transients (e.g. car pass-bys), or for fluctuating sounds, loudness exceeded 5% of the time (N_5) is a better measure of overall loudness (Zwicker and Fastl, 1999). Another issue that will be addressed in this research is how to determine the most appropriate loudness statistics given the temporal variations in the sounds, so that sounds like thunder, repeated gunfire, and door slams can be evaluated consistently.

In the previous work of Marshall (2007), summarized below, it was found that low amplitude sonic booms might evoke startle responses. People with certain neurological conditions, such as post-traumatic stress disorder or dyslexia, for example, may be more sensitive to startle than the normal population. Given the difficulty of studying these groups, this work will focus on modeling responses associated with neuro-typical populations. However, while this project lacks the resources necessary to address the impact of these sounds on people in those groups, populations with these conditions are likely to be more vulnerable to these sounds and thus must be considered when evaluating the acceptability of noise produced by supersonic aircraft.

This work is one of many efforts to examine the impact of sonic booms on people. Considerable investigation of this topic was completed both before and after the ban of supersonic aircraft was in place in 1972. In addition, there has been much research into related fields such as psychophysiology involving the study of startle which may aid in the construction of a startle model. In the remainder of this chapter, a discussion of the previous research on the human response to sonic boom will be given, as well as a brief review of sonic boom generation. It will be followed by a review of the startle response literature including studies of startle evoked by sonic booms. Finally, the hypotheses will be stated along with a discussion of some experimental methods that will be used in this research.

1.2 Basic Sonic Boom Generation and Propagation

When any object moves through a fluid, the fluid is affected and displaced (Warren, 1972b). This displaced fluid, caused by lift or drag forces on the object, creates pressure variations in the fluid around the object that can propagate away as waves in the fluid. A common example of this effect is found when a boat moves through water, creating a "V" of waves emanating from the bow. At subsonic velocities, these pressure variations are smooth. However, at higher and higher velocities, these pressure disturbances become more abrupt and of larger magnitude (Warren, 1972b). As these disturbances propagate, they become shock waves. These shock waves are sonic booms.

While attributing specific aspects of aircraft geometry to boom characteristics is a complicated task requiring computational fluid dynamics, some generalities can be made about aircraft shape and sonic booms. First, the positive pressure part of the boom waveform stems from the forward facing parts of the craft, while the rearward facing parts influence the negative pressure region of the waveform (Warren, 1972b). In addition, lift surfaces, such as the wings, contribute significantly to the pressure disturbances around the craft (Seebass and George, 1972). The shape of the boom can be influenced by the manipulation of these surfaces as well as by aircraft maneuvers.

Due to the large pressure magnitudes and the long propagation distances involved, the assumption of infinitesimal small pressure amplitudes used in linear acoustics is invalid. Thus, nonlinear acoustic effects are observed in boom propagation. One of the main non-linear effects observed is shock steepening. This effect, first derived by Riemann (Hamilton and Blackstock, 1998), is caused by the variation of sound speed with pressure. Unlike in linear acoustics, for finite amplitude sound waves, which include sonic booms, the peaks have a faster sound speed than the valleys of the waveform. As the wave propagates, this difference in speeds causes the positive and negative pressure peaks to separate. This separation of the peaks and valleys of the waveform produce the classic capital "N" shape (Warren, 1972b). For a simple sine-wave, the propagation can be described by the formula:

$$p(x,t) = P_0 sin(\omega(t - x/c) + x\beta/\rho c^2(\omega/c)P_0 sin(\omega(t + x/c))), \qquad (1.1)$$

where c is the linear sound speed and β is the nonlinearity coefficient. In Figure 1.3, the nonlinearly propagated 20 millisecond duration, 50 Hz sine-wave and its spectrum are shown for several propagation distances without including attenuation effects. In the time-domain, the rise time of the waveform decreases as the propagation distance increases. In the frequency domain, this effect manifests itself in the transfer of energy from the 50 Hz component to higher "harmonics" at multiples of 50 Hz (most noticeable at 100 and 150 Hz).

The energy transfer of low frequencies to high frequencies caused by steepening is one of the reasons why the impact of sonic booms on communities is important over the entire flight-path of the supersonic aircraft and not only near airports. For a linearly propagating disturbance, such as that produced by a conventional aircraft, propagation through the atmosphere during flight significantly reduces the noise levels at audible frequencies. This is because the atmosphere attenuates high frequencies more efficiently than lower frequen-

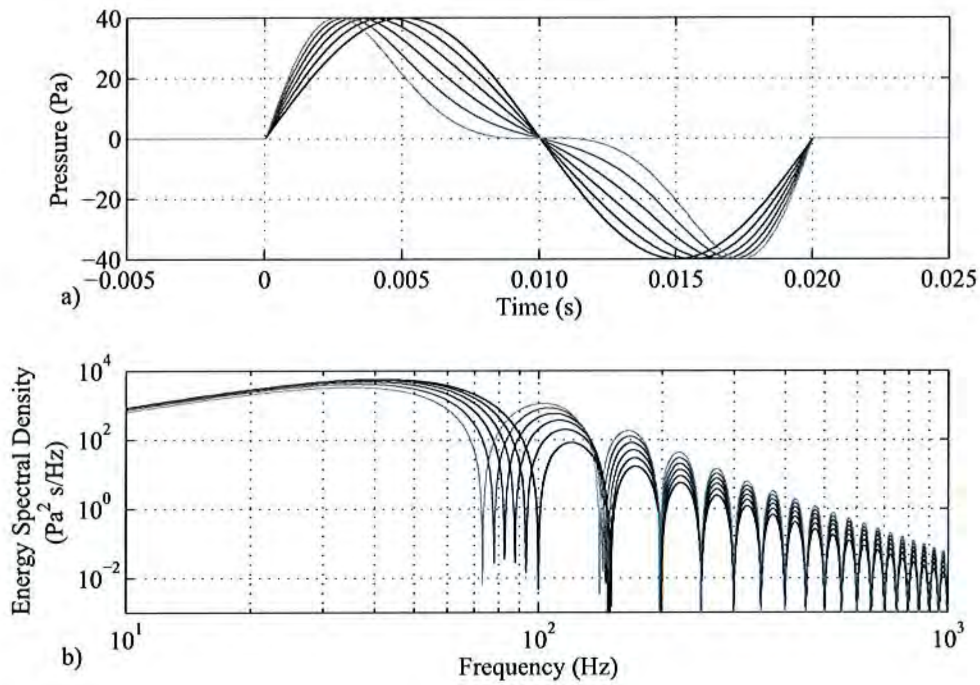

Figure 1.3. a) The time-history and b) corresponding energy spectral density of non-linearly steepened 20 ms, 50 Hz sine wave at propagation distances of 0 (black), 2*c*, 4*c*, 6*c*, 8*c* and 10*c* (lightest gray). *c* is the linear sound speed (343 m/s).

cies. However, because nonlinear steepening "replenishes" high frequency energy as the waveform propagates, sonic booms maintain significant high frequency energy even at long propagation distances. Thus, supersonic aircraft noise can be heard along the entire flight-path of the aircraft over a width of 75-250 km under the flight path (Heimann, 2001). This is one of the main reasons why the noise regulation of supersonic aircraft is very different from that of conventional aircraft and, in particular, more people are affected by a single supersonic flight.

Aside from steepening, there are other acoustic phenomena that are important for boom propagation. These are due to features of a real atmosphere (such as turbulence) and the effect of the ground. The details of these effects and the methods of numerically simulating booms are not included in this document because that is not the main focus of the research described here. The reader is referred to Bass, Raspet, Chambers, and Kelly (2002) and Plotkin (2002) for more information on sonic boom simulation and propagation. For more information about the propagation of booms through turbulence, the reader is referred to the work of Locey and Sparrow (2007) and Lipkens (2002).

1.3 Sonic Boom Reproduction and Low Frequency Perception

Using real aircraft to produce sonic booms for experimentation is expensive and creating consistent sonic booms from flights is very difficult. Because of this, there have been many efforts focused on the synthetic reproduction of sonic booms. In the first attempts to simulate sonic booms explosive charges (as described in (Warren, 1972a)0 or mechanical pistons (Thackray, Touchstone, and Bailey, 1974) were used to generate the waveforms. The main limitations of these simulation methods was their inability to generate repeatable stimuli and inability to reproduce subtle changes in the pressure waveform. More recent methods of synthetic sonic boom production utilize electric loudspeakers. Currently, there are three sonic boom simulators designed to reproduce sonic booms heard outdoors in the US: the one developed by NASA Langley, the portable one developed by Gulfstream and one developed by Lockheed Martin. In addition, NASA has developed a simulator called the indoor effects room which simulates booms and other aircraft noise heard indoors. The original simulator, the one developed by NASA Langley, was the inspiration for the two developed by the aircraft industry. All three of the simulators are composed of an array of speakers positioned in front of the subject. In two of the simulators, the NASA and Lockheed simulators, the speakers and subject are in an airtight booth with carpet on the floor and walls (Leatherwood, Shepherd, and Sullivan, 1991). The Gulfstream simulator, in contrast, has subjects placed in a waveguide. There are speakers in a wall in front of the subjects and an anechoic termination behind the subjects that suppresses reflections (Salamone, July 18-22, 2005). All three simulators are digitally equalized to have flat frequency responses from about 7 Hz to 6000 Hz. These three simulators have been found to be realistic by subjects experienced in listening to sonic booms (Sullivan, Davies, Hodgdon, Salamone III, and Pilon, 2008).

The fourth simulator, the indoor effects room, is different from the other simulators because it was designed to simulate sonic booms heard indoors. The simulator itself consists of a room manufactured to have a similar construction and materials to those found in houses. To produce sonic booms, two arrays of speakers, each arranged in a grid, are attached to two of the outside walls. The two arrays allow for the reproduction of sonic booms at different incidence angles relative to the house. At the time of this writing, initial subject testing is being completed with this simulator to verify its accuracy in reproducing indoor sonic booms (Klos, Loubeau, and Rathsam, 2011).

1.3.1 Tactile Responses Evoked by Low Frequency Sound

One question that arises from the construction of systems to reproduce sonic booms is how accurately these systems need to be for human subject testing. As low frequency components of sonic booms require special simulators to be reproduced, it is reasonable to ask how are these low frequency components perceived by people. Due to the large amount of low frequency energy components found in these sounds some authors, such as von Gierke (1966), have claimed that tactile responses are an important aspect of sonic boom perception and have criticized early simulators for not having the capability to accurately reproduce these effects.

It is not clear what the importance of tactile perception is compared to auditory perception of sonic booms and auditory perception may be influenced by tactile perception and visa versa. The human body can be excited by these sounds causing vibration of various body parts. Leventhall (2006) found that the average resonance frequency of the chest is in the range of 50-80 Hz and it is excited by levels above 80 dB. This resonance was found to be associated with chest compression (Leventhall, 2006). However, it does not appear that this human body vibration is easily perceived by people. Yamada, Ikuji, Watanabe, and Kosaka (1983) exposed profoundly deaf subjects and subjects with normal hearing to 8-1000 Hz tones. Thresholds for the profoundly deaf population were found to be 30-50 dB higher than the threshold founds for the hearing population for the tones at 63 Hz and below (Yamada et al., 1983). This means that the hearing system was found to be more sensitive to the low frequency tones than the tactile system.

This result is consistent with the physiology of tactile sensors in the skin. There are three types of tactile sensors in the skin: Merkel cells, Meissner's corpuscles and Pacianian corpuscles. Each has a different skin displacement threshold and highest sensitivity occurs over different frequency ranges: low (5 to 15 Hz), medium (20-50 Hz) and high frequency (60-400 Hz), respectively (Johnson, 2001). These thresholds range from 0.2 mm displacement for Merkel cells, 0.1-0.01 mm for Meissner's corpuscles to around 0.001 mm for Pacianian corpuscles. However, when skin displacement occurs through excitation by low frequency sounds, differences between the stimulus level and the tactile threshold are smaller than the corresponding hearing threshold, primarily due to the mis-matched impedance between the air and the body (Leventhall, 2006). This evidence led Leventhall and others (e.g. Berglund and Lindvall (1995)) to the conclusion that the primary method of perceiving low frequency sound is via the ears and the auditory system.

In the context of this research, it is clear that most low amplitude booms should be above the thresholds necessary to produce some tactile sensation (> 80 dB). However, considering that these sounds are also heard and that the hearing system is more sensitive, it is unclear how important the presence of tactile effects would be on people's judgments of sonic booms, and the dominant pathway is the auditory pathway.

1.4 Human Response to Sonic Booms and Similar Sounds

The majority of investigations into the impact of sonic booms on people have been focused on quantifying the level or loudness of sonic booms. This effort first began with the investigation of simple time-history based metrics (such as overpressure and rise-time) and, more recently, focused on the use of psychoacoustic models such as Steven's Mk7 perceived loudness level (Beranek, 1988) or Moore and Glasberg's time-varying loudness (Glasberg and Moore, 2002) to quantify subjects' judgments of sonic booms.

The first metrics used to quantify the impact of sonic booms on people were based on properties of the sonic boom waveform. Peak pressure amplitude alone was initially used to predict response (von Gierke, 1966). While the peak pressure of a boom was found to be correlated to people's perception of loudness, the rise time and duration of these sounds were also found to be important. One of the studies in which these issues were examined was conducted by (Niedwiecki and Ribner, 1978) who played a series of booms with various rise times, durations and overpressures to subjects. It was found that rise time strongly influenced subject-rated loudness. Duration was also found to affect the loudness of these sounds, but only for booms longer than 250 ms. Models consisting of a function of all of these properties, including additional parameters for shaped sonic boom were also developed (see, for example, May and Sohn, 1990). These models, while capable of predicting the response to a particular type of sonic boom, were limited in that the sonic boom waveform had to follow a certain shape in order for the models to be applicable.

The efforts in determining the impact of sonic booms on people mirrored other research conducted on noise produced by blast and arms fire. This work is relevant because Schomer, Sias, and Maglieri (1997) found that peoples' judgments of blast-noise and sonic booms were similar. Most of this research was conducted by the US Army's Construction Engineering Research Laboratory. As such, this work was focused on the community response to military training noise around installations. In general, in most of this work Schomer compared predictions of A-weighted metrics with penalties for impulsive noise and C-weighted metrics (see, for example, Schomer (1986), Schomer and Sias (1998), and Schomer and Wagner (1995)). However, more recently, some of the data collected in these studies were reanalyzed using Zwicker's stationary loudness model and an "equal-loudness" weighting function (Schomer, Suzuki, and Saito, 2001). It was found that both the "equal-loudness" metric and Zwicker's stationary loudness predicted responses better than the A-weighted metrics (Schomer et al., 2001).

The recent efforts to model the loudness of shaped sonic booms have focused on using models based on the human hearing system. Much of this research was conducted as part of NASA's high speed research program (Leatherwood and Sullivan, 1994) and, more recently, the Federal Aviation Administration (FAA)/National Aeronautics and Space Administration (NASA)/Transport Canada PARTNER Center of Excellence . The research conducted at NASA was focused mostly on the developments of models to predict the loudness level of higher level (49-114 Pa) shaped booms that would likely be generated by a commercial supersonic airliner. This research project differed from earlier research because a thorough examination of the human response to shaped sonic booms was conducted. One boom shaping strategy investigated in this research study was front-shock minimization, which is based around rounding the initial rise times of the boom. A-weighted sound exposure level, Steven's Mk7 perceived level (PL) and Zwicker's stationary loudness were found to be highly correlated to subject responses for front-minimized sonic booms (Leatherwood and Sullivan, 1994). However, all of these metrics were highly correlated to each other, rendering the selection of the "best" metric difficult. Asymmetry in the sonic boom waveform, in which the levels of the peak positive and peak negative pressure are at different levels, was also investigated. Steven's Mk7 PL was found to be the best metric, but it required a correction factor to adjust for boom time-history asymmetry (Leatherwood, Sullivan, Shepherd, McCurdy, and Brown, 2002). Leatherwood et al. (2002) suggested that the correction factor was necessary due to the temporal growth of loudness, which is not accounted for in Steven's Mk7 PL.

In addition to the high speed research program, a single study on low level sonic booms was also conducted at NASA by Sullivan (2004). In this study, participants were asked to judge the loudness of a variety of recorded and simulated low level shaped sonic booms. A-weighted sound exposure level (ASEL), Steven's Mk7 and Moore's stationary loudness predicted subject-rated loudness fairly well ($R = 0.931$, 0.93 and 0.91, respectively). However, these metrics were also highly correlated to each other, as was the case in the front shock study (Sullivan, 2004).

Using the previous low boom study as a guide, Marshall (2007) conducted a series of experiments to examine both the characteristics of sonic booms as well as other impulsive transient sounds. In one of these experiments, a lexicon of words used to describe sonic booms and other transients was assembled. In this experiment, participants heard a series of shaped low booms as well as recordings of distant gunfire, car door slams and thunder. After each sound was played, the participants wrote down words to describe the sound they heard. Once collected, the words were categorized into groups based on their meaning. Words describing the type of sound as well as its spectral and temporal content were the most common words given by subjects.

This lexicon of words was then used to develop a series of word pair scales for a semantic differential experiment. In this experiment, 25 subjects rated a subset of the signals used in the lexicon experiment on 20 word-pair scales. Impulsive sounds other than sonic booms were included to lower the correlation between metrics. A factor analysis of the data yielded four hidden factors: one containing primarily the loudness, annoyance and startle scales, one related to temporal elements (duration, "quickness", etc.), one related to the "naturalness" of the sounds and one related to spectral content. Average annoyance and startle ratings were highly correlated to each other and loudness ratings could not fully explain annoyance ratings alone. This result was evidence that perceptual characteristics in addition to loudness may be important in modeling the annoyance evoked by sonic booms.

Some metrics were also examined. It was found that the maximum value of Moore and Glasberg's time-varying loudness (Glasberg and Moore, 2002) was the best single-number predictor of subject-rated loudness, annoyance and startle. Steven's Mk7 PL, the maximum of Zwicker's time-varying loudness and A-weighted sound exposure level did not perform as well as Moore and Glasberg's model. Specifically, while these metrics were able to predict ratings of the booms well, they tended to predict the ratings of the non-boom sounds on a different trend-line. All of the metrics predicted annoyance and startle ratings less well. It is possible that including additional parameters such as other metrics derived from time-varying loudness models and sound metrics not related to loudness would improve prediction of average annoyance and startle ratings. This analysis will be discussed in Chapter 2.

Startle was found to be a potential issue with higher level booms in studies conducted before the ban on commercial supersonic flight (e.g. Patterson (Sept. 1970), May (1972) and Rylander, Sorensen, Chatelier, Espmark, Larsson, and Thackray (1974)) . Thus, a review of the literature related to physiological responses due to startle, startle-like physiological responses and sonic boom-related startle is presented in the next sections.

1.5 Startle

Given the sudden rise-times in sonic boom waveforms (typically 1 to 20 ms) and the relatively high peak levels, it is not surprising that sonic booms could evoke startle responses. In this section, a review of the general startle response and its properties will be presented, followed by startle research that was focused on sonic boom-evoked startle. While this research is focused on startle responses of neuro-typical adult populations, there has been considerable work focused on startle responses of people with neurological conditions, e.g. dyslexia and post traumatic stress disorder. The reader is referred to the work of Grillon (2002) for more information on this subject. It should also be noted that there are several definitions of startle used by psycho-physiological researchers in different research communities. In this work, the startle response will be defined as the full-body response as observed by Landis and Hunt (1939).

1.5.1 Basic Properties of Startle and its Anatomy

The startle reflex is a physiological response associated with an abrupt exposure to sensory information. Startle responses can be elicited by abrupt vestibular (feeling of falling), acoustic, and tactile sensory information as well as combinations of these presented simultaneously (Yeomans, Li, Scott, and Frankl, 2002). Visual stimuli alone cannot reliably cause startle responses (Yeomans et al., 2002). One of the original works in which a catalog of the basic startle response was given by Landis and Hunt (1939), who were the first to record the startle reflex using film. Landis and Hunt (1939) exposed subjects to the noise produced by small arms fire and recorded the subjects' responses with a movie camera. The main contribution of this work, aside from encouraging additional scholarship, was the observation of the cascade of responses that occurs during a startle response. The startle response originates (as observed in the movies) with involuntary muscle contractions near the face, head and neck. More distal muscles respond later. Landis and Hunt's (1939) "startle pattern" consisted of eye-closure, grimacing, as well as neck and trunk flexion. Brown, Rothwell, Thompson, Britton, Day, and Marsden (1991) and others (e.g. Jones and Kennedy (1951) and Delwaide and Schepens (1995)) have clarified the response properties of individual muscles during startle through the use of electromyography. Brown et al. (1991) exposed 12 subjects to a 50 ms, 124 dB, 1000 Hz tone roughly every 20 minutes until the startle response habituated. The muscles of the neck (such as the sternocleidomastoid) were the muscles which most consistently responded (41-75% of responses) and required the most exposures to habituate. These muscles were also the first to respond, roughly 39-136 ms after the stimulus onset. Muscles of the torso and limbs responded less frequently (8-39% of responses) and occurred later (60-200 ms after the stimulus onset). Of importance, the distribution of these delays (called latencies) for each muscle were not normal; most of the responses were skewed towards the lower-end of the range. The most common muscle response was that of the orbicularis oculi, which controls eye-blink. Earlier startle research, including Landis and Hunt's work (1939) had focused on the presence of eye-blink during a startle response. However, in Brown et al.'s work (1991), the recorded electrical activity of the eye-blink muscles (orbicularis oculi) was significantly shorter in duration and had a different distribution of latencies (time between sound event onset and measured physiological response) than the response of other nearby muscle groups which occured 25-69 ms after the stimulus onset. In addition, the response did not habituate as the other responses measured. Brown et al. (1991) conducted a smaller study with the same stimuli but with only one minute between stimuli and found that there was still eye-blink response despite rapid habituation of other responses. Brown et al. (1991) concluded that the eye-blink was part of the auditory blink response, which is not part of the general startle response.

In addition to the investigation of the prime-movers (or "agonist") muscles associated with the startle response, there has been some investigation of the opposing (or "antagonist") muscles. Delwaide and Schepens (1995) exposed 66 participants to a rectangular-windowed 30 ms, 700 Hz tone burst at a level of 90 dB. One goal of this study was to examine if voluntary muscle contraction influences the acoustic startle response. The electrical activity of three muscles were measured, the trapezius (responsible for neck extension and scapular elevation) and two muscles of the lower leg, the tibialis anterior and soleus. Subjects were exposed to the stimuli three times, once when seated at rest, and once with voluntary contractions of either measured leg muscle. It was found that responses under all three conditions were indistinguishable. Responses in the trapezius were the most common (94-96% of all subjects under each condition) and occurred roughly 50 ms (with a standard deviation of 18.1 ms) after the stimulus onset. These latencies were similar to the latencies of other neck muscle responses found in Brown et al.'s (1991) study. Lower limb muscle responses occurred more rarely: 37-41% of subjects had a response under each condition with latencies around 120 milliseconds. This was similar to the latencies and responses found in Brown et al.'s study (1991).

In addition to the visible outward response, additional physiological responses occur with startle. Francis and Graham (1966) proposed that heart-rate acceleration after a stimulus could be due to a startle response. Additional experiments followed, such as those conducted by Graham and Slaby (1973) and Turpin (1986), which provided further evidence for heart-rate responses during startle. After a startling stimulus, there is a heart-rate increase starting roughly 2 seconds after the end of the sound event. Aside from heart-rate, startle responses include electro-dermal (sweat gland) activity and blood vessel dilation (Turpin, 1986).

The neurological anatomy of the startle response has also been investigated. A basic framework is described here. The acoustic startle response originates in the sensory centers of the brain and is processed by neuro-circuitry located in the brain stem. A diagram of the main startle circuitry, compiled from (Koch, 1999) and (Yeomans et al., 2002) is shown in Figure 1.4. For acoustic startle, the circuit begins at the auditory nerve extending from the cochlea. These are connected to the caudal pontine reicular nucleus (PnC) directly via the dorsal cochlear nucleus or indirectly via one of the many routes that include the ventral cochlear nucleus (Koch, 1999). All of the connections to the PnC from the ear are shared by the conscious hearing system (Niparko, 2009). The PnC is connected to the motor neurons and other stress response systems (Koch, 1999). From results of lesion and electro-physiological studies, the enhancement or inhibition of startle by additional factors, discussed later, is known to be mediated by the PnC (Koch, 1999). However, to complicate this framework, there appears to be processing through additional circuitry that bypasses the PnC (Yeomans et al., 2002). The reader is referred to Koch (1999) for more information.

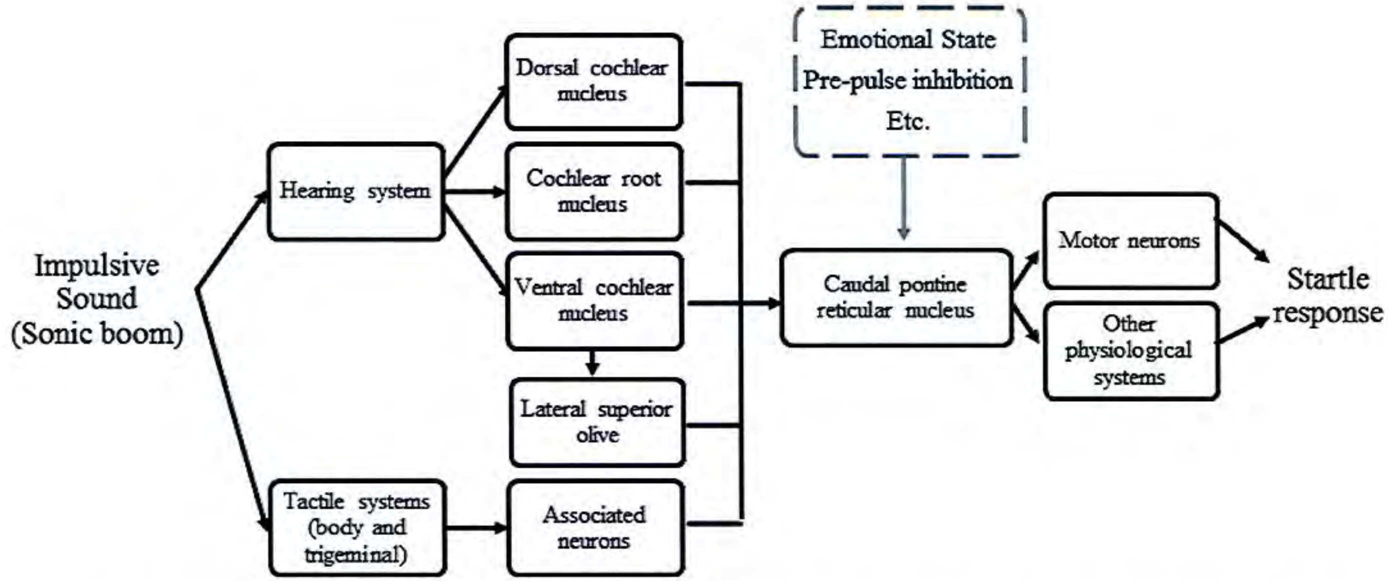

Figure 1.4. Diagram of startle neuro-circuitry. Startle modification circuitry is labeled in box with dashed outline.

In psycho-physiological studies, the sound pressure level (SPL) threshold of the acoustic startle response that is most cited is 80 dB unweighted SPL (Koch, 1999). This number is given assuming a white noise pulse stimulus. One of the main references cited for this threshold was a paper on the startle reaction of rats (Pilz, Schnitzler, and Menne, 1987) where the thresholds of startle for rats were obtained for tones at different frequencies. The minimum SPL thresholds for startle in rats, defined as the level of stimulus where a rat had a 50% likelihood of a response, were around 80 dB when rats were exposed to a 10 kHz tone (the lowest frequency investigated was 4.35 kHz) (Pilz et al., 1987). Aside from determining startle thresholds levels, it was found that the thresholds for startle were parallel with the thresholds of hearing for the rats. Pilz et al. (1987) claimed that this was evidence that acoustic startle has a similar frequency dependence to that of loudness. Startle threshold studies involving humans similar to those conducted on rats were not found in our literature search. The stimulus rise time has also been found to be important, (e.g. Rylander and Dancer (1978); Bjork (1999)), but there appears to be little agreement among researchers on how fast a rise time is needed to elicit startle. In studies with humans, startle responses have been evoked with some consistency when using 60 dB (B-weighted) white noise as long as the stimulus has a quick (around 5 ms) rise time (Turpin, Schaefer, and Boucsein, 1999).

The tactile means of evoking startle, potentially important for sonic boom evoked startle, has also been investigated. In particular, according to Yeomans et al. (2002), electrical thresholds to the trigeminal (facial skin) nerves are an order of magnitude smaller than those associated with the cochlear nerves. Thresholds for skin receptors on other parts of the body were not mentioned. However, it is unclear how important acoustically-excited tactile responses are for startle. The most common method of eliciting tactile startle responses in the literature, via air puffs, is not a pure tactile stimulus there is also an acoustic component. In rat studies, it was found that physiological responses associated with startle were reduced when the rats were deafened (via puncturing the eardrum) compared to when their hearing was intact (Taylor, Casto, and Printz, 1991). Interestingly, Taylor et al. (1991) also found that the lack of acoustic sensation eliminated involuntary muscle activity in the deafened mice. Given that the hearing system appears to be the dominant method of sensing low frequencies, (see Section 1.3), it is likely that startle evoked by booms will be influenced primarily by the hearing system.

One distinction made by the psychophysiology community is the difference between the startle response and other similar responses such as the defense response and the orientation response. Sokolov (1963) was one of the first authors to distinguish between orientation and defense responses. The work is also important for the extensive investigation of skin conductance as a physiological measure. The defense response is defined as a type of stress response that generally occurs with sounds of longer durations than sonic booms (>1 second). Orientation is a less severe response associated with acknowledgement of a novel stimulus (Turpin et al., 1999). Both were found by Sokolov to be accompanied by skin conductance responses. According to Sokolov (1963), the main distinction between orientation and defense is that defense is associated with vascular constriction (as opposed to dilation) and habituates slowly for repeated stimuli. Sokolov did not examine the startle response. As a result, in other experiments, such as those conducted by Graham and Slaby (1973) or Turpin et al. (1999), Sokolov's original definitions of both responses have been modified; these researchers also considered startle in their studies. Classically, orientation is associated with heart-rate deceleration, electro-dermal activity and fast habituation of both responses with repeated stimuli. Startle is associated with involuntary muscle activity, heart-rate acceleration shortly after the stimulus and habituation based on the intensity of the stimulus. Defense is similar to startle, but with heart-rate acceleration that peaks around 30 seconds after the stimulus onset.

This classical categorization of these three responses is not uncontroversial. For example, there is some criticism of this framework and arguments are based on cardiac responses. For a recent example, Ramirez, Sanchez, Fernadez, Lipp, and Vila (2005) claim that heart-rate acceleration that occurs shortly after the stimulus is not part of the startle response, as claimed by Turpin (1999), but actually part of the defense response. They exposed subjects to 105 dB white noise pulses with rise times of 0, 24, 48, 96 and 240 milliseconds with durations of around 1 second. It was found that the stimulus rise time did not significantly influence the magnitude of the short-latency heart-rate responses, but did significantly affect eye-blink EMG magnitude. Given the problems associated with the use of eye-blink as a measure of startle (Brown et al., 1991), it is difficult to assess the claims of this research. In addition, it is unclear how a zero millisecond rise-time stimulus could be generated; this signal probably had a longer rise time. Examining the mean heart-rate response over time observed by Ramirez et al. (2005), it appears that the 0, 24 and 48 ms rise times produce a larger response at 2 seconds than the stimuli with the two other longer rise times. It is possible that shortening rise times to be less than 48 ms does not cause an increase in response and it is possible that rise times of faster than 48 ms are required for startle. This conclusion is also supported by Turpin's (1999) research.

Some authors have attempted to determine the range of stimuli that cause each response, but startle responses and orientation responses have been observed with both high and low level stimuli (Turpin, 1986). In this work, Turpin (1986) suggests that it is possible that orientation and startle responses are, in fact, the output of two separate neurological systems that could act in parallel. Later, Cook and Turpin (1997) created a conceptual model than includes orientation, defense and startle responses as well as anxiety responses and transient stimulus registration (e.g. eye-blink) all in the same framework. The ramifications of such a result are that the responses associated startle and orientation are part of a continuum of responses as opposed to being two separate distinct responses. The latter definition is an implicit assumption in experimental design of most of the sonic boom startle studies (e.g. startle or orientation, not both). However, in defense of the previous boom studies, nearly all of these were conducted prior to the experiments that Turpin conducted and used to develop of this framework. In the context of this research, adherence to a strict differentiation of orientation, defense and startle responses is not of critical importance. If the reader is interested in this topic, the work of Cook and Turpin (1997) is recommended as a starting point. However, understanding the distinction between startle and orientation will be advantageous when attempting to understand the correspondence between physiological responses and ratings by subject on a "startle" scale.

1.5.2 Startle Modification

One area of active research in the psychophysiology community is aimed at gaining an understanding of how the startle response can be modified by additional factors. One of these factors is the presence of additional stimuli with emotional content, such as pictures, which manipulates the "affective state" of the person. For example, Bradley and Lang (2000) exposed subjects to pictures of various content while also exposing the subject to startling audio stimuli. They found that startle responses in the presence of more "extreme" pictures (such as spiders or mutilated corpses) were more severe than those responses to the same stimuli in the presence of "calming" pictures such as flowers. A similar study was conducted by Hawk and Cook (1997) who examined air puff (acoustic+tactile)-evoked startle responses when subjects were also exposed to pictures and they found similar results to Bradley and Lang.

Relatedly, pre-conditioned fear and ambient light levels have also been found to influence the severity of startle responses. Fear conditioning has been shown to reliably enhance startle responses and can be used to measure the magnitude of fear conditions (see, for example, (Spence and Runquist, 1958) and (Grillon and Davis, 1997)). Darkness has been shown to enhance the startle responses in humans (Grillon, Merikangas, Dierker, Snidman, Arriaga, and Kagan, 1999) and bright lights have been shown to enhance the startle response in rats (Walker and Davis, 1997), one of the few areas where rat and human startle responses differ. It is argued by Grillon and Baas (2003) that this difference is due to evolutionary adaptation. Rats are nocturnal creatures and are, thus, sensitive to bright lights. Darkness-enhanced startle is also different from the other forms of fear conditioning as it requires no learning or memory (Grillon and Baas, 2003).

Startle responses can also be modified by the level of background noise present during the startling stimulus. May (1971), who later was responsible for some of the early sonic boom startle research, exposed subjects to 40 ms white noise pulses at sound pressure levels of 105, 112, 114, and 117 dB (unweighted) under different background noise conditions (45 dBA, 72 dBA and 74 dBA). Startle was measured using performance during a line-tracking task and each participant was exposed to each stimulus at one background noise level. It was found that the task was more disrupted when participants were exposed to the stimuli when lower background noise levels were present.

Another important pair of startle modification phenomena are pre-pulse inhibition or facilitation. Either of these phenomena occur when a less intense event (called a "pre-pulse") precedes a startling stimulus (Schmajuk and Larrauri, 2006). Depending on the background noise, the intensity of the pre-pulse and the interval between the pre-pulse and the main event determines if the startle response will be larger or smaller than when the main stimulus is presented alone (Schmajuk and Larrauri, 2006). The inhibition effect occurs when the prepulse acts as a subconscious cue for the body to prepare for the startle response. Roughly, maximum inhibition is said to occur when the prepulse event occurs an interval of 50-100 ms before the onset of the main noise event. A pre-pulse occurring less than 50 ms before the main event can cause facilitation (Plappert et al. (2004) and Reijmers and Peeters (1994)). Inhibition has not been found to occur when the pre-pulse stimulus occurs more than 3 seconds before the main event onset for low level prepulses, but the separation between the pre-pulse and main event must be longer to suppress the inhibition responses for higher-level (>80 dB) prepulses or other startling stimuli (Swerdlow, Shoemaker, Stephany, Wasserman, Ro, and Geyer, 2002). Disruption of background noise, both before and during the startling stimulus can also cause pre-pulse inhibition (Schmajuk and Larrauri, 2006). However, pre-pulses only inhibit temporally-adjacent startle responses. It has been found that the responses to additional events that cause a startle response are unaffected by the original response and only by the most recent response (Swerdlow et al., 2002). In addition, the more intense the responses evoked by a stimulus, the less noticeable the pre-pulse inhibition effects are.

Studies have been conducted to examine the effect of age and gender on pre-pulse inhibition. However, there appears to be a lack of agreement among researchers as to the extent of these differences, especially those based on gender. Kumari, Aasen, Papadopoulos, Bojang, Poon, Halari, and Cleare (2008) have found that women have less pre-pulse inhibition than men, while Ludewig, Ludewig, Seitz, Obrist, Geyer, and Vollenweider (2003) have shown no difference between women and men. Complicating matters is that Kumari et al.

(2008) misrepresents the number of subjects used for the gender comparison in Ludewig et al. (2003) when discussing and ultimately dismissing their results. However, the data of both studies support the finding that there are no gender differences between startle response magnitudes and that the rate of habituation of startle responses tends to increase with age.

In the broader context of this research, pre-pulse inhibition may be important because the temporal spacing (50-100 ms) associated with pre-pulse inhibition is similar to the interval between loudness peaks of shaped sonic booms. It is possible that if these effects are important in predicting people's responses, they could be exploited to yield a sonic boom that is less startling for a given loudness. It may also be important to consider pre-pulse inhibition effects when designing experiments. Recall that changes in background noise can act as a pre-pulse. This means that to avoid unintended inhibition of startle responses, background noise must be strictly controlled during experiments.

1.5.2.1. Pre-pulse Inhibition Models

Schmajuk and Larrauri (2005) have created a mathematical model of startle focused on pre-pulse inhibition or facilitation. This model was created as an aid to understanding the neural circuitry of pre-pulse inhibition and facilitation, and has been shown to predict results of previous investigations into pre-pulse inhibition and excitation and its effect on the startle response. To aid the reader in understanding pre-pulse inhibition and facilitation, a brief description of the model and some example calculations are discussed here. The output of this model is "startle", in units unique to the model. This output has been fitted to the results of rat and human studies, including eye-blink magnitude, neck muscle responses and other physiological measures to determine model parameters. The input of this model (D_0) is the unweighted sound pressure level in decibels of white noise, as measured by a fast-averaging sound level (Schmajuk and Larrauri, 2005). It is unclear from the paper what the bandwidth of the instrumentation used to produce the "white" noise is because the instrumentation details are not given. However, given the variety of different methods to present "white" noise used in the research on which this model is based, it is possible that the model is not sensitive to the bandwidth of the "white" noise input assuming normal audio sampling rates (> 44.1 KHz). The reason for the specialized input is that this model was developed from studies where only white noise pulses were used as stimuli. One way to calculate the input to this model for an arbitrary sound is to create amplitude-modulated white noise sounds with the same loudness time-history as the original sound. Then the fast-averaged sound pressure level can be used as a surrogate for the original signal in Schmajuk's model. Details of specific methods for creating these sounds are described in Chapter 5.

In between the input (D_0) and output (S) are several internal variables in the model, as illustrated in Figure 1.5. These variable are denoted by those associated with startle and excitation (E_1 to E_6) and those associated with inhibition (L_1 to L_5). Larger subscript numbers denote internal variables that are closer to the output. The model consists of two main paths: the facilitation-excitatory path and the inhibition path (Schmajuk and Larrauri, 2005). The former path, which contains the neural "circuitry" of both the startle response and pre-pulse facilitation, begins by transforming the input signal (D_0, in decibels ref 2×10^{-5} Pa) into root-mean-squared pressure and normalizing it:

$$E_1 = p_{ref} 10^{D_0/20}, \tag{1.2}$$

$$E_2 = k_1 E_1/(E_1 + \beta_O), \tag{1.3}$$

where p_{ref} refers to the reference pressure ($2x10^{-5}$ Pa) and k_1 and β_0 are constants. The purpose of this part is ensure a that the range of E_2 is bounded from 0 to 1 and that its value increases drastically when E_1 is larger than 90 dB (Schmajuk and Larrauri, 2006). This value is an input to a first order differential equation of state E_3 given by:

$$dE_3/dt = k_2(E_2 - E_3), \tag{1.4}$$

which is a low pass filter whose output E_3 is a smoothed version of E_2. This state is compared to E_2 to yield E_4 , using the following:

$$E_4 = \begin{cases} E_2 - E_3 & \text{if } E_2 > E_3 \\ 0 & \text{if } E_2 < E_3 \end{cases}. \tag{1.5}$$

Note that the combination of equations (1.3) and (1.4) is a high pass filter followed by a half wave rectifier. State E_4 is an input to both the startle equation and the facilitation path. The latter is modeled as a nonlinear fast changing function of the excitatory path (E_4) followed by another smoothing, low pass filter as shown in the following equation:

$$E_5 = k_3 E_4^m / (E_4^m + \beta_1), \tag{1.6}$$

$$dE_6/dt = k_4(E_5 - E_6), \tag{1.7}$$

where k_3, k_4, β_1 and m are constants. The output of the model (S) is then calculated from the following equation:

$$dS/dt = E_4^p(1 + k_8 E_6) - S k_2 L_5 - S k_{10}, \tag{1.8}$$

$$S' = S^q / (S^q + \beta_4), \tag{1.9}$$

where k_8, k_9, $k_1 0$, β_4, p and q are constants and L_5 is the contribution from the inhibition path of the model. L_5 is calculated from L_1 which is equal to the input signal in dB (D_0). It is passed through a low pass filter to create L_2 and the next state, L_3, is calculated by comparing L_1 to slow changes in itself (L_2):

$$dL_2/dt = k_5(L_1 - L_2), \tag{1.10}$$

$$L_3 = |L_1 - L_2|. \tag{1.11}$$

Note that the net effect of equations (1.9) and (1.10) is a high-pass filter followed by a full-wave rectifier. To obtain L_5, L_3 is normalized and low-pass filtered:

$$L_4 = L_3/(L_3 + \beta_2), \tag{1.12}$$

$$dL_5/dt = k_7(L_4 - L_5). \tag{1.13}$$

All of the parameters of this model were selected by fitting the model predictions to data obtained from a series of pre-pulse inhibition and startle studies in the literature (Schmajuk and Larrauri, 2006). However, it is unclear from that paper what method was used to estimate the parameters values.

Schmajuk and Larrauri (2006) simulated the series of equations using a sixth order Runge-Kutta-Verner method to predict the response at 1 millisecond time increments. Schmajuk and Larrauri provided Fortran code that simulates the output of their model, and that was used in the research described later. The reader is referred to (Schmajuk and Larrauri, 2006) for more details.

While the system of equations that compose the model are complicated, the qualitative behavior of the model is simple. For a qualitative example, let us assume that the input to the model is a simple short pulse with constant background noise and no pre-pulse event. For a simple pulse, the effect of the excitation path (Equations 1.3 through 1.5) is isolated because other parts of the model do not respond fast enough to contribute significantly to the response. Ignoring the nonlinear normalization, the high frequencies of the input are emphasized because of the high pass filter. The high-pass filter behaves roughly as a differentiator well below the cut-off frequency of the low-pass filter. At higher frequencies than the cut-off, the magnitude of the response is flat with frequency. For low frequency oscillations in the input, this part of the model behaves as a differentiator which only outputs positive values (due to equation 1.5).

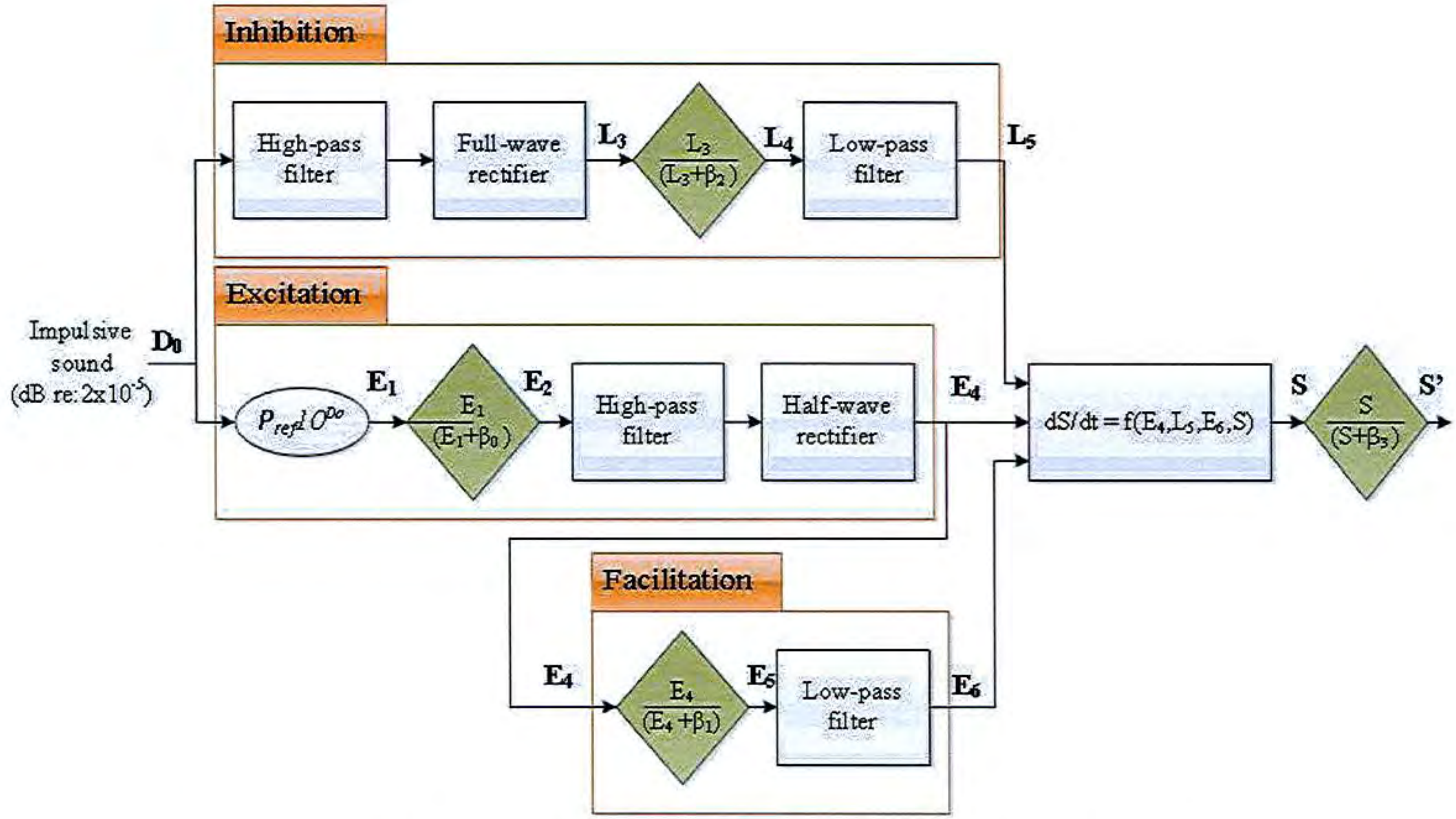

Figure 1.5. Signal flow of prepulse inhibition model of Schmajuk and Larrauri (2006).

To show a quantitative example of the output of this model for this simple case, a 10 ms duration, 80 unweighted dB white noise pulse was created using *randn* with a sampling rate of 32000 Hz in Matlab and it fast-averaged sound pressure level was input into the pre-pulse model. The response is shown along with the output of Moore and Glasberg's time-varying loudness model (Glasberg and Moore (2002)) in Figure 1.6. The instantaneous and short-term loudness (a) along with the derivative of the short-term loudness (b) are plotted. The derivative was calculated by using a 51-point finite impulse response filter that behaves as a differentiator from frequencies of 0-300 Hz and a low-pass filter at higher frequencies (200-500 Hz). Instantaneous loudness is not consciously perceived and short-term loudness is a prediction of the loudness that is perceived as the sound is heard (Glasberg and Moore, 2002). Aside from the differences in scales (Schmajuk's model output is measured in units of startle unique to the model), it is noticeable that the startle model increases and decreases at a more rapid rate than that for short-term loudness, but increases at a rate similar to instantaneous loudness or that of the derivative of short-term loudness. This difference is consistent with findings reported in the literature. Startle responses have been found to be about an "order of magnitude faster" than perceptions of loudness when the time constants of temporal summation are considered (Yeomans et al., 2002). However, there is a correlation between the maximum output of the startle model and the maximum of the short-term loudness derivative prior to the first peak in loudness. A series of white noise signals with similar loudness time-histories as sonic booms were created and their maximum short-term loudness derivative and the maximum of the output of Schmajuk's model's time-history were calculated and plotted against each other. This is shown in Figure 1.7. Maximum loudness derivative was determined from the initial rise of the short-term loudness time-history. Maximum loudness derivative and the maximum output of Schmajuk's model are highly correlated to one-another for this set of signals ($R^2 = 0.83$).

A practical example of where pre-pulse inhibition could be important is when a change in the level of background noise precedes the main sound during an experiment. This change may be due to artifacts in the playback system or the result of varying levels of background noise that are present in signals that are recorded in the environment. A schematic

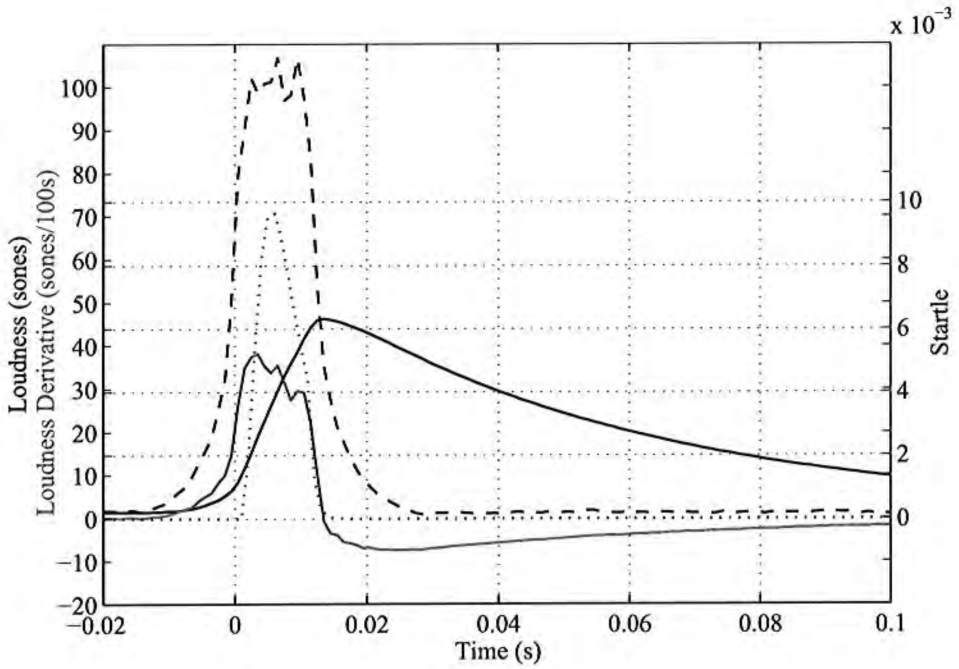

Figure 1.6. Output prepulse model (short dash, black), instantaneous loudness (long dash, black), short-term loudness (solid black) and loudness derivative (gray solid) for a 50 ms, 80 dB white noise pulse.

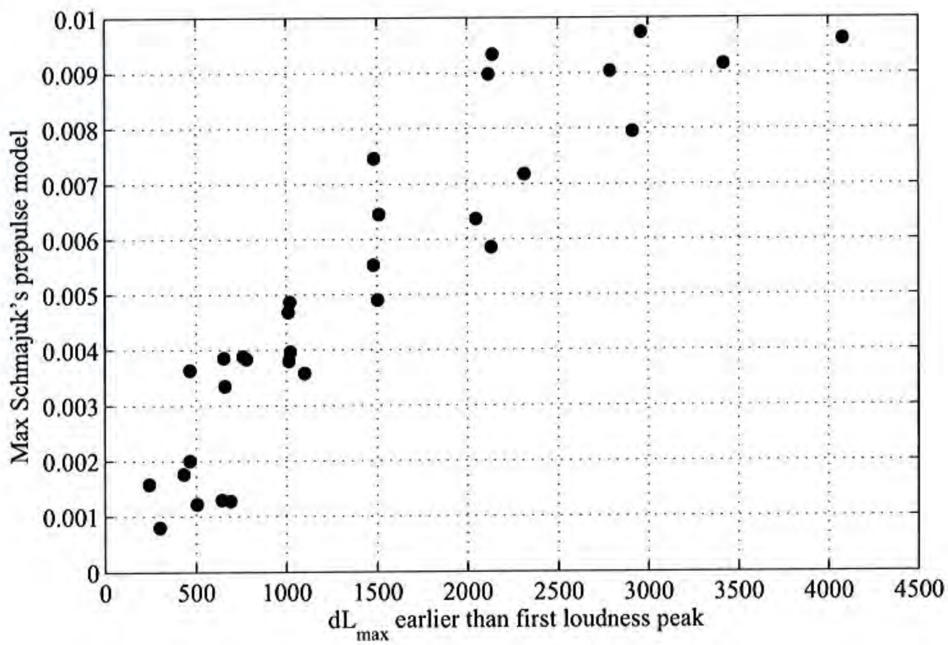

Figure 1.7. Maximum startle output of Schmajuk's model versus maximum short-term loudness derivative before first loudness peak (R^2 = 0.83) for boom-like white noise signals.

of this scenario is in Figure 1.8(a). To determine what would happen in this case, a simulation was conducted using this model. The output of Schmajuk's model was calculated for a 50 millisecond-duration, 90 dB white noise pulse with a 1 millisecond rise-time with 0-30 seconds of "recording" background noise at levels of 45, 50, 55, and 65 dB preceding it. Ten seconds of background noise preceded the "recording" and a half second of "recording" noise followed the pulse signal. The base background noise was assumed to be 35 dB. The maximum output of Schmajuk's model was then calculated for each case and normalized by the maximum output when no background noise changes are present. Thus, a value of 1 would be the same magnitude as expected from the pulse being played with constant 35 dB background noise. The results of this analysis are shown in Figure 1.8.

As expected from the study conducted by May (1971), lower recording noise levels result in larger startle responses. For the 45, 50 and 55 dB levels of recording noise the startle response is facilitated by its presence when $Tdur$ is less than 60 ms. For the highest level of recording noise (65 dB) there is very little facilitation and mostly inhibition for $Tdur > 20$ ms. For all levels of recording noise inhibition occurs when $Tdur > 60$ ms and for $Tdur >= 100$ ms responses are constant with lower levels of recording noise yielding larger responses. Facilitation is greatest at $Tdur = 30$ ms. Given the possible effect of these pre-event cues on the startle response, the background noise in startle experimentation needs to be highly controlled.

1.5.3 Sonic Boom Evoked Startle

Much of the previous sonic boom research on startle was conducted in the 1960s and 1970s (Thackray, 1972). Concerns about startle evoked by sonic booms stem from the early community noise surveys see, e.g., (Nixon and Borsky, 1966), in which participants claimed to be startled by these sounds. In the majority of this startle research, subjects' evaluations of startle or subjects' performance of a task were recorded. Only a few researchers investigated physiological responses for boom-related startle, see, e.g., (Thackray et al., 1974). The advantage of subjective measures of startle or annoyance is that they

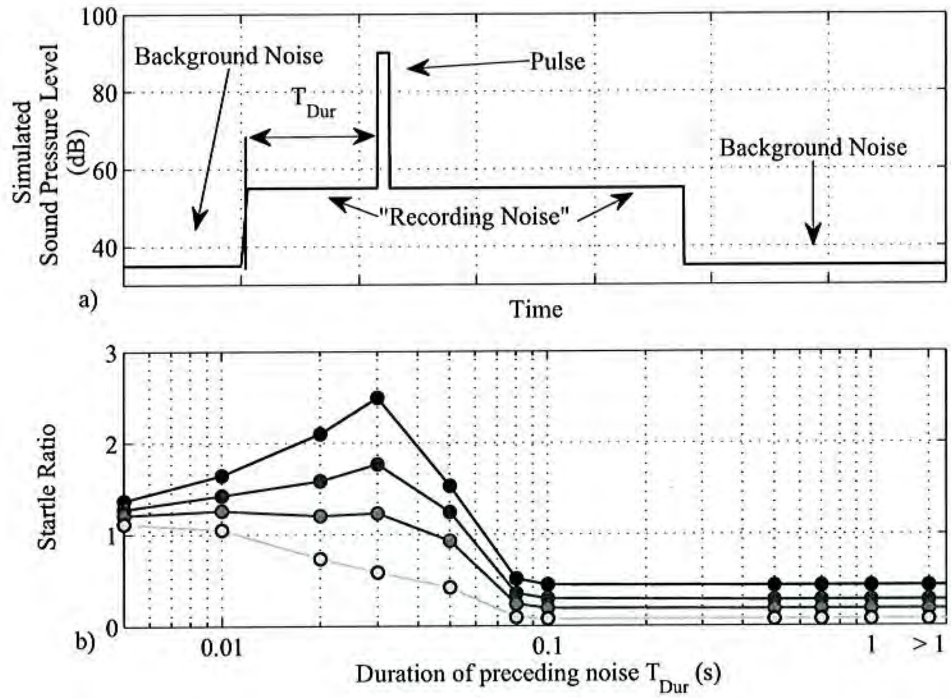

Figure 1.8. a) Schematic of sound presentation scenario for simulation of a 90 dB pulse with recorded background noise. b) Maximum startle due to 80 db pulse with recording noise present normalized by maximum startle when there is no recording noise, plotted against duration of recording noise prior to pulse event ($Tdur$). Results predicted using Schmajuk and Larrauri's (2006) model for recording levels of 45 (black), 50, 55 and 65 dB (lightest gray).

are measures of the subjects' perception of startle. While this statement might appear facetious at first, it becomes less so when one analyzes real physiological signals. Physiological signals can be measured more continuously and can capture subconsciousness response behavior. However, these measures also capture physiological processes that are not related to the stimulus presentation. This means that one part of the processing of physiological measures is identifying artifacts in the signal due to movement or responses to other non-acoustic stimuli. While the sources of variation present in subjective measures are more subtle, the act of sorting the "response" from the background is usually not necessary. The advantage of using both kinds of measures in human subject testing is that, when combined, the physiological measures may give insight into what people are responding to when they decide on the startle rating.

One of the critical factors that limited early research into startle evoked by sonic booms was limitations in sonic boom playback. Thus, in much of the early boom startle research real aircraft overflights were used to generate stimuli. May (1972) made one of the first attempts to relate subjective ratings of startle to overflight sonic boom characteristics. May had subjects estimate "how jumpy" the booms make them while going about their normal life (May, 1972). Subjects were exposed to overflights of an F-104G Starfighter aircraft over the course of ten working days. The wording of the paper implied that these overflights were part of some military exercise and their locations, magnitude and frequency were not directly controlled by experimenters. The booms were reported to have, on average, a peak overpressure of 75 Pa and rise-times from 2 ms to 39 ms. May developed a model that was a function of boom overpressure (P_{peak}) and rise-time (t_{rise}) and determined model parameters by fitting the model to the response data:

$$S = 1.55 \times 10^{-5}(P_{peak}^2/t_{rise}) + 0.166t_{rise} + 9.05, \tag{1.14}$$

where S is the prediction of subjects' responses. However, he cautioned against any conclusions about the general utility of the model since the population was inadequately sampled (May, 1972).

Another large field study was conducted by Rylander, Sorensen, Chatelier, Espmark, Larsson, and Thackray (1974). They examined the effects of sonic booms on humans and animals. Participants and animals were exposed to 5-12 booms and an equal number of subsonic overflights per day from a Saab Draken J-35 aircraft for a total of 47 boom exposures. Subjects were indoors when exposed to sounds. Booms ranged from 60-640 Pa peak pressures outside and 20-130 Pa inside. The construction of the building and its sound transmission properties were not discussed in detail. Rise times varied from 0.1-12.5 ms with most around 2 ms. Startle responses were measured via a hand steadiness apparatus, a tracking task, or via heart-rate measurements. The results of the hand-steadiness task and video recording showed that many subjects had visible startle responses to most of the stimuli. The frequency of startle responses increased as the boom overpressure increased (roughly 20% of subjects exhibited startle responses for the 100 Pa booms and up to 100% for the 300+ Pa booms). Notably, heart-rate acceleration was also evoked by the booms. Subject-to-subject variation in heart-rate was not discussed. A model was developed from the subject ratings from this study by Chatelier (1974):

$$K = 10log_{10}[P_{peak}^2/sqrt(t_{rise})] - 30, \quad (1.15)$$

where K is the prediction of subjects' responses. This model, like the one developed by May (1972), contains a peak pressure squared over rise time term. The lack of the additional rise time term in Chatelier's models is probably due to the lack of rise-time variation in the stimuli in that study.

In only a few of these studies were sonic booms at a level that remotely resembles what is termed a "low level" sonic boom. Rylander and Dancer (1978), for example, examined booms with up to a 200 Pa overpressure, roughly eight times greater overpressure than that of expected low booms. Aside from the study of Thackray, Touchstone, and Bailey (1974), in no study was a boom of overpressure less than 75 Pa used as a stimulus. In addition, the models developed from this data are composed of functions of pressure time-history characteristics of sonic booms (typically overpressure and rise time). The limitation of these types of models is that they were developed using stimuli of limited variety, most were simple N-waves. These models may not be appropriate for predicting responses to more complicated signals, such as shaped sonic booms or sonic booms that have been propagated through turbulence. Also the parameters used are sound pressure-based, i.e., they are not based on how sound characteristics are perceived. A possible improvement would be to use models of loudness (how humans perceive sound intensity) and generate loudness time-histories. Then derive parameters from theses such as maximum loudness and rate of change of loudness.

Thackray, Touchstone and Bailey (1974) examined both physiological and subjective responses in one of the few studies in which low overpresssure sonic booms were used as stimuli. In their experiment, they recruited thirty subjects and exposed them to sonic booms with overpressures of either 16, 30 or 50 Pa, all with approximate rise-times of 4 ms and durations of approximately 210 ms. The sonic boom waveforms were generated by using an electrical device which is not described in their paper (the booms used were not recordings). The N-waves were produced by this device and played to subjects by using an electric loudspeaker. The subject pool was split into 3 groups, one for each boom overpressure. Each subject in each of the three groups was exposed to sonic booms at no greater a frequency than 1 every 5 minutes. The subject heard a total of six booms for the entire test. After the boom exposures, the subjects were exposed to a 0.22 caliber pistol shot (103 dB peak measured using a fast setting on a sound level meter) and asked to rate the sonic booms heard previously. The subject was asked to give a number quantifying the "startle", given the reference that the pistol shot was rated at 100. The physiological measures examined were arm steadiness (via a stylus), heart-rate, eye-blink and skin conductance. They classified a response as a startle response if an eye-blink response occurred within a certain amount of time after the boom exposure. This time, called a response latency, was found by using the latencies observed from the pistol shot data. A "minimum" startle response was defined as only an eye-blink response, while a "normal" startle response was defined as an eye blink response plus a deviation in hand steadiness (Thackray et al., 1974). Interestingly, heart rate changes and skin conductance changes were not utilized in determining startle. It was found that the number and severity of "startle" responses to 30 Pa booms and

50 Pa booms were statistically indistinguishable. For the 16 Pa booms, very few "startle" responses were observed and all of these were categorized as "minimal startles". The skin conductance response magnitude was reduced with later exposures, which was attributed to habituation effects.

In general, the arm-steadiness and eye-blink measurement showed little evidence of habituation effects as the number of boom exposures is increased (Thackray et al., 1974). The lack of eye-blink habituation was consistent with the results observed by Brown et al. (1991). Heart-rate deceleration decreased systematically with additional boom exposures, which is also evidence of habituation. The base-line level of the skin conductance appeared to decrease over the course of the experience, but this did not affect the amplitude of these responses after boom exposures. For the other physiological measures and subjective measures there were no significant differences between exposure groups (Thackray et al., 1974), although 16 Pa booms were rated as slightly less annoying than 30 Pa and 50 Pa booms. The authors used the results of this analysis to conclude that booms of less than 30 Pa (possibly less than 16 Pa) would be required to avoid startle responses.

The lack of significant differences in skin conductance, heart rate, and subjective responses between each level of stimuli is likely a product of each subject being exposed to only one level of sonic boom stimulus. The decision to present the stimuli this way was done because Thackray et al. (1974) wanted to examine habituation effects. The only common sound exposure among all subjects was the pistol shot, which was much louder and had spectral characteristics that were different to those of the boom stimulus. It is possible that this difference in level was large enough to increase the difficulty of making judgments of startle for the sonic boom exposures.

1.5.4 Thresholds of Startle: Booms and Other Signals

Assuming that the conclusions of Thackray et al.'s (1974) work are sound, this would mean that a dividing line (if it exists) between non-startling and startling booms exists at a level between the 16 and 30 Pa overpressure booms. To compare this range to the range investigated in previous work conducted by the author, the maximum value of Moore and Glasberg's time-varying loudness model was calculated for estimates of the 16 and 30 Pa overpressure boom time-histories used in Thackray et al.'s 1974 experiment as well as for the noise pulses used in the psychophysical research discussed earlier. Utilizing Zwicker and Fastl's time-varying loudness model gives the same result. The boom waveforms were generated by using a frequency domain sampling technique often used in finite impulse response filter design. In Table 1.1, the metric values for the sonic booms used in previous studies and the stimuli used in psychoacoustic studies conducted at Purdue previously are shown. The range of values covered with stimuli used in previous experimentation extends across the range used in Thackray et al.'s 1974 experiment and the psychophysiological studies. The stimuli characteristics (pressure peaks and rise times) also cover the range for stimuli used in the psychophysiological research. Given this, it is possible that startle responses could have occurred with the range of low booms considered.

Table 1.1 Maximum of Moore and Glasberg's time-varying loudness of low-level signals that have been found to evoke startle responses compared to range used in author's previous study.

Author	Stimuli	Startle Responses by	Maximum Loudness
Thackray et al. (1974) Turpin et al. (1999) Pilz et al. (1987)	N-waves 60 dB-B Noise pulse 80 dB Noise pulse	hand-steadiness, blink heart-rate, EMG EMG	36.0 sones 19.2 sones 50.9 sones
Marshall (2007)	low N-waves and other transients	subject judgments	11-58 sones

1.6 Methods

In this section, a review of some of the physiological measures and their associated signal processing techniques used in startle research will be presented. The physiological measures used in this research are skin conductance, pulse-rate and electromyography (EMG). Some of the experimental designs used to gather subjective response data featured repeatedly in this research, such as the semantic differential (Osgood and Suci, 1955), will not be discussed here. The reader is referred to the author's MS thesis (Marshall, 2007) for an overview of these methodologies.

There are several key concepts in the processing and analysis that are common to all of these physiological measures. The main one is the concept of response latency. When a physiological response is evoked by some stimulus, there is a time delay before the physiological response changes from its no stimulus level. This delay is due the speed of the physiological systems involved in the response. The latency of a physiological signal is defined as the time delay between the stimulus exposure onset and the onset of the response. An illustration of two commonly used definitions of signal latency is in Figure 1.9. For real signals, either the delay is calculated based on the time after the event onset when the physiological signal breaches some threshold, called the onset latency, or the time after the event onset when the signal has reached a maximum, which is called the peak latency. The particular type of criterion used when calculating the latency is specific to the signal being measured. Signal-response latencies are important because they can be used to identify the type and origin of a particular response. Muscle activity, for example, can be due to voluntary movement or due to startle. By comparing the signal latencies to those observed in more controlled startle studies (such as Brown et al. 1991), the signal latency can be used to determine if a particular response was likely evoked by a startling stimulus.

Another application of latencies is to use them to examine the underlying physiological processes involved in producing the response. In rough terms, the longer response latencies are associated with longer and more complex neural circuits (Davis, Gendelman, Tischler, and Gendelman, 1982). By analyzing the distribution of latencies observed, observations can be made regarding the type of physiological process associated with its evoked response. This approach is analogous to experimental methods in other areas of acoustics, such as in ultrasonics. If the sound propagation path is known, one can estimate the sound speed of that path by knowing the time it took for the sound to propagate. Similarly, a latency is the speed of neurological "propagation" including the response apparatus (e.g. muscles). While the research presented in this dissertation was not focused on investigating the particular neurological geography of startle responses, the results of those studies have been used in the analysis of the physiological signals measured in the studies reported in this thesis. A summary of typical startle response latencies for the measures used in this research is in Table 1.2. A description of the three types of measurements is presented below.

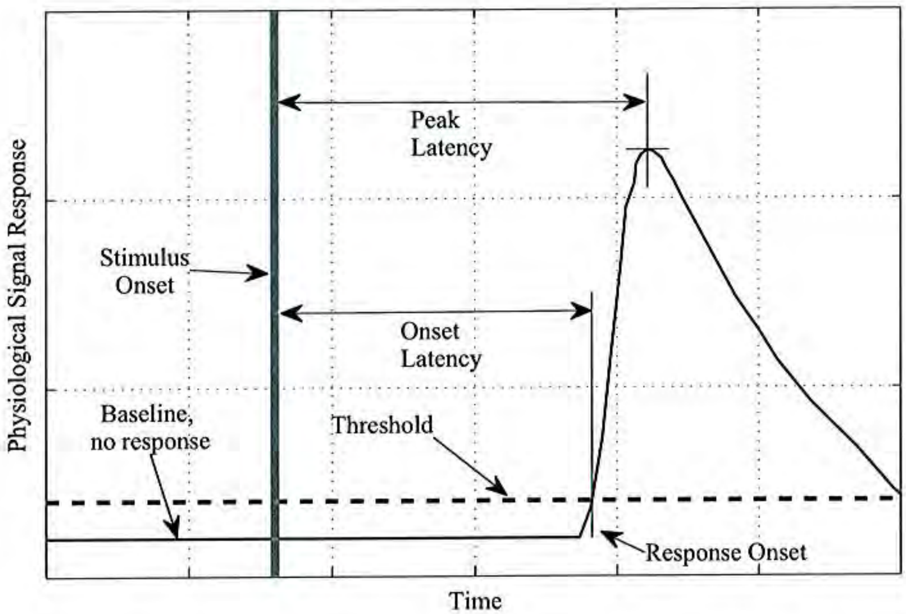

Figure 1.9. Illustration of onset and peak latency definitions for an arbitrary response to an event (gray vertical line).

Table 1.2 Summary of typical response latencies for physiological measures used in this research. Values tabulated from (Brown et al., 1991); (Turpin et al., 1999) and (Dawson et al., 2000)

Measure	Latency Type	Latency Range
Neck muscle EMG	Onset	39ms-136ms
Skin conductance	Onset	1-3 s
Heart rate (increase or decrease)	Peak	$>$ 1-2 s

1.6.1 Skin Conductance

One of the oldest and mostly commonly examined physiological measures is that of skin conductance. Known in the early literature as galvanic skin response, the first studies of this measure were conducted as early as 1878 (Dawson et al., 2000). This measure is the change in conductivity through the skin in response to some stimulus. The physiological origin of this response is the eccrine sweat glands found on the palms of the hands, which are called palmar surfaces (Dawson et al., 2000). While the primary function of most eccrine sweat glands is regulation of body temperature, the glands of the palms are more responsive to emotion-evoked stimuli than thermal stimuli (Dawson et al., 2000). When the skin is dry and the sweat glands are closed, the conductance is low. However, when the sweat glands dilate (due to some event), the sweat is allowed to flow to the surface of the skin and conductivity is increased. The sweat glands are controlled by the sudomotor nerves whose responses have been found to be strongly correlated with sympathetic nervous system activity (Alexander, Trengove, Johnston, Cooper, August, and Gordon, 2005). Thus, it is not surprising that skin conductance has been found to be an index of stress and attention, particularly orientation responses (Sokolov, 1963). Sokolov (1963) also observed that pain induced skin conductance responses.

Skin conductance measures can be divided into two major components based on the speed of responses observed, the skin conductance level (slow response characteristics) and the skin conductance response (fast response characteristics). An analogy often used in the literature is that the "skin conductance responses are the small waves superimposed over the tidal drifts of the skin conductance level" (Dawson et al., 2000). Skin conductance responses are also further divided into two main categories: specific and non-specific responses. Specific responses are those that are attributed to identifiable stimuli and are the part of the response associated with orienting. Non-specific responses, also called spontaneous responses, are responses that are due to background physiological events.

A typical skin conductance response to a sound is in Figure 1.10. The expected onset latency of a specific skin conductance response is expected to be between 1 to 3 seconds after the stimulus (Dawson et al. (2000); Kucera, Goldenberg, and Kurca (2004)). Most researchers examine the maximum response amplitude, but the rise time and half amplitude recovery time have also been used in some studies (Dawson et al., 2000). In addition, some researchers examine the presence or absence of a response (Kucera et al., 2004).

Note that during a skin conductance response, the signal is slow to return to its base-level, i.e., it has a very slow recovery time. After a response peak, it takes roughly 2-10 seconds for the signal to reach a value of half its peak magnitude (Dawson et al., 2000). The long duration of these responses can lead to response overlap if stimulus exposures are too close together. Aside from experimental means to remedy this, several methods have been proposed to isolate overlapped responses using signal processing techniques. Alexander, Trengove, Johnston, Cooper, August, and Gordon (2005) have proposed a method that uses an estimated model of the sudomotor nerve behavior. The method is based on the assumption that skin conductance response results from discrete events that cause the sudomotor nerves to fire. It is also assumed that the input signal to the sudomotor nerve is much faster than the measured skin conductance response. The authors also assume that the impulse response of the path from the sudomotor nerve to the skin conductance response can be modeled as the sum of two exponential functions. The authors then estimated the model parameters using a large database of skin conductance responses (Alexander et al., 2005). The time constants found were 2 and 0.75 seconds. To isolate overlapped skin conductance responses, the following procedure was used. First, an inverse filter was designed that removed the dynamics associated with the sudomotor nerve, yielding the estimate of the underlying input signal to the sudomotor nerve, called the "driver signal" (Alexander et al., 2005). Then the driver signal responses are isolated and each response is put through the sudomotor nerve motor model to estimate individual skin conductance responses. From these responses latencies and other measures can be calculated.

1.6.2 Pulse-rate

Another type of physiological response that is associated with startle is the response of the cardiac system. Most people are familiar with the "fight or flight" response to extremely

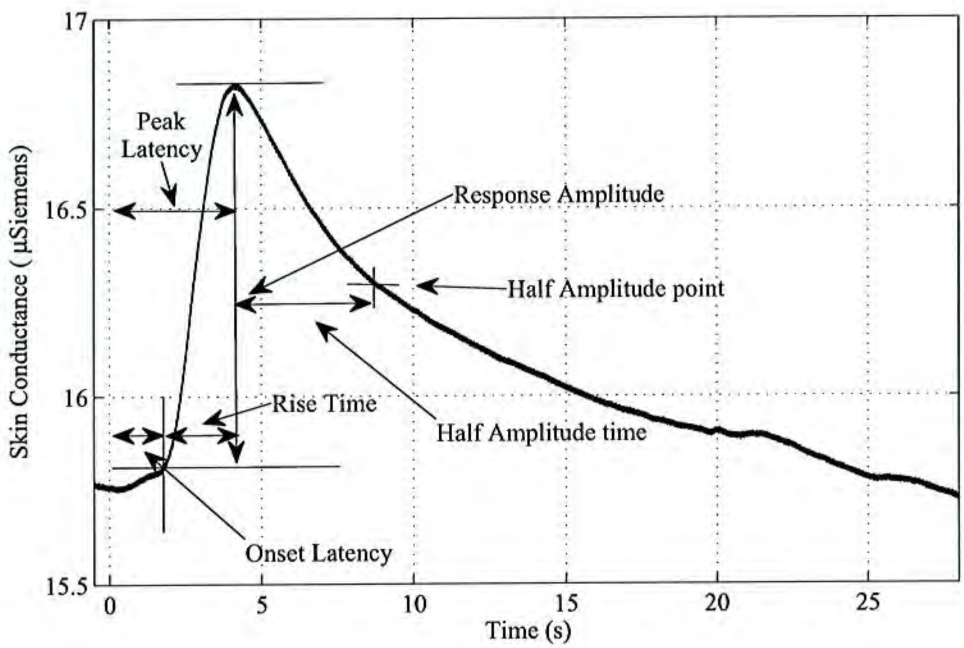

Figure 1.10. A typical skin conductance response to a sonic boom which started at t = 0 seconds. Relevant response properties are labeled.

stressful stimuli and associated heart-rate increase. There is also a similar reflex called the "vigilance" response, that is characterized by a slowed heart-rate, despite its being associated with stress.

Heart activity can be measured indirectly by using an optical sensor at the finger, or directly by using an electrocardiogram (ECG). As ECG was not be used in the research described in the following chapters and so processing of such signals will not be discussed here. See (Rangayyan, 2002) for more information. The optical sensor works by measuring the change in infrared reflectivity as the blood moves through the finger. While heart rate calculated from this signal is similar to that obtained from ECG, some particular differences between these measurements need to be addressed. It has been found that finger-pulse derived heart-rate is about a heart-beat behind the heart-rate derived from an ECG signal (Barry and Mitchell, 1987). This delay is due primary to the transit time from the heart to the arteries of the finger. Because many of the latencies calculated for heart-rate responses were derived from ECG-measured heart rate, this delay must be included in the processing if finger pulse heart-rate responses are to be correctly analyzed.

Aside from systematic differences between finger pulse and ECG measurements, there are noise issues in the finger pulse measurement. First, since the optical sensor is attached to the finger, movement of the subject's hand can cause changes in the response. This is caused by the vibration of the optical sensor relative to the finger to which it is attached. This motion changes the distance between the optical sensor and the finger, which changes the distance the infrared light travels, adding multiplicative movement noise to the measurement. Unfortunately, these motion artifacts have similar bandwidth to acceptable heart rates. In general, due to the multiplicative nature of the noise, these artifacts are best addressed by stopping them from occurring by firmly securing the sensor to the subject's finger and asking the subject to try to avoid moving their finger.

The output of the optical sensor during a sonic boom exposure is shown in Figure 1.11a). To calculate heart-rate, one calculates the number of temporal pulses that occur for a particular period. Typically the measured signal is first bandpass filtered to remove frequency information that is not associated with acceptable heart-rates. For this research, the range was between 30-180 beats per minute (0.5-3 Hz), the end-points of the range corresponding to the resting heart-rate of elite endurance athletes and the average anaerobic threshold which is unlikely to be breached while sitting. Once this filter has been applied to the signal, there are many possible ways to extract the heart-rate. A simple time-domain method to calculate heart-rate involves marking a particular time the signal crosses some portion of each pulse waveform and using the difference between these marks to calculate the heart-rate. Commonly the peak or some threshold, such as when the signal crosses the origin is used. The potential problem with this method is that it can be sensitive to distortions in the pulse waveform. Another method one could use to calculate heart rate is to calculate the instantaneous frequency (Rangayyan, 2002). In this method, the sensor output is assumed to be a single frequency-modulated tone that varies according to the heart-rate. A digital filter that behaves like a Hilbert transformer is used to estimate the phase information of that signal. The instantaneous phase is extracted by calculating the arctangent of the the Hilbert transform divided by the original signal. This is then unwrapped and differentiated to yield an estimate of the instantaneous frequency of signal through time. Naturally, deviations from the assumption that the pulse waveform is a simple frequency modulated and/or amplitude modulated sinusoid will introduce errors in the heart-rate calculation.

Heart-rate calculated via all three methods (peak-picking, zero crossing and instantaneous frequency) for one response measurement is shown in Figure 1.11b). The pulse sensor signal was sampled at 8192 Hz. To reduce noise in the calculation, the output of each heart-rate method was median filtered (3-pt.) and low-pass filtered (3rd Order Butterworth, f_c =0.3 Hz). Changes in heart-rate are expected to occur at frequencies less 0.3 Hz, so that cut-off was chosen. Notice that prior to the stimulus, the heart-rate responses are fairly stable and similar for all three calculation methods. After the boom onset, the heart-rate increases starting roughly 2-3 seconds after the boom. This latency is consistent with that of startle responses if one accounts for the delay caused by using an optical sensor at the finger ($\approx$ 1 s). The raw signal has a low-frequency variation around this time (15-25 seconds) which will be removed or greatly attenuated by the band-pass filtering (0.5-3 Hz) of the finger pulse signal. However the amplitude of the oscillation is smaller in this region.

This variation is probably due to changes in the vascular system (blood vessel constriction and dilation). This also occurs at 35 seconds. The zero-crossing method is very sensitive to any residual low frequency components in the signal which has an even greater effect when the amplitudes of the oscillations are small. There is a little distortion in the signal peaks at around 23-25 seconds, which results in a small increase in the heart-rate predicted by the peak-picking algorithm. The difference between the peak-peaking estimated heart-rate and the other estimates, however, is fairly small.

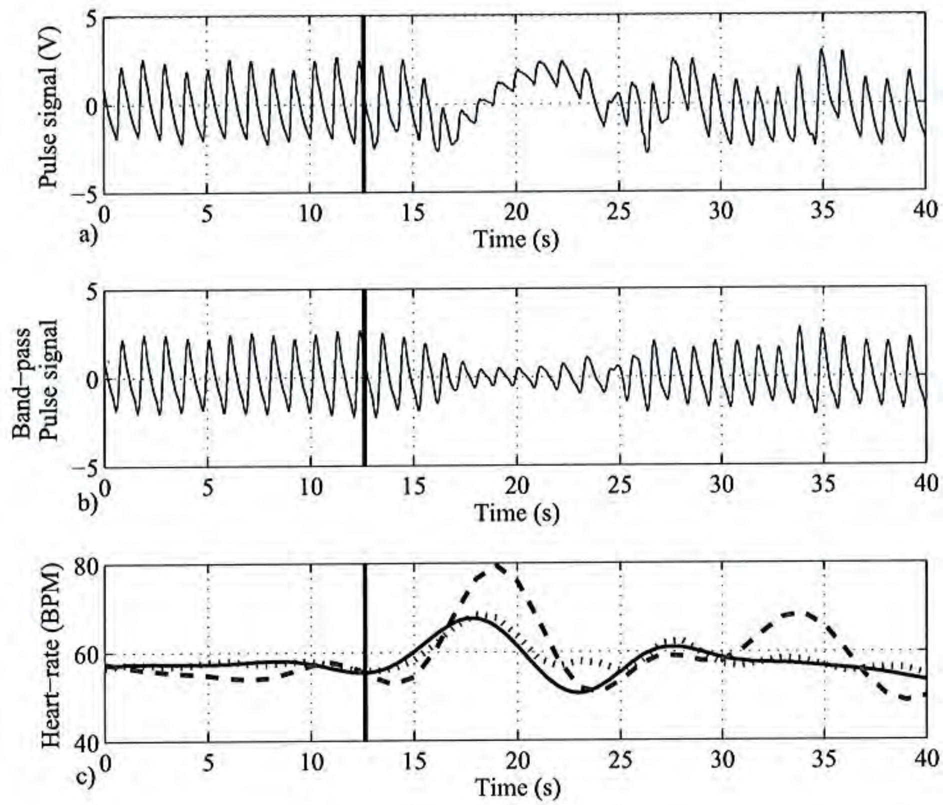

Figure 1.11. a) A typical signal produce by an optical heart-rate sensor in response to sonic boom. b) Band-pass filtered pulse signal. c) Associated heart-rate calculated via instantaneous frequency (black, solid), zero-crossing (dark gray, long dashes) and peak-picking (light Gray, short dashes). Boom onset denoted by vertical black line. Note presence of movement artifacts.

1.6.3 Electromyography

Electromyography, abbreviated as EMG, is the measurement of the electrical activity of skeletal (voluntary) muscles. This measurement is conducted by measuring the potential difference between two electrodes along the path of a muscle group. Electrodes can be placed on the surface of the skin or with needles (the latter are called indwelling electrodes). The former method, due to its non-invasiveness, is the one most commonly used in startle studies.

When a muscle contracts, the nerves send electrical current to the individual muscle fibers which are called motor units (Rangayyan, 2002). The motor units that contract are said be to recruited by that nerve impulse. The number of motor units recruited for a given event depends on the type of motor units available in the muscles (i.e., in broad terms, slow or fast twitch fibers) and the amount and duration of force required. Due to the variability with which motor units are recruited, it is nearly impossible to predict the contribution of individual motor units to the measured signal when surface electrodes are used. Measured EMG levels can depend on the size and relative strength of the muscle involved. EMG measurements are calibrated by normalizing the signal to some reference contraction. Usually, for large muscle groups (such as the muscles of the thigh), the EMG signal is normalized by the maximum isometric (no joint movement) voluntary contraction. However for smaller muscles, such as those of the neck, maximum contractions may be unsafe for the subject. In these cases, usually some other reference movement is used. There are many different analysis methods for EMG, depending on the application (see (Rangayyan, 2002)). However, for startle studies, typically the peak level of this normalized signal is analyzed.A sample EMG trapezius response evoked by a loud sonic boom is shown in Figure 1.12.

1.7 Hypothesis

It appears that modeling startle may be fruitful because annoyance judgments of low level sonic booms as heard outdoors appear to be highly correlated with startle judgments (Marshall, 2007). However, in previous attempts to model startle responses, stimuli that are

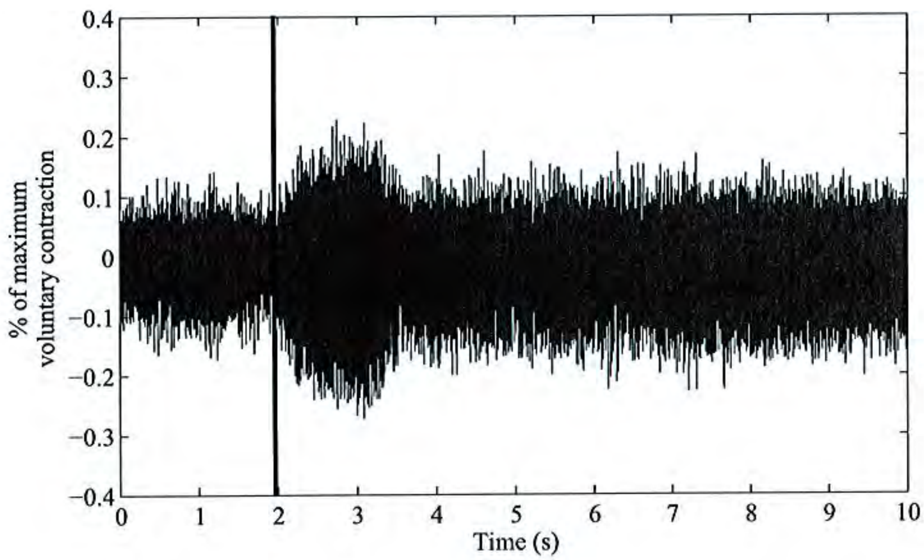

Figure 1.12. EMG measurement of a trapezius muscle response when exposured to a loud sonic boom. Vertical black line indicates onset of sonic boom.

radically different from low level shaped sonic booms, i.e. bursts of broadband noise, were the basis of the model development. Because startle is accompanied by well known physiological responses, it is possible to measure startle both subjectively and physiologically. The issue is that it is unclear if people's ratings of startle (rating startle on a scale from "not startling" to "startling") come from the same sound characteristics that give rise to the physiological responses associated with startle. Also the relationship between the intensity of the physiological startle response and where people rate on the scale is not known.

The startle model of Schmajuk and Larrauri (2006) takes as input the unweighted sound pressure level, yet we know that this may not be correlated with loudness because of the non-linear frequency response characteristics of the human hearing system. Time-varying loudness models may produce better inputs for a startle model because they utilize better models of what is perceived when people are exposed to the sounds.

There is processing in the brain that loudness models attempt to emulate. The nervous system takes this time-varying "loudness", or a precursor to loudness, input and then reacts accordingly; this is the startle model. Therefore, the startle model and the loudness model are not equivalent because each model has different machinery for processing the input. Because of this difference it is possible that other sound attributes in addition to loudness may be useful in predicting startle.

We want to understand the relationship between physiological responses and subjective ratings of startle and the characteristics of the sounds that play a role in startle responses for low amplitude sonic boom and similar impulsive sounds. We also want to understand the relationship between peoples' startle and annoyance ratings. The hypotheses that are tested in this research are:

1. Metrics based on characteristics of loudness, as predicted by the most recent models of time-varying loudness are more highly correlated to subjects' ratings of loudness and annoyance than metrics previously used to quantify people's responses to sonic booms and these loudness-based metrics can predict more accurately responses to a broader range of transient impulsive sounds than previously used sonic boom metrics.

2. Sound characteristics, other than peak loudness, influence people's startle and annoyance responses and the incorporation of these characteristics into predictive models will significantly improve the model performance.

3. Subjects' ratings of startle in response to transient environmental sounds reflect physiological responses associated with startle and orientation as defined by the work of Turpin et. al. (1999) and, thus, it is valid to use subjective evaluations as a measure of startle.

1.8 Organization and Summary

In this chapter, an outline of the recent history of sonic booms research was presented. This was followed by a review of recent efforts to quantify boom impact on people, physiological responses associated with startle and boom related startle research. The first task in the research reported in this dissertation was to verify the results obtained by the author in previous research where earphones were used for playback, in a playback environment that is capable of reproducing the low (<25 Hz) energy components of sonic booms measured outdoors so that tactile responses due to exposure to these sounds (Chapter 2) could be examined as well as any differences in judgments of the sounds when presented via loudspeakers versus over earphones. The results of this experiment demonstrated that subject-rated annoyance and startle are highly correlated for the stimulus sets used in the experiments. To examine startle more closely, a pilot experiment (Chapter 3) and a repeatability study (Chapter 4) were conducted to examine the relationship between physiological startle and subjective responses. Following on from this an experiment was designed to further develop a model that predicts startle from sound characteristics (Chapter 5). The last experiment, a startle model validation experiment was also designed to investigate annoyance and acceptability and their relationship to startle; this is reported in Chapter 6. A summary, conclusions and recommendations for future research are given in Chapter 7.

2. COMPARISON OF RESPONSES BETWEEN TWO SEMANTIC DIFFERENTIAL EXPERIMENTS WITH TWO METHODS OF PLAYBACK

In the last chapter, the results of a semantic differential experiment (Osgood and Suci, 1955) were described. In this study, subjects were asked to judge the characteristics of low booms and other environmental sounds. It was found that startle and annoyance ratings were highly correlated and could not be wholly explained by loudness judgments alone. However, the earphones used for presenting stimuli in that study were limited; frequency components of less than 25 Hz could not able to be reproduced at recorded levels. In addition, the use of earphones eliminated the ability to simulate tactile effects (i.e., being "hit" by the sound) that may occur when sonic booms are heard outdoors. To investigate the effect of these limitations in playback, another semantic differential experiment was designed and conducted. To examine the effects of low frequency components (<25 Hz) on playback, a sonic booms simulator (Salamone, July 18-22, 2005), designed by Gulfstream, was used to present signals to subjects. Aside from investigating playback effects, the utility of metrics to predict subject responses was also examined. In this Chapter, the design of the simulator semantic experiment is described, along with a brief review of the earphone semantic differential experiment. Then a comparison between the results of both experiments is presented along with some conclusions.

2.1 Methods

The simulator experiment was designed to be similar to the earphone experiment. In this section the experimental materials, apparatus, stimuli and participants are described for both the experiments. Prior to all research, Institutional Review Board approval was obtained.

2.1.1 Word Scales

Twenty word scales were used in each experiment and are listed in Table 2.1. The word pair scales were generated from a lexicon of words collected from subjects in an experiment described in Marshall (2007). In that experiment, subjects were asked to write down words to describe impulsive sounds. Word pair scales were developed from the 100 most used words and refined with informal testing. For the simulator semantic differential, two additional word-scales were added related to "feel" effects. Words used to describe this were given by participants in the high-level sonic boom lexicon studies (Chapter 2 of Marshall, 2007). The words used for the new scales were generated and refined from consultation with people at Gulfstream and at Purdue who had experience with low frequency noise and linguistics.

Table 2.1 Semantic Differential Word Pairs. $*$ and $+$ denote scales used only in the earphone and simulator tests respectively.

Loud	Soft	Acceptable	Not Acceptable
Sharp	Dull	Annoying	Not Annoying
High pitch	Low pitch	Startling	Calming
Peaked	Rounded	Agitated	Tranquil
Harsh/Rough	Smooth	Soothing	Disturbing
Fast	Slow	Distracting	Easily ignored
Close	Distant	Long	Short
Industrial	Natural	Rumbling	Clean
Violent	Peaceful	Deep	Shallow
$*$ Pleasant	Unpleasant	$+$ No Feel	Feel
$*$ Lots of echoes	No echoes	$+$ Hearing Dominant	Feel Dominant

2.1.2 Apparatus

A sonic boom simulator, developed by Gulfstream, was used for signal playback. This system, described by Salamone (July 18-22, 2005) and reviewed in Section 1.3, is a trailer with a subject area and a control room. The subject area consists of a quiet room facing a series of loudspeakers hidden behind acoustic screens. Each subject was seated in a chair in a location where the system is equalized. At this location, the frequency response of the system is relatively flat for frequencies greater than 7 Hz (Salamone, July 18-22, 2005). In previous studies, this system has been shown to reproduce sonic booms realistically (Sullivan, Davies, Hodgdon, Salamone III, and Pilon, 2008). For this experiment, the Gulfstream simulator was parked in a relatively quiet location, outdoors next to the Herrick laboratories. Despite this precaution and all attempts to minimize background noise, the noise levels in the simulator were higher than in the sound booth (63 dB versus 40 dB). In addition, the simulator had a speaker hiss that preceded the start of each signal playing. This could not be eliminated, so each sound had a random amount of hiss preceding it, which may have given subjects a cue to anticipate the signal presentation.

In the earphone experiment, Etymotics ER-2 earphones were used for playback. Due to the operating mode of the earphones, where the stimulus was reproduced at the eardrum rather than at the entrance of the ear canal, all sounds were compensated for their normal operating mode using a digital filter. Subjects in this experiment were placed in an IAC double-walled sound booth. To extend the effective playback range of the system, all of the sonic booms were high pass filtered using a zero-phase, 24 dB/octave roll-off Butterworth filter. The properties of this filter were selected based on an experiment described in the author's masters thesis (2007), where it was found that no subject could reliably distinguish between a boom and a high-pass filtered version of the same boom if the filter cut-off was higher than 28 Hz, assuming the filter had more than 24 dB/octave roll-off.

2.1.3 Signals

Twenty-four sounds were used in each experiment and are described in Table 2.2. All were selected from a database of signals provided by NASA and industry partners. Sounds common to both experiments were 10 recordings of low amplitude sonic booms, 5 simulated sonic booms, 1 recording of thunder, 1 recording of a car door slam, 1 recording of distant gunfire and a sonic boom and the door slam that were high-pass filtered (150 Hz) and amplified. In addition, two of the boom recordings were played twice, to examine the consistency of subject ratings. Two simulated booms were played at higher levels in the simulator than in the earphone experiment. In the simulator experiment, the two remaining signals were low-pass filtered (25 Hz cut-off) sonic boom recordings, one with a 4th order Butterworth filter and one with an 8th order Butterworth filter. These signals were added to examine the effect of removing these frequency components in the earphone experiment. For the earphone experiment, the two remaining sounds were repeats of two 150 Hz cut-off high-pass filtered sounds. In addition, all sonic boom recordings were high-pass filtered (25 Hz cut-off) to increase the range of playback over earphones.

Table 2.2 Description of signals used in each semantic differential test. * are signals used in the earphone experiment.

Signal Number	Description
1-10	Low amplitude boom recording
11-12	Simulated boom 1-2
13	High-pass filtered (150 Hz) car door slam
14	Car door slam
15	Gunfire
16-18	Simulated boom 3-5
19	High-pass filtered Dryden boom
20	Thunder Recording
21	Signal 6
22	Low-pass filtered (25 Hz) boom or Signal 13*
23	Signal 1
24	Low-pass filtered (25 Hz) boom or Signal 19*

2.1.4 Procedure

The experimental procedure for the simulator semantic differential was designed to be as similar as possible to the earphone experiment. Only one participant was tested at a time. First, the experiment was explained to the subject, informed consent was obtained and the subject's hearing was tested. The subject was then placed in the listening area (simulator) and the experiment instructions were given. Upon hearing a sound, each subject was asked to rate the sound she/he heard on 20 descriptive scales. Subjects were told to assume they were "outdoors in a park or garden and heard the sound intermittently throughout the day" when rating the sounds. This phrasing, used in both experiments, was intended to be vague. The rate of occurrence of these sounds is likely to influence their perceived annoyance. However, with no current plans to fly these aircraft, it was decided that a vague scenario would be appropriate.

Each sound was repeated at 5-10 second intervals while the subject completed his/her ratings. The sounds were repeated at random intervals to prevent the subjects from anticipating when the stimulus would be played. Plotkin and Bradley (1991) used a similar approach to this in a study of conventional aircraft noise. The subject moved on to the next sound once he/she completed the ratings for the current sound.

After being given the instructions, each subject heard two of the test stimuli (16 and 6) and then practiced rating the word pair scales for two additional sounds (12 and 13). After practicing, the subject completed the test.

2.1.5 Subjects

Twenty-six subjects were recruited for the simulator semantic differential experiment. Subjects were recruited from the area around Purdue University and their ages ranged from 19-75 years old. All except for one subject had less than 20 dB of hearing loss at frequencies from 125-8000 Hz. Thirteen of the subjects with normal hearing were male and twelve were female.

Twenty-five subjects were also recruited for the earphone experiment. The subjects had a similar age range as the simulator test. Eighteen of the subjects were male and seven were female in the earphone test. The data presented here for the earphone test includes responses from ten more subjects than the data presented in Marshall (2007). These additional subjects were recruited to minimize differences in variance between the two tests due to different numbers of subjects.

2.1.6 Metric Calculation

Sound quality and environment noise metrics were calculated from the recordings of test stimuli in the simulator experiment. Due to electric line noise, the simulator recordings were comb-filtered (60 Hz fundamental frequency) before metric calculation. A-weighted metrics (ASEL, LAmax), C-weighted metrics (CSEL, LCmax), Zwicker's time-varying loudness and Zwicker's Sharpness were calculated used Head Acoustic's Artemis software. Steven's Mk7 Perceived Level (PL) and Moore and Glasberg's time-varying loudness were calculated from software developed by NASA and the author's research group. Both time-varying loudness models were calculated at 1 ms intervals between samples. Sound exposure level (A & C) was calculated for the region of recording that included high-level echoes. This region corresponded to that which was within 10 dB of the peak level. PL was calculated using the entire sound event (including echoes).

In addition to single number metrics, linear models of two or more metrics were also examined: the maximum of a time-varying loudness model and either loudness rise time, a statistic of loudness derivative or a statistic of Zwicker's sharpness. Models were estimated using parameters from both time-varying loudness models. These models can be thought of as loudness-based extensions to the pressure time-history models discussed in Section 1.5.3.

The loudness rise time was defined as the interval between when the loudness breached the noise floor and the first loudness peak. The loudness derivative was calculated by applying a 51-pt. finite impulse response (FIR) filter to the loudness time-history. The filter was designed to behave as a differentiator from 0-350 Hz and as a low-pass filter at higher frequencies. The filter was designed using $firpm$ in Matlab which is an implementation of the algorithm designed by Rabiner, McClellan and Parks (1975). Statistics of the loudness derivative were determined from the portion of the loudness time-history associated with the initial rise of each sound event, corresponding to the interval used to calculate loudness rise time. Sharpness metrics were calculated using Zwicker's sharpness model. Statistics of sharpness were calculated for the potion of the signal associated with the main boom event. This region was defined as the portion of the signal where the loudness leaves the noise floor and returns to within 10% of the noise floor. This retains the sharpness of the main boom events, but not the echoes following. An example of these calculations is in Figure 2.1.

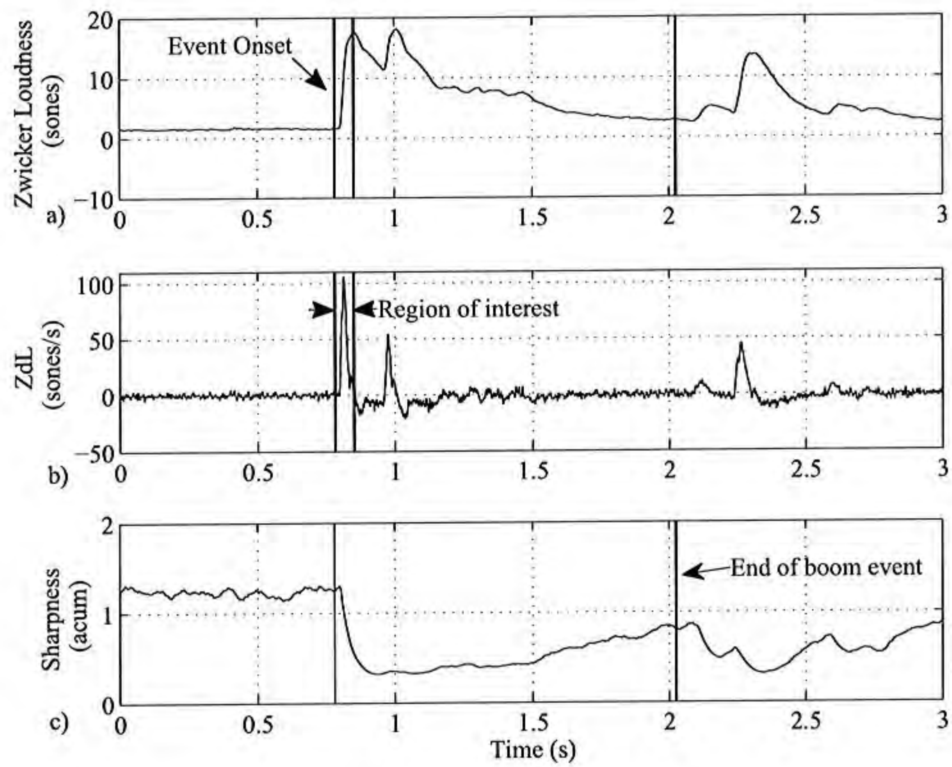

Figure 2.1. a) Zwicker loudness time history, b) Loudness derivative time history and c) Zwicker Sharpness time history of signal 8.

2.2 Results

The subjects' responses were given values on a scale from -16 to +16. Using the soft-/loud scale as an example, a score of -16 would correspond to the "soft" end, while a score of +16 would correspond to the "loud" end. To check for agreement between subjects, correlation coefficients (R) were calculated between each subject's responses and that of the mean responses of the rest of the subjects. For the simulator experiment, all subjects with normal hearing (25) had correlation coefficients of greater than 0.4. All of these responses were retained for analysis.

For the earphone experiment, 23 out of 25 subjects had correlation coefficients of greater than 0.4. The remaining two subjects had correlation coefficients of 0.29 and 0.15, respectively. The elimination of either or both of these subjects results in, at most, a difference of 0.6 in the mean of the ratings on the loudness, annoyance and startle scales and for most signals the difference was less than 0.3. Because the mean ratings and subsequent analysis were not greatly affected by the removal of these subjects, all subjects who participated in the earphone test were retained for analysis.

The distributions of the subject responses for each experiment appear to be uni-modal and without significant outliers. Thus, the analysis of the mean of the subject responses for each scale should be appropriate. Analysis of the median of the subject responses yields similar results. The standard deviation of the estimated mean for both experiments was similar (1.10-2 for all scales).

2.2.1 Comparison of Mean Ratings

A factor analysis (Kim and Mueller, 1978) was conducted on responses of each experiment using $factoran$ in Matlab. In both experiments, four factors were found and factor loadings were generated using the orthomax rotation algorithm (Kim and Mueller, 1978). To aid interpretation, after the rotation was applied, the factor loadings were rotated and normalized so that one factor is aligned with loudness and has a loading of 1. The factor loadings are plotted in Figure 2.2. Apart from the loudness scales, two of the three additional factors are common to both experiments: one related to temporal and duration sound characteristics (dark gray) and one related to spectral/feel characteristics (white). The main difference in loadings between the experiments is the relative strengths of these factors. In the simulator experiment, the spectral factor has larger loadings, and explains more of the variance than the temporal factor. In the earphone experiment, this situation was reversed.

It is possible that the better low-frequency playback in the simulator focused the subjects' attention on the spectral characteristics in the stimuli. In the earphone experiment, one factor was found whose loadings were concentrated on the "industrial-natural" scale. In the simulator experiment, this factor's loadings were more diffuse; other scales related to annoyance and those related to the temporal factor had large factor loadings. If the analysis were repeated with only three factors using simulator data, the "naturalness" factor would collapse into the loudness and temporal factors.

The average ratings of the loudness, annoyance and startle scales are plotted against each other for both tests in Figure 2.3. For both experiments, the mean loudness and startle scales are highly correlated to mean annoyance ratings ($R^2 = 0.75$ & $R^2 = 0.81$ for the earphone test and $R^2 = 0.78$ & $R^2 = 0.93$ for the simulator test). In Figure 2.3, the ratings of the low pass filtered sonic booms and the non-boom transient sounds appear to be rated as more annoying and startling than that predicted from their loudness. For the low-pass filtered booms, this result is not unexpected as the frequency content of these sounds are almost entirely composed of low frequencies (<25 Hz). These sounds also have the largest mean ratings on the "no feel/feel" scales (9.57 and 9.52). It is possible that exposure to these sounds caused tactile responses. However, the gunfire, thunder and door slam signals also are rated as more annoying than expected from loudness ratings, but have comparably little signal energy at those frequencies. It is possible that the spectral balance of these signals influenced subjects' response.

The mean loudness, annoyance and startle ratings of the earphone test are plotted against those of the simulator test for the signals common to both tests in Figure 2.4. Aside from a few signals, the average ratings of the three scales are correlated. As expected, the change in one of the ends for the annoyance scale, from "not annoying annoying" to "not annoying extremely annoying" resulted in an overall reduction of the average annoyance ratings. In the simulator experiment, subjects judged many of the low boom recordings and simulated booms as relatively more loud, annoying and startling than in the earphone experiment. However, for the annoyance scale, many of the boom recordings have nearly the same relative annoyance ratings in both experiments. Specifically, while the loudest and

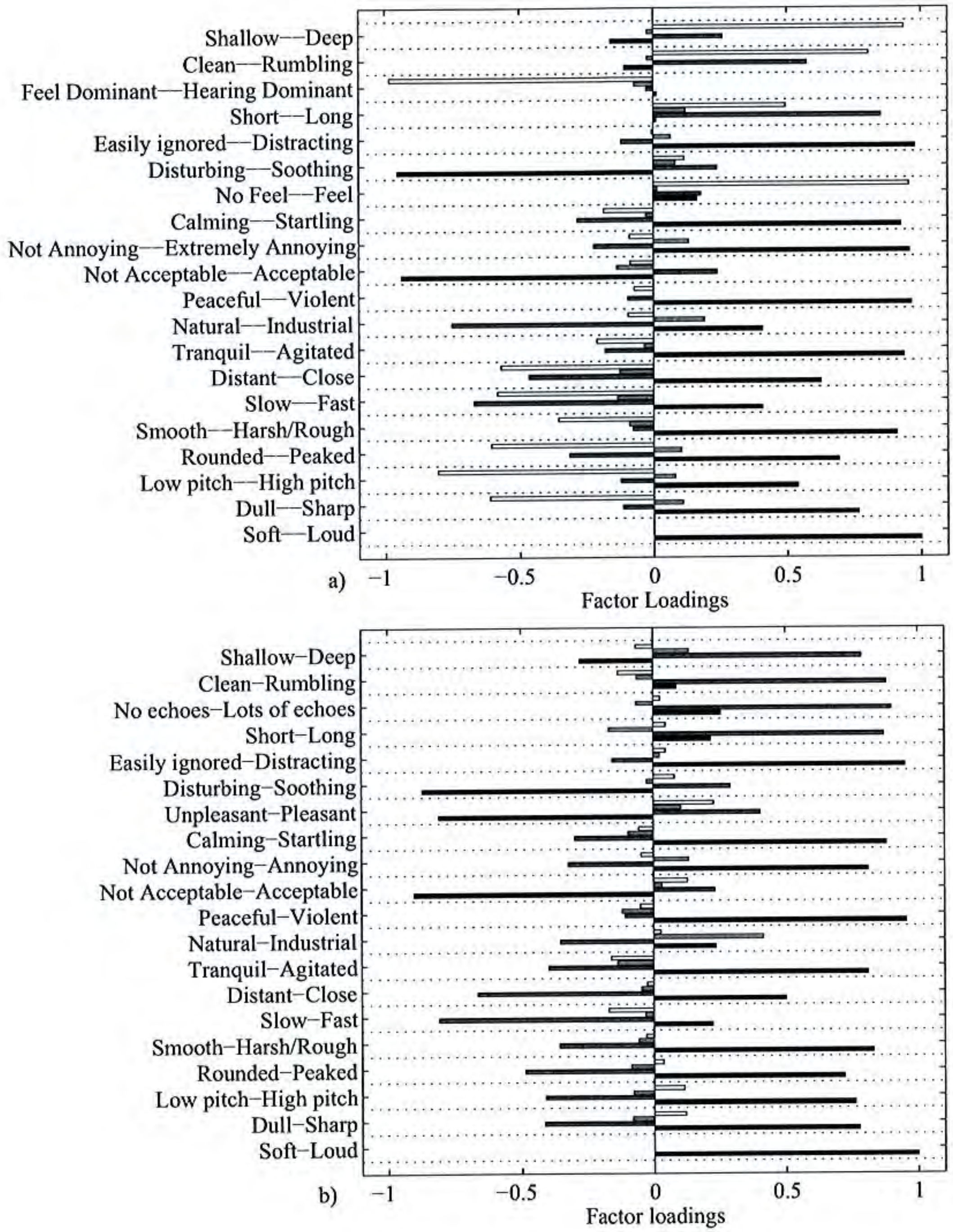

Figure 2.2. Factor loading of a) Simulator experiment and b) Earphone experiment. Black is loudness factor, dark gray is the temporal factor, white is the spectral factor and light gray is industrial factor (only present in earphone experiment).

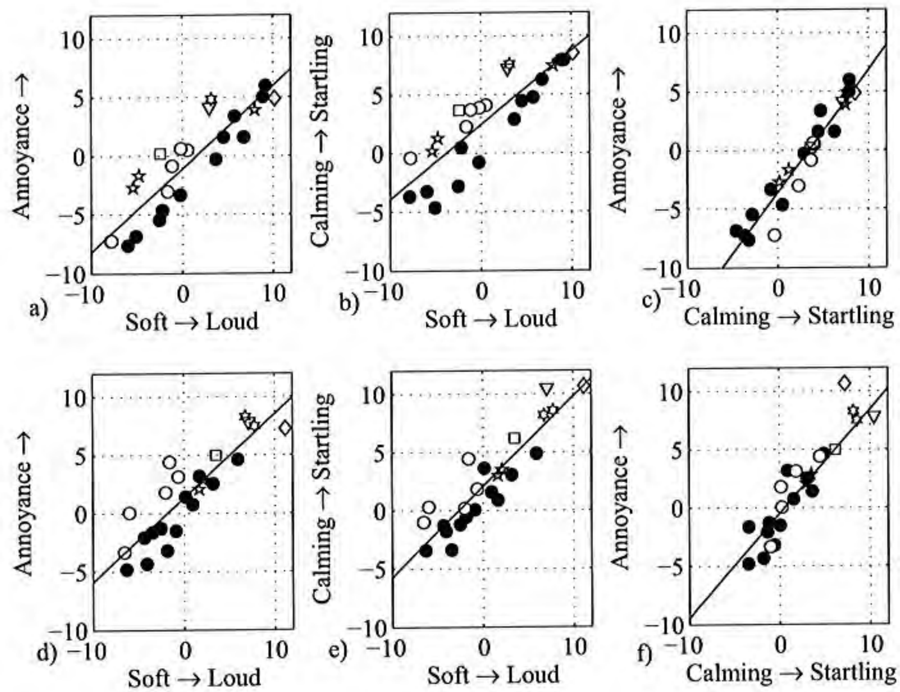

Figure 2.3. a) Annoyance versus loudness ratings ($R^2 = 0.80$), b) startle versus loudness ratings ($R^2 = 0.75$), and c) annoyance versus startle ratings ($R^2 = 0.93$) for simulator experiment. d) Annoyance versus loudness ratings ($R^2 = 0.78$), e) startle versus loudness ratings ($R^2 = 0.83$), and f) annoyance versus startle ratings ($R^2 = 0.87$) for earphone experiment. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder.

softest sounds are rated fairly similarly in both experiments, sounds with average ratings in the middle of the scale have more variation in average ratings. Similarly, the non-boom signals (gunfire and door slams) are rated as relatively quieter and slightly less annoying and startling in the simulator. In general, booms in this range were rated as louder in the simulator, while door slams were rated as louder over earphones. This trend of larger ratings for recorded booms in the simulator was also observed in the startle and annoyance scale. For the annoyance scale, however, these differences in ratings were less clustered by signal type.

One possible explanation for the differences in ratings is that the tactile effects influenced judgments of loudness, annoyance and startle in the simulator experiment. In Figure 2.4d)-f), the average ratings are plotted again with the size of each datapoint based on average "No Feel–Feel" ratings. For the sonic boom signals, both recorded and simulated, sounds with larger average "feel" ratings also have relatively larger loudness, annoyance and startle ratings in the simulator than over earphones. Similarly, the gunfire and door slam signals have the lowest "feel" ratings and relatively lower loudness ratings in the simulator. This effect was also noticeable for the ratings of the annoyance and startle scales, but this effect was smaller. This result is surprising because the ratings of the loudness scale are not strongly correlated to feel scale ratings ($R^2 = 0.22$). It is possible that subject felt/heard these frequencies, and used both characteristics to judge loudness. However, given the difficulty of correctly distinguishing between heard and felt responses at these frequencies (Leventhall, 2009), without more evidence it is difficult to determine if this actually occurred.

2.2.2 Metric Analysis

Metrics were calculated for each of the signals used in each test and a coefficient of determination (R^2) was calculated for each as a predictor of the average loudness, annoyance and startle ratings. Results of individual metrics are in Table 2.3 and linear models of two or more metrics in Table 2.5. For each metric examined, an R^2 value was calculated

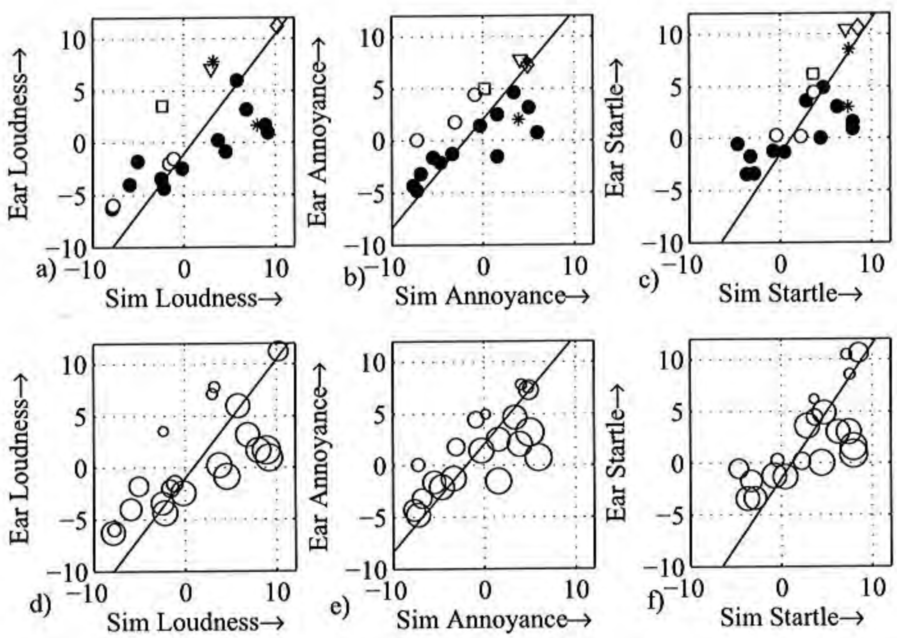

Figure 2.4. Earphone v. Simulator ratings of a) and d) loudness (R^2 = 0.55), b) and e) annoyance (R^2 = 0.75), and c) and f) startle (R^2 = 0.59). For d)-f) marker size is based on average ratings of the "No Feel-Feel" scale. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, filtered booms are stars and diamonds are thunder.

for the entire signal set and the signal set without outliers. A signal was designated as an outlier based on the magnitude of the residuals and the effect of its removal on the fitted parameters. In the first case, if the 95% confidence intervals of a signal's residual (prediction error) did not include a prediction error of zero it would be considered an outlier. In addition, if removal of the signal resulted in a significant ($\alpha = 0.05$) difference in the fitted parameters, then it would also be considered an outlier. For the models with multiple parameters, the statistical significance of the addition model terms was determined via the partial F-test, see for example, Montegomery (2005).

2.2.2.1. Single Number Metrics

Most metrics are better able to predict subject ratings in the simulator experiment. Improved predictions are greatest for the loudness scale ratings; most metrics predicted these ratings well in the simulator. However, nearly all of the metrics did not predict the loudness ratings of the low-pass filtered booms well (signals 22 and 24). The metric best able to predict the low-pass filtered booms was Steven's Mk7 PL. However, this model's predictions tended to under-predict the loudness ratings of the non-boom transients.

Most of the differences in the performance of metrics between both experiments likely stem from the differences in ratings of the non-boom transients (door slams and gunfire) compared to the rest of the stimuli. For the earphone results, many metrics tended to predict the ratings of these sounds on a different trend-line to the trend-line on which the boom lie. However, in the simulator experiment, the differences between the trend-lines for the average loudness ratings of these two groups of sounds is smaller. Despite improvements in the prediction of these sounds, many of these metrics still produced poor predictions of the the gunfire and door slam signals (numbers 13, 14 and 15) in the simulator test.

The best predictor of average loudness ratings in the simulator was the maximum value of Zwicker and Fastl's time-varying loudness; the best predictor of subject-rated startle and annoyance was the maximum value of Moore and Glasberg's time-varying loudness.

However, the difference in performance between these metrics is relatively small. The performance of Moore and Glasberg's model was very consistent across both experiments. Average loudness, annoyance and startle ratings are plotted against the maximum output of both loudness models in Figures 2.5 and 2.6.

Table 2.3 Coefficients of determination (R^2) for single number metrics predicting average loudness, annoyance and startle ratings. Coefficient of determination for Earphone experiment are in parenthesis. *denotes not significant at $\alpha = 0.05$. Mm is maximum of Moore loudness, Zm is maximum of Zwicker loudness.

Metric	Loudness			Annoyance			Startle		
	All	Without Outliers	Excluded	All	Without Outliers	Excluded	All	Without Outliers	Excluded
ASEL	0.85 (0.55)	0.90	20,22, 24	0.56 (0.34)	0.83	14,15, 22,24	0.56 (0.29)	0.80	13,14,15, 22,24
CSEL	0.08* (0.07)	None	None	0.13* (0.12)	None	None	0.16* (0.20)	None	None
LAmax	0.67 (0.68)	0.68	13,14,15, 16,22,24	0.52 (0.36)	0.74	13,14,15, 16,22,24	0.52 (0.31)	0.74	13,14,15, 16,22,24
LCmax	0.11* (0.03)	None	None	0.01* (0.14)	None	None	0.01* (0.12)	None	None
PL	0.76 (0.45)	0.92	13,14,15	0.40 (0.22)	0.91	13,14,15, 20,22	0.37	0.67	13,14,15
Zm	0.89 (0.62)	0.92	1,9,16	0.56 (0.37)	0.90	14,15, 22,24	0.56 (0.35)	0.79	13,14,15, 22,24
Mm	0.83 (0.88)	0.90	20,22,24	0.63 (0.60)	0.83	20,22, 24	0.64 (0.75)	0.77	20,22, 24

There is some similarity between the predictions of both time-varying loudness models in the simulator experiment. The signals with the largest differences between average ratings and model predictions for both experiments are the non-boom transients (door slams and gunfire). Moore and Glasberg's model predicts relatively higher loudness for these signals than Zwicker and Fastl's model. However, in the simulator experiment, these differences between ratings and model predictions are smaller. In Figure 2.7, maximum Zwicker's loudness is plotted against maximum Moore's loudness for signals used in both experiments. The differences in predictions, in general, are much smaller in the simulator experiment. One reason for this is that the sonic boom simulator does not reproduce some

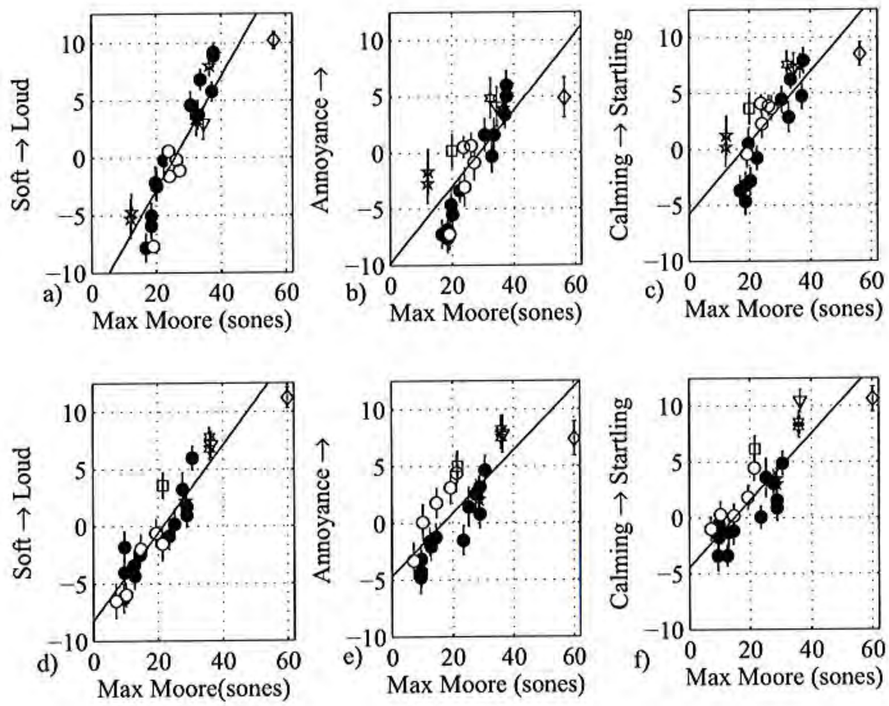

Figure 2.5. Subject-rated a) loudness (R^2 = 0.83), b) annoyance (R^2 = 0.63) and c) startle (R^2 = 0.64) versus Maximum Moore's loudness for Simulator Study results. Subject-rated d) loudness (R^2 = 0.88), e) annoyance (R^2 = 0.69) and f) startle (R^2 = 0.75) versus Maximum Moore's loudness for Earphone Study results. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder.

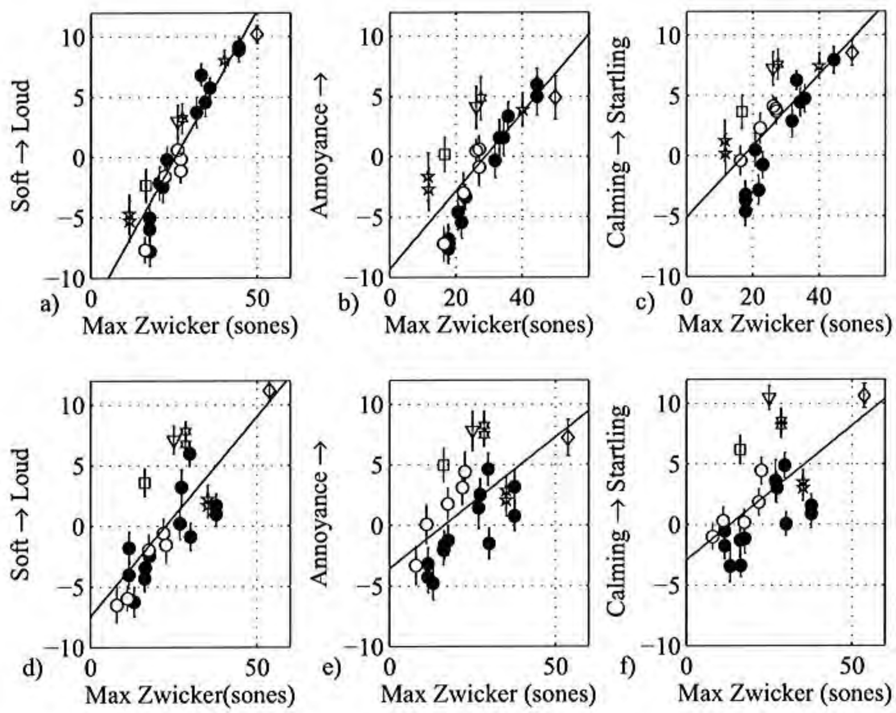

Figure 2.6. Subject-rated a) loudness (R^2 = 0.89), b) annoyance (R^2 = 0.61) and c) startle (R^2 = 0.59) versus Maximum Zwickers's loudness for Simulator Study results. Subject-rated d) loudness (R^2 = 0.51), e) annoyance (R^2 = 0.36) and f) startle (R^2 = 0.28) versus Maximum Zwicker's loudness for Earphone Study results. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder.

of these non-boom signals well due to high frequency limitations in the playback. This, combined with the relatively high outside background noise resulted in the difference in predictions and subject ratings.

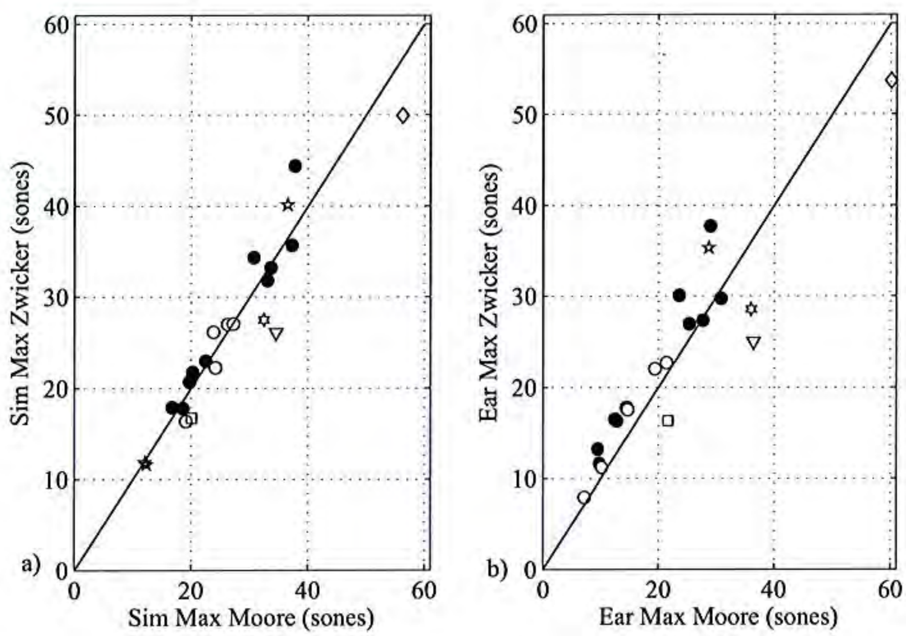

Figure 2.7. Maximum Zwicker's time-varying loudness versus maximum Moore and Glasberg's time-varying loudness for a) simulator experiment and b) earphone experiment.

2.2.2.2. Models of two or more metrics

It is possible that the maximum loudness does not capture additional features of these sounds that are important for annoyance or startle. To examine the effect of additional model parameters, a series of linear models based on maximum loudness (Moore (Mm) or Zwicker (Zm)) and either statistics of sharpness or the rate of change of loudness (loudness derivative or loudness rise time) were estimated and a coefficient of determination (R^2) calculated for the data obtained from each experiment. The results of this analysis are displayed in Tables 2.4 and 2.5. However, it should be noted that this analysis is exploratory, additional experiments would be required to validate the parameters of these models.

Average annoyance and startle ratings in both experiments are plotted again predictions of the best two-metrics models in Figures 2.8 and 2.9, for Mm-based models and Zm-based models, respectively. Very few models produced significant improvements in performance compared to just using maximum loudness alone (Partial F-test, $\alpha = 0.05$). Including maximum loudness derivative with maximum loudness for either model results in the best predictions of average annoyance ratings. The model with the inverse of loudness rise time ($1/MRT$) and Mm, did not predict subject ratings in the simulator experiment as well as in the earphone experiment. However, these models as well as the model based on Zm and $ZdLmax$ produced poor predictions of the average ratings of the low-pass filtered booms (signals 22 and 24). If these signals are considered outliers and are not included in the regression, the models with maximum loudness and rate of change of loudness metrics perform similarly in both experiments.

The models that include statistics of sharpness generally produce the fewest outliers for the ratings in the simulator. The sharpness metrics for the low pass filtered booms were stratified away from the other stimuli. Thus, including statistics of sharpness corrected for the lower level of annoyance and startle predicted by Zm and Mm for these sounds. Interestingly, including sharpness with Mm in models with either experiment resulted in much smaller improvements than when sharpness was included with Zm. Out of the models that include Mm and sharpness metrics, only three were significant: Mm and $Smean$ and Mm and $Smin$ in the earphone experiment and Mm and $Smax$ in the simulator experiment. If outliers are removed, however, the model with Mm and $Smean$ is the best predictor of startle ratings in the simulator out of the models that include Mm.

Using the components of the best 2-metric models, several 3-metric models were estimated. Coefficients of models using Zm or Mm, a measure of the rate of change of loudness and a statistic of sharpness were estimated. For the simulator data, there was no difference in performance of any of the 3-metric models if $Smax$ or $Smean$ was used as the statistic of sharpness. Including one of these metrics was enough to prevent the low-pass filtered booms being outliers. However, none of the 3-metrics models that included Mm in either experiment were significant (partial F-test, $\alpha = 0.05$). The only 3-metric model that was significant was that based on Zm, $ZdLmax$ and $Smean$. This model appears to combine the best features of the two best Zm-based 2-metric models.

Table 2.4 Coefficients of determination (R^2) for linear models predicting subject-rated annoyance and startle utilizing Zwicker's time-varying loudness in simulator experiment. Coefficients of determination for the models in the earphone experiment are in parentheses. * denotes models with partial F-test p-values of >0.05 i.e., not significant at the 5% level. Zm denotes maximum loudness, ZRT denotes loudness rise time, ZdL denotes loudness derivative and S denotes sharpness.

Metrics	Annoyance All	Annoyance Without Outliers	Annoyance Excluded	Startle All	Startle Without Outliers	Startle Excluded
Zm	0.56 (0.37)	0.90	14,15,22, 24	0.56 (0.35)	0.79	13,14,15, 22,24
Zm, ZRT	0.62* (0.37*)	0.76	13,15, 20	0.61* (0.40*)	0.71	13,15, 20
Zm, $1/ZRT$	0.60* (0.38*)	0.71	13,15	0.59* (0.36*)	0.69	13,15
Zm, $ZdLmax$	0.76 (0.74)	0.87	22,24	0.72 (0.68)	0.72	22,24
Zm, $ZdLmean$	0.60* (0.38*)	0.89	13,14,15, 22,24	0.59* (0.37*)	0.81	13,14,15, 22,24
Zm, $Smax$	0.70 (0.43*)	0.71	22,24	0.74 (0.50)	0.77	22,24
Zm, $Smean$	0.71 (0.56)	0.78	20	0.76 (0.65)	0.82	20
Zm, $Smin$	0.70 (0.50)	0.79	20	0.68 (0.66)	0.73	20
Zm, $ZdLmax$, $Smean$	0.79* (0.73*)	None		0.81 (0.75)	None	

2.3 Discussion

While there were some differences in subject ratings in both experiments, the results of each factor analysis and the comparisons between loudness, annoyance and startle were

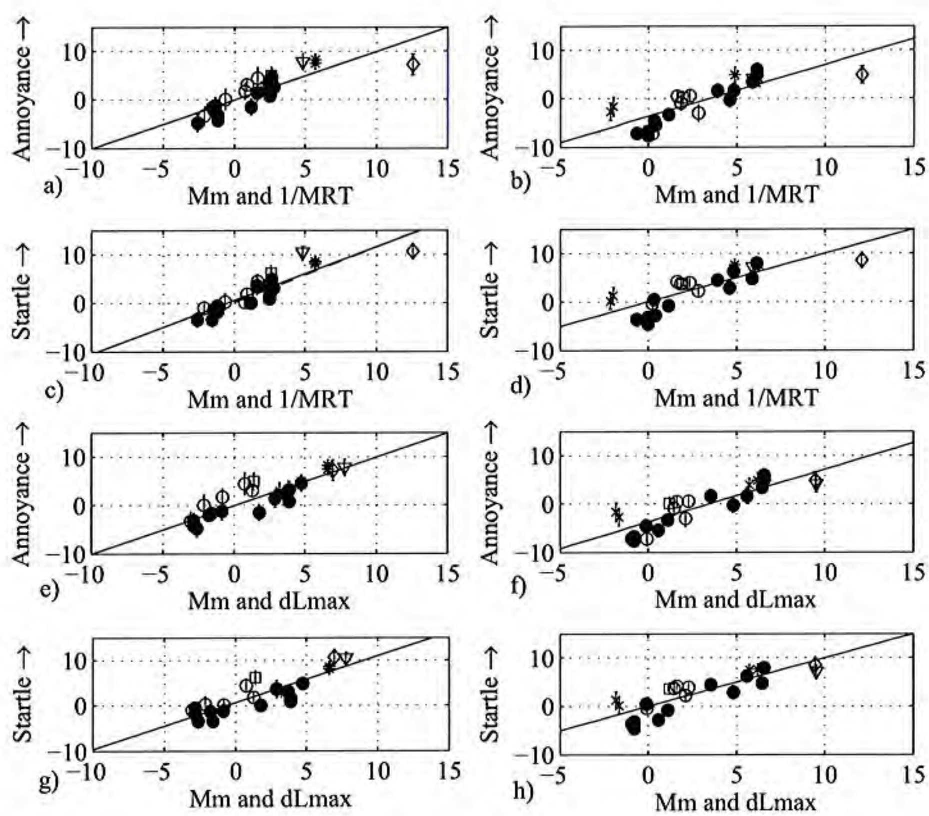

Figure 2.8. Annoyance and startle ratings versus predictions of linear model of Mm and 1/MRT in a) & c) the earphone and b) & d) the simulator experiments, respectively. Annoyance and startle ratings versus predictions of linear model of Mm and MdLmax in e) & g) the earphone and f) & h) the simulator experiments, respectively. Refer to Table 2.5 for R^2 values. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder.

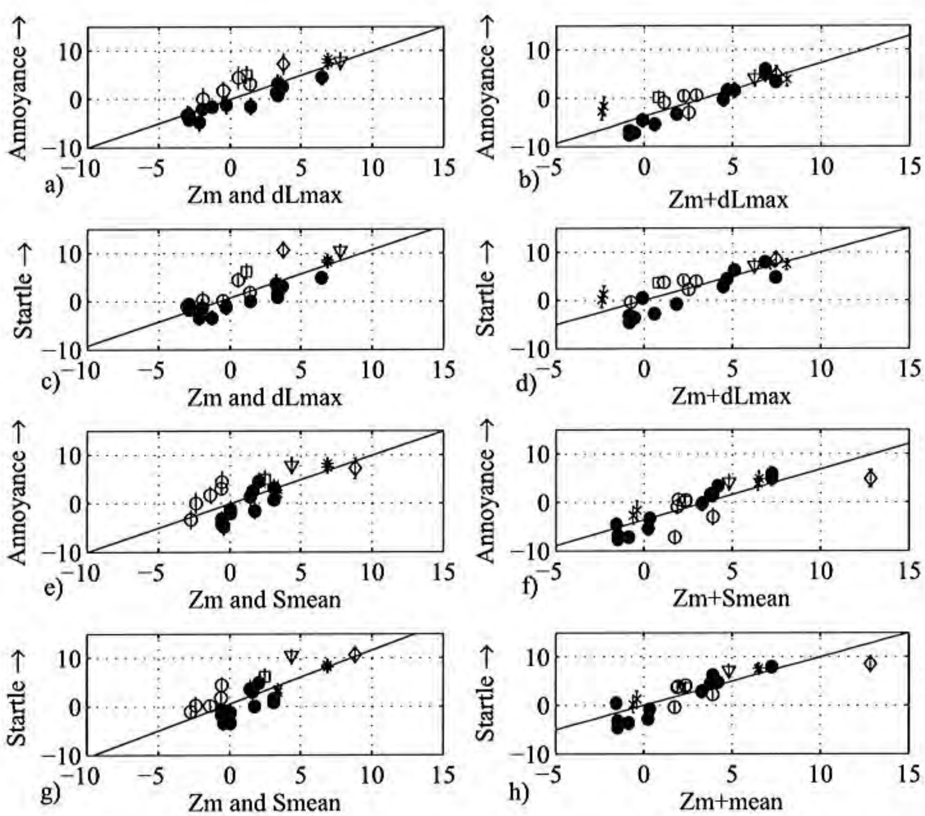

Figure 2.9. Annoyance and startle ratings versus predictions of linear model of Zm and ZdLmax in a) & c) the earphone and b) & d) the simulator experiments, respectively. Annoyance and startle ratings versus predictions of linear model of Zm and Smean in e) & g) the earphone and f) & h) the simulator experiment, respectively. Refer to Table 2.4 for R^2 values. Filled circles are recorded booms, open circles are simulated booms, squares are door slams, triangles are gunfire, stars are filtered booms and diamonds are thunder.

Table 2.5 Coefficients of determination (R^2) for linear models predicting subject-rated annoyance and startle utilizing Moore's time-varying loudness for both experiments. Coefficients of determination for the models in the earphone experiment are in parentheses. * denotes models with partial F-test p-values of >0.05, i.e. not significant at the 5% level. Mm denote the maximum loudness, MRT denotes loudness rise time, MdL denotes loudness derivative and S denotes sharpness.

	Annoyance			Startle		
Metrics	All	Without Outliers	Excluded	All	Without Outliers	Excluded
Mm	0.63 (0.60)	0.83	20,22,24	0.64 (0.75)	0.77	20,22,24
Mm, MRT	0.65* (0.68*)	0.86	20,22,24	0.67* (0.75*)	0.81	20,22,24
$Mm, 1/MRT$	0.65* (0.74*)	0.86	20,22,24	0.67* (0.82)	0.81	20,22,24
$Mm, MdLmax$	0.74 (0.76)	0.85	20,22,24	0.73 (0.78*)	0.79	20,22,24
$Mm, MdLmean$	0.67* (0.68*)	0.86	20,22,24	0.67* (0.75*)	0.79	20,22,24
$Mm, Smax$	0.68* (0.68*)	0.79	20,22,24	0.73* (0.75*)	0.74	20
$Mm, Smean$	0.66* (0.70*)	0.70	20	0.66* (0.79)	0.83	20
$Mm, Smin$	0.67* (0.67*)	0.74	20	0.68* (0.79)	0.73	20
$Mm, 1/MRT, Smean$	0.67* (0.74*)	0.74	20	0.70* (0.84*)	0.75	20
$Mm, MdLmax, Smean$	0.74* (0.78*)	None		0.74* (0.78*)	None	

similar across both experiments although spectral balance played a greater role in the simulator experiment. Startle judgments are highly correlated to annoyance judgments for low booms and impulsive transients. In addition, while the startle and annoyance scales both have strong loudness factor loadings, both also had temporal and spectral factor loadings as well. This was the reason for investigating models with maximum loudness and rate of change of loudness or sharpness metrics.

All of the metrics (which are models of loudness) were found to predict loudness ratings better than they predicted annoyance or startle ratings. Both time-varying loudness models appear to generate statistics (e.g. maximum) that were the best predictors of subject ratings. While Moore and Glasberg's model was not the best single metric predictor of subject rated loudness in the simulator, it was the best predictor of startle and annoyance when results from both experiments were considered and the most consistent predictor of subject ratings. Including the derivative of loudness and statistics of sharpness in the model appears to improve predictions of annoyance and startle, but more data is needed to confirm this result.

2.4 Summary

In this chapter, the results of two semantic differential experiments were described and compared. In one, earphones were used for playback and in the other a sonic boom simulator was used for playback. The main difference between the two playback methods was the ability of the simulator to reproduce low frequency content (<25 Hz). The results of the experiments were found to be similar. By using factor analysis, three out of four factors were found to be similar in each experiment: one related to loudness, annoyance and startle; one related to spectral characteristics; and one related to duration/speed characteristics.

A metric analysis was conducted. Statistics of time-varying loudness models (both Zwicker and Moore) were found to be the best predictors of subject ratings in both experiments. However, most metrics produced better predictions of loudness ratings in the simulator than in the earphone experiment. Including additional factors such as statistics of sharpness or rate of change of loudness with maximum loudness in a linear model improved predictions of annoyance and startle ratings. However, due to the fact that many of these additional parameters are correlated with maximum loudness over the stimulus set, additional experiments are needed to confirm the model structure and parameters. Further experiments to examine these parameters are discussed in Chapters 4, 5 and 6.

Ratings of annoyance and startle are highly correlated to each other in both the earphone and simulator experiments. However, it is unclear if physiological responses associated with startle are evoked when subjects hear the stimuli. To investigate whether subjective judgments of startle are correlated with physiological responses typically associated with startle, a series of studies was conducted. The first of these studies is described in Chapter 3.

3. A COMPARISON OF SUBJECTIVE RATINGS AND PHYSIOLOGICAL MEASURES OF STARTLE

In the last chapter, the results of two experiments were presented in which the judgments of annoyance and startle were found to be highly correlated to each other. However, it was unclear what criteria subjects used when determining their ratings on the startle scale. In particular, it was unclear if subjects were experiencing physiological responses consistent with startle when exposed to these sounds. As mentioned in Chapter 1, physiological measures have been investigated as a means to measure startle; however in nearly all of this research, white-noise tone pulses or high level sonic booms were used as stimuli. The stimuli used in the experiments described in Chapter 2 were at levels near those found in previous research to elicit startle responses (see Table 1.1). However, extrapolating the results of previous research to draw conclusions about startle responses to low amplitude sonic booms and other sounds in that experiment is difficult.

A pilot experiment was designed to examine physiological responses and subjective responses to low level sonic booms and other impulsive transients. This experiment was designed to be an exploratory study to highlight issues that need to be investigated in additional experiments and to identify any problems in the experimental methodology. The experiment consisted of two parts. In the first part, the subjects were exposed to impulsive sounds while completing an arithmetic and memory task. The goal of this part was to expose people to these sounds in a situation where they were not expecting them. In the second part, the subjects completed a semantic differential test with a subset of the stimuli used in the tests described in Chapter 2. The purpose of this part of the experiment was to obtain physiological measures under similar conditions experienced by subjects in earlier experiments.

3.1 Experimental Methodology

Approval from Purdue's IRB was obtained for this study (Protocol #0904007961). The signals were played binaurally (diotic) to subjects in an IAC double-walled sound booth. Signals were presented to subjects using Etymotics ER-2 earphones connected to a Tucker-Davis HB7 amplifier and a LynxOne sound card. All signals were pre-processed to account for the operating mode of the earphones. The physiological measures consisted of skin conductance, pulse-rate (optically measured at the finger) and electrical activity (electromyography or EMG) of the sternocliedomastoid (SCM), anterior scalene and upper trapezius. A diagram of these locations is in Figure 3.1. These EMG locations were selected due to the quality of the signals that could be obtained in preliminary testing and from locations used by other experiments reported in the literature (e.g., Brown et al., 1991). Ocular EMG (eye EMG) was not used, despite its prevalence of its use in previous startle experiments for fear that it would be uncomfortable for subjects (requiring electrodes near the eyes) and thus, affect subjects' ratings.

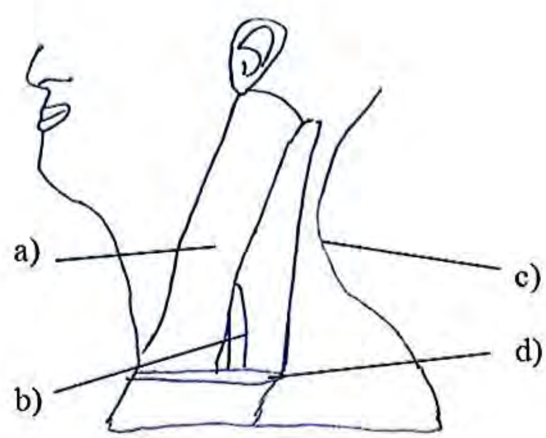

Figure 3.1. Diagram of EMG locations: a) sternocliedomastoid (SCM), b) anterior scalene, and c) upper trapezius. For anatomical reference, d) is the collarbone.

Skin conductance and pulse rate were measured using tranducers and an amplifier made by James Long Company. Two channels of skin conductance were recorded. One channel was the raw skin conductance and the other was the raw signal, amplified and high-pass filtered to highlight changes in skin conductance. The former channel will be referred to as skin conductance and the latter channel will be referred to as filtered skin conductance. The EMG amplifier was made by Bortec. Earphone Voltage (for event timing) and the physiological measures were recorded by using a Bruel & Kjaer PULSE data-acquisition system, and were sampled at 8192 samples per second. To avoid eliminating long-term physiological responses, time-series data for all channels were recorded continuously for each of the two parts of the experiment.

3.1.1 Task

For the first part of the experiment, each subject completed an operation span task developed by Unsworth, Heitz, Schrock, and Engle (2005). The task was divided into segments, called letter blocks. For each segment, the subject would be shown a simple arithmetic problem (e.g. (3 x 7)-4) to solve. All arithmetic problems had integer terms and integer solutions. Once the subject solved the problem, she/he clicked "ok", and another screen was displayed. This screen would display a potential answer. The subject was then asked if the answer was correct (yes/no) and after answering, received feedback. After answering each problem, the subject would receive a letter to remember (e.g. F). After a total of six problems, the subject would be asked to recall the last six letters in order of presentation. This was done via a screen where the subject would place the order number by the letters she/he remembered (Figure 3.2). This process would then be repeated for a total of 23 letter blocks.

The primary goal of the task was to distract the subject from anticipating the sound stimuli. However, as an additional benefit, this particular task was selected because performance of this task has been correlated to the performance of other cognitive functions, particularly reading comprehension (Unsworth et al., 2005). Thus, the impairment in the

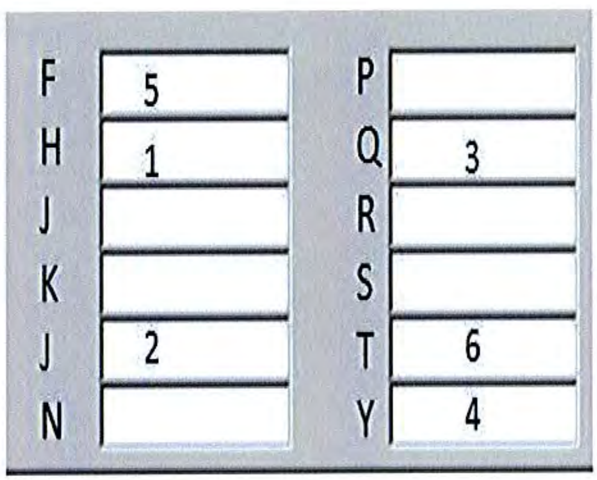

Figure 3.2. Letter recall screen under perfect recall of letter sequence: H,J,Q,Y,F, and T.

performance of this task may be related to the impairment in the performance of other tasks, similar to what the subject might be doing while exposed to sonic booms in the real world.

3.1.2 Semantic Differential Component

After completing the memory task, the subject completed a semantic differential experiment where she/he rated all of the signals on twenty word-pair scales. These scales were the same as those used in the earphone experiment (Table 2.1) described in Chapter 2. Scale ends and scale orders were randomized with a different order for each subject. The subject rated each sound by marking, in pencil, on each of the paper scales. Each page contained all twenty of the word-pair scales; one page of scales was rated for each sound. Subjects used the computer mouse to indicate when they had finished rating the scales for each sound. While completing the scales for one sound, the sound would repeat with a random interval between 5-10 seconds. The order of the stimuli was also randomized, a different order was used for each subject.

3.1.3 Signals

Nine sounds were used in this experiment. They were recordings of low amplitude sonic booms and impulsive transients sounds. As in the earphone experiment described in Chapter 2, all boom signals were high-pass filtered (25 Hz cut-off). Five recordings of low amplitude sonic booms were used, with one recording each of distant gunfire and thunder. The two remaining signals were a boom recording and a car door slam recording that were high-pass filtered with a cut-off of 150 Hz. All recordings were used in the experiments described in Chapter 2 and selected to cover the span of average startle ratings in those tests. In the task-based experiment, three of the boom recordings were played one additional time (for a total of twelve). A description of the stimuli is in Table 3.1.

Table 3.1 Stimuli used in both parts of the experiment. In the task part of the test signals 1-12 were used and in the semantic differential test signals 1-9 were used. The maximum of Moore and Glasberg's time-varying loudness for each signal is given as a reference.

Signal #	Ch 2 Signal #	Description	Max Loudness (sones)
1	1	Low boom	9.7
2	2	Low boom	25.3
3	3	Low boom	24.0
4	4	Low boom	29.0
5	5	Low boom	14.6
6	13	Filtered door slam	35.4
7	15	Distant gunfire	36.4
8	19	Filtered low boom	28.7
9	20	Thunder	60.1
10	3	Signal 3	24.0
11	4	Signal 4	29.0
12	5	Signal 5	14.6

3.2 Procedure

After greeting the subject, informed consent was obtained and the subject's hearing was tested. The subject then sat in the sound booth and the physiological sensors were attached. The finger cuffs (for skin conductance and pulse-rate) were attached on the left hand so as not to interfere with the operation of the computer mouse. After all sensors were connected, a base-line measurement (30 secs) was obtained while the subject was sitting quietly in front of the computer. Next, another measurement was made while the subject moved his/her head in a slow circle at least five times. This result was used to normalize the subsequent EMG measurements for each subject. A slow head circle was used instead of a maximum voluntary contraction to avoid over-taxing the muscles of the neck.

Once the sensors were attached and the baseline values attained, the experiment was explained to the subject. The subject was then asked to complete the memory span task. The subjects were told that they might hear sounds through their earphones and that they should ignore them. No direction was given as to the kinds of sounds to expect. After three minutes of completing the tasks, where subjects typically completed between 2-3 out of 23 letter blocks, the first sounds was presented. Sounds were presented to the subject after that at random intervals, at a rate of presentation not exceeding one sound every 30 seconds. Sound orders were randomized and played in a different random order for each subject. A total of twelve sounds was played.

For the second half of the experiment, the subject completed a semantic differential experiment similar to that described in Chapter 2, except that fewer stimuli (9) were used. Signal presentation order was also randomized for this part of the experiment. After completing the experiment the subject was thanked for her/his time and compensated ($10).

3.2.1 Subjects

Ten subjects were recruited for this experiment. Four subjects were male and six were female. Ages ranged from 19-52 years. Subjects were recruited from the area surrounding Purdue University and included students and faculty. Six subjects were involved in technical fields (e.g. Engineering, Computer Science, etc.) and four subjects were involved in the humanities (e.g. History, Anthropology, etc.). All subjects had less than 20 dB of hearing loss at frequencies between 125-8000 Hz.

3.3 Results

After the data was collected, the ratings scales from the semantic differential experiment were given numerical values of -16 to +16 and averaged across subjects. Average ratings of the loudness, annoyance and startle scales were highly correlated to the results in the previous earphone semantic differential experiment (R^2 = 0.75, 0.87 and 0.87 for these three scales respectively). As observed in the experiments described in Chapter 2, the startle ratings and the annoyance ratings were highly correlated ($R^2 = 0.79$) .

The task was graded based on the number of correct letters recalled in the correct order (the memory span) and the time taken to finish each set of letters. A profile of each subject's performance through time, measured via the number of letters correctly recalled, is plotted along with the mean values in Figures 3.3. Some subjects heard more than one sound in a single letter block. Performance varied from subject to subject with the worst performance for many subjects occurring during the first two letter blocks. Of note, two of the subjects (2 & 4) did not miss a single letter during the sound exposures and one of these (4) did not miss a single letter during the entire experiment. For most of the rest of the subjects, there appears to be a reduction in the number of letters correctly recalled during the first few initial sound exposures. However, after that, performance of the task increased to about the level associated with the letter blocks without sound exposures.

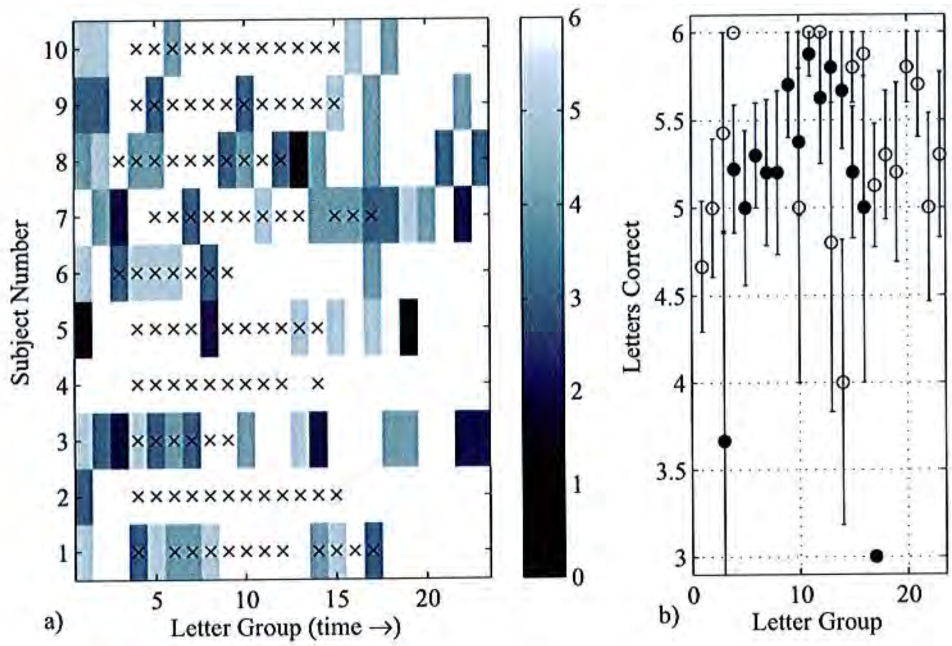

Figure 3.3. a) Letter correct for each subject through time. Shades of gray indicate number of letters correctly. White indicates all correct. Letter blocks with sound exposures are marked with X's. b) Mean letters correct through time for sounds (black) and no sounds (white). Errorbars are standard deviation of the estimated mean.

Profiles of the time taken to complete each letter block along with the average time to complete each letter block are in Figure 3.4. Timing errors occurred with subject 1, and for the 23rd letter blocks for many subjects. These responses were not considered when calculating the mean. The amount of time it took each subject to complete each letter block varied considerably. Generally, subjects took the most time for the initial few blocks, then, as their familiarity with the test increased, they performed faster. Subjects with only a few missed letters (2, 4, 5 and 10) generally performed the task faster than those who missed more letters.

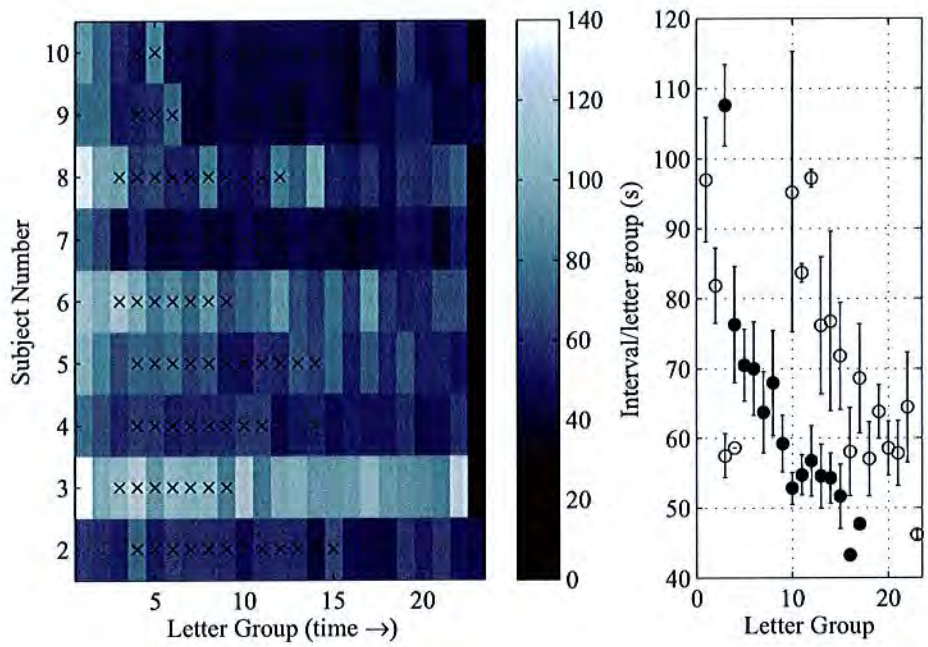

Figure 3.4. a) Time taken by each subject (in seconds, 0-140 s) for each letter block. Letter task blocks with sound exposures are marked with X's. b) Mean time per letter block over the course of the experiment for blocks with sounds (black) and without sounds (white). Errorbars are standard deviation of the estimated mean.

3.3.1 Post-processing of physiological measures

The EMG signals were full-wave rectified and normalized by the maximum response during the calibration head circle. In general, the head circle results procedure varied across both sessions (before and after the experiment) for each subject. This is probably due to the inability to regulate the speed of the head-circle motion across both head-circle procedures.

The data from the pulse-rate sensor was converted to heart-rate by using a peak-picking algorithm. Other methods of calculating heart-rate (zero-crossings and instantaneous frequency) tended to produce similar heart-rate values as the peak-picking approach but with more artifacts. Typically 2 out of the 3 methods gave very close results.

3.3.2 Analysis of Physiological Measures

Once the physiological data was obtained it needed to be processed in some way to reduce the measures into a suitable form for comparison with the subjective measures. Unfortunately, there were noise issues present in the data which made this analysis difficult. To understand the issues associated with the processing of this data, a case study of subject 7 will be presented. Subject 7's responses were selected because this subject exhibited most of the noise and artifact issues present in other subjects, while also exhibiting some very straightforward startle responses.

In Figure 3.5, a profile of the physiological data for the task portion of the experiment is displayed. Of importance, there are transient physiological responses before, during and after the portion of the task in which sounds were played. Both of the skin conductance responses appear to respond to the presence of letter recall screens during the task portion of the test. This result is not surprising, as skin conductance is an index of stress (Sokolov, 1963). This result is evidence that the task itself was causing stress responses. The presence of responses evoked by both the task and the sound signals reduces the effective interval between responses. Thus, some physiological responses overlapped as there was not enough time for the subject to return to baseline values. While this is readily noticeable with the skin conductance measure, it also to a less degree affect the analysis of the heart-rate responses. Long latency responses, occurring 20-30 seconds after the stimuli, are possible (e.g. Ramirez et al. (2005)). It is possible that these responses would overlap with shorter latency responses. The heart-rate measure also spikes rapidly toward the end of the experiment. This increase coincides with long duration elevated EMG signals. This is associated with the subject completing the experiment and stretching.

To examine some responses in more detail, the time around the sound exposures at 270 and 532 seconds are in Figures 3.6 and 3.7. Both of these correspond to periods in the task after a letter recall screen was displayed. At 270 seconds, the high-pass (150 Hz cut-off) filtered sonic boom was played to the subject. The electrical activity of the neck muscles is elevated prior to the sound exposure. All of the muscles show activity a little more than a second prior to the sound event. The trapezius signal shows spiked behavior after the peak pressure, but it also had elevated levels before the sounds. It is possible that this response is due to the onset of the recording playback or due to body adjustment associated with completing part of the task. However, regardless of what initiated the early response, it is likely that the response to the sound itself was inhibited. The skin conductance shows a response starting about 1-1.5 seconds after the peak pressure of the sound. The latency between the peak pressure and the response was fairly consistent when responses were evoked by sound exposures (between 1 and 2 seconds for most signals).

The heart-rate response shows acceleration, a classical sign of a startle response. However, the acceleration appears to begin before the peak pressure itself. This response could be due to the prior event that caused the EMG responses. However, the response is fairly long in duration; perhaps there is some contribution from the sound event present in the response.

At 532 seconds the subject was exposed to the thunder signal, the loudest of the stimuli presented. The responses are displayed in Figure 3.7. In the EMG signals, there is a clear sign of a startle response in all three muscles. The trapezius signal has a response, but it was much closer to the pre-event levels of the other two EMG signals. There were slightly elevated EMG levels near 537 seconds, right after the math problem and at the tail end of the signal. This response could be due to subject movement associated with completing the task. The EMG responses are of a lower magnitude than those observed in Figure 3.6, but background muscle activity is lower. The heart-rate increased beginning roughly 2-3 seconds after the sound event. This response is very similar to the "classical" heart-rate response associated with startle. The skin conductance response showed a similar latency as with the previous example, but the magnitude of the response was much higher.

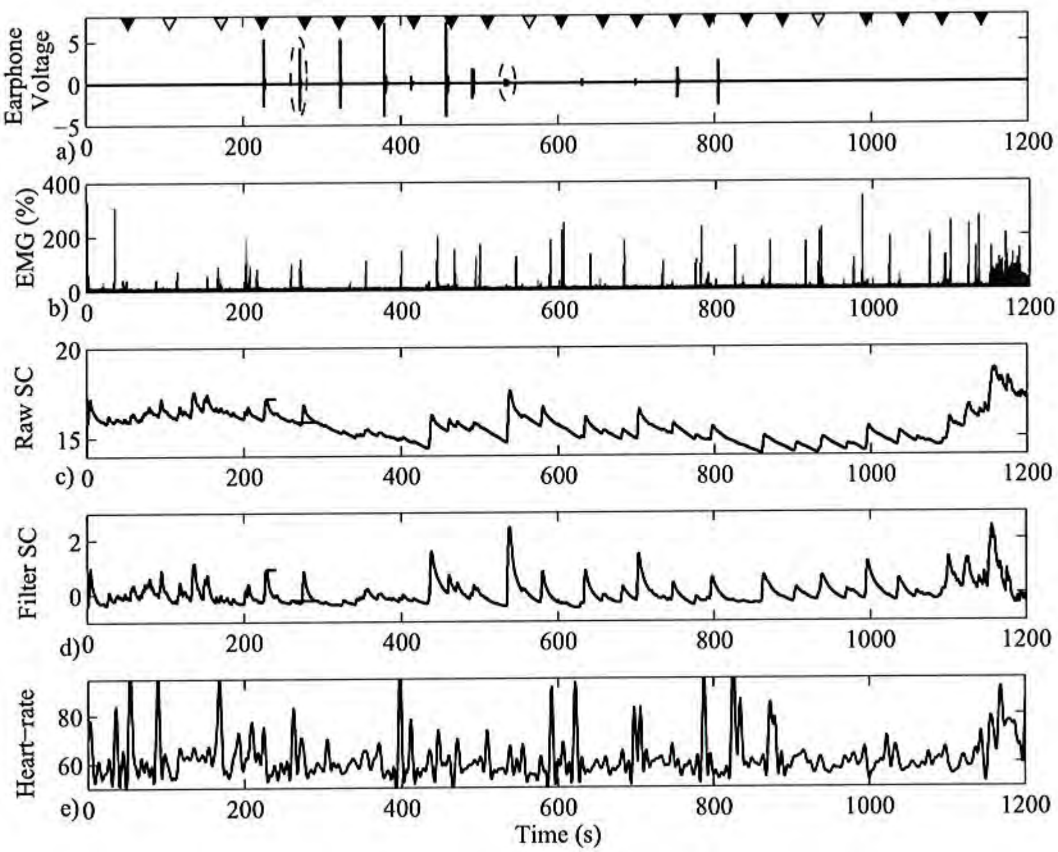

Figure 3.5. Profile of Subject 7's task experiment. Triangles correspond to letter recall screens. (6 correct = black, 5 correct = gray, 4 correct = light grey). a) Earphone voltage: dashed circles denote signals of interest, b) EMG: SCM is black, scalene is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute.

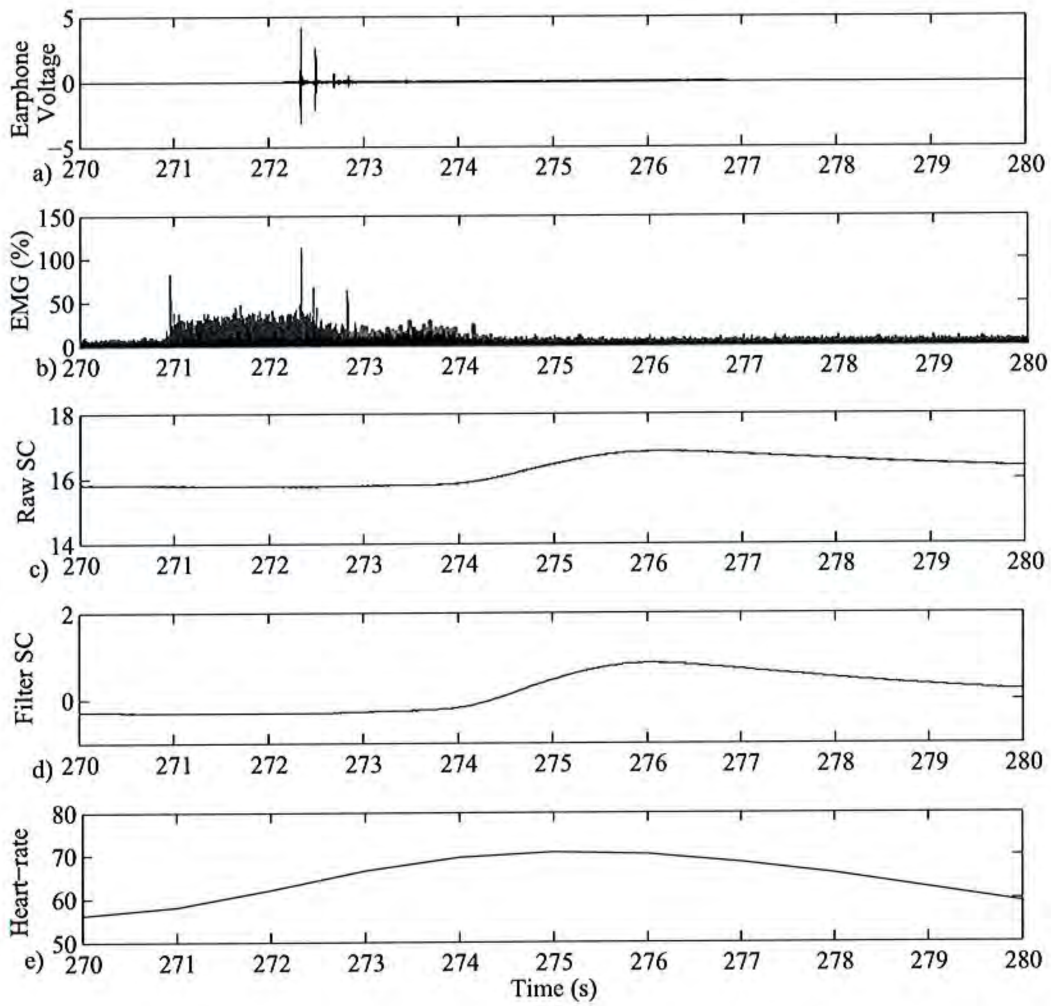

Figure 3.6. A closer examination of the sound exposure of subject 7's task around 270 seconds. a) Earphone voltage, b) EMG: SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute.

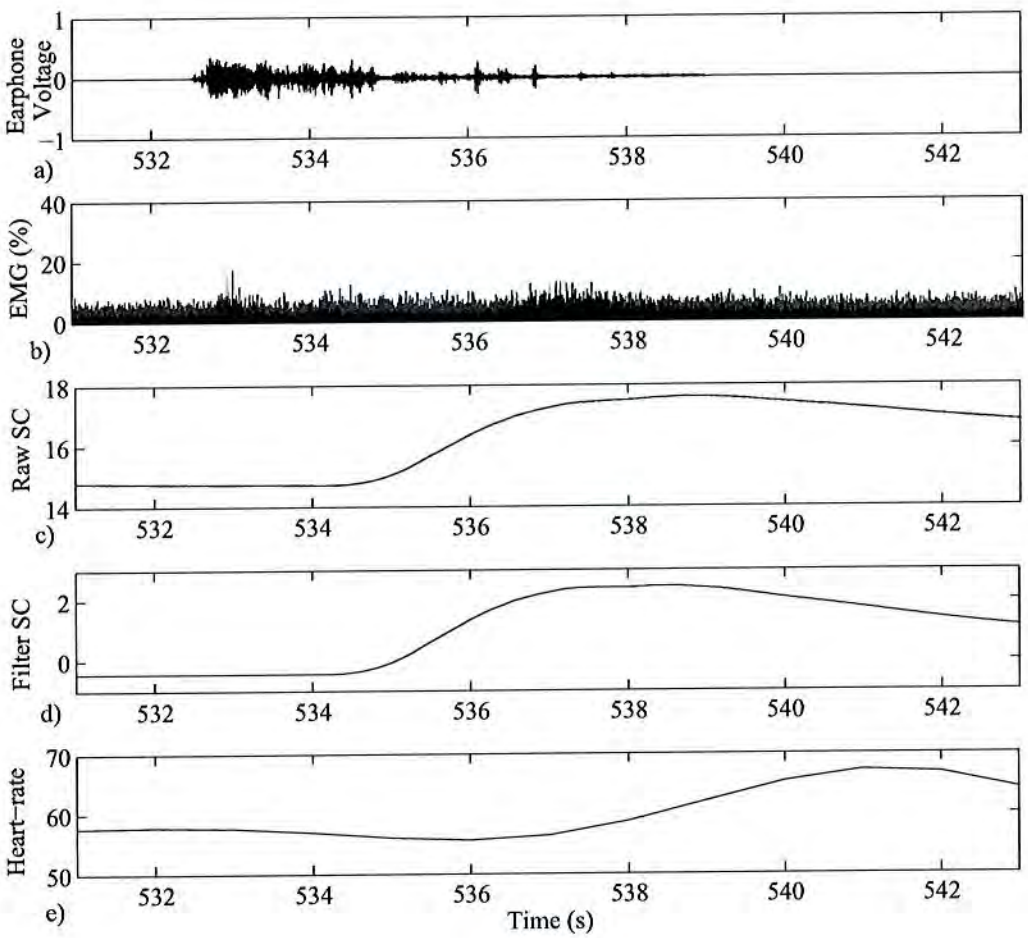

Figure 3.7. A closer examination of the sound exposure of subject 7's task around 532 seconds. a) Earphone voltage, b) EMG , SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μSiemens), d) filtered skin conductance (μSiemens), and e) heart rate in beats-per-minute. Note that due to the high frequency content of this signal, the earphone voltage levels are low.

A profile of the physiological responses measured during the semantic differential portion of the test is displayed in Figure 3.8 for subject 7. Skin conductance and filtered skin conductance responses change after the presentation of stimuli. However, the time interval between stimulus exposures during the semantic differential test is even shorter than that in the task portion: 5-10 seconds compared to around 45 seconds. Skin conductance responses, in particular, do not have enough time to return to baseline before another sound is played. This resulted in a series of combined responses. The heart-rate signal also appears to contain overlapped responses in addition to some movement artifacts. One example of this occurs around 800 seconds where heart-rate oscillates rapidly between low to high values. Given the relatively slow changes in heart-rate observed previously, it is probable these effects are due to movement of the sensor hand.

For the EMG measures, the results are mixed. First, while there are some responses to sounds, there are periods of background movement and noise events that are present in the measurement. One visible artifact is the low level, large duration response in the trapezius signal, from approximately 200 to 550 seconds. Since the subject completed the ratings on a sheet of paper in front of her/him (as opposed to on the computer, like in the task section), it is possible that this response is due to the subject's posture. Leaning over the desk would cause isometric contractions (contraction not associated with movement) in the trapezius. Similarly, the peaked responses observed after the last time each signal is played, are probably due to the subject looking up towards the computer screen to click on the button to play the next sound. Identifying the cause of other artifacts in this signals is more difficult.

3.3.3 Latency Analysis

While there are artifacts present in the data, it is possible to quantitatively analyze the signals by counting the number of responses evoked for each type of measurement. An advantage of this approach is that it is less sensitive to artifacts or variations in responses due to processing methods. The EMG signals were processed and each response was de-

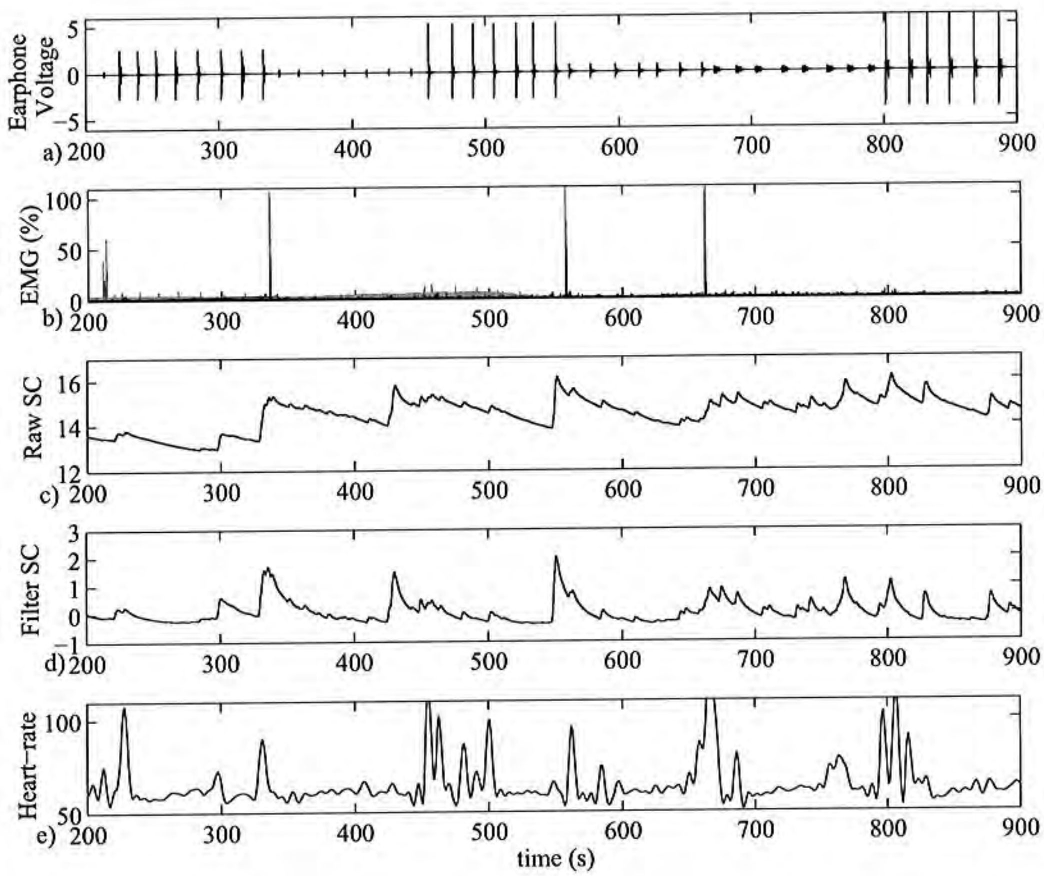

Figure 3.8. A profile of Subject 7's semantic differential physiological measures. a) Earphone voltage, b) EMG , SCM is black, trapezius is gray, and scalene is light grey, c) raw skin conductance (μ- Siemens), d) filtered skin conductance (μ- Siemens), and e) heart rate in beats-per-minute. For the EMG plot, SCM is black, trapezius is gray, and scalene is light grey. Both Skin conductance measures are in μSiemens and the heart rates are in beats per minute.

termined via visual inspection. An EMG response was considered to be evoked by a sound if its onset occurred 40-140 ms after the sound onset and had a magnitude of greater than 2 standard deviations of the mean of the preceding 5 seconds of signal. For the semantic differential test responses, only the responses to the first exposure to each signal were analyzed. In general, the determination of which responses were caused by the stimuli was difficult due to voluntary movements (especially in the semantic differential) and artifacts caused by loose electrodes. The same analysis procedure was repeated for the skin conductance and finger pulse measures. For skin conductance a response was considered to be evoked by a sound if its onset began 1-3 seconds after the sound event or its peak occurred within 3-6 seconds after the sound event. The dual definition was adopted because of the presence of overlapped responses. For heart-rate, the responses were classified based on their latency (short: 2-4 s, medium: 4-7 s and long: 7-10 s) and if the heart-rate was increasing, decreasing or both. To examine changes in the number of responses due to signal exposures over time, the probability of a response occurring in EMG, pulse-rate or skin conductance versus each sound in signal presentation order is plotted in Figure 3.9.

All subjects had a skin conductance response and many subjects had heart-rate acceleration in response to the first stimuli presented in the task experiment. Trapezius and SCM responses were the most common muscle responses. After the first signal presentation, both the number of skin conductance responses and the number and severity of heart-rate responses decrease slightly with additional sound exposures. There are fewer EMG responses than skin conductance and heart-rate responses.

Fewer heart-rate and skin conductance responses are observed in the semantic differential part of the test. This is probably due to the subject being familiar with the stimuli. There is, however, not much variation over the course of the semantic differential. EMG responses occurred at about the same frequency as in the task portion of the experiment. It is possible that these responses are due to the subject rating the scales on paper (requiring neck movement as observed in the last section).

The probability of a response occurring for each individual stimulus (rather than presentation order) is in Figure 3.10 for all of the physiological measures. For the task portion,

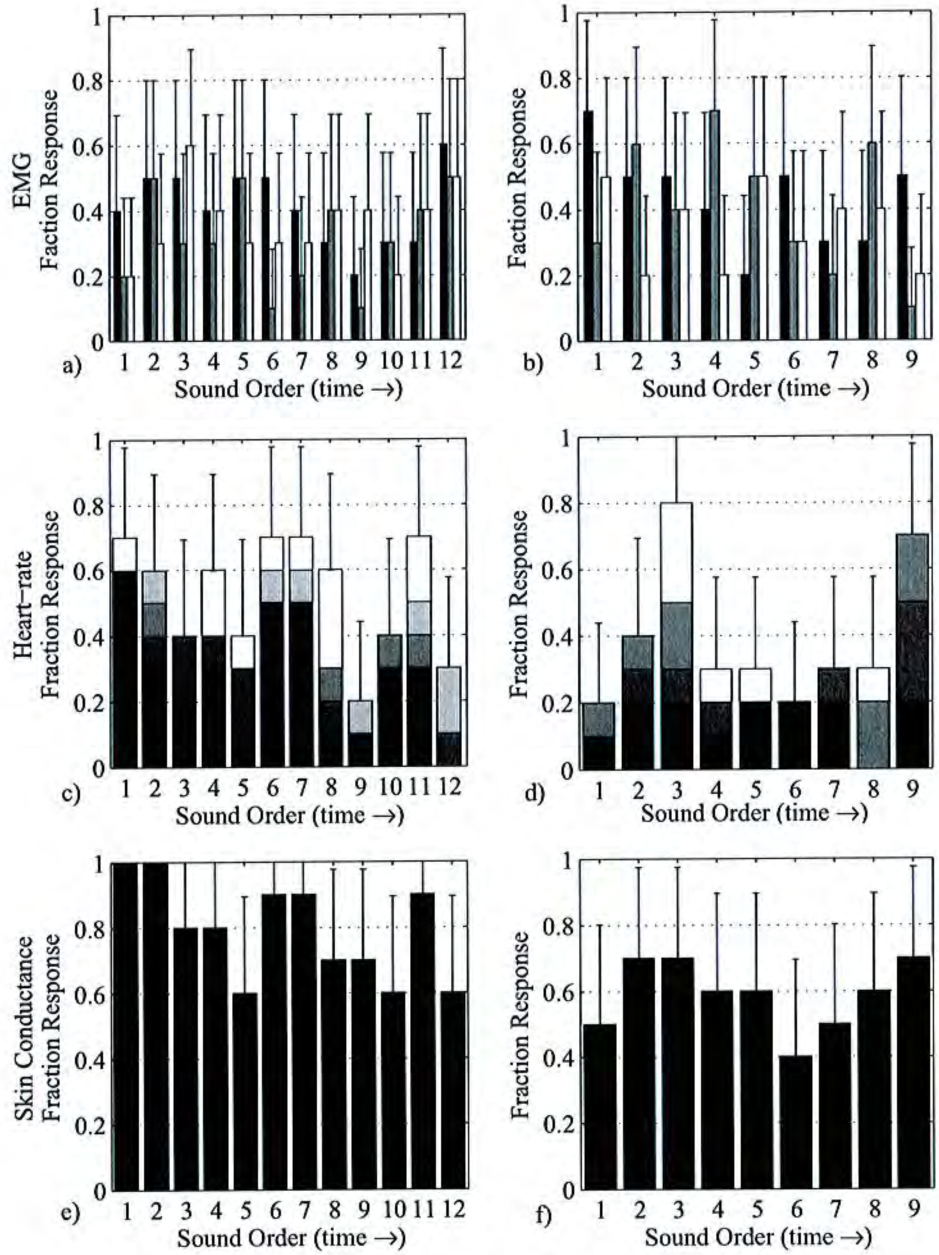

Figure 3.9. Fraction of EMG responses, heart-rate responses and skin conductance responses observed for each sound exposure for a), c) & e) the task portion and b), d) % e) the semantic differential portion, respectively. For EMG, black is SCM, gray is scalene and white is trapezius. For heart-rate: black is short latency increase; dark gray is decrease, then increase; gray is medium latency increase; light gray is long latency increase and white is medium latency decrease.

all of the signals yielded roughly a similar number of responses in all three measures, even the signals with the lowest maximum loudness (1, 5 and 12). This result may be due to ordering effects, since signals presented earlier in the experiment (for heart-rate and skin conductance) were more likely to cause a response and this affects the average value. Despite this effect, for skin conductance, signals with a maximum loudness of greater than 26 sones (signals 4 and 6-9) were very likely to cause a skin conductance response. For the semantic differential portion of the test, the three loudest signals (6, 7 and 9) evoked the greatest number of responses for heart-rate and skin conductance.

3.3.4 Subjective Ratings versus Physiological Responses

While fully understanding relationships between physiological responses and subjective measures was difficult due to the noise present in the measurements and faults in the experimental design, some preliminary effects can be observed. The probability of each physiological response occurring and the mean startle ratings from the semantic differential part of the experiment were compared to the startle ratings and a coefficient of determination (R^2) was calculated, this is listed in Table 3.2.

Table 3.2 Coefficient of determination between probability of response and mean startle ratings. * denote p-values >0.05, i.e., not significant at the 5% level. SD = Semantic differential test.

Response	R^2 Task	R^2 SD
Skin conductance	0.32	0.60
All heart rate (HR)	0.30*	0.68
HR short latency increase	0.07*	0.18*
HR short & medium latency increases	0.37	0.77
All heart rate increases	0.15*	0.59
All EMG	0.04*	0.16*
All SCM	0.02*	0.03*
All anterior scalene	0.02*	0.00*
All trapezius	0.00*	0.47

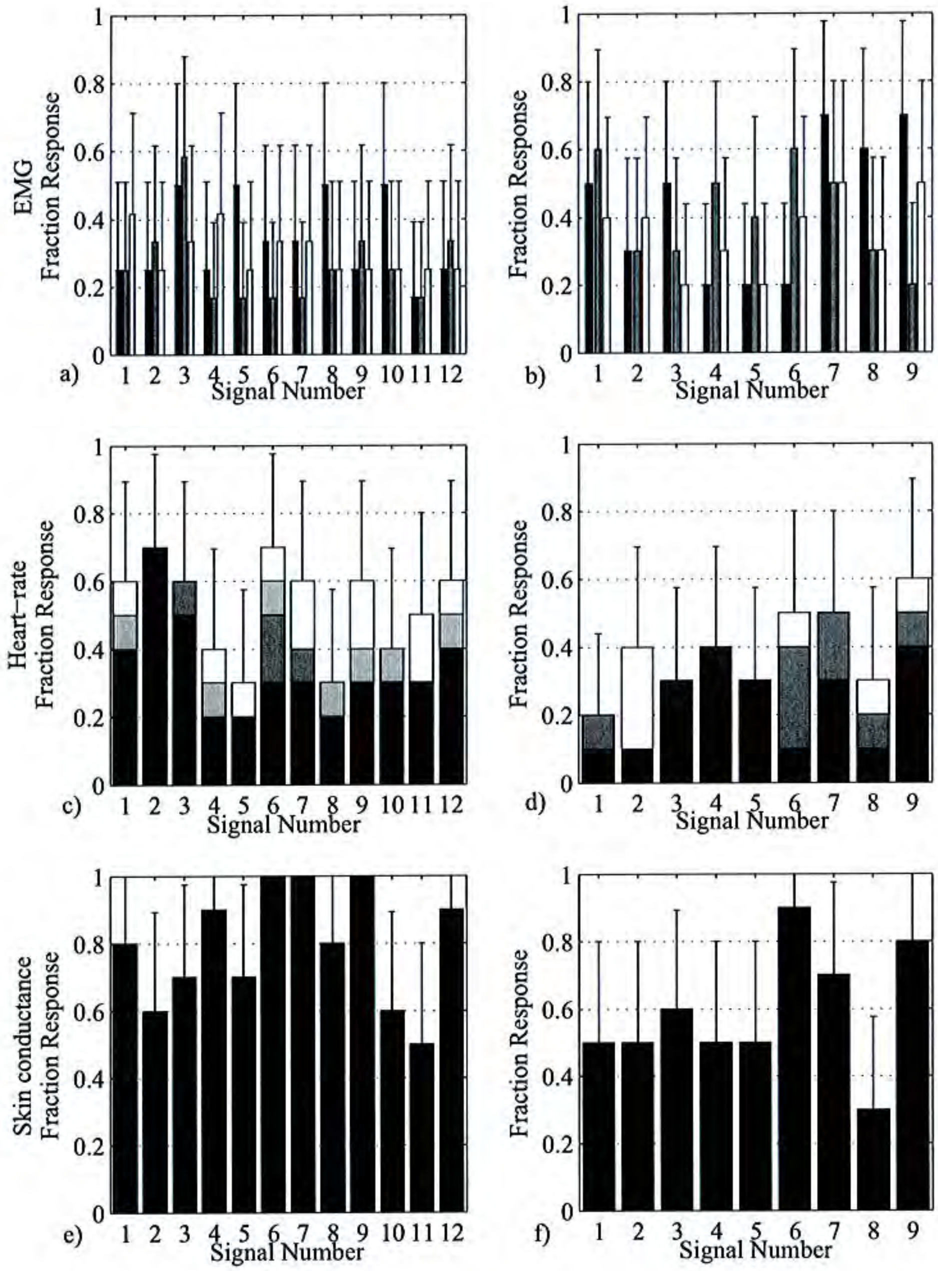

Figure 3.10. Fraction of EMG responses, heart-rate responses and skin conductance responses observed for each signal for a), c) & e) the task portion and b), d) % e) the semantic differential portion, respectively. For EMG, black is SCM, gray is scalene and white is trapezius. For heart-rate: black is short latency increase; dark gray is decrease, then increase; gray is medium latency increase; light gray is long latency increase and white is medium latency decrease.

The measures recorded in the task were not highly correlated to the startle ratings. It is possible that the task, having the potential to cause stress-base responses, influenced this. For example, a particularly stressful math problem combined with a signal presentation could have resulted in a more response than the same signal presented alone.

In the semantic differential test, the probability of skin conductance responses or any type of heart-rate response were correlated with average startle ratings. Mean startle ratings were most strongly correlated to the presence of short or medium latency heart-rate responses. Most of the EMG measures were not correlated to startle ratings. This may be due to the relative infrequency of EMG responses compared to the other two measurements. Similarly, with a resolution of 10% in the probability estimates (only 10 subjects), there is not enough data to distinguish between signals with small differences in average startle ratings. If more subjects were recruited, perhaps, the probability of an EMG response may be correlated with subject ratings.

One difficulty in analyzing the data stems from the short interval between stimulus exposures and the noise present in the measures. If the experiment used longer inter-stimulus intervals, it is possible that some of the slower responses, such as skin conductance, would be more highly correlated with in startle ratings. Most of these issues can be reduced with better experimental design. Using a headrest to minimize postural responses, for example, could potentially result in less noisy EMG responses.

3.4 Summary

In this chapter, an exploratory experiment was described. In this experiment physiological responses to sonic booms and other transients were investigated. The main goal of this study was to understand issues related the to design of boom startle experiments where physiological responses are measured. This experiment was split into two parts, an arithmetic & memory task and a semantic differential test. Despite noise issues, classical startle physiological responses were observed when subjects were exposed to low booms.

In addition, startle judgments were highly correlated to short and medium latency increases in heart-rate, a marker of startle.

Several aspects of the experimental design complicated analysis of the data obtained from this experiment. The main one was the duration of time between stimulus responses which caused response overlap in some of the physiological measures. Part of this was due to the task itself because the arithmetic and memory task also caused stress responses. In addition, there appears to be large variation in physiological responses when people are exposed to these sounds repeatedly. It is unclear how much the responses would vary if the subject was exposed to the sound on a different day. In the next chapter, an experiment designed to address the shortcomings of this experiment will be presented.

4. REPEATABILITY OF PHYSIOLOGICAL MEASURES EVOKED BY LOW AMPLITUDE SONIC BOOMS

In the experiment described in the previous chapter, physiological responses associated with startle were evoked by low booms and other impulse transient sounds. The probability of a skin conductance response or a heart-rate response occurring in response to a sound was found to be correlated with subjects' evaluations of startle. It also was found that there was some subject-to-subject variability and exposure-to-exposure variability in the physiological responses. However, the experimental design limited insight to the extent of this variability. Primarily, the intervals between stimuli in that experiment were too short. Recall that one part of the experiment was an arithmetic and memory task. It was found that the task itself also caused stress responses, further shortening the effective interval between responses. This meant that responses of some of the physiological measures overlapped, which complicated the analysis and the accuracy of measures derived from the physiological response data.

To address these issues another experiment was designed to examine the repeatability of physiological responses to sonic booms. The experiment was split into two parts: a picture-viewing task, and a paired comparison experiment. To examine the repeatability of responses, each participant completed the experiment twice, once a day on adjacent days. Details are described below.

4.1 Methods

Institutional Review board approval was obtained for this study (IRB # 0904007961). The experimental apparatus was very similar to the one described in Chapter 3. As before, the subjects completed the experiment in an IAC double-walled sound booth. The sounds were stored on a computer and were presented using a LynxOne soundcard which was connected to a Tucker-Davis HB-7 amplifier and Etymotics ER-2 earphones. Sounds were pre-processed to compensate for the earphones operation mode whereby the stimulus is reproduced at the eardrum rather than at the entrance to the ear canal. The computer was placed outside of the sound booth and only the keyboard and screen were inside. The subject's progress was observed through an additional screen outside of the booth. Only one subject was tested at a time.

Finger pulse, skin conductance and neck muscle activity were measured. Skin conductance was measured using silver/silver-chloride finger cuffs, attached to the index and middle finger of the left hand. On the ring (3rd) finger of the subject's left hand, an optical finger pulse sensor was placed. Both the skin conductance and finger pulse sensors were attached to a battery-operated bio-amplifier made by James Long Company. Both raw and high-pass filtered skin conductance were measured. Electrical activity of the sternocliedomastoid (SCM), anterior scalene and trapezius were measured with sticky surface electrodes. To reduce issues with loose electrodes described in Chapter 3, the electrodes were taped to the subject's skin. The locations of these muscles are displayed in Figure 3.1. In addition to the physiological measures, the voltage applied to the earphones was also recorded, to synchronize stimulus presentation with the recorded physiological responses. The measures were sampled at a rate of 4096 Hz with a 7-channel Bruel & Kjaer PULSE system using the PULSE time-data recorder software. The responses and earphone voltage were recorded continuously throughout the test.

4.1.1 Signals

Two sets of stimuli, all synthetic sonic booms, were used in this experiment. The stimuli were designed using a frequency-domain sampling technique used in finite impulse response (FIR) filter design. Briefly, a continuous-time Fourier transform of the candidate signal was calculated and sampled in frequency (at an interval of 1.9 Hz). Next, the signal was windowed in frequency so that there was no signal energy at frequencies greater than half the desired sampling rate (16 kHz for the 32 kHz signals). The window was flat in frequency from 0-8000 Hz and a half cosine from 8000 Hz to 16000 Hz and zero above 16000 Hz. Finally, the windowed spectrum was "mirrored" to have complex conjugate symmetry about half the sample rate and the inverse discrete Fourier transform was performed, yielding the sampled boom signal (16384 points) in the discrete time domain. The signal was rearranged so that the acausal parts were placed before the causal parts. The details of the signal's frequency characteristics were checked by zero-padding to 64000 points and performing a discrete Fourier transform. This result was compared to the spectrum of the continuous signal. This frequency domain sampling technique was used to avoid artifacts induced by direct sampling the candidate signal in the time-domain.

Because the main purpose of this experiment was to examine the repeatability of startle responses evoked by sonic booms, it was necessary to select signals that could evoke startle responses. As described in detail in Chapter 1, Thackray et al. (1974) found in a study that 30 Pa and 50 Pa overpressure booms evoked significantly more responses than 16 Pa booms. However, in their study, the rise time of all stimuli was constant (approximately 4 ms), due to limitations in the kinds of signals they could play. As the time-histories of expected low booms can vary substantially from classical N-waves, it was decided to use synthetic sonic booms with similar time-varying loudness statistics as the waveforms used by Thackray et al. (1974). In Chapter 2, it was found that one of the most consistently performing models to predict startle was a linear combination of maximum loudness and maximum loudness derivative (dL_{max}) prior to the first loudness peak. Five signals were created, all at the levels of the 16 and 30 Pa overpressure signals used by Thackray et al. (1974). One signal was designed to be the same 30 Pa signal used in that study and the rest were designed to have the highest and lowest dL_{max} expected from shaped low booms. In Figure 4.1a), a summary of the stimuli metric values is displayed. In Figure 4.1b)-f), the short-term loudness time-histories of the stimuli are presented (calculated using Moore and Glasberg's time-varying loudness). In order to obtain the low dL_{max} values, it was necessary to make the first loudness peak of signals 2 and 3 much lower than the second peak level.

Due to the limitations in the reproduction of low frequencies with the playback system, all signals were high passed filtered (3rd Order, Butterworth). Signals 1, 4 and 5 were high pass filter with a cut-off of at 25 Hz and signals 2 and 3 were high pass filtered with a cut-off of 35 Hz. The higher cut-off frequency for signals 2 and 3 was necessary to play these signals at the correct maximum loudness. This filtering was accounted for in the calculation of metrics.

To examine how responses would change if subjects were exposed to booms at a higher level that these five stimuli, a small additional study was conducted. A single high level stimulus (denoted by the square in Figure 4.1) was designed and used as part of a small study to examine responses to a higher-level boom.

4.1.2 Procedure

The experimental procedure was the same on both days the participants completed the test. The test began by explaining the experiment to the participant, obtaining informed consent and testing the participant's hearing. She/he was then placed in the sound booth and the finger cuffs (two for skin conductance and one for heart-rate) and the EMG electrodes were attached. At this time, the chin-rest was adjusted for the participant comfort. The participant then completed a calibration procedure to obtain baseline levels of the physiological measures and calibration information for the EMG measures. For this part, a pre-recorded set of instructions was played to the subject over her/his earphones. First, the subject was asked to sit relaxed and complete several head-circles without being in the chin-rest. Then the subject was asked to sit relaxed, tilt his/her head from side-to-side and simulate coughing while using the chin-rest. The head-circles, head-tilting and relaxed sitting were selected both to obtain baseline and elevated measures for the recorded signals. The simulated coughing was added to the instructions to obtain examples of one of the common artifacts that were found in the pilot study (Chapter 3).

After the subject completed the calibration procedure, the test instructions were read to them. The participants were asked to sit and use the chin-rest while looking at pictures.

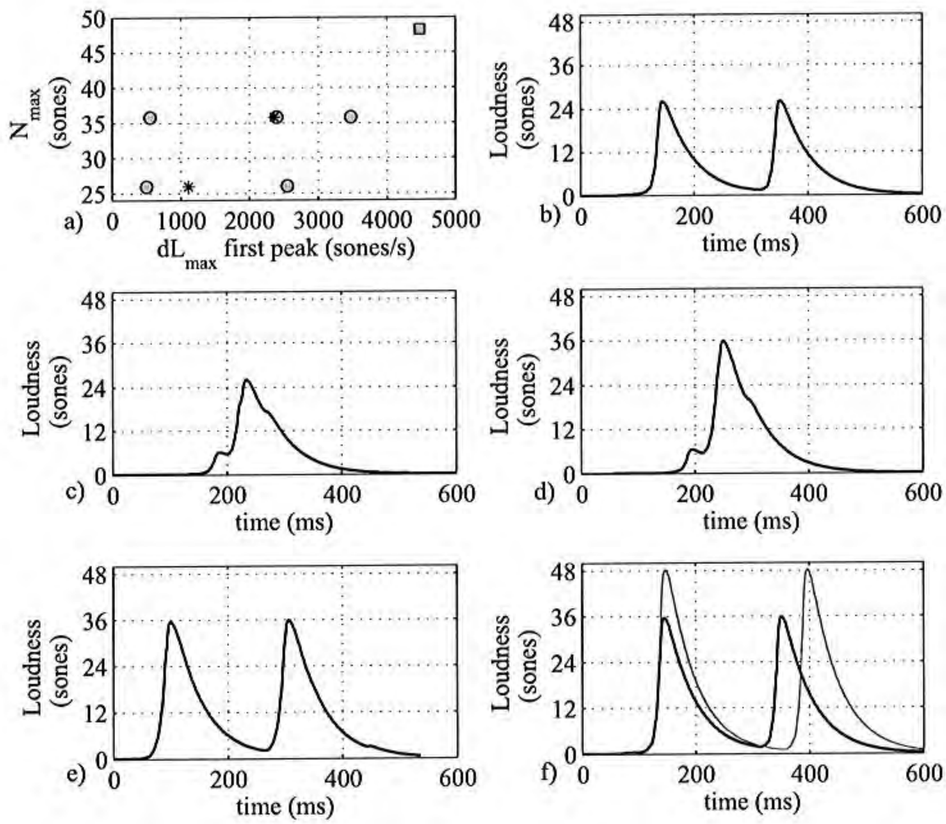

Figure 4.1. a) Maximum loudness versus maximum loudness derivative of first loudness peak for stimuli. Levels of stimuli used in Thackray et. al. (1974) are stars. High amplitude stimulus is a square. The five stimuli in the main study are circles. Loudness time-histories of each of the signals: b) signal 1, c) signal 2, d) signal 3, e) signal 4 and f) signal 5 (black) and signal 6 (gray).

They were also told that they might hear sounds through their earphones. They were instructed to ignore them and continue looking at the pictures. At no point in the instructions were the subjects given any indication that the experiment was about startle. The subject then completed the picture viewing part of the study.

Each picture was displayed for 20 seconds and the order of pictures was kept constant for all subjects and sessions. Recall that in Bradley and Lang's study (2002), the content of pictures presented alongside startling acoustic stimulus has been found to modify startle responses. Pictures of landscapes were selected for this experiment because these were found to be neutral, neither increasing or decreasing the severity of responses in that study. As an unintentional bonus, landscapes were also consistent with the scenario in the subject instructions for the experiment presented in Chapter 2 ("outdoors in a park or garden...").

The background noise, nominally 40 dB unweighted low pass filtered white noise, was constant throughout the entire test. After five minutes of pictures, the subject was exposed to the first boom stimulus and additional exposures followed at a rate of approximately one exposure every 2 minutes. The timing of exposures was the same for all sessions and subjects. The subjects were exposed to the stimuli ten times each day. For the main study, each of the five signals was played twice. For the mini-study using the high-amplitude stimulus, the same signal was played all ten times. During the course of this experiment, the subjects' movements were monitored by the author to denote times of potential movement artifacts (such as sneezing, coughing and stretching).

After this part of the experiment, the subject then completed a paired comparison study with the same stimuli they heard in the first part of this experiment. Subjects who completed the high-amplitude mini-study did not complete this part of the test. For each trial, the subject heard a pair of sounds and was asked to judge which of the two was "more startling." Subjects heard all combinations of the five stimuli (both A-B and B-A pairs) at some time in the test and the ordering of the pairs was randomized, a different order for each day and subject. A one second delay was used between each sound of a pair. The subject completed the paired comparison test three times, 20 pairs per test, for a total of 60 pairs per session.

After the subjects completed all of the tests for the day, pictures and paired comparison for the five main stimuli or just pictures with high-level stimuli, the subjects' hearing was tested again and the subjects were compensated ($10 for each day). The next day, the subjects repeated the experiment again, following the same procedure as described above.

4.1.3 Subjects

Nineteen subjects were recruited for the main study and five subjects were recruited for the high-amplitude study. All subjects had less than 20 dB of hearing loss at frequencies of 125-8000 Hz. Subject ages ranged from 18-52 years. For the main study, nine of the subjects were female and ten were male. For the high-amplitude study, four subjects were male and one was female. Each subject in both tests completed two sessions, one each day, over the course of two days.

4.2 Results

Once the data was obtained, the measured physiological responses were post-processed. The EMG signals were full-wave rectified and normalized by the maximum value of the calibration procedure. This maximum occurred for all subjects during the head circles. The output of the pulse-rate signal was converted to heart rate via a peak-picking algorithm. For each physiological signal, the time-histories were visually inspected for responses. Responses for the sounds and the pictures prior to the sound exposures were examined. Pictures in the first minute of the experiment were excluded from analysis due to the presence of artifacts in the signals due to the subject settling into the test. Responses to pictures were examined to understand what responses would look like without the presence of sounds. For the repeatability study, the picture timings were not synchronized with stimulus presentation. To obtain picture transition times, a delay parameter was estimated so that the greatest number of skin conductance responses would have latencies consistent with responses evoked by the pictures. Pictures changed every 20 seconds so once this delay was estimated the timing of all picture changes could be determined.

An EMG response was considered to be evoked by an event if its onset occurred 40-140 ms after the sound onset and had a magnitude of greater than 2 standard deviations of the mean of the preceding 5 seconds of signal, with no response above that threshold for at least 100 ms prior to the event. As the EMG signals have expected latencies that were similar to the duration of sound signals, latencies were calculated using both the first pressure peak of the sound and the end of the sound as the definition of when the sound event occurred. The same analysis procedure was repeated for the skin conductance and finger pulse measures. For skin conductance a response was considered to be evoked by an event if its onset began 1-3 seconds after the sound event. For heart-rate, a responses was considered to be evoked by an event is the heart-rate magnitude was 2 standard deviations above or below the mean heart-rate of the 10 seconds preceding the event. Heart-rate responses were also classified based on their peak latency (short: 1-3 s, medium: 4-7 s and long: 7-10 s) and if the heart-rate was increasing, decreasing or both.

The probability of a response to the 5 main stimuli, high amplitude stimuli and pictures as well as the probability of each subject having a response are shown for skin conductance and heart-rate in Figures 4.2 and 4.3. The probability of a response to the sound stimuli pictures and subjects for the EMG signals is in Figures 4.4 and Figure 4.5.

4.2.1 Skin Conductance

Skin conductance responses were the most common responses observed on both days. Signals 1 and 2, the signals with the lowest maximum loudness, appear to have the least skin conductance responses on each day. Signals 4 and 5, the signals expected to be the most startling, evoked the highest number of responses. There are fewer responses on Day 2 than on Day 1 (average probabilities of 0.76 ± 0.04 versus 0.62 ± 0.05) overall probability of response). There is a slight reduction in responses of the course of the experiment for some of the sounds, however, this result is not statistically significant. The sounds evoked many more skin conductance responses than the pictures (Figure 4.2).

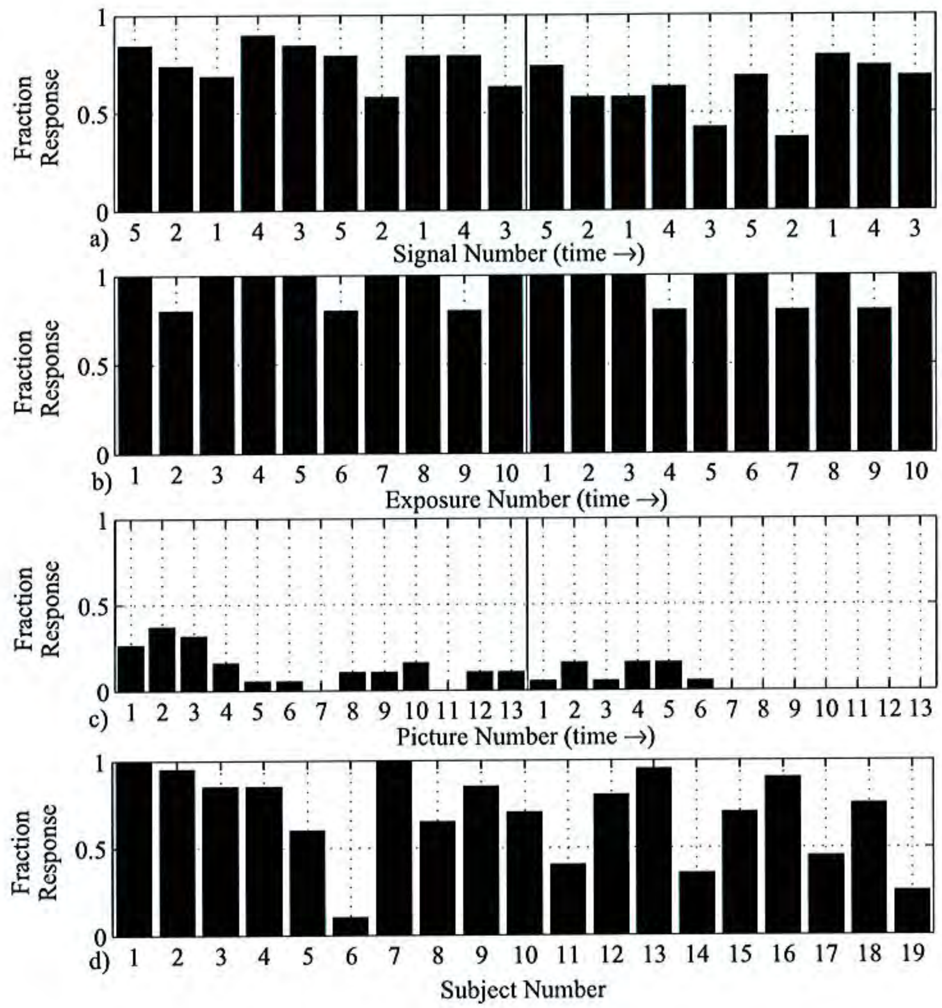

Figure 4.2. Fraction of skin conductance responses obtained from a) 5 sound stimuli, b) high amplitude stimuli and c) pictures on both days. d) Fraction of responses for each subject on both days. for a) skin conductance, b) heart-rate, c) SCM, d) scalene and e) trapezius. Vertical line separates Day 1 and Day 2 responses.

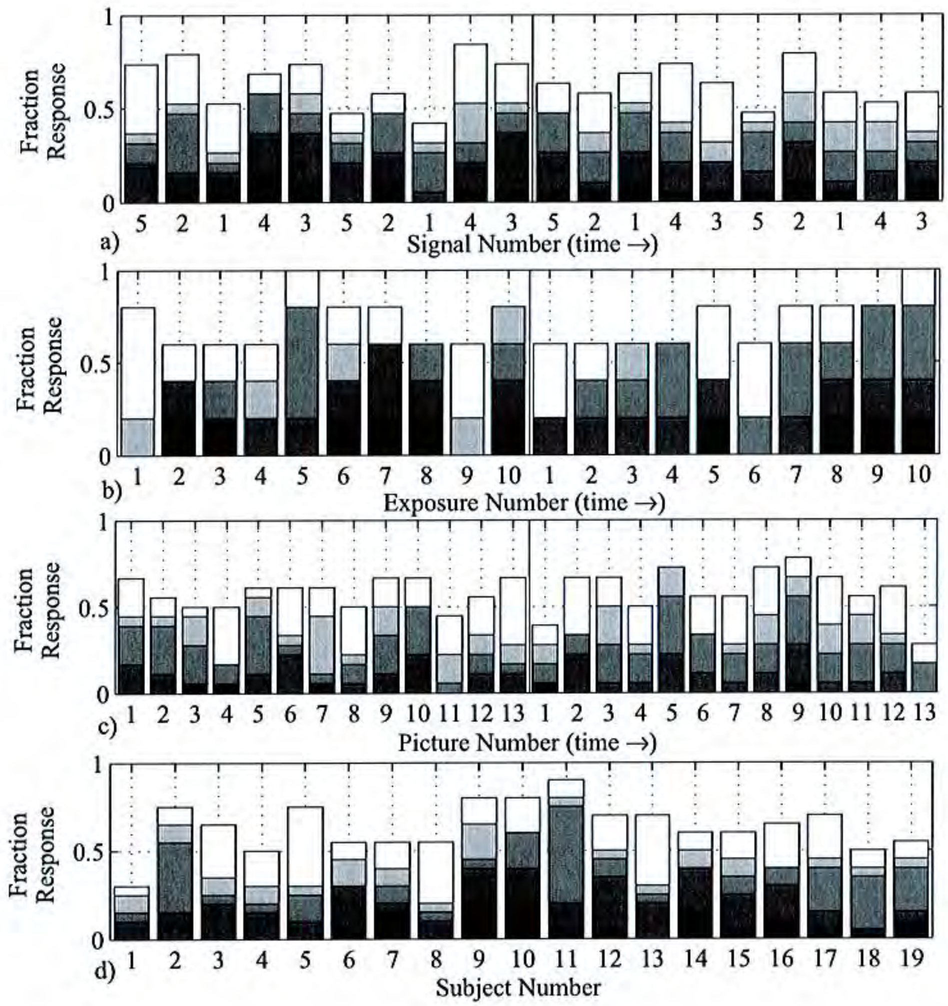

Figure 4.3. Fraction of heart-rate responses obtained from a) 5 sound stimuli, b) high amplitude stimuli and c) pictures on both days. d) Fraction of responses for each subject on both days. Short-latency heart-rate increases, decrease then increases, medium latency increases, long latency increases and decreases only are denoted by black though white respectively. Vertical line separates Day 1 and Day 2 responses.

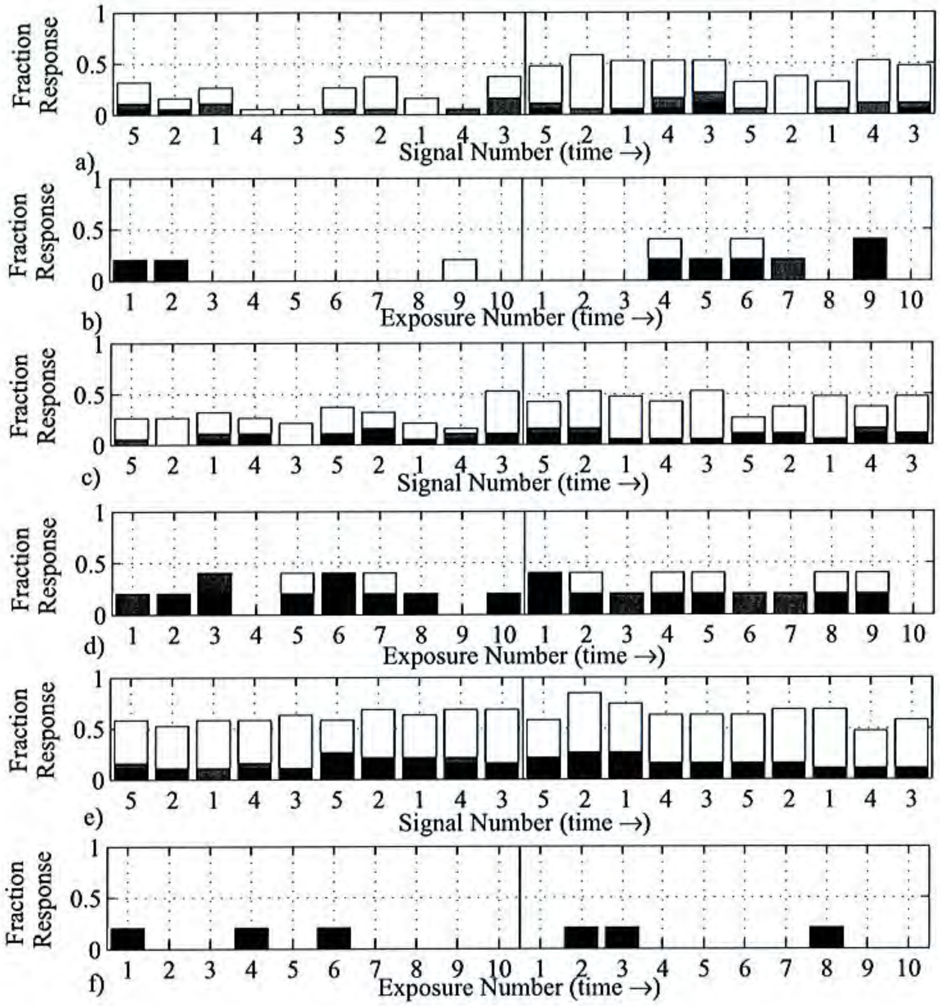

Figure 4.4. Fraction of a) SCM, c) scalene and e) trapezius responses obtained for 5 main stimulis. Fraction of b) SCM, d) scalene and f) trapezius responses for high amplitude stimuli. For Black denotes responses with a magnitude of greater than 10% of head circle, gray denotes all responses from the sound start, white denotes responses from the end of the sound. Vertical line separates Day 1 and Day 2 responses.

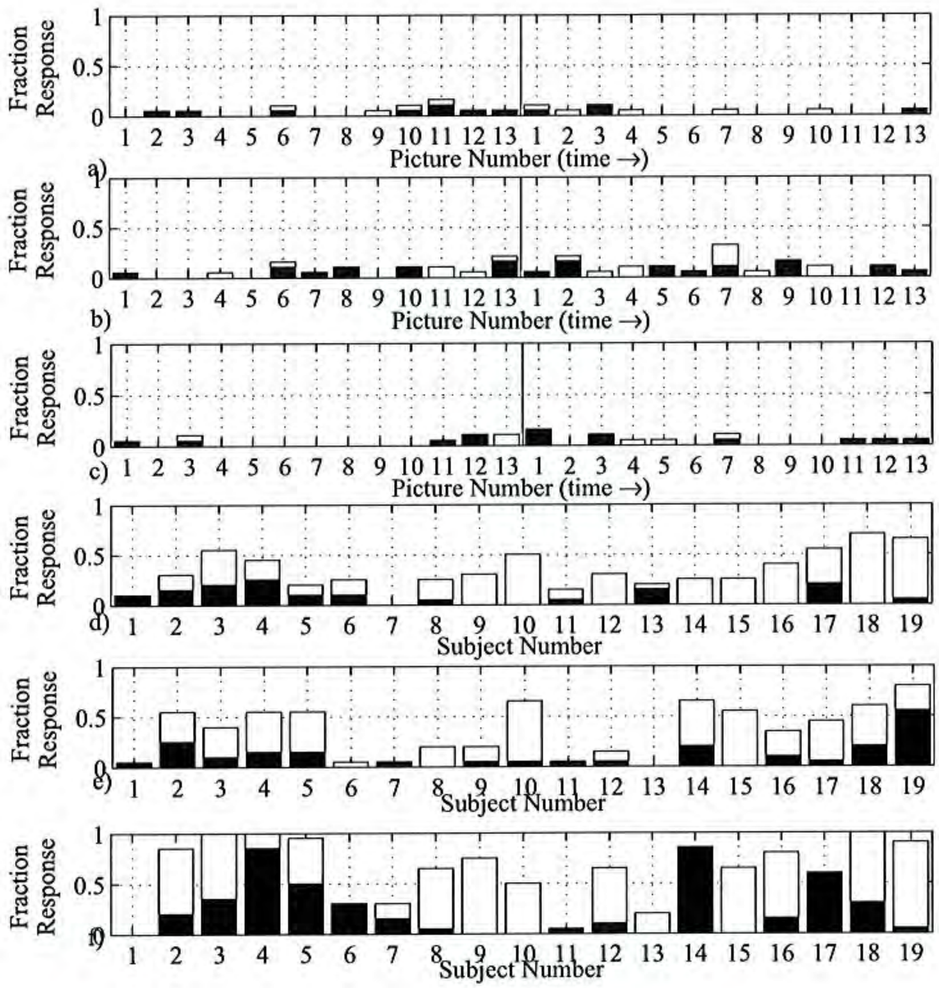

Figure 4.5. Fraction of a) SCM, c) scalene and e) trapezius responses obtained for pictures. Fraction of b) SCM, d) scalene and f) trapezius responses for each subjects on both days. For Black denotes responses with a magnitude of greater than 10% of head circle, gray denotes all responses from the sound start, white denotes responses from the end of the sound. Vertical line separates Day 1 and Day 2 responses.

As shown in Figure 4.2d), there is variation in the number of responses obtained from each subject. Five out of the 19 subjects had responses to nearly all of the sound exposures (subjects 1,2,7,13 and 16) and an additional 5 subjects had responses to more than half the sounds (subjects 3,4,5,8,9,10,15 and 18). A few subjects (4/19), however, did not have many responses at all (subjects 6, 11, 14, and 19). The high-pass filtered skin conductance time-histories were plotted for a subject from each of these three groups (Subjects 2, 4 and 6) in Figure 4.6.

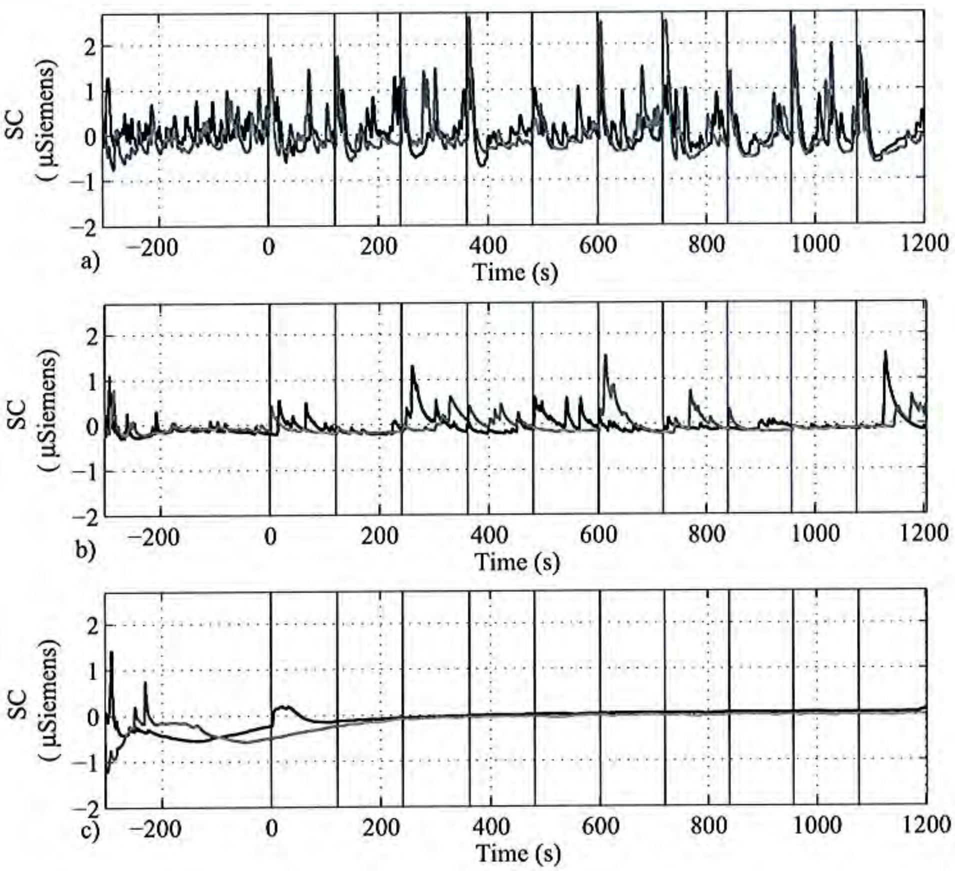

Figure 4.6. Filtered skin conductance time-histories for subject a) 2, b) 4 and c) 6. Black is the first day and gray is the second day.

There does not appear to be a lot of day-to-day variation in skin conductance time-histories for these subjects. The main difference in skin conductance time-histories appears to be the number of responses that were not associated with the sound exposures. Subject 2, for example, has many responses to both the sounds and to other events (pictures and others). Subjects with fewer responses generally had fewer responses to pictures as well. It is not clear what the subjects in the most sensitive and least sensitive groups have in common with the other members of their groups. Both groups contain subjects both genders and a mix of ages. The subject whose skin conductance responses is shown in Figure 4.6c), had only one large response and a few very small responses (e.g., the sound exposure between 400 and 600 seconds). While subjects 11, 14, and 19 had more responses than subject 6, all of these subjects, based on skin conductance responses, were not particularly sensitive.

The high amplitude signal produced significantly more skin conductance responses than the five repeatability study stimuli (Figure 4.2)b). The probability of responses do not significantly decrease over time. The mean response amplitudes of the skin conductance responses were obtained and plotted for the high amplitude study and for the repeatability study in Figure 4.7. As expected, the high amplitude stimulus caused much higher magnitude responses than evoked from exposure to the other 5 sounds. Both sets of responses were similar from day-to-day. Both showed a reduced response magnitude over the duration of the test. The responses to the 5 stimuli, however, do not decrease as noticeably as those from the high amplitude boom.

4.2.2 Heart-rate

A variety of heart-rate responses were evoked by each of the main 5 sounds (Figure 4.3)a). Fewer responses were observed on day 2 than day 1, but this was not significant (average probabilities were 0.65 ± 0.04 and 0.62 ± 0.04). The distribution of types of responses observed for day 1 were also different to those observed on day 2. For example, signal 5 on day 2 produced a greater number of more severe, "darker", responses on day 2 than day 1. The sounds that evoked the most short-latency heart-rate increases, the most severe heart-rate response, were signals 1 and 4, which have similar dL_{max} values. Signal 5, the other high dL_{max} signal, did not evoke many of these types of responses. Heart-

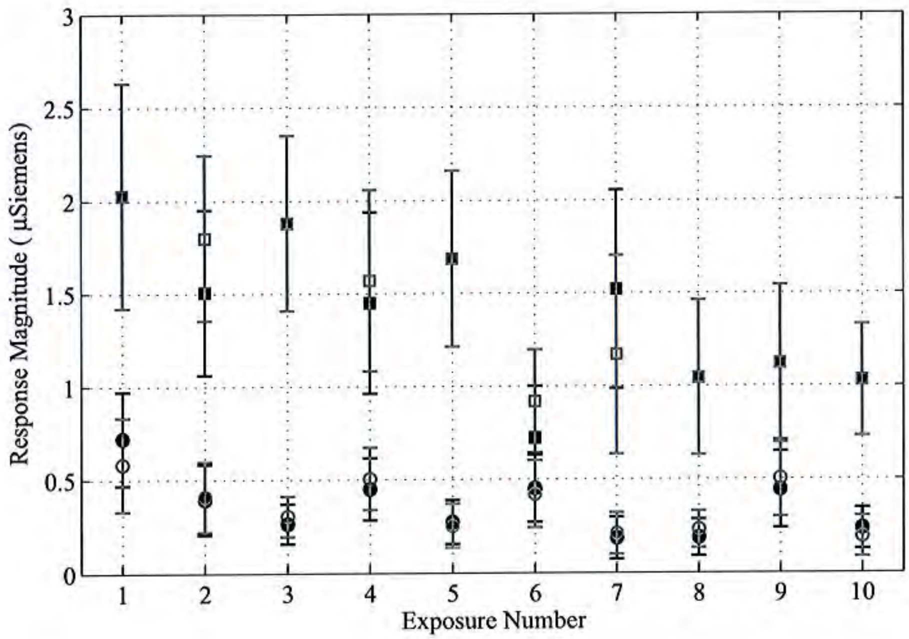

Figure 4.7. Average skin conductance response magnitude for both high amplitude stimuli (squares) and 5 original stimuli (circles). Black is day 1 and gray is day 2.

rate responses to sounds were more likely to occur for sounds than pictures (0.64 ± 0.03 compared to 0.58 ± 0.03). Heart-rate acceleration responses occurred more frequently in response to sounds than pictures (0.37 ± 0.03 compared to 0.28 ± 0.03).

One source of this variability is differences between individual subjects (Figure 4.3d). Most subjects had some type of heart-rate response at least half of the time. In general, the more heart-rate responses a subject had, the greater number of severe responses were present (see, for example subjects 9-12). However, several subjects (2, 3 and 4) did not have any short-duration heart-rate increases, despite a large number of total responses.

The responses in the high-amplitude study were slightly more severe than that those in the five repeatability stimuli study (Figure 4.3b)). However, there is a similar level of variation in the types of responses evoked. The first exposure to the high-amplitude stimulus resulted in mostly less severe responses. This was similar to the heart-rate responses to the first exposure of signal 5 on the first day. It is possible that these sounds are not severe enough to consistently evoke startle responses, hence the lower severity responses to the first signal exposure in both tests.

The average heart-rate response profiles for each exposure to 5 main sounds and the high-amplitude stimulus are plotted in Figure 4.8. The method for calculating profiles was used by Ramirez et al. (2005) and Turpin et al. (1999). The profile was created by calculating the median of heart-rate for intervals at 1-3, 3-6, 6-11, 11-16, 16-22, 22-28 and 28-35 second after the beginning of the stimulus exposure for each subject' heart-rate responses to each sound stimulus. The 0-1 second interval was ignored because of the transit time associated with the use of a finger pulse sensor. The mean heart-rate of the 15 seconds preceding each exposure was subtracted from these values for each response. Then the values were averaged across subjects with a heart-rate response to that sound to yield the heart-rate profiles for each stimulus exposure on each day. Heart-rate time-histories with no response (determined via analysis of latencies) were not included in this analysis.

For the five main stimuli, responses were very small in magnitude. A few signals have increases in heart-rate at the 1-3 second interval. However, due to the variation present, these responses are not significantly different from each other. In addition there appears to

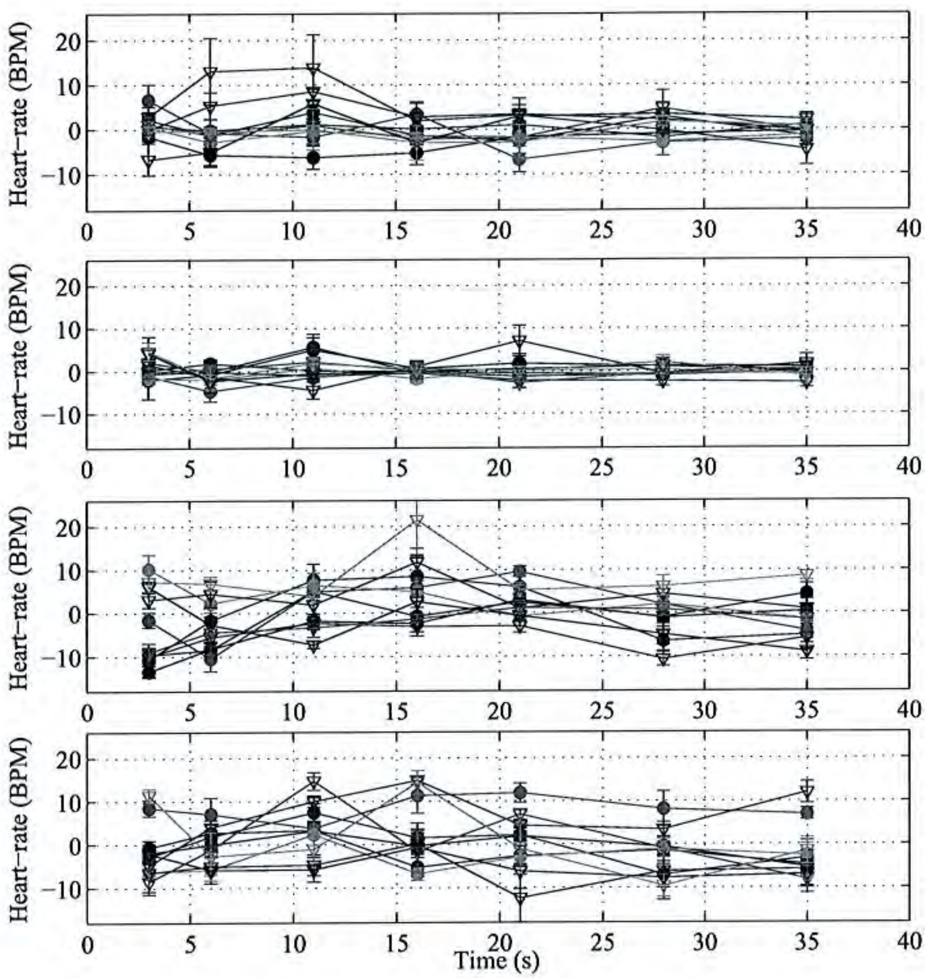

Figure 4.8. Heart-rate profiles of a) Day 1 main stimuli, b) Day 2 main stimuli, c) Day 1 high stimulus and d) Day 2 high stimulus. Black through light gray filled circles correspond to 1st through 5th sound exposures. Black through light gray triangles correspond to 6th through 10th sound exposures. For a) and b) exposure order was signal 5, 2, 1, 4, 3 for the 1st-5th and 6th-10th exposures respectively. Error bars are standard deviation of the estimated mean.

be a long latency increase in heart-rate for some of these signals occurring around 6 and 11 seconds. This type of response is commonly attributed to a defense response (e.g. Ramirez et al. (2005)). This response is also present in day 2. The large variation in responses is expected given the results of the latency analysis; there was a lot of subject-to-subject variability in the responses.

The heart-rate responses were larger in magnitude and more severe for the high-amplitude stimuli. First, there were heart-rate increases at 3, 11 & 16 seconds and some of these were significant. After more exposures, the magnitude of heart-rate changes becomes smaller. On Day 2, more of the heart-rates decrease the 1-3 second intervals than Day 1 (8 on Day 2 verses 4 on Day 1). The magnitude of heart-rate changes at 1-3 seconds is larger than that observed for the main stimuli as well. However, unlike much of the startle literature, heart-rate increases were not consistently evoked by either set of the stimuli. It is possible that if these responses were present, they are small in magnitude and would only be discernible if many more subjects were recruited so that mean estimates were more accurate.

4.2.3 Electromyography

The fewest responses observed for any physiological measure examined in this study were the EMG signals (Figure 4.4). In addition, many of the responses to the 5 stimuli were very low in magnitude, especially for the SCM and AC muscles. The trapezius responded the most of the muscles and some of these responses were of higher levels. However, no more than 25% of subjects had a response of greater a tenth of the maximum of the calibration procedure. It is possible that this result is due to the motion restriction caused by the presence of the chin-rest. Background levels of this muscle tended to be larger than for the other two muscles (average background levels of the trapezius were, on average 2-4 times higher than for the SCM and AC), thus stimulus responses tended to be larger. These higher background levels may have been due to posture demands imposed by the chin-rest. Due to its location in front of the subject, many of the subjects leaned into the chin-rest. This situated their heads with the neck flexed over the desk and into the chin-rest when taking the test. This meant that the trapezius (which extends the neck) was the only muscle that was resisting the weight of the head, thus yielding larger background levels and responses. The majority of responses were found to be within 140 ms of the end of the sound exposure (white in Figure 4.4); fewer, in comparison, were found within 140 ms from the onset of the signals. This result is possibly due to the less stringent criteria for these later responses. More of the later responses were found on the second day than the first.

For the high-level stimulus, there are fewer responses evoked than evoked by the five main stimuli (Figure 4.4b),d) and f)). However, most of these responses are larger in magnitude. For the SCM and the trapezius, many of these responses occurred towards the beginning of the experiment. Very few EMG responses were evoked by the pictures and none of them were of high magnitude.

It is possible that the large differences in responses between those observed for the high amplitude signal (Figure 4.4) and for the main stimuli were due to variations between individual subjects. The number of EMG responses evoked by each subject varied considerably. In figure 4.5, subjects 1, 2, 11, and 13 had very few responses. Subjects 4, 5, 14, 18 and 19, however, had many high level responses, especially for the trapezius.

Despite the small number of responses evoked by the high amplitude stimulus, some comparisons can be still be made between the response magnitudes. The average normalized response magnitudes are plotted versus exposure number in Figure 4.9. While some exposures did not yield a response, the high amplitude stimulus had larger response magnitudes than the original stimuli, especially for the SCM and trapezius. The scalene responses were of roughly similar magnitudes between the two experiments. In general, response magnitudes tended to decrease over the course of the experiment for the high amplitude stimuli to roughly the levels associated with the five stimuli. For the five stimuli, responses had smaller magnitudes the second time the sounds were played (exposures 6-10). However, the average magnitude of these responses are very small.

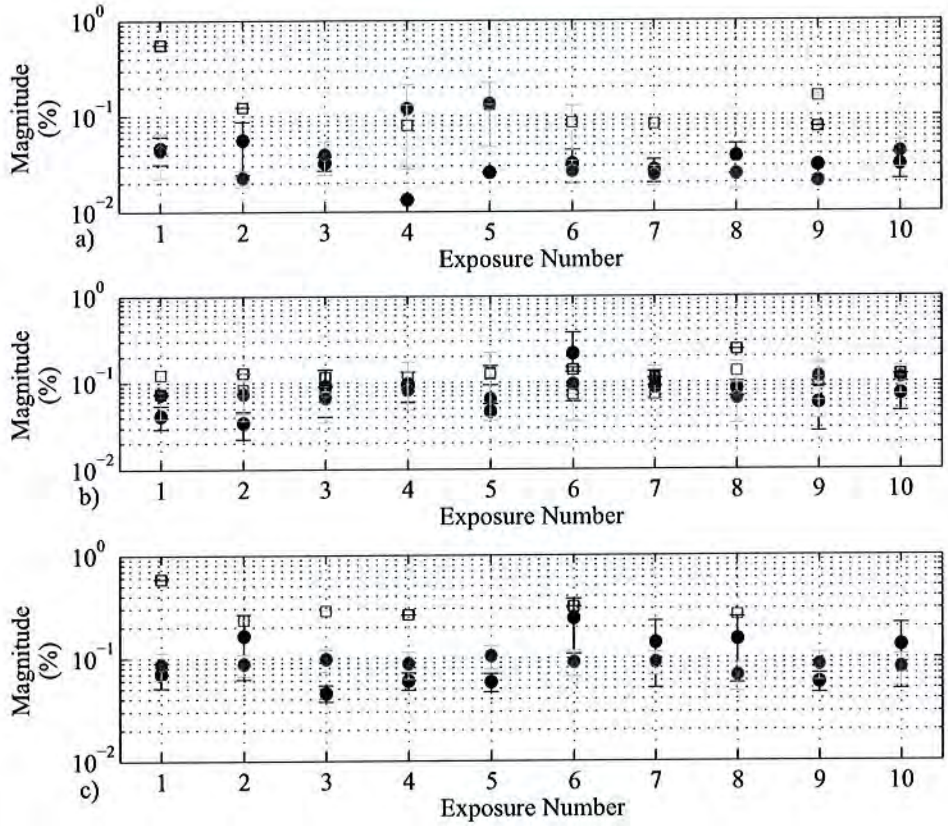

Figure 4.9. Average magnitude of EMG responses (in % of maximum obtained in calibration procedure) for both high amplitude stimuli (squares) and 5 original stimuli (circles) measured from the a) SCM, b) scalene and c) trapezius. Black is day 1 and gray is day 2.

4.2.4 Paired-comparison Analysis

The paired comparison data were converted into probabilities and transformed into startle scores using the Bradley Terry Luce (BTL) model (Guilford, 1954). To aid in analysis, the scores were normalized so that the score of signal 1 was zero. The variance of each score was estimated using the following resampling procedure. First, a series of probability matrices were created by assuming that each element of the original probability matrix varied according to a binomial distribution with its value as the mean. A series of probability matrices were sampled from this distribution. A distribution of BTL scores were then generated from these permuted probability matrices, and the standard deviation of the scores was determined taking the standard deviation of that distribution. To examine the effect of individual subjects on the BTL startle scores, BTL scores were calculated for the case when each individual subject was removed from analysis. The results of this analysis are plotted in Figure 4.10.

The removal of an individual subject's data from the groups resulted in small changes to the BTL scores. The largest variations observed were due to the removal of subject 9 (diamonds in Figure 4.10), but these were relatively small. The relative distance in ratings between signals 1 and 4 did not have much subject-to-subject variability. The relative difference in ratings (compared to signal 1) for the signals with the highest and lowest loudness derivatives of the first loudness peak (dL_{max}), signals 2, 3 and 5, had more subject-to-subject variation.

In Figure 4.11, BTL scores estimated from the first test results on each day were plotted alongside BTL scores estimated from the 2nd and 3rd tests and BTL scores from the entire dataset. The most variation appears in the scores of Signal 5 on the first day and in the differences between BTL values for signals 4 and 5 and the BTL values for signals 1, 2 and 3 on Day 2. The range of BTL values is slightly bigger on Day 2 than on Day 1. Overall, there does not appear to be a large amount of variation in the BTL values based on subject or session differences.

The BTL scores are plotted against maximum loudness (N_{max}) and maximum loudness derivative of the first peak (dL_{max}) in Figure 4.12. The BTL scores for signals 1 and 3

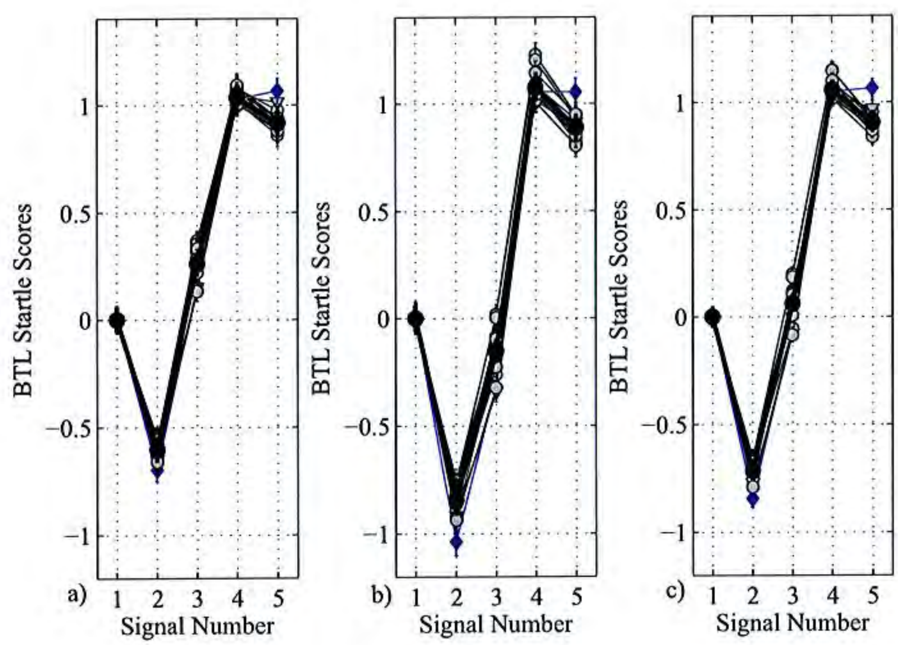

Figure 4.10. BTL startle scores for a) first day, b) second day and c) both days. Black are scores calculated using all subjects' data, gray are scores calculated for all but one subject's data. Dark gray diamond are scores from after removing subject 9.

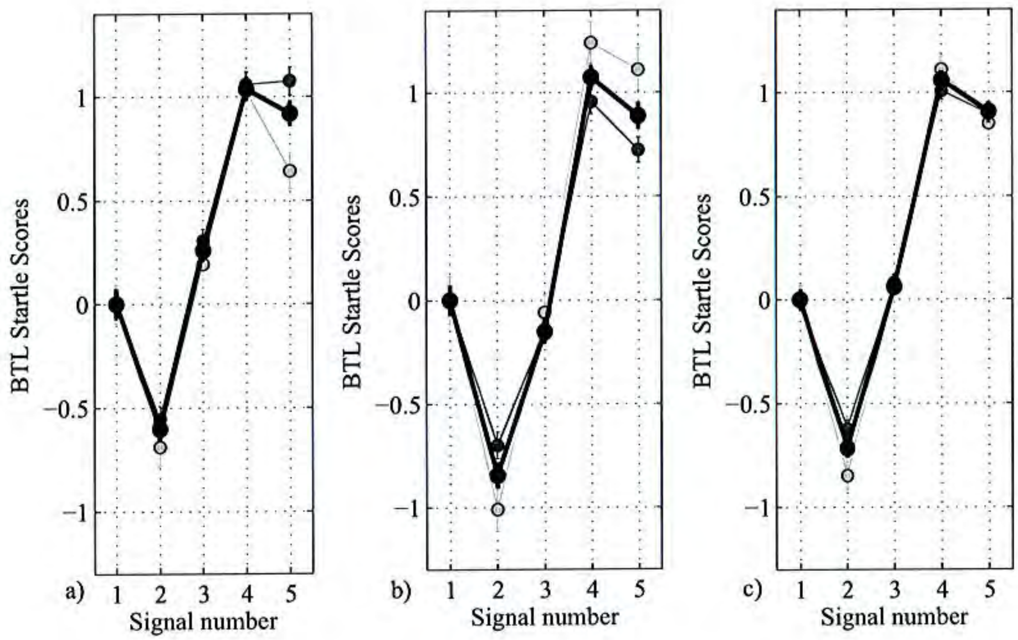

Figure 4.11. BTL startle scores calculated for 1st session (light gray), 2nd and 3rd session (dark gray) and all sounds (black) for a) first day, b) second day and c) both days.

appear to be nearly identical. From this result, it appears that a change of 10 sones in maximum loudness produces an effect equivalent to a change of 2000 sones/s of dL_{max}. Compared to the results in Chapter 2 (10 sones equivalent to 3320 sones/s of dL_{max}), dL_{max} has a stronger impact (relative to that of loudness) on startle scores in this experiment. However, this result may be due to the larger variation of dL_{max} in stimuli used in this experiment.

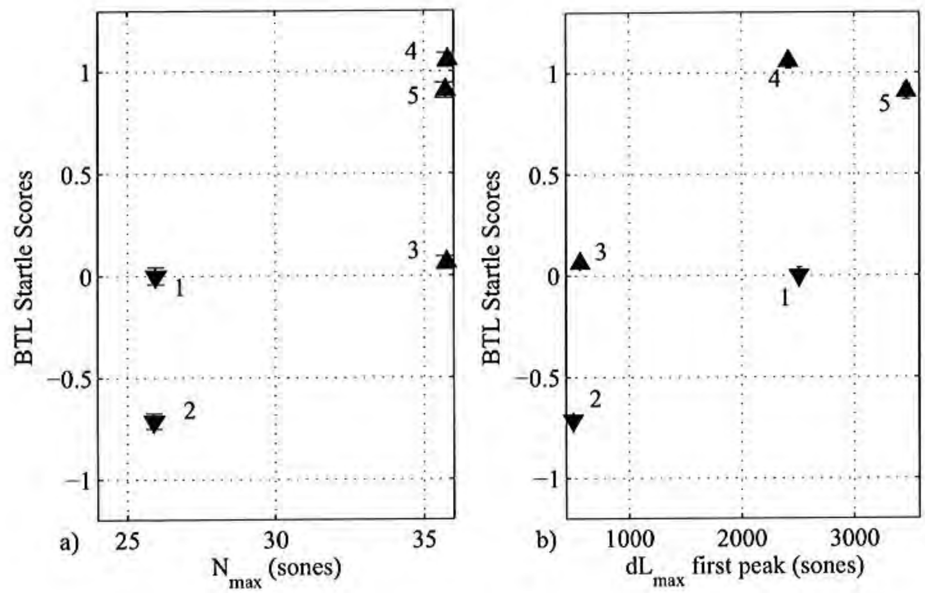

Figure 4.12. BTL startle scores versus a) N_{max} ($R^2 = 0.61$) and b) dL_{max} ($R^2 = 0.58$).

As there are only five signals in this experiment, there is not enough statistical power to estimate parameters for linear models of two or more metrics. However, it is possible to examine how previously estimated models of startle, developed in Chapter 2, predict the BTL startle scores. In addition to the model that includes dL_{max} as a factor, models incorporating maximum loudness with either the reciprocal of loudness rise time ($1/MRT$) or mean Zwicker Sharpness (S_{mean}) were considered. Loudness rise time and mean sharpness were calculated in the manner described in Chapter 2. For each linear model, the parameters were estimated from the average startle scores for the experiments described in Chapter

2. Scores ranged from -16 (Calming) to +16 Startling. The output of these models were compared with the BTL scores and a coefficient of determination (R^2) was calculated. The models that included $1/MRT$ or S_{mean} as a factor predicted BTL values as well as maximum loudness or dL_{max} alone (0.61 and 0.55, respectively). The poor performance of these models is likely caused by the limited amount of variation in MRT and S_{mean} in the stimuli used in the paired comparison tests. The model of maximum loudness and dL_{max} together performed better (R^2 = 0.88). BTL scores are plotted versus predictions of this model in Figure 4.13.

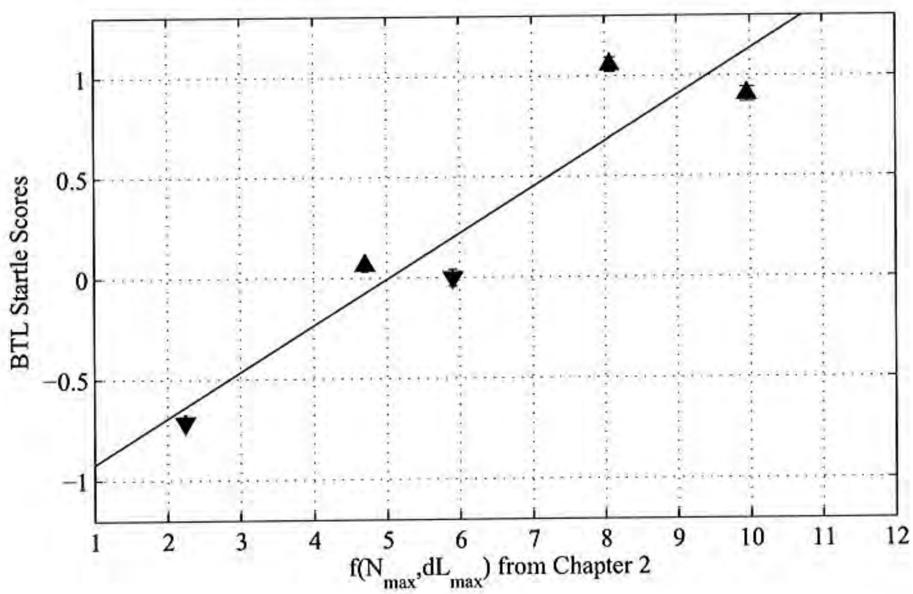

Figure 4.13. BTL startle scores versus estimates of the linear model of maximum loudness and maximum loudness derivative (R^2 = 0.88).

In addition to the statistics of hearing models, the output of Schmajuk and Larrauri's prepulse inhibition model (2006) was compared with the BTL startle scores. As mentioned in Chapter 1, the input of this model is the fast-averaged sound pressure level of white noise. To calculate the output of this model, windowed white noise signals were created to have loudness time-histories within 2% of that associated with the boom stimuli used in this experiment. The fast-averaged sound pressure level was calculated using Artemis by Head Acoustics for these signals. Then the sound pressure level time-history of these sounds was used as a surrogate input to Schmajuk's pre-pulse model. The maximum value of the output of the pre-pulse model was used as a metric.

Schmajuk's model performed about as well as the other linear models and dL_{max} alone (R^2 = 0.60). The maximum of Schmajuk's model is also strongly correlated to dLmax (R = 0.97). This result is not surprising given the similarities between these two metrics that was discussed in Section 1.5.2.1.

4.2.5 Comparison of Physiological Measures and Subjective Ratings

While there was a large amount of variability in the physiological responses evoked by the stimuli, it is possible to compare the two kinds of measures. To compare the measures, the probability of each response occurring (including the probability of two or more responses occurring for the same stimuli) were compared to the BTL scores for each day and for both days together and a coefficient of determination (R^2) was calculated. The results of this analysis are listed in Table 4.1 for individual responses and Table 4.2 for co-occurring responses. However, because of both the limited number of stimuli and because the stimulus order in this experiment was not randomized, the generalizability of these results is limited.

The p-values of this analysis are large due to the limited number of signals. However, several measures were highly correlated to the BTL startle scores for each day's data and both days' data combined ($\alpha = 0.05$). Among these, both the trapezius and the short latency heart-rate responses were negatively correlated with BTL scores: startle scores decreased as the number of responses increased. This was also the case when comparing BTL startle scores with the probabilities of 2 or more muscle responses occurring with the same signal.

The probability of having a skin conductance response was highly correlated (positively) with BTL startle scores. In addition, the presence of skin conductance responses combined with any muscle response or, to a lesser extent, any heart-rate response was correlated to BTL scores as well. However, the presence of all three measures occurring for the same exposure was only nearly significant ($p = 0.12$) in the case with a skin conductance response, any heart-rate responses and any muscle response. Although the coefficients of determination in this experiment are less than that between physiological measures and startle ratings found in the experiments described in Chapter 3, the results are similar. The probability of a skin conductance response occurring appears to be the most stable physiological measure that correlates well with subject ratings of startle in both experiments.

Table 4.1 Coefficient of determination (R^2) between probability of physiological responses and BTL startle scores. p-values are displayed in parenthesis. *'s denote negative correlations. HRA - heartrate increase, HRD - heartrate decrease, SCM - sternocliedomastiod, AC - anterior scalene, TRP - trapezius and SC - skin conductance.

Measure	Day 1	Day 2	Both Days
Short HRA	*0.03 (0.77)	*0.76 (0.06)	*0.42 (0.23)
HRD, then HRA	0.18 (0.48)	0.45 (0.22)	0.18 (0.48)
Medium HRA	*0.44 (0.23)	0.14 (0.53)	*0.07 (0.67)
Long HRA	0.55 (0.15)	*0.34 (0.31)	0.00 (0.92)
HRD	0.62 (0.11)	*0.18 (0.47)	0.00 (0.91)
Any HR	0.05 (0.71)	*0.44 (0.22)	*0.01 (0.87)
SCM	*0.21 (0.44)	0.00 (0.94)	*0.47 (0.20)
AC	*0.05 (0.71)	*0.51 (0.18)	*0.40 (0.26)
TRP	0.00 (1.00)	*0.71 (0.07)	*0.95 (0.00)
SC	**0.98** (0.00)	0.75 (0.06)	**0.88** (0.02)
SCM % AC	*0.41 (0.25)	*0.27 (0.37)	*0.37 (0.28)
SCM & TRP	*0.32 (0.32)	*0.88 (0.02)	*0.72 (0.07)
AC & TRP	*0.16 (0.50)	*0.83 (0.03)	*0.68 (0.09)
All EMG	*0.38 (0.27)	*0.55 (0.15)	*0.49 (0.19)
Any EMG	0.01 (0.88)	*0.13 (0.55)	*0.56 (0.15)

4.3 Summary and Conclusions

An experiment was conducted to examine the repeatability of physiological measures evoked by low level sonic booms. While viewing pictures of landscapes, subjects were exposed to sonic booms while their skin conductance, heart-rate and neck muscle electrical

Table 4.2 Coefficient of determination (R^2) between probability of co-occuring physiological responses and BTL startle scores. p-values are displayed in parenthesis. *'s denote negative correlations. HRA - heartrate increase, HRD - heartrate decrease, SCM - sternocliedomastiod, AC - anterior scalene, TRP - trapezius and SC - skin conductance.

SC & SCM	*0.06 (0.69)	**0.85** (0.03)	0.31 (0.33)
SC & AC	0.00 (0.96)	*0.02 (0.82)	*0.04 (0.74)
SC & TRP	0.47 (0.20)	0.14 (0.54)	0.61 (0.12)
SC & Any EMG	**0.92** (0.01)	**0.76** (0.05)	**0.94** (0.01)
SC & Short HRA	0.13 (0.55)	*0.06 (0.69)	0.06 (0.70)
SC & Short & Medium HRA	0.07 (0.67)	0.60 (0.12)	0.48 (0.20)
SC & Any HRA	0.27 (0.37)	0.48 (0.20)	0.56 (0.15)
SC & Any HR	0.46 (0.21)	0.89 (0.02)	0.60 (0.13)
SC & All EMG & Short HRA	0.00 (0.93)	*0.53 (0.16)	*0.27 (0.37)
SC & All EMG & Short, Medium HRA	0.14 (0.53)	*0.74 (0.06)	*0.22 (0.42)
SC & Any EMG & Short, Medium HRA	0.01 (0.85)	0.24 (0.40)	0.06 (0.68)
SC & Any EMG & Any HRA	0.33 (0.31)	*0.10 (0.61)	0.19 (0.46)
SC & Any EMG & Any HR	0.61 (0.12)	0.52 (0.17)	0.61 (0.12)

activity was measured. After completing the picture viewing task, participants completed a paired comparison test. Then, the next day, they completed all the tasks again. In addition, for an additional five subjects, the picture viewing task was completed using a single high-level stimulus. It was found that there was a large amount of subject-to-subject and day-to-day variation in the physiological measures. In contrast, BTL scores from the paired comparison test had relatively little subject-to-subject or day-to-day variation. BTL scores were also found to be highly correlated with the probability of a skin conductance response occurring after a sound.

It appears that these low level sonic booms are at levels below that associated with consistent startle responses for the population studied. While some noticeable startle responses were observed to be evoked by the stimuli (skin conductance plus heart-rate increase and muscle activity), these were relatively uncommon. Thus, it appears that most of the responses are orientation responses. Subjects' judgments of startle were found to have less variance among participants over each day and to be correlated with physiological responses. Subjective measures appear to be ideal for the creation of the startle model. Development and verification of this startle model is discussed in the next two Chapters.

5. DEVELOPMENT OF A STARTLE MODEL

In the last chapter, people's evaluations of low boom evoked startle were compared to physiological responses. It was found that results from a paired comparison study were correlated to subjects' physiological responses. In addition, there was less subject-to-subject variability in the psychophysical ratings than there were in the physiological responses. Thus, it appears that the use of subjects' judgments is a valid method of collecting data to build a startle model.

Linear models of metrics used to predict startle and annoyance have been developed in previous studies described in Chapter 2. The metrics in these models that predict subject responses best were statistics of Moore and Glasberg's (2002) time-varying loudness and the time-derivative of loudness, and statistics of Zwicker sharpness (Zwicker and Fastl, 1999). The best of these models was based on maximum loudness (N_{max}) and either maximum loudness derivative (dL_{max}) or mean sharpness (S_{mean}). However, these three parameters were partially correlated with one another making it difficult to examine the unique influence of each. In the pair comparison experiment described in Chapter 4, it was again found that maximum loudness derivative and maximum loudness both affected average startle ratings. Again, statistics of sharpness did not vary significantly in that study that involved very few signatures.

In order to develop a startle model, it was necessary to collect data that could be used to gain a deeper understanding of how maximum loudness (N_{max}), maximum rate of change of loudness (dL_{max}) and mean sharpness (S_{mean}) affected startle. Another experiment was conducted and the results are described here.

5.1 Methods

Institutional review board approval was obtained for this study (IRB#0904007961). A description of the test set-up, subjects recruited, stimuli and procedure are presented in the following sections. The sounds were presented to subjects using Etymotics ER-2 insert earphones while the subject sat in an IAC double-walled sound booth. The earphones were connected to a Tucker-Davis HB-7 amplifier that was driven by a LynxOne sound card. All sounds that were presented to subjects were filtered to compensate for the operating condition of the earphone. The subject completed the experiment using a keyboard, mouse and LCD monitor placed in the booth. The computer tower was placed outside of the booth.

All signals were played over low-pass filtered white noise (1st order Butterworth, 200 Hz cut-off) with an unweighted level of around 40 dB. The noise was presented continuously throughout the test and was included to minimize anticipation effects.

5.1.1 Signals

Twenty-four synthetic sonic booms, 94 boom-like noises and three low boom signal provided by industrial partners were used as stimuli in this study. The characteristics of the synthetic low boom signals and boom-like noises chosen to span the range of metric values that were used in previous tests (Chapters 2-4). The stimuli were designed so that the three sound metric parameters (N_{max}, dL_{max} and S_{mean}) were weakly correlated over the test signal set.

The synthetic booms were generated using the frequency domain sampling technique used in finite impulse response (FIR) filter design. All signals were high-pass filtered (f_c = 25 Hz, 3rd order Butterworth) to extend the range of sounds that could be played. Loudness derivative through time was calculated by using a 51-point FIR filter that behaves as a differentiator from 0-350 Hz and as a low-pass filter at higher frequencies. The filter was convolved with the short-term loudness time-history from Moore and Glasberg's time-varying loudness. dL_{max} was defined as the maximum value of the loudness derivative time-history between the onset of the signal and its first loudness maximum. S_{mean} was calculated over the duration of the main boom event. This duration (Dur) was defined as the duration between when the loudness time-history leaves the noise floor and when it returns for the last time, so post-boom noise is included. The methods utilized for calculating metrics were the same as described in Chapter 2. The synthetic sonic booms were designed to so that N_{max} and dL_{max} were weakly correlated. Synthetic boom time-histories included classical N-waves as well as shaped sonic booms. N_{max} of the stimuli is plotted against dL_{max} in Figure 5.1a).

To increase the range of sharpness in the signal set, five booms-like noise signals were generated for each boom signal. The boom-like noise signals were composed of windowed white or colored noise that was designed to have nearly identical loudness time-histories as some of the boom signals. A full description of the algorithm used to produce these signals is in the Appendix. Recall that there are three outputs of Moore and Glasberg's time-varying loudness: instantaneous, short-term and long-term loudness. These outputs correspond loudness that is not conscientiously perceived, loudness that is perceived through time, and the memory of loudness through time, respectively. N_{max} is calculated from the maximum of the short-term loudness time-history. LN_{max} is calculated from the maximum of the long-term loudness time-history. The largest difference in instantaneous loudness between any boom and its corresponding boom-like noise signals was 3% (i.e. 10.3 sones for a target of 10 sones). Most of these signals had a maximum loudness difference of around 0.05%. The largest differences in loudness time-histories were observed for the signals composed with pink-noise and were created for the signals at the lowest maximum loudness level and the largest dL_{max} values. Because of this, only three boom-like noise signals were generated for this boom. This procedure resulted in signals with similar N_{max} and dL_{max} characteristics as the synthetic sonic booms, but with different mean sharpness. The result of this design was that N_{max}, dL_{max} and sharpness metrics for the test stimulus were weakly correlated ($R < 0.08$ for S_{mean} and other metrics; $R < 0.20$ for S_{max} and other metrics). Mean sharpness is plotted against N_{max} and dL_{max} in Figures 5.1b) and c). A table with correlation coefficients (R) between all calculated metrics is in Table 5.1.

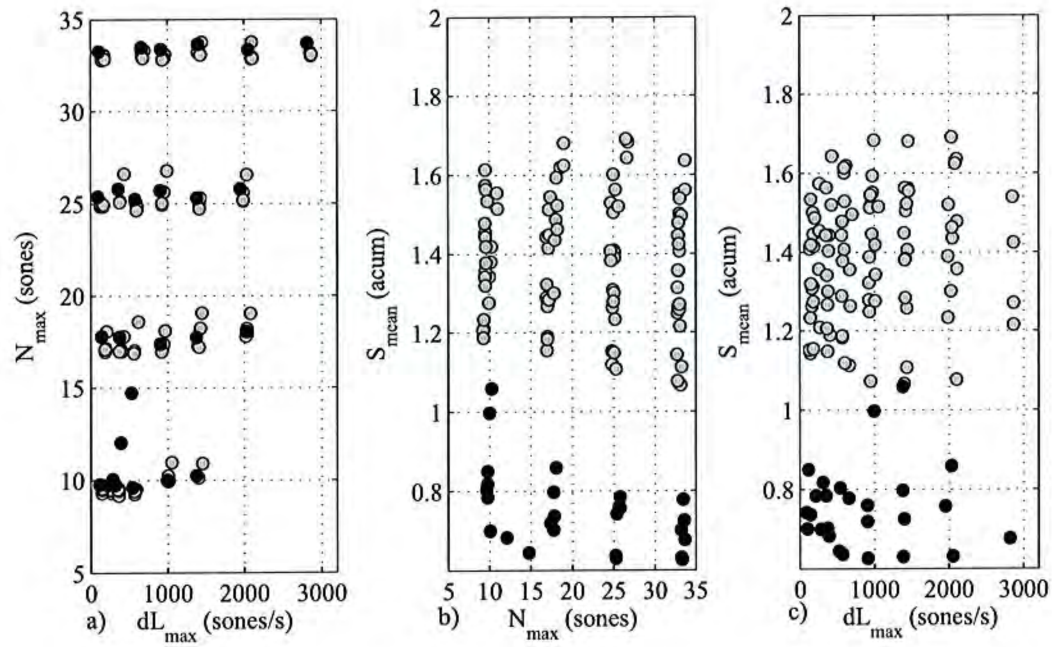

Figure 5.1. a) N_{max} versus dL_{max}, b) S_{mean} versus N_{max} and c) S_{mean} versus dL_{max} for test stimuli in the presence of background noise. Synthesized and industry simulated boom signals are dark gray. Boom-like noise signals are light gray.

Table 5.1 Correlation coefficient (R) between metrics of stimulus set. LN_{max} is maximum Moore's long-term loudness, Dur is duration.

	N_{max}	dL_{max}	S_{mean}	S_{max}	LN_{max}	Dur
N_{max}	1	0.38	-0.07	0.16	0.96	0.67
dL_{max}	0.38	1	0.07	0.20	0.47	0.61
S_{mean}	-0.07	0.07	1	0.89	-0.16	-0.16
S_{max}	0.17	0.20	0.89	1	0.09	0.07
LN_{max}	0.96	0.47	-0.16	0.09	1	0.78
Dur	0.67	0.61	-0.16	0.07	0.78	1

The signals were divided into five blocks. In the first four blocks, signals in each were all of approximately the same maximum loudness: Block 1 $N_{max} \approx 32$ sones, Block 2 $N_{max} \approx 24$ sones, Block 3 $N_{max} \approx 16$ sones and Block 4 $N_{max} \approx 8$ sones. The fifth block of signals were repeats of 23 of the 24 boom signals from the previous block plus three simulated boom waveforms provided by industry partners. The types of signals composing each block is in Table 5.2. Signals were separated into the first four blocks based on N_{max} to help subjects focus on characteristics in the sounds other than loudness within each block. The presentation order of the first four blocks was randomized with a different block order for each subject. All subjects heard the boom-only block last. The presentation order of individual signals in each block was also randomized with a different order for each subject.

Table 5.2 Description of block content by signal type.

	Number of			
Block	Boom signals	Boom-like noises	Industry simulated booms	Total
1	6	24	0	30
2	6	24	0	30
3	6	24	0	30
4	6	24	0	28
5	23	0	3	26
Total	24 + 23 repeats	94	3	144

5.1.2 Procedure

The experiment was explained to the participant and informed consent was obtained. Next, the participant sat in the sound booth and had her/his hearing tested. After the hearing test, the participant was read the test instructions. The participant was asked, upon hearing a sound, to rate on a scale how startling she/he would find it if she/he was "outdoors in a park or garden and heard this sound intermittently throughout the day". The startle scale was displayed on the computer screen and was marked by using a mouse. The scale was divided into quarters and labels of "Not likely to be startling", "Somewhat startling", "Moderately startling", "Very startling" and "Extremely startling" were used. After being read the instructions, the participant completed a familiarization session where she/he heard a series of twenty sounds that were representative of the stimuli used in the test. These sounds were selected to encompass the full range of metrics used in the entire test stimulus. Then the participant heard these sounds again and practiced rating these sounds on the scale. Finally, the participant rated sounds in all five of the stimulus blocks. Between each block, there was a two-minute break where the participant was asked to write down her/his comments about the test. After completing all of the blocks, the participant's hearing was checked again and the participant was awarded $10 for her/his time.

5.1.3 Subjects

Forty-three people volunteered to be participants in this study. Participants' ages ranged from 18-46 years and they were recruited from the area surrounding Purdue University. Twenty-one subjects were male and twenty-two were female. All subjects had less than 20 dB of hearing loss at frequencies from 125-8000 Hz. Many of the participants were students and faculty at Purdue University. However, none of the participants had a background in environmental noise or sound quality.

5.2 Results

Numbers (0-100) were assigned to the scale positions so that a score of 5 corresponds to "Not likely to be startling" and a score of 95 corresponds to "Extremely startling". A correlation coefficient (R) was calculated between each participant's ratings and the mean of the rest of the group. Three out of forty-three participants have correlation coefficients of less than 0.3. One of these (number 29) has a coefficient of -0.01. If this analysis is repeated without participant 29, then the lowest correlation coefficient is 0.29. All responses aside from those associated with participant 29 were retained for the following analysis. The distribution of the responses appears to be uni-model; hence, the analysis of mean values for each signal should be appropriate. Analysis of the median does not greatly affect the subsequent analysis. Average startle ratings ranged from "Not likely to be startling" (5) to "Very startling" (77.5). The standard deviation of the estimated mean ranges from 2.3%-4.2% of the total scale length.

To check for ordering effects, the mean startle rating was calculated for each signal in the order each was presented (e.g. first signal heard, second signal heard etc.). The results are plotted in Figure 5.2 along with the average N_{max} presented. The ratings of the signals between the dash and solid vertical lines have higher average ratings and higher N_{max} levels than the rest of the signals. Recall that the signals in the lowest loudness block (N_{max} = 8 sones) had two fewer signals than the other non-boom only blocks. Because of this, due to the random presentation order of the first four blocks, signals between the dashed and solid lines have higher loudness levels than the other signals. Aside from those signals, there does not appear to be any large ordering effects based on presentation order alone.

In Figure 5.3a), the mean ratings of the boom signals played once in Blocks 1-4 and the ratings of the same signals played in Block 5 are plotted. For Blocks 2-4, the mean ratings of each signal was not significantly different compared to the mean of the same signal in Block 5; mean ratings were within one standard deviation of the estimated mean. However, for the signals with the maximum loudness (N_{max}) near 32 sones (signals 1-6 in Block 1), the signals were judged to be significantly more startling when presented with other boom signals than when these same signals were presented with boom-like noises of the same

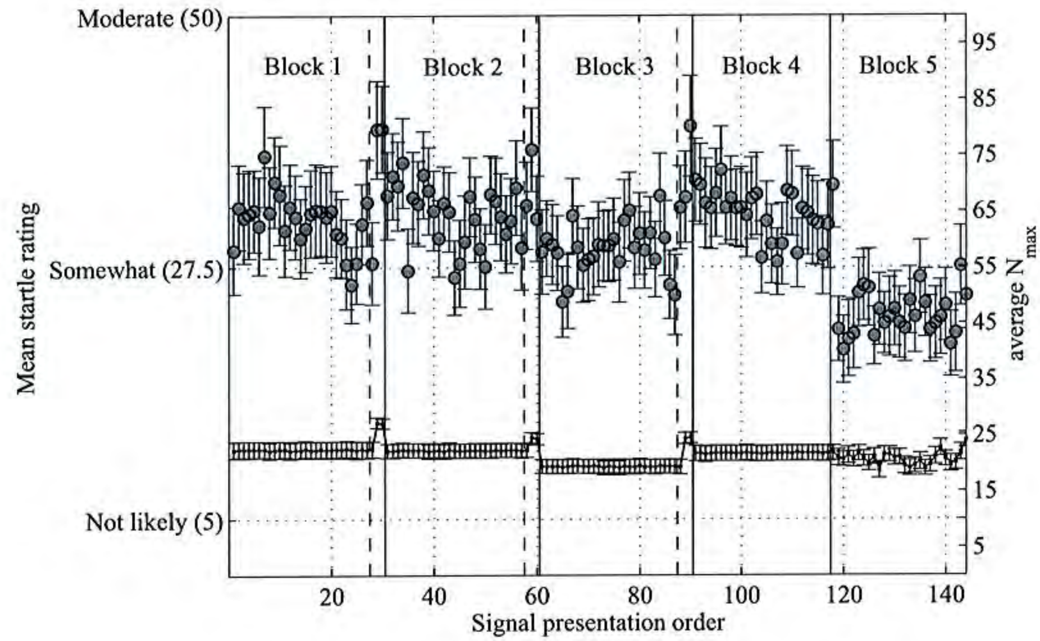

Figure 5.2. a) Average startle ratings in signal presentation order (Gray circles). Black dots are average N_{max} for each signal presentation. Errorbars are standard deviation of the estimated mean. Solid vertical lines correspond to boundaries of blocks with 30 signals. Dashed vertical lines correspond to boundaries of blocks with 28 signals.

maximum loudness in Block 1. Of interest, the difference in mean ratings of each of the boom signals in Block 1 and the same signals in Block 5 are similar. It is possible that judgments of startle of both booms and boom-like noises was difficult in Block 1; boom-like noise signals in this block had the highest average startle ratings and it is possible that participants rated the booms relative to the boom-like noise rather than giving an absolute rating of startle. The mean startle ratings across all signals in each block for booms and boom-like noise as well as the booms presented in Block 5 are plotted versus maximum loudness in Figure 5.3b).

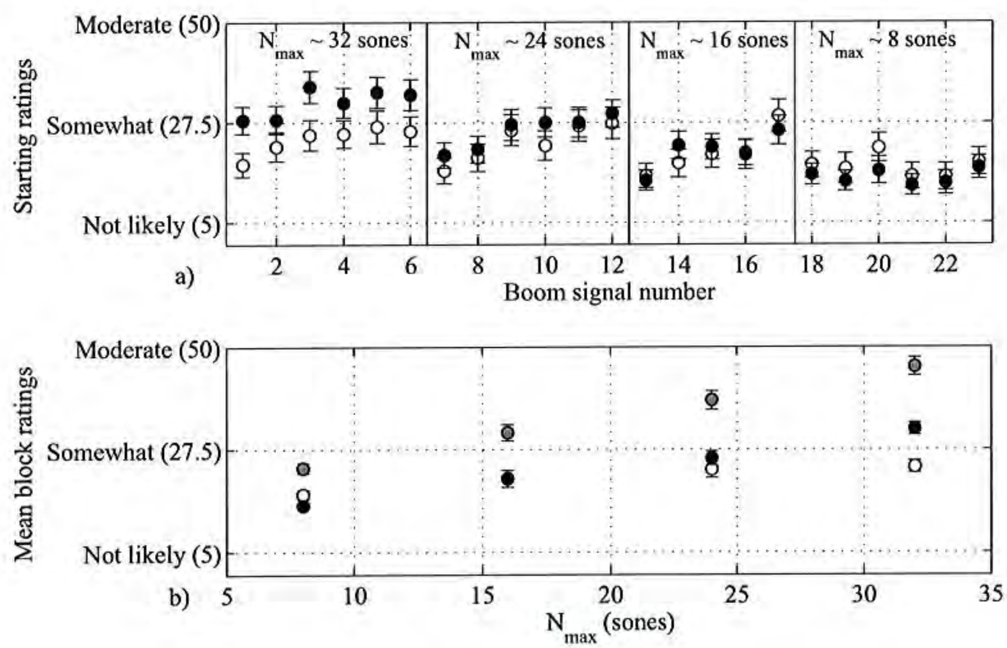

Figure 5.3. a) Average startle ratings of booms presented in blocks 1-4 and booms presented in block 5 . b) Block averaged startle ratings for booms in 1-4, boom-like noises and booms in block 5 with equivalent N_{max}. Booms in blocks 1-4 are white, Booms in block 5 are black and boom-like noises are gray. Error bars are standard deviation of the estimated mean.

The ratings of the boom-like noise signals and the booms in Block 5 both increase linearly with increasing maximum loudness values. The booms presented in Blocks 2-4 received nearly identical average ratings as those booms presented in Block 5. However, from Figure 5.3b), it is clear that the average ratings of the booms presented in Block 1-4 appear to plateau with increasing maximum loudness. The signals in Block 1 have almost the same average ratings as the boom signals in Block 2. This does not occur with the ratings of the same signals in Block 5 (white circles in Figure 5.3c)). It appears that ratings of the boom signals in Block 1 were lower than would be expected from trends in the ratings of the other boom signals in that figure. This appears to be due to over-compensation when comparisons were made with the non-boom signals presented in that block. While it is difficult to determine what exactly is happening, only the ratings of the boom signals presented in Block 5 were retained for the regression analysis described below because it was felt that the increase in startle with increasing N_{max} was more consistent with what had been found in previous research.

5.2.1 Metrics Analysis: Single Metric Models

Signal metric values were compared with average ratings for each sound, and coefficients of determination (R^2) were calculated across the entire signal set, as well as for the boom signals only. Mean scores are plotted against the three metrics used in the stimulus design as well as maximum sharpness (S_{max}), maximum Moore and Glasberg's long term loudness (LN_{max}) and event duration (Dur) in Figure 5.4. The low R^2 values are expected because of the variation in the other sound metrics at each level of the metrics on the abscissa. In this analysis, we only expect to see general trends. Aside from S_{mean}, there were general trends of increasing metric values with increasing startle scores. Startle ratings appear to increase as S_{mean} increases or decreases away from around 1.0 acum. The U-shaped trend with spectral balance changes has been observed by other researchers (e.g. Zwicker and Fastl, 1999). The event duration, appears to be strongly correlated with the boom signals but not for the boom-like noises. N_{max} and LN_{max} are similarly correlated to the startle ratings. This was expected because these metrics are strongly correlated to each other.

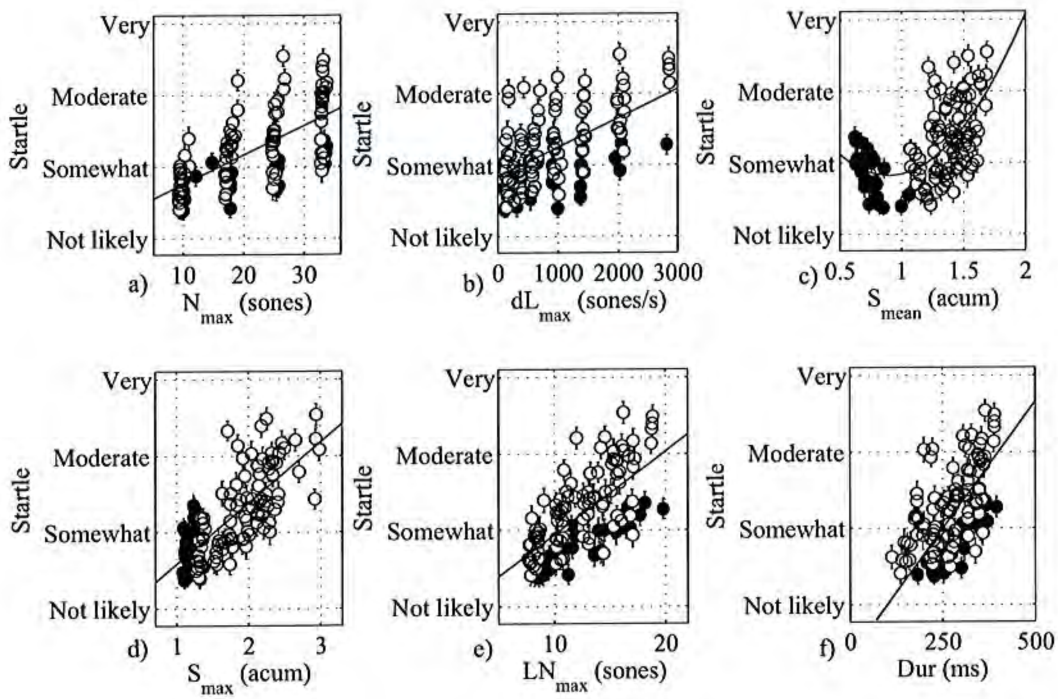

Figure 5.4. Average startle ratings versus a) N_{max} ($R^2 all = 0.45$, $R^2 booms = 0.69$), b) dL_{max} ($R^2 all = 0.31$, $R^2 booms = 0.31$), c) S_{mean} ($R^2 all = 0.34$, $R^2 booms = 0.41$), d) S_{max} ($R^2 all = 0.56$, $R^2 booms = 0.08$), e) LN_{max} ($R^2 all = 0.46$, $R^2 booms = 0.77$) and f) Dur ($R^2 all = 0.32$, $R^2 booms = 0.73$). Black circles are booms, white circles are boom-like noises. Error bars are standard deviation of the estimated mean.

5.2.2 Metrics analysis: models of two or more metrics

A series of candidate startle models composed of two or more metrics were estimated. These models consisted of a loudness metric (N_{max} or LN_{max}), dL_{max}, a sharpness metrics (S_{mean} or S_{max}) and/or Dur. In addition to these parameters, squared, cubed and interactions of these metrics were considered. The coefficients of the linear sum of these terms in each model were estimated using the data (metrics and startle ratings) across the entire signal set without the booms signals presented in Blocks 1-4. Coefficients of determination (R^2) were calculated for the mean ratings of all of these signals and for the mean ratings of the boom signals in Block 5 only. The significance of each model term was examined by using the partial F-test (α = 0.05). Models with terms that were not-significant were removed from the analysis. The models with the best performance are displayed in Table 5.3. For comparison, the results of all two term linear models of these metrics with significant terms (α = 0.05) are also listed. It should be noted that this model generation process is exploratory. The estimated parameters of these models need to be confirmed in another study. This confirmation is discussed in the next chapter.

Models that utilized LN_{max}, dL_{max} and S_{max} combined (models 14, 15, 16) were the best predictors of the examined signals. Including more than 2 estimated parameters in a model results in, at most, 3% more variance explained for all of the signals and, 7% more variance explained for the boom signals. Among the linear 2 metric models, three models, all containing S_{max} and either N_{max}, LN_{max} or Dur as a parameter (models 5, 7, 8, 9), predicted both types of signals well. The performance of the model including Dur and S_{max} is expected given that Dur is correlated to LN_{max} (R^2 = 0.61). Using LN_{max} instead of N_{max} resulted in slightly better predictions of startle ratings.

The slightly better performance of the models with LN_{max} may be due to the fact that LN_{max} is more correlated to dL_{max} than N_{max} (R = 0.47 versus 0.38). Models that included dL_{max} or Dur had the best predictions of the boom signals for the 2 metric models. Models that included statistics of sharpness were the among the best at predicting the entire signal set. Perhaps spectral balance is less important for the prediction of these booms than temporal aspect (rate of change of loudness or duration) of these sounds. However, if the

Table 5.3 Coefficient of determination for selected models. N is N_{max}, dL is dL_{max}, S is S_{max}, Dr is Dur and LN is LN_{max}.

#	Model	Notes	R^2 all	R^2 booms
1	$a_0 + a_1 dL + a_2 Dr$		0.39	0.62
2	$a_0 + a_1 N + a_2 Dur$		0.48	0.78
3	$a_0 + a_1 LN + a_2 dL$		0.53	0.78
4	$a_0 + a_1 N + a_2 dL$		0.56	0.73
5	$a_0 + a_1 dL + a_2 S$		0.72	0.27
6	$a_0 + a_1 N + a_2 S_{mean}$		0.75	0.55
7	$a_0 + a_1 S + a_2 Dr$		0.78	0.74
8	$a_0 + a_1 N + a_2 S$		0.84	0.70
9	$a_0 + a_1 LN + a_2 S$		0.88	0.78
10	$a_0 + a_1 N + a_2 dL + a_3 S$		0.89	0.70
11	$a_0 + a_1 NS + a_2 dLS^2$	Best 2 term model including N_{max}	0.91	0.73
12	$a_0 + a_1 S + a_2 NS + a_3 dL^2 S$	Best 3 term model including N_{max}	0.92	0.68
13	$a_0 + a_1 S + a_2 NdL + a_3 dL^2 + a_4 N^2 dL + a_5 N^2 S + a_6 dL^2 S + a_7 dL^3 + a_8 NdLS$	Best 8 term model including N_{max}	0.94	0.73
14	$a_0 + a_1 LNS + a_2 dL^2 S$	Best 2 term model including LN_{max}	0.92	0.74
15	$a_0 + a_1 LNS + a_2 dL^2 xS + a_3 LNSDr$	Best 3 term model including LN_{max}	0.93	0.69
16	$a_0 + a_1 S + a_2 Dr + a_3 LN^2 S + a_4 dL^2 LN + a_5 dL^2 S + a_6 LNSDr$	Best 6 term model including LN_{max}	0.95	0.81

boom-like noises are included, some measure of spectral balance (sharpness) is necessary for good predictive performance. Values of estimated parameters for the best 2 and 3 metric models are in Table 5.4.

Table 5.4 Estimated parameters for selected models. N is N_{max}, dL is dL_{max}, S is S_{max}, Dr is Dur and LN is LN_{max}.

Intercept	1st Term	Value	2nd Term	Value	3rd Term	Value	R^2 all	R^2 booms
6.7 ± 6.6	N	0.7 ± 0.2	Dr	0.04 ± 0.03			0.48	0.78
3 ± 5	LN	1.9 ± 0.5	dL	0.005 ± 0.002			0.53	0.78
11 ± 4	N	0.7 ± 0.2	dL	0.006 ± 0.002			0.56	0.73
-2 ± 4	dL	0.007 ± 0.002	S	16 ± 2			0.72	0.27
-20 ± 5	S	16 ± 2	Dr	0.08 ± 0.02			0.78	0.74
-11 ± 4	N	0.7 ± 0.1	S	15 ± 2			0.84	0.70
-22 ± 4	LN	2.1 ± 0.2	S	16 ± 2			0.88	0.78
13 ± 1	NS	0.40 ± 0.04	dLS^2	0.0011 ± 0.0002			0.91	0.73
7 ± 2	LNS	1.06 ± 0.08	dL^2S	$6\times10^{-7} \pm 2\times10^{-7}$			0.92	0.74
6 ± 2	S	5 ± 2	NS	0.36 ± 0.04	dL^2S	$9\times10^{-7} \pm 1\times10^{-7}$	0.92	0.68
5 ± 2	LNS	1.4 ± 0.3	dL^2S	$8\times10^{-7} \pm 2\times10^{-7}$	$LNSDr$	-0.0009 ± 0.0005	0.93	0.69

Estimated parameter values for N_{max}, dL_{max} and S_{max} were very similar for all of the models. For the models that contain N_{max} and dL_{max}, an increase in N_{max} of 1 sone was roughly equivalent to an increase of 100-400 sones/s in dL_{max}. This was similar to the ratio found in Chapters 2 and the paired comparison experiment in Chapter 4.

Of the models with two or three term, none of the 3 term model predict both the booms and the rest of the signals as well as the two term models. However, it is difficult to determine which of the 2-term models is best. LN_{max} and N_{max} are highly correlated to each other for this dataset (R = 0.97). Additional data is needed to determine which model is better able to predict startle.

5.3 Summary and Conclusions

An experiment was conducted to develop a model of subject-rated startle applicable to sonic booms and other impulsive sounds. Test stimuli were designed so that it would be possible to examined contributions of loudness, rate of change of loudness and spectral balance. Models that included maximum loudness (short-term N_{max} or long term LN_{max}), maximum loudness derivative of the first loudness peak (dL_{max}) and maximum sharpness (S_{max}) in some combination had significantly improved predictions of mean startle ratings for all sounds. Candidate startle models were estimated and the best performing model was $a_0 + a_1 LN_{max} S_{max} + a_2 dL^2_{max} S_{max}$. However, a similar model with N_{max} instead of LN_{max} also predicted startle ratings well.

When only booms are considered, the variation in S_{max} is small. In this case, the model is effectively: $a_0 + a_1 N_{max} + a_2 dL_{max}$. This is the form of the model found in Chapter 2 and Chapter 4. It is possible that including S_{max} increases the validity of model predictions for a wider range of stimuli. In the next chapter, verification of the candidate models will be described.

6. GENERALIZATION OF STARTLE MODEL WITH DIVERSE IMPULSIVE STIMULI

In the last chapter, several candidate startle models were developed. However, the regression analysis used to estimate these models was exploratory. It was unclear how well the models developed from that test would predict startle judgments for other stimuli because the stimulus set of the modeling experiment was, by necessity, limited to booms and boom-like noise. In addition, it is unclear if this model of startle could also predict annoyance judgments. In our previous studies, while ratings of both annoyance and startle were correlated, it was unclear where startle and annoyance judgments differ. To address these issues, another experiment was designed and conducted. The results are described below.

6.1 Methods

Institutional review board approval was obtained for this study (IRB #0904007961). The apparatus for playing sounds to subjects was identical to that of the startle modeling study that was described in Chapter 5. Sounds were presented to subjects with Etymotics ER-2 earphones that were driven by a Tucker-Davis HB7 amplifier and a LynxOne sound card. As with the Chapter 5 study, signals were presented binaurally and pre-filtered to account for the operating mode of the earphones. Sounds were presented over continuous low-pass filtered noise (1st order Butterworth, 200 Hz cut-off) at levels around 40 dB unweighted throughout the experiment.

6.1.1 Signals

A total of 63 signals were used in this experiment. Care was given to include a wide variety of signal types with diverse metric values. Of these signals, 15 were synthetic booms and 22 were boom-like noises from the experiment described in Chapter 5. Four additional signals were created by using finite impulse response filters (FIR) to simulate ground effects for four of these booms. Five signals were simulated sonic booms provided by industry, 7 were car door slams and 8 were distant gunfire or blast noise. In addition, 3 indoor boom recordings that contained rattle were also included. The signals were selected to span or be just outside the range of metrics used in previously described experiments studies. Two of the car door slam signals, 1 gunfire signal and 2 industry booms were part of the stimuli used in the semantic differential experiments described in Chapter 2. In Figure 6.1, the values of the 3 main metrics of the stimuli are displayed. As in the startle modeling experiment, many of the metrics were weakly correlated. A table of correlation coefficients (R) between all metrics examined is in Table 6.1.

Table 6.1 Correlation coefficient (R) between metrics of stimulus set. LN_{max} is maximum Moore's long-term loudness, Dur is duration, Z_{max} is maximum Zwicker's time-varying loudness and PL is Steven's Mk7 perceived level.

	N_{max}	dL_{max}	S_{max}	LN_{max}	Dur	Z_{max}	PL
N_{max}	1	0.56	-0.06	0.86	0.10	0.95	0.88
dL_{max}	0.56	1	0.09	0.43	-0.05	0.54	0.54
S_{max}	-0.06	0.09	1	-0.18	-0.20	-0.09	-0.22
LN_{max}	0.86	0.43	-0.18	1	0.32	0.82	0.74
Dur	0.10	-0.05	-0.20	0.32	1	-0.01	0.05
Z_{max}	0.95	0.54	-0.09	0.82	-0.01	1	0.94
PL	0.88	0.54	-0.22	0.74	0.05	0.94	1

6.1.2 Procedure

The experiment was explained to the participants, informed consent was obtained, and the participant's hearing tested. For the first part, the participants were asked, upon hearing a sound, to rate on a scale how startling they would find it if they were "outdoors in a park or garden and heard this sound intermittently throughout the day". The startle scale was displayed on the computer screen and was manipulated using a mouse. The scale

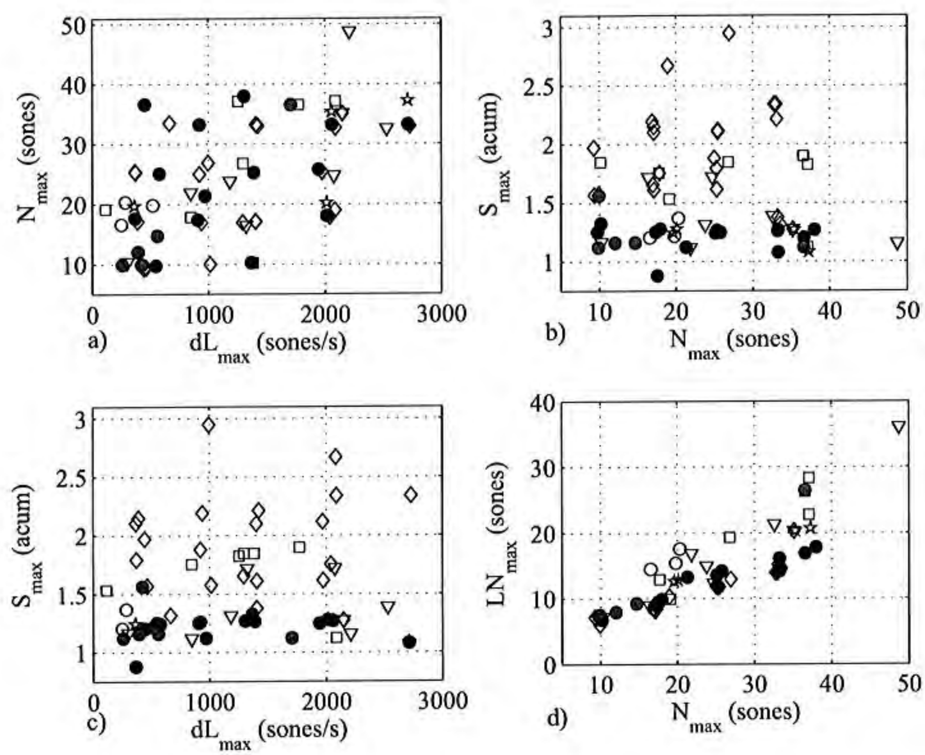

Figure 6.1. a) N_{max} versus dL_{max}, b) S_{max} versus N_{max}, c) S_{max} versus dL_{max} and d) LN_{max} versus N_{max}. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

was divided into quarters and labels of "Not likely to be startling", "Somewhat startling", "Moderately startling", "Very startling" and "Extremely startling" were used.

After the instructions, the participant completed the same familiarization and practice session that was used in the startle modeling experiment described in Chapter 5. Then the participant completed a block of 59 sounds, consisting of all signals except for the four booms with ground effects. Next, after a 2 minute break, the participants were asked to rate how the annoying each of the 63 sounds (59 sounds plus four booms with ground reflections) were on a similar scale under the same scenario. The scale in this part of the experiment ranged from "Not at all Annoying" to "Extremely Annoying". After another 2-minute break, the participant judged the acceptability of all 63 sounds. For this last part, participants could choose either "Acceptable" or "Not Acceptable" for each sound. After completing all of the sections, the participant's hearing was checked again and the participant was compensated ($10) for her/his time.

6.1.3 Participants

Thirty subjects with ages from 18-58 years volunteered to be participants in this experiment. 15 were male and 15 were female. All subjects had less than 20 dB of hearing loss at frequencies of 125-8000 Hz. Participants were recruited from the area surrounding Purdue University. Many of the subjects were students or faculty at Purdue, but none studied sound quality or environmental noise.

6.2 Results and Discussion

After collecting the data, the annoyance and startle scores were given values of 0-100, corresponding to 5, 27.5, 50, 77.5 and 95 for "Not likely" through "Extremely". A correlation coefficient (R) was calculated between each subject's data (startle ratings, annoyance ratings and acceptability ratings) and the mean of the rest of the group. The lowest value was 0.32 and all subject's responses were retained for analysis. Due to an error in the test software, some subjects selected neither "acceptable" or "not acceptable" for some signals; these responses were not considered in the subsequent analysis. The distributions of both the startle and annoyance scores were uni-modal without significant outliers, thus the responses were averaged across subjects for each scale. The standard deviation of the estimated mean ranged from 2-4.5% of the scale for both the annoyance and startle ratings. Ninety-five percent confidence intervals of the acceptability fractions ranged from 5.5-9.3% of the scale.

To examine if ordering effects are present in the data, responses were averaged in signal presentation order and plotted in Figure 6.2. For annoyance, average ratings of the first two presentations were lower than the rest of the signals. If these signals are removed before calculating the mean signals, two out of the 63 signals have there average annoyance ratings change by 2.2% of the scale and most do not change more than 1% of the scale. Removal of 1st two ratings has little effect on the subsequent analysis.

Average subject ratings and acceptability fractions are plotted against each other in Figure 6.3. All three sets of subject ratings were highly correlated to one another. This result is similar to that found in the semantic differential experiments described in Chapter 2. Average startle ratings of signals used in Chapter 5 were correlated to average startle ratings this this test ($R^2 = 0.74$).

6.2.1 Analysis of Single Metric Models

A coefficient of determination (R^2) was determined between single number metrics and either average startle ratings, average annoyance ratings or acceptability fractions. The three indoor boom signals were not included in this analysis. These signals were outliers for all of the models examined. A summary of the single metric analysis is displayed in Table 6.2.

Due to the weak correlation between many of the metrics, coefficients of determination are low for these single metric models, illustrating that more than one sound attribute is contributing to the response rating. Subject ratings are plotted against N_{max} (maximum of Moore and Glasberg's short-term loudness) and LN_{max} (maximum of Moore and Glas-

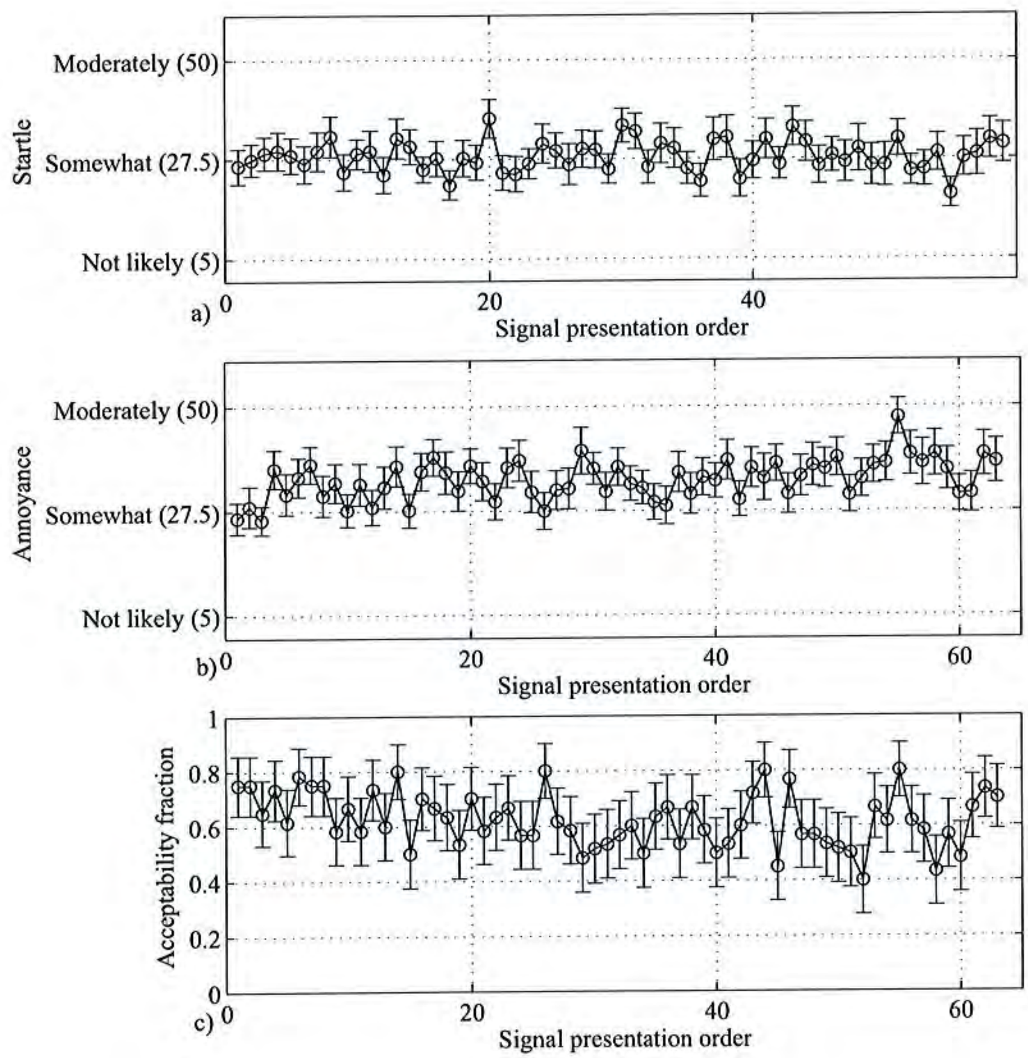

Figure 6.2. Signal presentation order-averaged a) startle ratings, b) annoyance ratings and c) acceptability fractions.

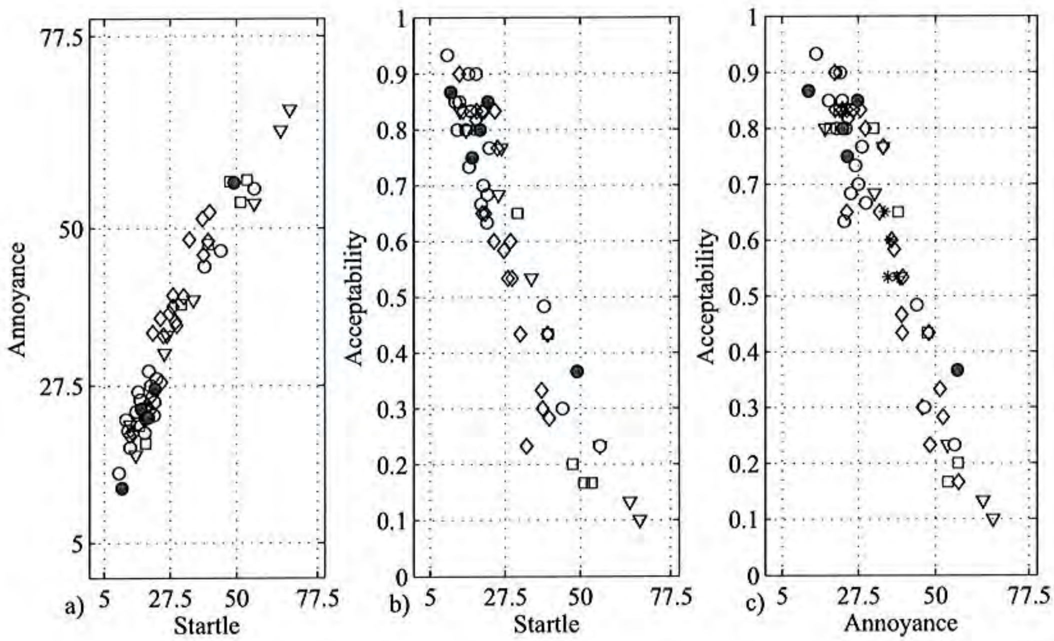

Figure 6.3. a) Average annoyance ratings versus average startle ratings (R^2 = 0.93). b) Acceptability fraction versus average startle ratings (R^2 = 0.87). c) Acceptability fraction versus average annoyance ratings (R^2 = 0.90). Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

Table 6.2 Coefficient of determination (R^2) between metrics of stimulus set and average subject responses. LN_{max} is maximum Moore's long-term loudness, D is duration, Z_{max} is maximum Zwicker's time-varying loudness and PL is Steven's Mk7 perceived level. Results from Chapter 5 experiment are in parenthesis.

Metric	R^2 Startle	R^2 Annoyance	R^2 Acceptability
N_{max}	0.42 (0.45)	0.45	0.43
dL_{max}	0.21 (0.31)	0.22	0.15
S_{max}	0.04 (0.56)	0.10	0.09
LN_{max}	0.62 (0.46)	0.58	0.54
Dur	0.45 (0.06)	0.31	0.30
Z_{max}	0.32	0.34	0.34
PL	0.20	0.21	0.22

berg's long-term loudness) in Figure 6.4. LN_{max} predicts average startle ratings better than N_{max}. N_{max} appears to predict many of the door slams and blast noise signals on a different trend-line than the sonic booms. LN_{max}, in contrast, predicts relatively higher loudnesses for the loudest of these signals, resulting in better prediction. However, neither was able to predict the ratings of the boom-like noises well.

In Figure 6.5, average subject ratings are plotted against the maximum of Zwicker and Fastl's time-varying loudness (Z_{max}) and Steven's Mk7 Perceived level (PL). Neither of these metrics predict average subject ratings well. Furthermore, like N_{max}, neither of these metrics could predict ratings of the loudest blast noises or the car door slams particularly well. For many signals, particularly the non-boom sounds, these metrics could not predict differences in ratings between sounds.

dL_{max}, S_{max} and Dur are plotted against subject ratings in Figure 6.6. To aid in interpretation, the startle ratings were fitted to N_{max} and the residuals were plotted against dL_{max}, S_{max} and D for groups of signals (Figure 6.7). The group ratings were determined by splitting the signals into five equal groups based on their metric values (dL_{max}, S_{max} or Dur). The ratings and metric values were averaged and plotted in that figure.

dL_{max} predicted subject ratings about as well in this experiment as it did in the previous test ($R^2 = 0.21$ for startle). dL_{max} appears to track ratings at 2000 sones/s and above.

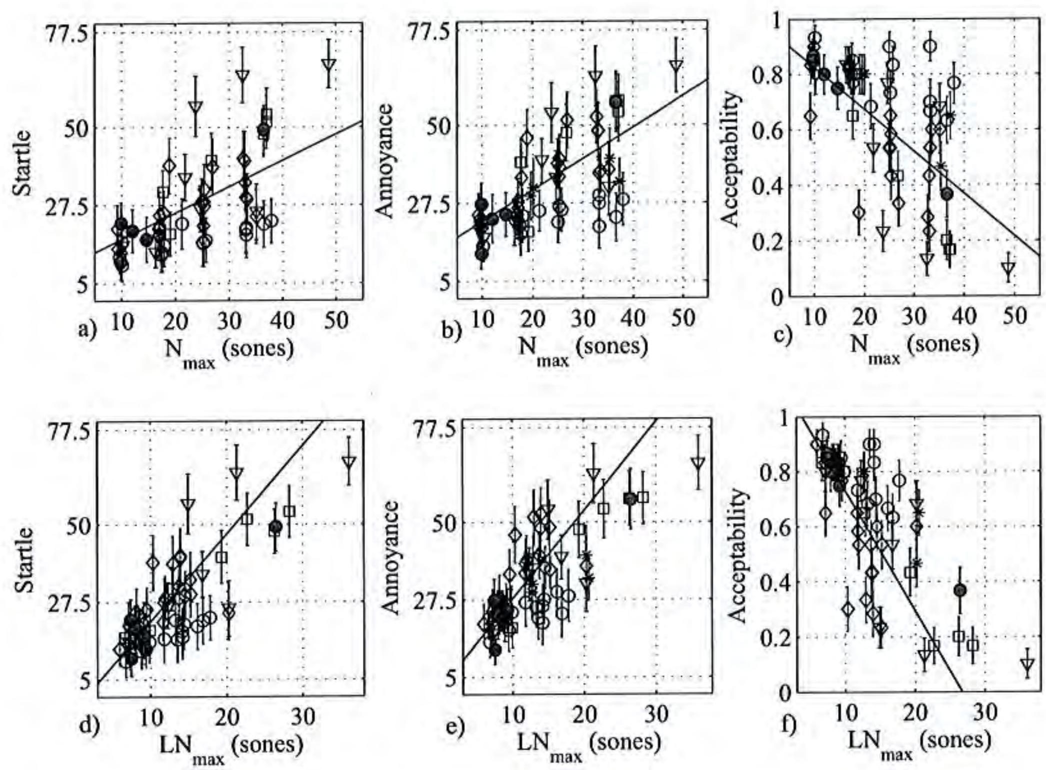

Figure 6.4. a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus maximum of short-term loudness (N_{max}). d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus maximum long-term loudness LN_{max}. R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

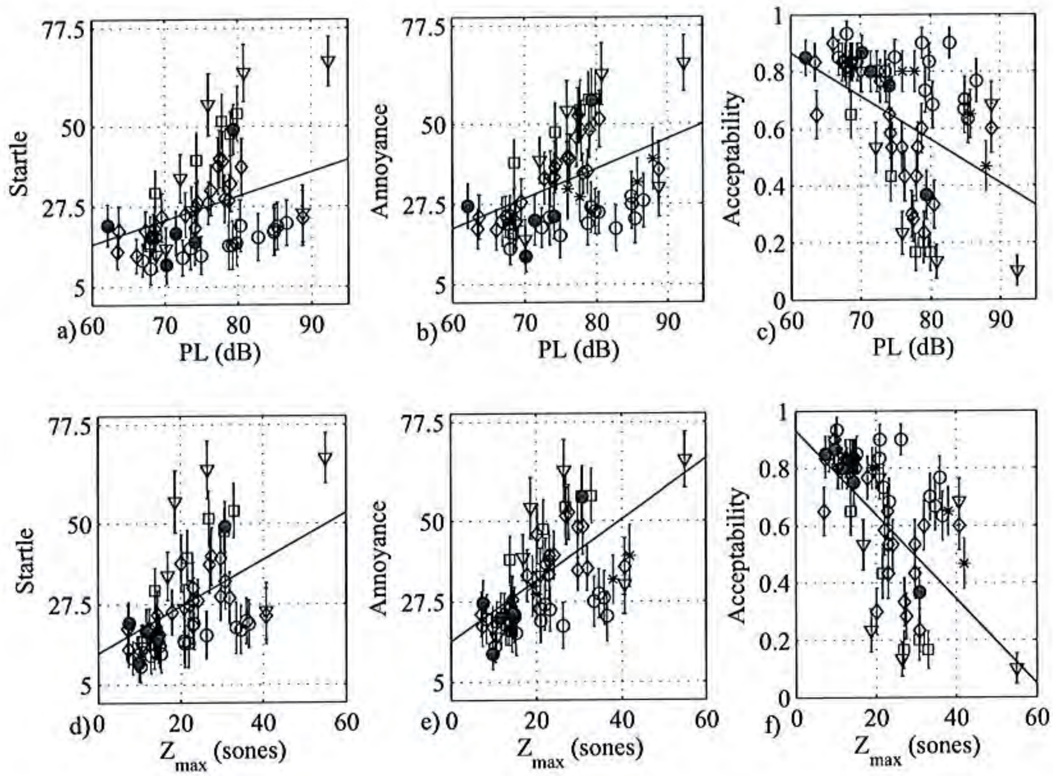

Figure 6.5. a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus perceived level (PL). d) Average startle ratings, e) Average annoyance ratings and f) Acceptability fractions versus maximum of Zwicker's time-varying loudness (Z_{max}). R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

For S_{max}, very little of the variance was explained when all signals were examined. This was because many more signals in this experiment have low sharpness value and high N_{max} values. The R^2 value, if only the signals used in the Chapter 5 test are included, is 0.57, similar to the R^2 value for that previous test results. Signals with S_{max} of greater than 2.1 acum are not very acceptable. This result is consistent with results found by Zwicker and Fastl (1999).

Startle and annoyance ratings increase with increasing duration when Dur is larger than 500 ms. Ramirez et al. (2005) found that signals with a duration of 500 ms or longer evoke a heart-rate response in subjects peaking at about 30 seconds after the stimulus onset. Long latency heart-rate responses, such as those, are associated with defense responses. Perhaps for stimuli with durations of 500 ms or longer, startle ratings reflect the combination of startle and defense responses.

6.2.2 Analysis of Candidate Models

Coefficients of determination (R^2) were calculated between the best performing models in Chapter 5 and the average startle ratings, average annoyance ratings and acceptability fractions for all of the signals and also for only the signals used as stimuli in the experiment described in Chapter 5. The results of all models examined are listen in Table 6.3.

Many of the models that performed well in the analysis presented in Chapter 5 (and for the signals common to both experiments) performed poorly when predicting subject responses for all signals in this experiment. In general, models that included LN_{max} performed better than ones that included N_{max}. This improvement in prediction is expected given the differences in performance when each of these metrics were individually compared with subjects' responses. The best 3-term model (#10) from the Chapter 5 experiment did not perform as well as the best two-term model. It is likely that this result was due to the inclusion of a Dur term because there was a larger range of Dur values in this test compared to the one described in Chapter 5 ($<$500 ms versus $<$2800 ms).

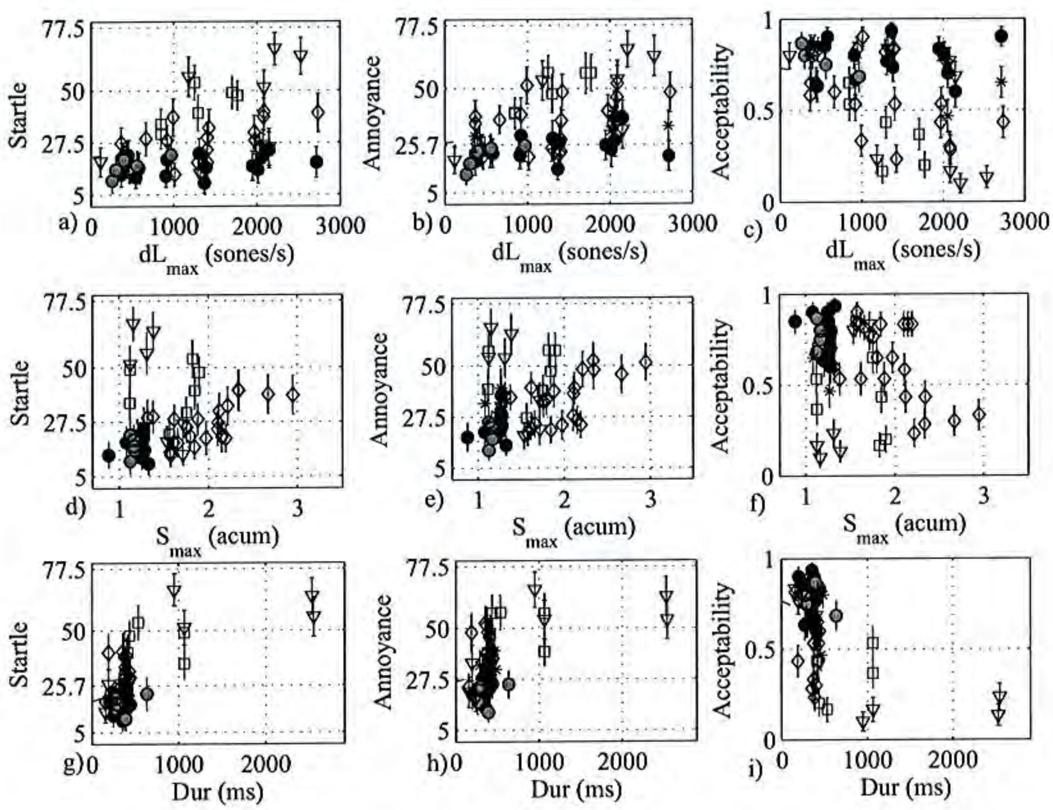

Figure 6.6. a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus dL_{max}. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus S_{max}. g) Average startle ratings, h) average annoyance ratings and i) acceptability fractions versus Dur. R^2 values are in Table 6.2. Black circles are synthetic booms, gray circles are simulated booms from industry, squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

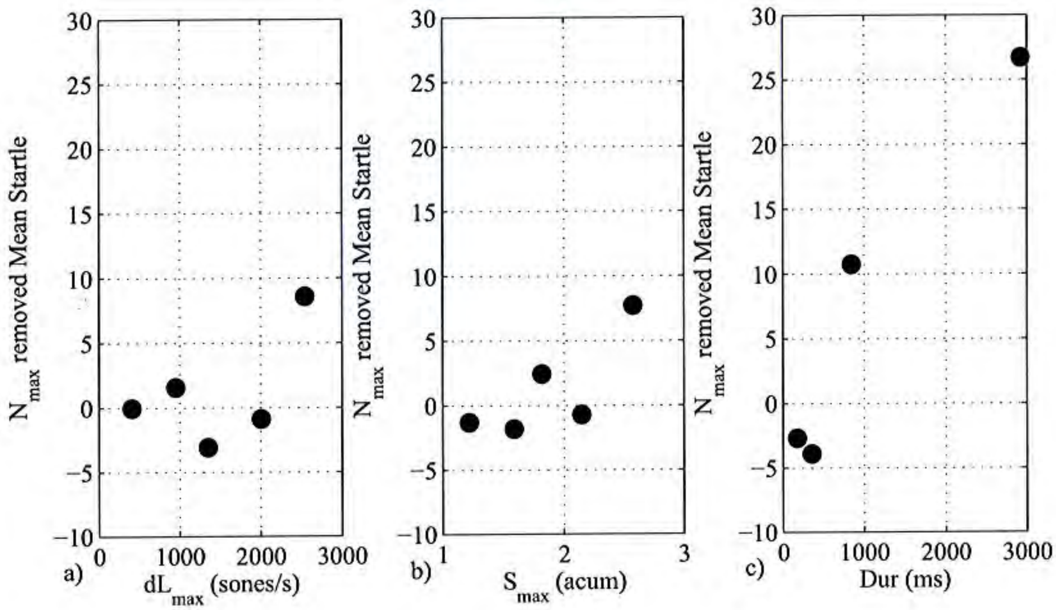

Figure 6.7. Group averaged startle residuals (startle ratings without N_{max}) versus a) dL_{max}, b) S_{max} and c) Dur.

Table 6.3 Coefficient of determination (R^2) between metrics of stimulus set and average subject responses. LN is LN_{max}, N is N_{max}, S is S_{max}, dL is dL_{max} and Dr is duration. M_8 is model 8 and M_6 is model 6. Coefficients of determination (R^2) for the set of signals also used in the test described in Chapter 5 are in parenthesis.

Model #	Startle Model	R^2 Ch5	R^2 Startle	R^2 Annoyance	R^2 Acceptability
1	$a_0 + a_1N + a_2dL$	0.56	0.43(0.38)	0.45(0.38)	0.40(0.29)
2	$a_0 + a_1LN + a_2dL$	0.53	0.63(0.36)	0.60(0.38)	0.52(0.30)
3	$a_0 + a_1S + a_2dL$	0.72	0.21(0.17)	0.22(0.14)	0.15(0.09)
4	$a_0 + a_1S + a_2Dr$	0.78	0.54(0.51)	0.41(0.58)	0.40(0.49)
5	$a_0 + a_1N + a_2S$	0.84	0.40(0.82)	0.54(0.82)	0.50(0.68)
6	$a_0 + a_1LN + a_2S$	0.88	0.71(0.82)	0.79(0.84)	0.72(0.72)
7	$a_0 + a_1NS + a_2S^2dL$	0.91	0.37(0.87)	0.52(0.86)	0.49(0.73)
8	$a_0 + a_1LNS + a_2dL^2S$	0.92	0.64(0.86)	0.75(0.85)	0.69(0.73)
9	$a_0 + a_1NS + a_2SdL^2 + a_3S$	0.92	0.36(0.85)	0.51(0.82)	0.46(0.68)
10	$a_0 + a_1LNS + a_2dL^2S + a_3LNSDr$	0.93	0.04(0.84)	0.11(0.81)	0.11(0.69)
11	$a_0 + a_1M_8 + a_2Dr$	0.91	0.90(0.85)	0.88	0.82
12	$a_0 + a_1M_6 + a_2Dr$	0.88	0.92(0.80)	0.87	0.81
13	$a_0 + a_1M_8 + a_2Nb$	0.68	0.71(0.78)	0.64	0.61
14	$a_0 + a_1M_6 + a_2Nb$	0.67	0.76(0.80)	0.70	0.66
15	$a_0 + a_1M_8 + a_2LN$	0.85	0.72(0.77)	0.76	0.70
16	$a_0 + a_1M_6 + a_2LN$	0.87	0.73(0.79)	0.77	0.71
17	$a_0 + a_1M_8 + a_2LNDr$	0.89	0.92(0.84)	0.87	0.80
18	$a_0 + a_1M_6 + a_2LNDr$	0.87	0.91(0.79)	0.84	0.78
19	$a_0 + a_1M_8 + a_2LNNb$	0.90	0.66(0.85)	0.73	0.67
20	$a_0 + a_1M_6 + a_2LNNb$	0.77	0.74(0.81)	0.72	0.67
21	$a_0 + a_1M_8 + a_2Th(LNDr)$	0.92	0.91(0.84)	0.87	0.80

The models that predicted both tests' stimuli well (Models 6 and 8) are plotted against startle ratings in Figure 6.8. Both models under-predict the ratings of the gunfire and blast noise rated between moderately and very startling. From Figure 6.6, it appears that these signals are among those with a much longer duration and are somewhat louder than the rest of the stimulus set.

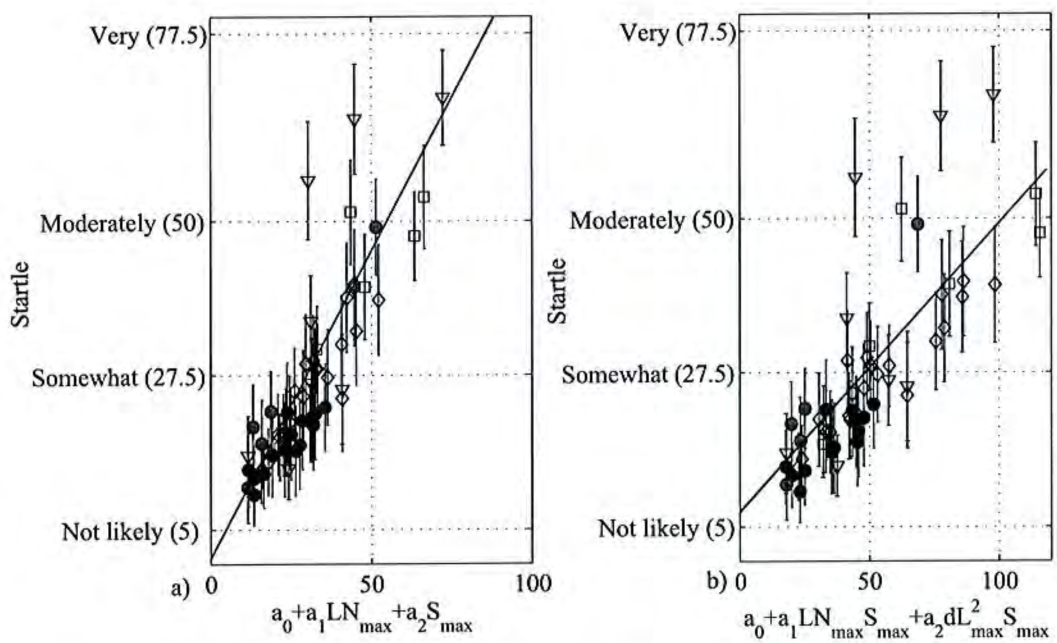

Figure 6.8. Average startle ratings versus a) linear model of LN_{max} and S_{max} and b) model of $LN_{max}S_{max}$ and $dL^2_{max}S_{max}$. R^2 values are in table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

It is possible that including duration-based metrics with these models would result in better prediction of subject ratings. A series of additional models was estimated using Models 6 and 8 as the first term and using duration-based metrics (Dur, number of sound events, Nb, and LN_{max} and combinations including LN_{max}) as the additional terms. Number of sound events was calculated by counting the number of peaks in the loudness time-history of each signal. A loudness peak was counted if it was local maximum in the loudness time-history and separated from other loudness maxima by a minimum loudness of, at most, 50% of its value. Coefficients of determination for these extended models (models 11-20) for all three types of responses are in Table 6.3.

The largest improvement in prediction in startle scores occurred when Dur or $LN_{max}Dur$ terms were added (Models 12 and 17). Startle ratings are plotted against predictions of these four models in Figures 6.9 and 6.10 for models with Dur and $LN_{max}Dur$, respectively. Including an additional Dur-based term results in better prediction of the blast noises rated above moderately startling. All of these extended models predicted acceptability fractions well, especially for the highly acceptable signal.

The predictions of the best of these extended models (12 and 17) both predicted the startle ratings from the experiment described in Chapter 5 well. Including the additional term resulted in very small differences in performance for Model 12 (less than 1% difference in variance explained). For Model 17, including the $LN_{max}Dur$ term resulted in slightly poorer predictions of Chapter 5 average startle ratings ($R^2 = 0.89$ compared to $R^2 = 0.92$ for the model without $LN_{max}Dur$). Including dL_{max} with the Model 12 (Model 21) slightly improved predictions for the Chapter 5 test data. However, the inclusion of this additional term was not significant (partial F-test, $\alpha = 0.05$).

Dur did not vary considerably in the Chapter 5 dataset; Dur was less than 400 ms for all of the sounds in that test. It is possible to improve the prediction of this models for both datasets by thresholding the $LN_{max}Dur$ term so that it is equal to zero when the duration is less than 500 ms. The reason for having the threshold at 500 ms is because this is where a transition occurs in subject ratings (see Figure 6.6) and because of the results found in Ramirez et al. (2005), described above. The thresholded model (21) was then compared

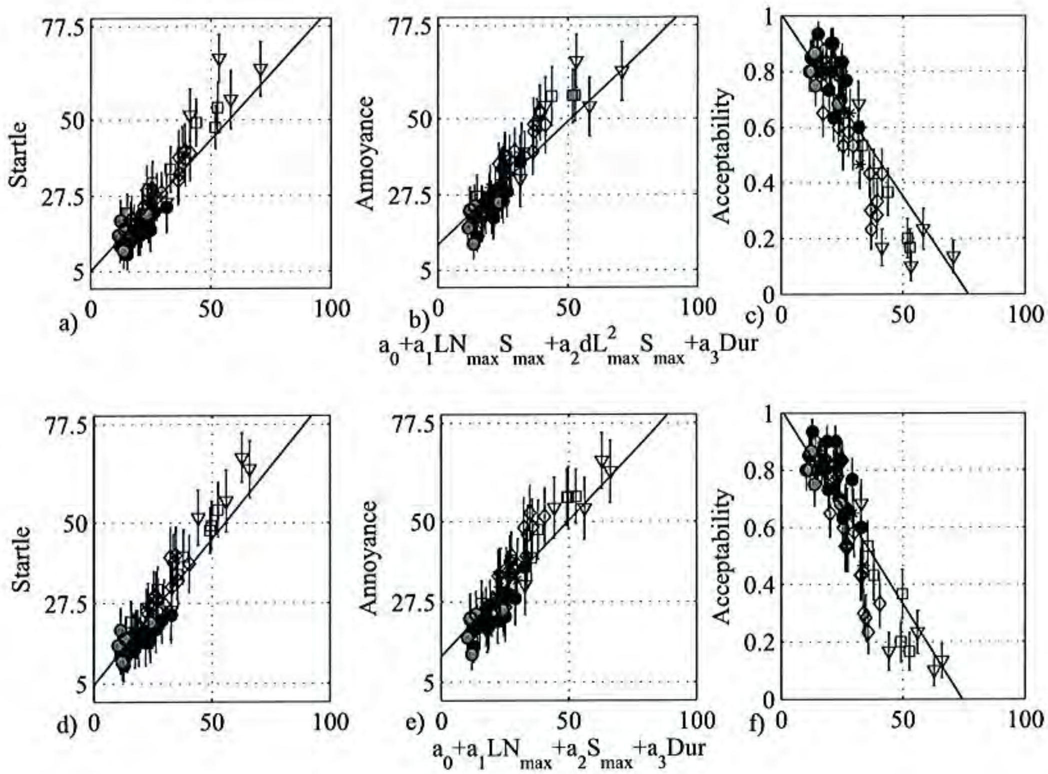

Figure 6.9. a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus model 8+Dur. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus model 6+Dur. R^2 values are in Table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

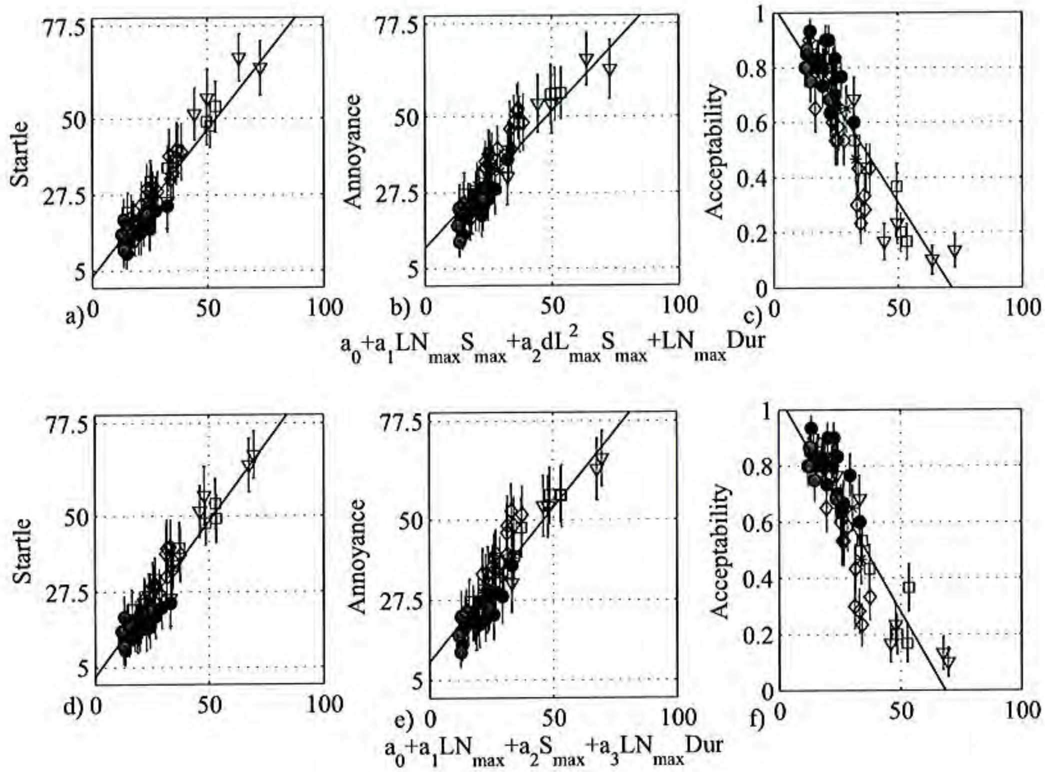

Figure 6.10. a) Average startle ratings, b) average annoyance ratings and c) acceptability fractions versus model 8+$LN_{max}Dur$. d) Average startle ratings, e) average annoyance ratings and f) acceptability fractions versus model 6+$LN_{max}Dur$. R^2 values are in Table 6.3. Black circles are synthetic booms, gray circles are simulated booms from industry, white circles are indoor booms,squares are car door slams, triangles are distant gunfire and blast noise, diamonds are boom-like noises and stars are booms with reflections.

to the subjective data obtained in tests described in Chapters 5 and 6. The coefficients of determination (R^2) from this analysis are compiled in Table 6.3 and the parameters of these models (12 and 21) are in Table 6.4.

Table 6.4 Variables and their coefficients in Models 12 and 21. The third-term coefficient is zero if Dur < 500 ms for Model 21.

	Model 21		Model 12	
	Term	Value	Term	Value
Intercept		5 ± 2		-18 ± 5
1st parameter	$LN_{max}S_{max}$	0.7 ± 0.1	LN_{max}	1.4 ± 0.2
2nd parameter	$dL^2_{max}S_{max}$	$(5 \pm 2) \times 10^{-7}$	S_{max}	12 ± 2
3rd parameter	$LN_{max}Dur$	0.0008 ± 0.0002	Dur	0.014 ± 0.002

Both models predict average startle ratings of both datasets well. The response predictions of the two models are near identical for the signals used in the Chapter 5 and 6 tests. Model 21 predictions explain 4% more of the startle response variance for the Chapter 5 dataset and 1% less startle and annoyance response variance for the Chapter 6 ratings. Model 12, however, has a simpler model structure than model 21: it contains no thresholded terms and no interaction terms (e.g. $LN_{max}S_{max}$) as parameters. Without thresholding the $LN_{max}Dur$ in Model 17, the performance between the two models is very similar (a difference of 1% of variance explained). In addition, the main variables in Model 12 are less correlated with each other than the main variables of Model 21 ($R^2 = 0.03$, for LN_{max} and S_{max} compared to $R^2 = 0.28$ for $LN_{max}S_{max}$ and $dL^2_{max}S_{max}$). Furthermore, the independence of loudness and sharpness in model 12 is also found in other researchers' annoyance model such as Zwicker and Fastl's (1999) psychoacoustic annoyance model. For these reasons, Model 12 is best model by parsimony.

6.3 Summary and Conclusions

An experiment was conducted to verify the performance of the candidate startle models in Chapter 6. In this experiments, participants were asked to rate the startle, annoyance and then acceptability of low amplitude sonic booms, door slams, boom-like noise and blast noise. It was found that average startle ratings, average annoyance ratings and acceptability fractions were all highly correlated (R^2 = 0.93, 0.87 and 0.90 for startle & annoyance, startle & acceptability and annoyance & acceptability, respectively). Candidate models were compared with participant ratings and previously estimated models that include the maximum of long-term Moore and Glasberg's time-varying loudness (LN_{max}) performed the best. However, these models did not predict all of the signals well. To extend the models to improve predictions of these currently poorly predicted responses, the inclusion of additional term was investigated. It was found that including the D with LN_{max} and S_{max} in a linear model resulted in good predictions of average startle ratings, average annoyance ratings and acceptability fractions in this study as well as startle ratings in the test described in Chapter 5.

7. CONCLUDING COMMENTS AND RECOMMENDED FUTURE WORK

The purpose of this research was to develop a model that could be used to predict the annoyance evoked by low amplitude sonic booms and other similar impulsive sounds when heard outdoors. After an examination of previous research on human responses to sonic booms, including startle, it was hypothesized that modeling startle would be yield a model that could predict sonic boom annoyance when heard outdoors.

7.1 Summary of Work Completed

A semantic differential experiment, described in Chapter 2, that used a sonic boom simulator for signal playback was conducted. The goal of this study was to examine how the results from a previous study, conducted over earphones, would change if low frequency (< 25 Hz) components are present in the stimuli, which can cause subjects to "feel" the sound. Subject ratings in both the earphone and simulator experiments were similar, but subjects did appear to respond to low frequency components in the stimuli. In both studies, judgments of annoyance and startle were highly correlated. Statistics of time-varying loudness models were highly correlated to average loudness, annoyance and startle ratings. Linear models of two or more metrics were examined and the best performing models in both tests included maximum Moore and Glasberg's short-term time-varying loudness (N_{max}) and either maximum rate of change of loudness (dL_{max}) or maximum sharpness (S_{max}). However, these additional terms were correlated with maximum loudness and each other. In addition, it was unclear if subjects' judgments of startle corresponded with subjects' physiological responses to these sounds.

Therefore, two studies were conducted to examine how subjects' judgments of startle compared to physiological responses. The first was the pilot study and is described in Chapter 3. The goal of this study was to examine the physiological responses evoked by stimuli used in the semantic differential tests. The experiment was divided into two parts: an arithmetic and memory task and a semantic differential. Participants' skin conductance, heart-rate and electrical activity of three neck muscles (sternocliedomastoid, anterior scalene and trapezius) were measured. Physiological responses consistent with startle were observed and the probability of a physiological response occurring was correlated with average startle ratings for skin conductance and heart-rate. However, the interval between stimulus exposures was too short. The arithmetic and memory task also caused physiological responses, further reducing the interval between responses. Both of these issues resulted in overlap between some of the physiological responses. In addition, there was considerable subject-to-subject variation in physiological responses. However, due to the aforementioned issues, it was unclear how much of this variance was due to the experimental design.

The subject-to-subject variation was investigated in an additional study and the results are described in Chapter 4. Participants were exposed to sonic booms while viewing pictures of landscapes. A longer inter-stimulus interval was used (2 minutes compared to 45 seconds). After the picture-viewing task, the subject completed a paired comparison test with the stimuli used in the picture exposure task. Participants completed the test twice, one session per day on adjacent days. Physiological responses associated with startle were found, but were fairly rare. In particular, there were relatively few neck muscle responses and many of these were of low magnitude. In addition, responses varied from day-to-day and from subject-to-subject. A mini-study, consisting of the same test with one repeated high-level sonic boom as the stimulus, resulted in slightly more severe responses, but a similar amount of subject-to-subject and day-to-day variability was found. Physiological measures evoked by the stimuli appeared to be consistent with orientation responses. The stimuli were found to be below the threshold of consistent startle responses. In contrast to the physiological measures, startle scores obtained from the paired comparison test were very similar from day-to-day and highly correlated to the probability of a skin conductance response occurring. Because of this, it was decided to use subject ratings to develop the startle model.

An experiment, described in Chapter 5, was conducted to obtain data to develop a startle model. Participants rated the startle evoked by each sound on a scale ranging from "Not likely to be startling" to "Extremely startling". Synthetic sonic booms, boom-like noises and industry-provided simulated sonic booms were used as stimuli. The stimuli were designed so that maximum Moore and Glasberg's short-term time-varying loudness (N_{max}), maximum rate of change of loudness before N_{max} is reached (dL_{max}) and mean sharpness (S_{mean}) were weakly correlated to each other. It was found that maximum sharpness (S_{max}) was more highly correlated with average subject ratings than mean sharpness. A series of models were estimated from the data, and it was found that several two-metric models predicted average startle ratings well. These models contained N_{max} or long-term maximum loudness (LN_{max}) and either S_{max} or dL_{max}.

To examine the performance of the candidate models another experiment, discussed in Chapter 6, was designed and conducted. Participants rated the startle, then annoyance, then acceptability of a series of sounds. These sounds consisted of synthetic sonic booms, door slams, blast noise, distant gunfire and industry-provided simulated booms. As with the semantic differential experiments, startle, annoyance and acceptability were highly correlated to each other. The previously developed candidate models were compared with average subject ratings and the models that contained LN_{max} or the interaction of LN_{max} and another metric as a parameter performed best. However, responses to signals with longer durations were predicted poorly by these models. To remedy this, additional terms that included loudness time-history-based duration (Dur) were investigated. The best of these extended models was

$$-18 + 1.4LN_{max} + 12S_{max} + 0.014Dur. \tag{7.1}$$

This model's predictions were highly correlated to participant ratings in both experiments (R^2 = 0.87 for the Chapter 5 data and R^2 = 0.91, 0.84, 0.81 for the startle, annoyance and acceptability data in Chapter 6).

7.2 Contributions of this Work

The contributions of the research described here are as follows:

1. The results of the earphone semantic differential experiment were verified in the sonic boom simulator.

2. Subjects' annoyance judgments were found to be highly correlated with startle judgments in multiple experiments.

3. Low amplitude sonic booms were found to produce physiological responses consistent with orientation and startle that are correlated to subject's judgments of startle. However, physiological responses have more day-to-day and subject-to-subject variability than psychophysical measures. Subjective measures were found to be a valid method of determining startle.

4. Low booms are at a level below the threshold necessary to consistently evoke startle responses. While physiological responses were evoked by these sounds, these responses were often less severe than those associated with startle. The majority of responses appear to be associated with orientation.

5. Metrics based on characteristics of loudness, as predicted by the most recent models of time-varying loudness were found to be more highly correlated to subjects' ratings of loudness and annoyance than metrics previously used to quantify people's responses to sonic booms and these loudness-based metrics can predict more accurately responses to a broader range of transient impulsive sounds than previously used sonic boom metrics.

6. Sound characteristics, other than peak loudness, were found to influence people's startle and annoyance responses and the incorporation of these characteristics into predictive models were found to significantly improve model performance.

7. A model of subject-rated startle was developed and validated with low amplitude sonic booms, distant gunfire, blast noise and door slams.

7.3 Recommendations for Future Work

The examination of how low amplitude sonic booms impact communities is ongoing. It is suggested that future work should focus on:

1. Understanding the role of spectral balance on startle.
2. Generalization of the startle model for low booms heard indoors.
3. Investigating how the frequency and number of events affects annoyance evoked by low amplitude sonic booms.
4. Investigating community impact of supersonic aircraft including sleep disturbance.
5. Investigating the impact of low booms on sensitive populations including people with sensorineural hearing loss.

All of these suggestions are natural extensions of this research. For 1), it was found in the experiments described in Chapters 5 and 6 that spectral balance is important for predicting subject-rated startle. Sounds with similar loudness characteristics were rated as significantly more startling when sharpness was high. It is unclear what the mechanism of this effect is because much of the fundamental startle research has used exclusively white noise or tone pulse stimuli. It is possible that understanding this effect would yield further insight into how startle works, particularly how it differs from loudness perception, which, like startle, shares similar nerve connections with the cochlear.

As the the startle model was generated and verified for booms heard outdoors, it is unclear how well this model would perform for booms heard indoors. When a sonic boom propagates through a house, the exposure environment can become more complex. Additional sound sources may be present from rattling objects and because of room reflections. This can result in a person inside a house feeling "surrounded by sound". As multiple paths of sound can reach the listener from different directions, understanding binaural hearing may be important. In order to understand how subjects perceive this situation, further investigation into booms under these conditions is needed.

Aside from limitations in predicting the response to low boom indoors, it is unclear how well this model would perform if used to predict the impact of these sounds in communities. The model was generated using laboratory studies and only the impact of individual sounds on neuro-typical populations with normal hearing were examined. Thus, it is necessary to verify this model in actual communities using noise surveys and extend it to be able to handle more than one aircraft event. It is also necessary to examine how well this model predicts the impact of these sounds on vulnerable populations because some vulnerable populations have heightened startle responses (such as those with Dyslexia) or perceive sound differently (such as those with hearing loss).

Furthermore, in a noise survey, people would be exposed to these sounds in an environment which is closer to that experienced by them if supersonic aircraft were allowed to operate. This would also include examining the effect of night-time flights, which may cause sleep disturbance. Noise surveys in a community are challenging because it is more difficult to control than a laboratory study as additional factors can affect a subject's opinion of these sounds. Even though the startle model developed here were able can be used to predicte trends in annoyance ratings in laboratory experiments, it does not incorporate complex socio-cultural factors associated with the operation of supersonic aircraft. Examples of these factors are a person's or community's views on the economic and environmental trade-offs associated with the operation of these craft. It is possible that booms that were found to be acceptable in a laboratory study would be considered unacceptable given the effect of these additional factors or visa versa. However, if sounds are found to be more annoying than others in a laboratory experiment, it is likely that when heard in a community they would also be more annoying than the other sounds. So the model developed here may be useful when incorporated in a test to simulate potential community response which in turn can be used to develop robust survey sampling techniques to gather sufficient data to validate (or reject) the model as well as to identify model deficiencies and identify socio-cultural factors.

LIST OF REFERENCES

LIST OF REFERENCES

D.M. Alexander, C. Trengove, P. Johnston, T. Cooper, J.P August, and E. Gordon. Separating individual skin conductance responses in a short interstimulus-interval paradigm. *Journal of Neuroscience Methods*, 146:116–123, 2005.

R.J. Barry and F.H Mitchell. A comparison of phasic cardiac responses derived from the electrocardiogram and the peripheral pulse. *International Journal of Psychophysiology*, 5:73–78, 1987.

H.E. Bass, R. Raspet, J.P. Chambers, and M. Kelly. Modification of sonic boom wave forms during propagation from the source to the ground. *Journal of the Acoustical Society of America*, 111(1):481–486, 2002.

L.L. Beranek. *Noise and Vibration Control.* Institute of Noise Control, revised edition, 1988.

B. Berglund and T. Lindvall. *Community Noise.* Centre for Sensory Research, Karolinska Institute, 1995.

E. A. Bjork. Startle, annoyance and psychophysiological responses to repeated sound bursts. *Acustica/Acta Acustica*, 85:575–578, 1999.

M.M Bradley and P.J. Lang. Affective reactions to acoustic stimuli. *Psychophysiology*, 37:204–215, 2000.

P. Brown, J. Rothwell, P. Thompson, T. Britton, B. Day, and C. Marsden. New observations on the normal auditory startle reflex in man. *Brain*, 114:1891–1902, 1991.

E. Cook and G. Turpin. *Attention and Orienting: Sensory and Motivational Processes.* Cambridge University Press, 1st edition, 1997.

US Supreme Court. City of Burbank v. Lockheed air terminal. Vol 411 U.S. 624,1973.

M. Davis, D.S. Gendelman, M.D. Tischler, and P.M. Gendelman. A primary acoustic starlte circuit: lesion and stimuluation studies. *Journal of Neuroscience*, 2:761–805, 1982.

M.E. Dawson, A.M Schell, and D.L. Filion. *Handbook of Psychophysiology*, chapter The electrodermal system, pages 200–223. Cambridge University Press, 2nd edition, 2000.

P.J. Delwaide and B. Schepens. Auditory startle (audio-spinal) reaction in normal man: Emg responses and h reflex changes in antagonistic lower limb muscles. *Electroencephalography and clinical Neurophysiology*, 97:416–423, 1995.

B. R. Glasberg and B. C. J. Moore. A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50(5):331–42, May 2002.

F.K. Graham and R.K. Clifton. Heart-rate change as a component of the orienting response. *Psychological Bulletin*, 65:305–320, 1966.

F.K. Graham and D.A. Slaby. Differential heart-rate changes to equally intense white noise and tone. *Psychophysiology*, 10:347–362, 1973.

C. Grillon. Startle reactivity and anxiety disorders: aversive conditioning, context, and neurobiology. *Biological Psychiatry*, 52:958–975, 2002.

C. Grillon and J. Baas. A review of the modulation of the startle reflex by affective states and its applications in psychiatry. *Clinical Neurophysiology*, 114:1557–1579, 2003.

C. Grillon and M. Davis. Fear-potentiated startle conditioning in humans: explicit and contextual cue conditioning folowing paired vs. unpaired training. *Psychophysiology*, 34: 451–458, 1997.

C. Grillon, K.R Merikangas, L. Dierker, N. Snidman, R.I. Arriaga, and J. Kagan. Startle potentiation by threat of aversive stimuli and darkness in adolescents: a multi-site study. *International Journal of Psychophysiology*, 32:63–73, 1999.

J.P. Guilford. *Psychometric Methods*. McGraw-Hill, 1954.

M.F Hamilton and D.T. Blackstock. *Nonlinear acoustics*. Academic Press, 1998.

L.W. Hawk and E.W Cook. Affective modulation of tactile startle. *Psychophysiology*, 34: 23–31, 1997.

D. Heimann. Effects of long-term atmospheric variability on the width of a sonic boom carpet produced by high-flying supersonic aircraft. *Acoustic Research Letters Online*, 2: 73–78, 2001.

K.O. Johnson. The roles and functions of cutaneous mechanoreceptors. *Current Opinion in Neurobiology*, 11:455–461, 2001.

F.P. Jones and J.L. Kennedy. An electromyographic technique for recording the startle pattern. *Journal of Psychology*, 32:63–68, 1951.

J.O. Kim and C.W. Mueller. *Factor Analysis: Statistical Methods and Practical Issues*. Sage, 1978.

J. Klos, A. Loubeau, and J. Rathsam. Overview of an indoor sonic boom simulator at nasa langley research center. *Journal of the Acoustical Society of America*, 129, 2011.

M. Koch. The neurobiology of startle. *Progress in Neurobiology*, 59:107–128, 1999.

P. Kucera, Z. Goldenberg, and E. Kurca. Sympathetic skin response: review of the method and its clinical use. *Bratisl Lek Listy*, 105:108–116, 2004.

V. Kumari, I. Aasen, A. Papadopoulos, F. Bojang, L. Poon, R. Halari, and A.J. Cleare. A comparison of prepulse inhibition in pre- and postmenopausal womean and age-matched men. *Neuropsychopharmacology*, 33:2610–2618, 2008.

C. Landis and W. Hunt. *The Startle Pattern*. Holt, Rinehart and Winston, 1st edition, 1939.

Jack D. Leatherwood, Brenda M. Sullivan, Kevin P. Shepherd, David A. McCurdy, and Sherilyn A. Brown. Summary of recent NASA studies of human response to sonic booms. *Journal of the Acoustical Society of America*, 111(1):586–598, January 2002.

J.D. Leatherwood and B.M. Sullivan. Laboratory study of effects of sonic boom shaping on judged loudness and acceptability. *Noise Control Engineering Journal*, 42(2):59–69, 1994.

J.D. Leatherwood, K.P. Shepherd, and B.M. Sullivan. *A New Simulator for Assessing Subjective Effects of Sonic Booms*. Technical Report NASA-TM-104150, National Aeronautical and Space Administration, Sept 1991.

H.G Leventhall. Somatic responses to low frequency noise. In *The proceedings of the 12th International Meeting: Low Frequency Noise and Vibration and its Control.*, Bristol, England, 2006.

H.G. Leventhall. Low frequency noise. what we know, what we do not know, and what we would like to know. *Journal of Low Frequency Noise, Vibration and Active Control*, 28:79–104, 2009.

B. Lipkens. Model experiment to study sonic boom propagation through turbulence. part iii: Validation of sonic boom propagation models. *Journal of the Acoustical Society America*, 111(1):509–519, 2002.

L. Locey and V. Sparrow. Modeling atmospheric turbulence as a filter for sonic boom propagation. *Noise Control Engineering Journal*, 55:459–495, 2007.

K. Ludewig, S. Ludewig, A. Seitz, M. Obrist, M.A. Geyer, and F.X. Vollenweider. The acoustic startle reflex and its modulation: effects of age and gender in humans. *Biological Psychology*, 63:311–323, 2003.

A. Marshall. A preliminary investigation into the perceptual characteristics of low level sonic booms heard outdoors, 2007.

D.N. May. Startle in the presence of background noise. *Journal of Sound and Vibration*, 17:77–81, 1971.

D.N May. Sonic boom startle: A field study in Meppen, West Germany. *Journal of Sound and Vibration*, 23:337–347, 1972.

D.N. May and R.A Sohn. Simple response metrics for minimized and conventional sonic booms. *Journal of Sound and Vibration*, 145(2):225–238, 1990.

D.C. Montegomery. *Design and Analysis of Experiments*. John Wiley and sons, 2005.

A. Niedzwiecki and H. S. Ribner. Subjective loudness of N-wave sonic booms. *Journal of the Acoustical Society of America*, 64(6):1617–1621, 1978.

J.K Niparko. *Cochlear Implants: Principles and Practices*. Lippincott, Williams and Wilkins, 2009.

C.W. Nixon and P.N. Borsky. Effects of sonic boom on people:St. Louis, Missouri, 1961-1962. *The Journal of the Acoustical Society of America*, 39(5,part 2):S51–S58, 1966.

C.E. Osgood and G.J. Suci. Factor analysis of meaning. *Experimental Psychology*, 50: 325–338, 1955.

H.P Patterson. *Public Reactions to Sonic Booms*. Technical report, NASA Contractor Report, NASA-CR-1665, National Aeronautics and Space Administration, Washington, D.C., Sept. 1970.

P.K. Pilz, H. Schnitzler, and D. Menne. Acoustic startle threshold of the albino rat (rattus norvegicus). *Journal of Comparative Psychology*, 101:67–72, 1987.

C. Plappert, P. Pilz, and H. Schinitzler. Factors governing prepulse inhibition and prepulse faciliation of the acoustic startle response in mice. *Behavioral Brain Research*, 152:403–412, 2004.

K. Plotkin and B. Bradley. *The effect of onset rate on aircraft noise annoyance Volume 1: Laboratory Experiments*. Technical report, Wyle Laboratories, Report A183982, November 1991.

K. Plotkin, D. Graham, P. Coen, E. Haering, D. Maglieri, A. Pilon, J. Salamone, J. Page, J. Pawlowski, D. Schein, D. McCurdy, J. Murraym, L. Ehernberger, and P. Bobbit. Ground measurements of a shaped sonic boom. In *10th American Institute of Aeronautical and Astronautics Aeroacoustics Conference Proceedings, Manchester, UK*, page 2923, May 2004.

Kenneth J. Plotkin. State of the art of sonic boom modeling. *Journal of the Acoustical Society of America*, 111(1):530–536, 2002.

L.R. Rabiner and B. Gold. *Theory and Application of Digital Signal Processing*. Prentice-Hall, 1975.

I. Ramirez, M. Sanchez, M. Fernadez, O. Lipp, and J. Vila. Differentiation between prtective reflexes: cardiac defense and startle. *Psychophysiology*, 42:732–739, 2005.

R. Rangayyan. *Biomedical Signal Analysis: A Case-Study Approach*. Wiley and Sons, 1st edition, 2002.

L. Reijmers and B. Peeters. Effects of acoustic prepulses on the startle reflex in rats: A parametric analysis. *Brain Research*, 661:174–180, 1994.

R. Rylander and A. Dancer. Startle reactions to simulated sonic booms: Influence of habituation, boom level and background noise. *Journal of Sound and Vibration*, 61(2): 235–243, 1978.

R. Rylander, S. Sorensen, G. Chatelier, Y. Espmark, T. Larsson, and R.I. Thackray. Sonic boom exposure effects - a field study on humans and animals. *Journal of Sound and Vibration*, 33:471–486, 1974.

J. Salamone. Portable sonic boom simulation. *Proceedings of the 17th International Symposium on Nonlinear Acoustics*, pages 667–670, July 18-22, 2005.

N. Schmajuk and J. Larrauri. A neural network model of prepulse inhibition. *Behavioral Neuroscience*, 119:1546–1562, 2005.

N. Schmajuk and J. Larrauri. Prespulse inhibition mechanism and congitive processes: a review and model. *Neurotransmitter interactions and congnitive function*, pages 245–278, 2006.

P. D. Schomer, J. W. Sias, and D. Maglieri. A comparative study of human response, indoors, to blast noise and sonic booms. *Noise Control Engineering Journal*, 45(4):169–82, July 1997.

P.D. Schomer. High-energy impulsive noise assessment. *Journal of the Acoustical Society of America*, 79(1):182–186, 1986.

P.D. Schomer and J.W. Sias. On spectral weightings to assess human response, indoors, to blast noise and sonic booms. *Noise Control Engineering Journal*, 46(2):57–71, March - April 1998.

P.D. Schomer and L.R. Wagner. Human and community response to military sound - part 2: Results from field-laboratory tests of sounds of small arms, 25 mm cannons, helicopters, and blasts. *Noise Control Engineering Journal*, 43:1–13, 1995.

P.D. Schomer, Y. Suzuki, and F. Saito. Evaluation of loudness-level weightings for assessing the annoyance of environmental noise. *The Journal of the Acoustical Society of America*, 110(5):2390–2397, November 2001.

R. Seebass and A. George. Sonic boom minimization. *The Journal of the Acoustical Society of America*, 51:686–694, 1972.

Y.N. Sokolov. *Perception and the Conditioned Reflex*. Pergamon Press, 1st edition, 1963.

K.W. Spence and W.N. Runquist. Temporal effects of conditioned fear on the eyelid reflex. *Journal of Experimental Psychology*, 55:613–616, 1958.

B.M. Sullivan. Human response to simulated low intensity sonic booms. In *Proceedings of Noise-Con 2004, Baltimore*, pages 541–550, July 2004.

B.M. Sullivan, P. Davies, K.K. Hodgdon, J.A Salamone III, and A. Pilon. Realism assessment of sonic boom simulators. *Noise Control Engineering Journal*, 56:141–157, 2008.

N. Swerdlow, J. Shoemaker, N. Stephany, L. Wasserman, H. Ro, and M. Geyer. Prestimulus effects on startle magnitude: Sensory or motor? *Behavioral Neuralscience*, 116: 672–681, 2002.

B.K Taylor, R. Casto, and M.P. Printz. Disssociation of tactile and acoustic components in air puff startle. *Physiology and Behavior*, 49:527–532, 1991.

R.I. Thackray. Sonic boom exposure effects II.3: Startle responses. *Journal of Sound & Vibration*, 20:519–526, 1972.

R.I. Thackray, R Touchstone, and J.P. Bailey. *Behavioral, autonomic and subjective reactions to low and moderate-level simulated sonic booms: A report of two experiments and a general evaluation of sonic boom startle effects*. FAA-AM-74-9, FAA Civil Aeromedical Institute, November 1974.

G. Turpin. Effects of stimulus intensity based on autonomic responding: the problem of differentiating orienting and defense reflexes. *Psychophysiology*, 23, 1986.

G. Turpin, F. Schaefer, and B. Boucsein. Effects of stimulus intensity, risetime, and duration on autonomic and behavioral responding: Implications for the differentiation of orienting, startle, and defense responses. *Psychophysiology*, 36:453–463, 1999.

N. Unsworth, R. Heitz, J. Schrock, and R. Engle. An automated version of the operation span task. *Behavior Research Methods*, 37:498–505, 2005.

H. E. von Gierke. Effects of sonic boom on people – review and outlook. *Journal of the Acoustical Society of America*, 39(5, Part 2):43–50, 1966.

D.L Walker and M. Davis. Anxiogenic effects of high illumination levels assessed with the acoustic startle response in rats. *Biological Psychiatry*, 42:461–471, 1997.

C.H.E. Warren. Sonic boom exposure effects II.6:sonic boom generators. *Journal of Sound and Vibration*, 20:535–539, 1972a.

C.H.E. Warren. Sonic boom exposure effects I.2:the sonic boom–generation and propagation. *Journal of Sound and Vibration*, 20:485–497, 1972b.

R. Widdle. A brief history of the ban on civil overland supersonic flight in the united states. August 2005.

S. Yamada, M. Ikuji, T. Watanabe, and T. Kosaka. Body sensations of low frequency noise of ordindary persons and profoundly deaf persons. *Journal of Low Frequency Noise and Vibration*, 2:32–36, 1983.

J.S. Yeomans, L. Li, B.W. Scott, and P.W. Frankl. Tactile, acoustic and vestibular systems sum to elicit startle reflex. *Neuroscience and biobehavior reviews*, 26:1–11, 2002.

E. Zwicker and H. Fastl. *Psychoacoustics, Facts and Models*. Springer, 2nd edition, 1999.

APPENDICES

Appendix A: Semantic Differential Experiments

The instructions for both tests are below. The text in the () sections were different between the two experiments.

The purpose of this experiment is to determine the characteristics of a set of transient sounds. In this experiment, you will here a variety of transient sounds from various sources. Upon hearing a sound, please mark on the scales provided how you would feel about the sound if you were outdoors in a park or garden and heard this sound intermittently throughout the day. The sound will be repeated while you complete the ratings. Some of the sounds had noticeable background noise. Please consider only the main impulsive sound event when making on the scale. When finished marking all of the scales for one sound, (click "next sound") OR (signal the camera) to move on to the next sound. Remember there are no right or wrong answers. We are interested in your opinion.

Statistics of the responses for all scales found in the earphone semantic differential experiment are in Table A.1. Statistics of the responses found in the simulator semantic differential experiment are in Table A.2.

Table A.1: Statistics of subject responses in the earphone semantic differential experiment. Signal numbers correspond to Table 2.2 and Scale numbers correspond to Table 2.1. Scales ranged from -16 to +16.

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
1	1	-15	12	6.67	-1.80	-3	11	1	-11	15	6.62	3.12	4
1	2	-13	13	6.62	0.20	2	11	2	-13	15	8.47	0.24	-1
1	3	-4	15	5.18	5.96	5	11	3	-16	12	6.67	-4.48	-3
1	4	-13	16	7.37	3.20	4	11	4	-14	11	6.50	-0.40	-3
1	5	-12	10	5.66	-0.88	-1	11	5	-7	15	6.36	3.28	5
1	6	-7	10	4.81	1.68	2	11	6	-14	10	6.86	-1.72	-1
1	7	-11	13	5.29	-2.48	-3	11	7	-10	13	6.14	3.16	4
1	8	-16	5	5.36	-4.36	-4	11	8	-10	10	5.95	1.64	3
1	9	-15	5	6.11	-6.28	-6	11	9	-10	15	6.89	5.84	8
1	10	-13	14	6.16	-3.44	-4	11	10	-13	13	5.91	4.04	6
1	11	-16	15	7.47	-6.52	-8	11	11	-12	16	7.95	3.04	3
1	12	-13	13	5.91	-0.60	-1	11	12	-11	14	7.25	0.16	-1
1	13	-2	15	4.17	7.80	9	11	13	-14	4	4.79	-5.72	-6
1	14	-4	16	5.64	3.56	3	11	14	-16	12	6.82	-2.68	-4
1	15	-9	16	5.69	7.12	7	11	15	-16	14	8.14	-6.64	-10
1	16	-16	4	5.04	-5.98	-6	11	16	-16	16	7.87	1.77	1
1	17	-15	11	6.41	-2.00	-3	11	17	-10	15	6.46	0.64	1
1	18	-15	13	7.36	-1.56	-3	11	18	-16	15	8.32	-0.84	-1
1	19	-6	10	5.10	1.68	2	11	19	-12	13	6.43	-2.04	-4
1	20	-3	16	4.54	11.20	12	11	20	-16	13	7.53	-7.68	-9
1	21	-7	14	5.13	0.96	1	11	21	-11	11	6.15	0.28	-2
1	22	-9	14	6.18	2.16	2	11	22	-10	16	6.20	0.32	1
1	23	-14	9	5.48	-4.04	-4	11	23	-7	14	6.42	4.68	6
1	24	-7	14	4.61	6.76	7	11	24	-13	10	5.89	-5.40	-7
2	1	-15	6	4.67	-6.20	-7	12	1	-12	6	4.61	-3.20	-4
2	2	-12	12	6.69	-2.12	-2	12	2	-15	16	7.91	1.40	2
2	3	-12	14	7.46	0.44	1	12	3	-6	15	6.13	4.60	3
2	4	-12	12	5.95	-0.80	-1	12	4	-12	13	6.49	2.52	4
2	5	-12	9	4.52	-5.28	-5	12	5	-16	9	6.10	-1.52	-2
2	6	-12	12	6.51	-3.52	-5	12	6	-10	12	5.80	3.16	4
2	7	-13	6	5.13	-5.60	-6	12	7	-14	13	6.71	-1.28	-2
2	8	-13	1	4.05	-7.64	-7	12	8	-12	10	5.87	-2.08	-2
2	9	-14	13	6.97	-6.88	-9	12	9	-15	12	7.20	-4.80	-5
2	10	-14	14	5.70	-4.72	-5	12	10	-14	15	8.57	-1.64	-2
2	11	-16	5	5.23	-4.56	-3	12	11	-15	10	8.65	-3.33	-6
2	12	-14	16	7.69	-2.00	-2	12	12	-12	15	6.98	3.12	5
2	13	-10	16	5.72	9.72	11	12	13	-6	15	5.63	7.48	8
2	14	-12	15	7.28	4.88	4	12	14	-9	15	6.61	4.96	5
2	15	-3	16	5.38	10.04	12	12	15	-8	16	7.11	7.80	10

Table A.1 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
2	16	-16	12	7.39	-2.75	-4	12	16	-16	16	8.58	0.06	2.25
2	17	-15	13	7.87	-2.32	-3	12	17	-15	13	8.13	1.76	4
2	18	-14	12	7.89	-0.76	1	12	18	-12	16	7.93	4.40	6
2	19	-14	11	5.61	-4.60	-5	12	19	-11	12	6.07	2.04	2
2	20	-9	16	5.94	9.40	11	12	20	-8	16	6.63	7.24	9
2	21	-15	10	7.47	-2.08	-3	12	21	-10	9	5.68	0.76	1
2	22	-13	7	5.61	-3.28	-4	12	22	-12	13	6.59	2.84	3
2	23	-14	12	6.04	-5.80	-7	12	23	-15	6	5.74	-4.32	-4
2	24	-10	14	5.82	7.36	9	12	24	-2	16	4.47	8.24	9
3	1	-16	11	5.33	-6.84	-6	13	1	-11	12	5.71	-0.56	1
3	2	-15	12	6.61	-3.32	-4	13	2	-12	16	8.63	3.60	5
3	3	-16	14	7.63	-0.80	1	13	3	-5	13	5.15	4.88	4
3	4	-14	5	5.79	-5.32	-6	13	4	-11	12	6.21	3.04	3
3	5	-14	6	5.15	-7.32	-9	13	5	-12	10	4.96	0.04	1
3	6	-12	14	8.33	-3.20	-5	13	6	-12	10	5.83	0.88	1
3	7	-15	3	4.76	-8.12	-10	13	7	-13	14	6.13	-1.20	-3
3	8	-15	0	4.22	-8.96	-10	13	8	-10	13	6.80	-1.32	-2
3	9	-16	12	6.69	-7.20	-7	13	9	-14	12	6.58	-3.44	-5
3	10	-16	13	7.91	-6.04	-7	13	10	-13	5	4.70	-3.40	-4
3	11	-15	16	6.75	-5.92	-7	13	11	-12	9	5.37	-1.00	0
3	12	-14	6	5.58	-4.20	-3	13	12	-10	16	5.46	1.84	1
3	13	-7	14	5.09	7.24	9	13	13	-6	16	5.06	8.56	10
3	14	-14	16	6.75	3.80	4	13	14	-5	15	5.90	6.17	7
3	15	-15	15	7.04	4.80	6	13	15	-3	16	4.97	10.48	12
3	16	-15	12	7.56	-3.23	-3.5	13	16	-15	13	5.73	0.29	1
3	17	-16	12	7.19	-3.68	-5	13	17	-16	9	5.44	0.16	1
3	18	-13	14	8.37	-1.68	-5	13	18	-8	16	5.57	4.44	4
3	19	-15	4	5.85	-5.00	-5	13	19	-7	11	5.18	3.00	3
3	20	-8	15	5.96	7.52	8	13	20	-3	16	5.07	10.64	12
3	21	-16	12	7.08	-4.16	-4	13	21	-6	10	4.68	1.56	3
3	22	-15	11	6.52	-4.32	-4	13	22	-11	13	4.95	3.48	4
3	23	-16	2	5.42	-7.76	-7	13	23	-14	6	5.53	-1.76	-2
3	24	-12	14	5.61	7.28	7	13	24	-3	14	4.81	8.16	9
4	1	-13	13	6.77	-3.12	-3	14	1	-12	9	6.41	0.52	1
4	2	-15	12	7.29	0.44	2	14	2	-13	12	7.25	-2.00	-4
4	3	-11	14	6.19	2.12	3	14	3	-16	3	5.33	-5.60	-5
4	4	-13	12	7.49	0.16	-1	14	4	-15	10	6.17	-3.60	-3
4	5	-13	7	5.02	-2.80	-3	14	5	-9	14	5.53	0.32	-2
4	6	-10	9	4.81	-2.48	-3	14	6	-12	9	5.50	-2.20	-2
4	7	-15	14	5.86	-2.96	-4	14	7	-13	12	6.52	1.20	2
4	8	-11	3	3.37	-6.20	-6	14	8	-9	7	4.36	1.04	1
4	9	-15	9	5.72	-6.20	-5	14	9	-9	15	6.57	4.36	4

Table A.1 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
4	10	-13	12	6.73	-4.12	-5	14	10	-12	13	5.96	2.16	3
4	11	-16	13	7.81	-3.17	-4.5	14	11	-15	13	6.55	-1.36	-1
4	12	-14	15	8.58	0.88	2	14	12	-12	3	4.69	-3.96	-4
4	13	-7	15	5.20	9.20	11	14	13	-13	1	3.62	-7.04	-6
4	14	-11	14	6.24	5.20	6	14	14	-16	4	5.09	-5.72	-5
4	15	-2	16	5.55	9.44	10	14	15	-16	3	5.28	-9.00	-11
4	16	-10	13	6.37	1.56	3	14	16	-16	12	6.76	-2.25	-1
4	17	-14	16	7.90	-0.60	-2	14	17	-13	16	7.03	-2.28	-2
4	18	-12	16	8.56	3.68	4	14	18	-16	13	7.77	-3.64	-5
4	19	-15	12	7.14	-1.48	-2	14	19	-11	13	5.64	-2.32	-3
4	20	-11	16	7.05	8.96	11	14	20	-16	6	5.60	-8.60	-10
4	21	-15	12	6.43	-3.16	-4	14	21	-14	11	6.35	-1.76	-4
4	22	-13	11	6.26	-2.56	-2	14	22	-14	6	4.92	-2.96	-2
4	23	-15	12	6.33	-5.04	-6	14	23	-8	11	5.33	2.20	3
4	24	-14	15	7.93	6.16	9	14	24	-14	2	4.27	-7.60	-7
5	1	-13	9	5.30	-1.88	-3	15	1	-13	8	5.66	0.80	2
5	2	-16	14	6.80	1.64	2	15	2	-15	12	6.51	-2.88	-3
5	3	-9	15	6.46	2.96	3	15	3	-15	2	5.18	-6.92	-9
5	4	-10	12	5.93	2.04	2	15	4	-14	11	6.54	-4.04	-5
5	5	-10	9	5.36	-1.00	-2	15	5	-10	14	6.14	-0.04	-1
5	6	-7	9	4.89	0.08	1	15	6	-14	11	5.51	-2.76	-4
5	7	-13	14	7.08	-1.40	-3	15	7	-14	14	6.39	0.20	-1
5	8	-12	6	5.60	-3.68	-4	15	8	-12	14	5.82	1.16	3
5	9	-12	6	5.68	-4.48	-4	15	9	-11	15	7.16	3.32	4
5	10	-15	12	6.68	-4.72	-6	15	10	-12	12	6.75	1.36	2
5	11	-14	12	6.28	-2.68	-2	15	11	-15	12	6.61	1.00	2
5	12	-11	13	5.87	1.92	3	15	12	-13	6	5.66	-4.00	-4
5	13	-11	12	5.11	7.28	10	15	13	-14	5	4.78	-8.08	-9
5	14	-16	15	7.23	3.16	3	15	14	-14	9	6.17	-4.56	-5
5	15	-4	15	4.70	9.60	11	15	15	-16	10	6.55	-8.36	-10
5	16	-15	16	6.45	-0.56	-0.5	15	16	-16	9	5.92	-1.33	-2
5	17	-14	13	6.25	0.44	1	15	17	-14	13	6.68	-2.84	-4
5	18	-12	14	7.08	4.28	6	15	18	-16	11	7.23	-4.56	-5
5	19	-16	11	7.35	-0.76	0	15	19	-11	12	5.67	-2.28	-2
5	20	-14	16	6.64	9.00	10	15	20	-16	1	5.34	-9.68	-11
5	21	-12	10	6.49	0.12	2	15	21	-10	13	6.68	-0.96	-3
5	22	-12	13	5.86	1.44	2	15	22	-12	11	5.15	-3.36	-4
5	23	-12	10	6.63	-4.16	-5	15	23	-10	11	5.92	0.96	1
5	24	-9	14	5.73	6.64	9	15	24	-14	1	4.44	-8.36	-10
6	1	-14	7	6.62	-2.64	-3	16	1	-16	12	7.12	-0.68	-1
6	2	-11	12	6.71	0.96	3	16	2	-12	16	7.46	3.24	4
6	3	-7	14	5.35	2.44	3	16	3	-5	16	5.00	6.96	7

Table A.1 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
6	4	-12	12	6.92	0.24	2	16	4	-13	12	6.24	3.28	4
6	5	-15	11	5.82	-2.80	-3	16	5	-14	14	7.52	-0.20	1
6	6	-13	12	6.37	-0.60	-3	16	6	-10	11	5.30	1.76	2
6	7	-15	9	6.10	-3.28	-3	16	7	-7	13	5.09	-0.44	-1
6	8	-10	9	5.18	-3.60	-5	16	8	-12	10	6.05	-0.72	1
6	9	-16	1	4.45	-6.32	-6	16	9	-15	11	6.99	-3.64	-4
6	10	-15	12	7.03	-2.84	-4	16	10	-12	11	6.73	-1.24	-3
6	11	-9	15	6.20	7.40	9	16	11	-16	12	8.88	-3.64	-5
6	12	-13	15	6.06	7.32	8	16	12	-16	15	6.96	1.88	2
6	13	-12	16	5.64	10.44	11	16	13	-5	15	4.77	9.00	10
6	14	-1	16	3.81	9.84	11	16	14	-14	14	6.93	4.72	7
6	15	-12	16	5.52	11.76	13	16	15	-12	16	6.76	9.28	11
6	16	-6	15	6.86	7.93	11	16	16	-13	16	7.05	-0.80	1
6	17	-13	15	7.25	6.16	6	16	17	-14	16	8.19	1.08	2
6	18	-9	16	5.76	8.76	10	16	18	-13	16	9.25	2.24	4
6	19	-13	14	6.79	0.16	2	16	19	-15	12	6.42	2.12	3
6	20	-12	16	7.30	1.88	2	16	20	4	16	3.57	11.56	13
6	21	-13	9	6.13	-1.48	-2	16	21	-12	14	6.82	1.24	1
6	22	-14	12	7.47	-0.40	-1	16	22	-13	13	5.85	3.56	3
6	23	-13	11	5.90	-3.84	-4	16	23	-12	10	6.10	-1.44	-2
6	24	-7	15	5.71	9.44	10	16	24	-4	16	5.53	9.24	11
7	1	-16	11	7.16	-5.84	-6	17	1	-13	14	7.32	4.80	7
7	2	-12	16	7.86	-2.20	-3	17	2	-12	14	6.89	3.96	4
7	3	-12	16	7.98	-1.84	-5	17	3	-10	13	6.33	1.24	2
7	4	-15	14	8.93	-0.40	-1	17	4	-13	12	6.95	0.80	1
7	5	-15	14	7.70	-3.33	-5	17	5	-14	14	6.60	2.40	3
7	6	-12	11	6.45	-2.52	-2	17	6	-11	14	5.95	2.48	2
7	7	-15	11	6.82	-5.48	-6	17	7	-9	14	6.28	2.80	3
7	8	-12	12	5.79	-5.48	-7	17	8	-10	15	5.88	2.12	2
7	9	-16	11	7.43	-6.60	-9	17	9	-13	15	7.29	2.60	2
7	10	-16	7	6.94	-4.88	-6	17	10	-16	14	8.44	1.40	3
7	11	-16	12	9.19	-0.44	1	17	11	-15	12	6.36	-8.52	-11
7	12	-6	13	5.66	5.44	4	17	12	-16	15	6.75	-8.28	-10
7	13	-11	14	5.96	8.64	11	17	13	-15	10	7.46	-6.16	-9
7	14	-11	14	5.49	8.32	10	17	14	-16	8	5.23	-8.96	-10
7	15	-12	16	6.02	10.32	12	17	15	-16	-6	2.94	-12.96	-14
7	16	-15	16	8.88	2.85	4	17	16	-16	0	4.85	-9.79	-11
7	17	-12	14	7.05	2.80	3	17	17	-15	6	5.96	-7.80	-9
7	18	-10	16	6.74	5.92	6	17	18	-16	10	5.80	-10.40	-12
7	19	-15	11	7.42	-2.64	-2	17	19	-13	15	6.46	0.92	0
7	20	-12	16	7.51	8.40	11	17	20	-12	16	6.74	7.36	7
7	21	-14	13	7.21	-4.48	-6	17	21	-9	15	6.44	3.88	4

Table A.1 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
7	22	-14	13	8.56	-2.20	-4	17	22	-12	15	6.57	0.68	1
7	23	-15	10	7.89	-6.12	-9	17	23	-15	14	8.32	2.20	3
7	24	-10	15	6.77	8.48	10	17	24	-16	9	6.79	-8.60	-12
8	1	-11	10	6.36	-1.20	-2	18	1	-13	15	7.98	2.92	4
8	2	-12	11	6.88	0.68	2	18	2	-13	14	6.92	3.56	5
8	3	-6	15	5.87	4.68	5	18	3	-6	15	6.78	5.08	5
8	4	-10	12	5.77	2.80	3	18	4	-11	14	7.30	4.12	5
8	5	-12	7	4.94	-0.08	1	18	5	-4	12	3.65	5.48	5
8	6	-10	13	5.44	2.12	2	18	6	-16	15	7.28	4.88	4
8	7	-11	12	6.73	-1.20	-2	18	7	-7	15	5.89	4.68	4
8	8	-10	10	5.25	-1.24	-1	18	8	-6	15	5.89	3.92	4
8	9	-13	11	5.93	-2.80	-3	18	9	-13	16	7.41	2.96	3
8	10	-15	5	5.60	-3.08	-3	18	10	-9	13	5.79	2.84	3
8	11	-13	14	7.95	1.00	1	18	11	-15	14	6.94	-6.96	-9
8	12	-16	13	7.14	2.40	3	18	12	-16	12	7.50	-6.48	-9
8	13	1	16	4.46	8.48	10	18	13	-16	9	8.24	-4.80	-7
8	14	-16	15	7.15	5.28	5	18	14	-16	16	7.90	-6.48	-9
8	15	-4	16	5.41	9.80	11	18	15	-16	16	8.23	-9.36	-13
8	16	-12	14	5.95	2.79	3	18	16	-16	4	5.81	-7.93	-9
8	17	-14	15	6.21	2.56	3	18	17	-16	10	7.39	-6.44	-8
8	18	-8	16	5.86	4.04	4	18	18	-16	12	8.26	-7.20	-10
8	19	-10	14	5.41	2.96	3	18	19	-16	12	6.71	1.68	2
8	20	3	16	4.32	10.24	11	18	20	-13	16	7.61	6.76	10
8	21	-9	13	6.02	2.20	3	18	21	-4	16	4.55	5.08	4
8	22	-7	13	5.05	3.96	4	18	22	-10	12	6.05	3.28	4
8	23	-14	6	6.05	-3.12	-5	18	23	-14	13	7.73	3.16	5
8	24	-1	15	4.04	7.92	9	18	24	-16	10	8.43	-5.76	-9
9	1	-15	15	8.74	0.32	-1	19	1	-8	15	6.13	5.24	6
9	2	-15	15	9.31	1.16	3	19	2	-12	14	6.61	3.84	4
9	3	-15	15	9.52	3.12	6	19	3	-14	16	6.22	6.08	6
9	4	-15	13	8.45	1.44	3	19	4	-10	14	7.12	4.40	4
9	5	-12	13	7.85	2.12	3	19	5	-10	14	6.00	5.80	6
9	6	-7	15	6.50	3.52	4	19	6	-11	15	6.35	7.40	9
9	7	-15	12	8.39	0.32	2	19	7	-7	16	6.50	5.88	7
9	8	-15	12	7.66	-0.36	-1	19	8	-7	15	6.59	5.92	6
9	9	-15	11	7.19	-2.68	-4	19	9	-13	16	8.67	2.76	4
9	10	-14	11	7.58	-1.52	-2	19	10	-11	15	6.57	4.28	5
9	11	-14	15	9.02	0.64	1	19	11	-12	14	7.46	-1.88	-3
9	12	-14	16	7.31	0.72	0	19	12	-16	16	8.07	-1.64	0
9	13	-13	14	7.56	6.12	7	19	13	-14	12	7.15	-3.36	-4
9	14	-13	16	7.93	5.75	6.5	19	14	-16	16	7.93	-3.32	-2
9	15	-10	16	6.94	7.80	10	19	15	-16	13	6.46	-8.24	-10

Table A.1 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
9	16	-16	12	8.64	-0.56	2	19	16	-15	10	8.20	-2.77	-2
9	17	-14	15	8.26	1.72	3	19	17	-16	13	7.83	-2.24	-3
9	18	-15	15	9.47	1.68	2	19	18	-15	12	7.69	-2.04	-2
9	19	-9	16	6.95	4.52	5	19	19	-4	16	4.51	6.04	5.5
9	20	-16	16	10.4	-5.68	-9	19	20	-13	15	8.08	5.68	6
9	21	-14	15	8.47	1.60	3	19	21	-10	14	5.69	7.44	8
9	22	-10	14	6.10	5.04	5	19	22	-11	13	5.87	5.56	6
9	23	-16	16	8.49	-0.92	-3	19	23	-6	15	6.07	6.72	7
9	24	-7	16	5.95	7.40	9	19	24	-15	12	7.49	-3.28	-3
10	1	-10	10	5.91	-0.48	1	20	1	-16	13	7.04	5.12	6
10	2	-12	15	7.11	3.96	5	20	2	-11	14	6.54	3.12	4
10	3	-14	15	7.00	5.00	6	20	3	-9	15	5.91	5.00	4
10	4	-12	15	7.40	3.00	3	20	4	-10	15	6.41	4.56	5
10	5	-12	12	5.72	1.20	1	20	5	-4	15	5.61	7.68	9
10	6	-6	13	4.93	1.56	3	20	6	-2	14	4.14	6.68	7
10	7	-14	13	7.22	-1.16	1	20	7	-6	15	5.30	6.04	6
10	8	-9	9	4.88	-0.56	-1	20	8	-5	14	5.87	7.20	9
10	9	-15	6	6.03	-3.44	-3	20	9	-12	16	8.33	4.36	6
10	10	-14	6	4.87	-3.32	-3	20	10	-12	16	6.90	6.52	7
10	11	-15	12	6.34	-1.68	-3	20	11	-15	16	9.44	-0.16	3
10	12	-16	7	6.02	-1.32	0	20	12	-13	14	6.05	0.68	2
10	13	-10	13	5.35	5.20	6	20	13	-15	11	7.08	-5.68	-6
10	14	-5	13	3.96	4.33	4	20	14	-14	15	7.89	-2.52	-4
10	15	-11	16	6.58	7.64	9	20	15	-15	11	6.96	-6.48	-9
10	16	-15	9	6.20	-1.31	1	20	16	-12	15	7.83	-2.08	-4
10	17	-16	10	6.31	0.76	2	20	17	-11	14	7.94	0.88	2
10	18	-11	13	5.78	0.44	0	20	18	-13	12	8.21	-0.44	-1
10	19	-10	13	5.75	1.44	3	20	19	-10	15	6.29	5.00	4
10	20	-6	16	5.40	10.00	11	20	20	-13	16	7.94	0.04	-2
10	21	-9	7	4.65	1.72	3	20	21	-9	15	6.19	5.28	5
10	22	-15	12	5.90	1.32	2	20	22	-10	14	7.00	3.12	2
10	23	-10	9	4.96	-2.00	-3	20	23	-16	15	7.14	6.00	7
10	24	-9	15	5.22	8.24	9	20	24	-15	11	6.73	-3.32	-3

Table A.2: Statistics of subject responses in the simulator semantic differential experiment. Signal numbers correspond to Table 2.2 and Scale numbers correspond to Table 2.1. Scales ranged from -16 to +16.

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
1	1	-15	12	6.67	-1.80	-3	11	1	-11	15	6.62	3.12	4
1	2	-13	13	6.62	0.20	2	11	2	-13	15	8.47	0.24	-1
1	3	-4	15	5.18	5.96	5	11	3	-16	12	6.68	-4.48	-3
1	4	-13	16	7.37	3.20	4	11	4	-14	11	6.51	-0.40	-3
1	5	-12	10	5.67	-0.88	-1	11	5	-7	15	6.36	3.28	5
1	6	-7	10	4.81	1.68	2	11	6	-14	10	6.86	-1.72	-1
1	7	-11	13	5.29	-2.48	-3	11	7	-10	13	6.15	3.16	4
1	8	-16	5	5.37	-4.36	-4	11	8	-10	10	5.95	1.64	3
1	9	-15	5	6.11	-6.28	-6	11	9	-10	15	6.90	5.84	8
1	10	-13	14	6.16	-3.44	-4	11	10	-13	13	5.92	4.04	6
1	11	-16	15	7.47	-6.52	-8	11	11	-12	16	7.96	3.04	3
1	12	-13	13	5.92	-0.60	-1	11	12	-11	14	7.26	0.16	-1
1	13	-2	15	4.17	7.80	9	11	13	-14	4	4.79	-5.72	-6
1	14	-4	16	5.64	3.56	3	11	14	-16	12	6.83	-2.68	-4
1	15	-9	16	5.69	7.12	7	11	15	-16	14	8.14	-6.64	-10
1	16	-16	4	5.04	-5.98	-6.25	11	16	-16	16	7.88	1.77	1.5
1	17	-15	11	6.42	-2.00	-3	11	17	-10	15	6.47	0.64	1
1	18	-15	13	7.36	-1.56	-3	11	18	-16	15	8.33	-0.84	-1
1	19	-6	10	5.10	1.68	2	11	19	-12	13	6.43	-2.04	-4
1	20	-3	16	4.55	11.20	12	11	20	-16	13	7.53	-7.68	-9
1	21	-7	14	5.14	0.96	1	11	21	-11	11	6.15	0.28	-2
1	22	-9	14	6.18	2.16	2	11	22	-10	16	6.21	0.32	1
1	23	-14	9	5.48	-4.04	-4	11	23	-7	14	6.42	4.68	6
1	24	-7	14	4.61	6.76	7	11	24	-13	10	5.89	-5.40	-7
2	1	-15	6	4.67	-6.20	-7	12	1	-12	6	4.62	-3.20	-4
2	2	-12	12	6.69	-2.12	-2	12	2	-15	16	7.92	1.40	2
2	3	-12	14	7.46	0.44	1	12	3	-6	15	6.13	4.60	3
2	4	-12	12	5.95	-0.80	-1	12	4	-12	13	6.49	2.52	4
2	5	-12	9	4.52	-5.28	-5	12	5	-16	9	6.10	-1.52	-2
2	6	-12	12	6.51	-3.52	-5	12	6	-10	12	5.81	3.16	4
2	7	-13	6	5.13	-5.60	-6	12	7	-14	13	6.71	-1.28	-2
2	8	-13	1	4.05	-7.64	-7	12	8	-12	10	5.87	-2.08	-2
2	9	-14	13	6.97	-6.88	-9	12	9	-15	12	7.20	-4.80	-5
2	10	-14	14	5.70	-4.72	-5	12	10	-14	15	8.58	-1.64	-2
2	11	-16	5	5.24	-4.56	-3	12	11	-15	10	8.66	-3.33	-6
2	12	-14	16	7.69	-2.00	-2	12	12	-12	15	6.98	3.12	5
2	13	-10	16	5.72	9.72	11	12	13	-6	15	5.64	7.48	8
2	14	-12	15	7.28	4.88	4	12	14	-9	15	6.61	4.96	5
2	15	-3	16	5.39	10.04	12	12	15	-8	16	7.12	7.80	10

Table A.2 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
2	16	-16	12	7.39	-2.75	-4	12	16	-16	16	8.59	0.06	2.25
2	17	-15	13	7.87	-2.32	-3	12	17	-15	13	8.13	1.76	4
2	18	-14	12	7.90	-0.76	1	12	18	-12	16	7.93	4.40	6
2	19	-14	11	5.61	-4.60	-5	12	19	-11	12	6.08	2.04	2
2	20	-9	16	5.94	9.40	11	12	20	-8	16	6.64	7.24	9
2	21	-15	10	7.47	-2.08	-3	12	21	-10	9	5.69	0.76	1
2	22	-13	7	5.61	-3.28	-4	12	22	-12	13	6.59	2.84	3
2	23	-14	12	6.04	-5.80	-7	12	23	-15	6	5.75	-4.32	-4
2	24	-10	14	5.82	7.36	9	12	24	-2	16	4.47	8.24	9
3	1	-16	11	5.33	-6.84	-6	13	1	-11	12	5.72	-0.56	1
3	2	-15	12	6.61	-3.32	-4	13	2	-12	16	8.63	3.60	5
3	3	-16	14	7.63	-0.80	1	13	3	-5	13	5.15	4.88	4
3	4	-14	5	5.79	-5.32	-6	13	4	-11	12	6.21	3.04	3
3	5	-14	6	5.15	-7.32	-9	13	5	-12	10	4.96	0.04	1
3	6	-12	14	8.33	-3.20	-5	13	6	-12	10	5.83	0.88	1
3	7	-15	3	4.76	-8.12	-10	13	7	-13	14	6.13	-1.20	-3
3	8	-15	0	4.22	-8.96	-10	13	8	-10	13	6.81	-1.32	-2
3	9	-16	12	6.69	-7.20	-7	13	9	-14	12	6.58	-3.44	-5
3	10	-16	13	7.91	-6.04	-7	13	10	-13	5	4.71	-3.40	-4
3	11	-15	16	6.75	-5.92	-7	13	11	-12	9	5.37	-1.00	-0.5
3	12	-14	6	5.58	-4.20	-3	13	12	-10	16	5.47	1.84	1
3	13	-7	14	5.09	7.24	9	13	13	-6	16	5.07	8.56	10
3	14	-14	16	6.75	3.80	4	13	14	-5	15	5.91	6.17	7
3	15	-15	15	7.04	4.80	6	13	15	-3	16	4.98	10.48	12
3	16	-15	12	7.56	-3.23	-3.5	13	16	-15	13	5.74	0.29	1
3	17	-16	12	7.19	-3.68	-5	13	17	-16	9	5.44	0.16	1
3	18	-13	14	8.37	-1.68	-5	13	18	-8	16	5.58	4.44	4
3	19	-15	4	5.85	-5.00	-5	13	19	-7	11	5.19	3.00	3
3	20	-8	15	5.96	7.52	8	13	20	-3	16	5.07	10.64	12
3	21	-16	12	7.08	-4.16	-4	13	21	-6	10	4.68	1.56	3
3	22	-15	11	6.52	-4.32	-4	13	22	-11	13	4.96	3.48	4
3	23	-16	2	5.43	-7.76	-7	13	23	-14	6	5.54	-1.76	-2
3	24	-12	14	5.62	7.28	7	13	24	-3	14	4.81	8.16	9
4	1	-13	13	6.78	-3.12	-3	14	1	-12	9	6.42	0.52	1
4	2	-15	12	7.30	0.44	2	14	2	-13	12	7.26	-2.00	-4
4	3	-11	14	6.19	2.12	3	14	3	-16	3	5.34	-5.60	-5
4	4	-13	12	7.49	0.16	-1	14	4	-15	10	6.17	-3.60	-3
4	5	-13	7	5.02	-2.80	-3	14	5	-9	14	5.54	0.32	-2
4	6	-10	9	4.81	-2.48	-3	14	6	-12	9	5.51	-2.20	-2
4	7	-15	14	5.86	-2.96	-4	14	7	-13	12	6.53	1.20	2
4	8	-11	3	3.37	-6.20	-6	14	8	-9	7	4.36	1.04	1
4	9	-15	9	5.72	-6.20	-5	14	9	-9	15	6.58	4.36	4

Table A.2 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
4	10	-13	12	6.73	-4.12	-5	14	10	-12	13	5.97	2.16	3
4	11	-16	13	7.81	-3.17	-4.5	14	11	-15	13	6.55	-1.36	-1
4	12	-14	15	8.58	0.88	2	14	12	-12	3	4.69	-3.96	-4
4	13	-7	15	5.20	9.20	11	14	13	-13	1	3.62	-7.04	-6
4	14	-11	14	6.24	5.20	6	14	14	-16	4	5.10	-5.72	-5
4	15	-2	16	5.55	9.44	10	14	15	-16	3	5.28	-9.00	-11
4	16	-10	13	6.37	1.56	3	14	16	-16	12	6.76	-2.25	-1
4	17	-14	16	7.90	-0.60	-2	14	17	-13	16	7.04	-2.28	-2
4	18	-12	16	8.56	3.68	4	14	18	-16	13	7.77	-3.64	-5
4	19	-15	12	7.14	-1.48	-2	14	19	-11	13	5.65	-2.32	-3
4	20	-11	16	7.06	8.96	11	14	20	-16	6	5.61	-8.60	-10
4	21	-15	12	6.43	-3.16	-4	14	21	-14	11	6.35	-1.76	-4
4	22	-13	11	6.27	-2.56	-2	14	22	-14	6	4.92	-2.96	-2
4	23	-15	12	6.33	-5.04	-6	14	23	-8	11	5.33	2.20	3
4	24	-14	15	7.93	6.16	9	14	24	-14	2	4.27	-7.60	-7
5	1	-13	9	5.30	-1.88	-3	15	1	-13	8	5.66	0.80	2
5	2	-16	14	6.80	1.64	2	15	2	-15	12	6.51	-2.88	-3
5	3	-9	15	6.46	2.96	3	15	3	-15	2	5.18	-6.92	-9
5	4	-10	12	5.93	2.04	2	15	4	-14	11	6.54	-4.04	-5
5	5	-10	9	5.36	-1.00	-2	15	5	-10	14	6.15	-0.04	-1
5	6	-7	9	4.89	0.08	1	15	6	-14	11	5.52	-2.76	-4
5	7	-13	14	7.08	-1.40	-3	15	7	-14	14	6.39	0.20	-1
5	8	-12	6	5.60	-3.68	-4	15	8	-12	14	5.82	1.16	3
5	9	-12	6	5.69	-4.48	-4	15	9	-11	15	7.16	3.32	4
5	10	-15	12	6.69	-4.72	-6	15	10	-12	12	6.76	1.36	2
5	11	-14	12	6.28	-2.68	-2	15	11	-15	12	6.61	1.00	2
5	12	-11	13	5.87	1.92	3	15	12	-13	6	5.66	-4.00	-4
5	13	-11	12	5.11	7.28	10	15	13	-14	5	4.79	-8.08	-9
5	14	-16	15	7.24	3.16	3	15	14	-14	9	6.18	-4.56	-5
5	15	-4	15	4.70	9.60	11	15	15	-16	10	6.55	-8.36	-10
5	16	-15	16	6.46	-0.56	-0.5	15	16	-16	9	5.92	-1.33	-2
5	17	-14	13	6.26	0.44	1	15	17	-14	13	6.68	-2.84	-4
5	18	-12	14	7.09	4.28	6	15	18	-16	11	7.23	-4.56	-5
5	19	-16	11	7.35	-0.76	0	15	19	-11	12	5.68	-2.28	-2
5	20	-14	16	6.65	9.00	10	15	20	-16	1	5.34	-9.68	-11
5	21	-12	10	6.50	0.12	2	15	21	-10	13	6.68	-0.96	-3
5	22	-12	13	5.86	1.44	2	15	22	-12	11	5.15	-3.36	-4
5	23	-12	10	6.63	-4.16	-5	15	23	-10	11	5.92	0.96	1
5	24	-9	14	5.74	6.64	9	15	24	-14	1	4.44	-8.36	-10
6	1	-14	7	6.63	-2.64	-3	16	1	-16	12	7.12	-0.68	-1
6	2	-11	12	6.72	0.96	3	16	2	-12	16	7.46	3.24	4
6	3	-7	14	5.35	2.44	3	16	3	-5	16	5.00	6.96	7

Table A.2 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
6	4	-12	12	6.92	0.24	2	16	4	-13	12	6.24	3.28	4
6	5	-15	11	5.82	-2.80	-3	16	5	-14	14	7.52	-0.20	1
6	6	-13	12	6.38	-0.60	-3	16	6	-10	11	5.30	1.76	2
6	7	-15	9	6.10	-3.28	-3	16	7	-7	13	5.09	-0.44	-1
6	8	-10	9	5.18	-3.60	-5	16	8	-12	10	6.05	-0.72	1
6	9	-16	1	4.45	-6.32	-6	16	9	-15	11	6.99	-3.64	-4
6	10	-15	12	7.03	-2.84	-4	16	10	-12	11	6.73	-1.24	-3
6	11	-9	15	6.20	7.40	9	16	11	-16	12	8.88	-3.64	-5
6	12	-13	15	6.07	7.32	8	16	12	-16	15	6.96	1.88	2
6	13	-12	16	5.64	10.44	11	16	13	-5	15	4.77	9.00	10
6	14	-1	16	3.82	9.84	11	16	14	-14	14	6.93	4.72	7
6	15	-12	16	5.52	11.76	13	16	15	-12	16	6.76	9.28	11
6	16	-6	15	6.87	7.94	10.75	16	16	-13	16	7.05	-0.80	1
6	17	-13	15	7.26	6.16	6	16	17	-14	16	8.19	1.08	2
6	18	-9	16	5.76	8.76	10	16	18	-13	16	9.25	2.24	4
6	19	-13	14	6.80	0.16	2	16	19	-15	12	6.42	2.12	3
6	20	-12	16	7.30	1.88	2	16	20	4	16	3.57	11.56	13
6	21	-13	9	6.13	-1.48	-2	16	21	-12	14	6.82	1.24	1
6	22	-14	12	7.47	-0.40	-1	16	22	-13	13	5.85	3.56	3
6	23	-13	11	5.90	-3.84	-4	16	23	-12	10	6.10	-1.44	-2
6	24	-7	15	5.71	9.44	10	16	24	-4	16	5.53	9.24	11
7	1	-16	11	7.16	-5.84	-6	17	1	-13	14	7.32	4.80	7
7	2	-12	16	7.86	-2.20	-3	17	2	-12	14	6.89	3.96	4
7	3	-12	16	7.99	-1.84	-5	17	3	-10	13	6.33	1.24	2
7	4	-15	14	8.93	-0.40	-1	17	4	-13	12	6.95	0.80	1
7	5	-15	14	7.70	-3.33	-5	17	5	-14	14	6.60	2.40	3
7	6	-12	11	6.45	-2.52	-2	17	6	-11	14	5.95	2.48	2
7	7	-15	11	6.82	-5.48	-6	17	7	-9	14	6.28	2.80	3
7	8	-12	12	5.79	-5.48	-7	17	8	-10	15	5.88	2.12	2
7	9	-16	11	7.43	-6.60	-9	17	9	-13	15	7.29	2.60	2
7	10	-16	7	6.94	-4.88	-6	17	10	-16	14	8.44	1.40	3
7	11	-16	12	9.19	-0.44	1	17	11	-15	12	6.36	-8.52	-11
7	12	-6	13	5.66	5.44	4	17	12	-16	15	6.75	-8.28	-10
7	13	-11	14	5.96	8.64	11	17	13	-15	10	7.46	-6.16	-9
7	14	-11	14	5.49	8.32	10	17	14	-16	8	5.23	-8.96	-10
7	15	-12	16	6.02	10.32	12	17	15	-16	-6	2.94	-12.96	-14
7	16	-15	16	8.88	2.85	4	17	16	-16	0	4.85	-9.79	-11
7	17	-12	14	7.05	2.80	3	17	17	-15	6	5.96	-7.80	-9
7	18	-10	16	6.74	5.92	6	17	18	-16	10	5.80	-10.40	-12
7	19	-15	11	7.42	-2.64	-2	17	19	-13	15	6.46	0.92	0
7	20	-12	16	7.52	8.40	11	17	20	-12	16	6.74	7.36	7
7	21	-14	13	7.22	-4.48	-6	17	21	-9	15	6.44	3.88	4

Table A.2 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
7	22	-14	13	8.57	-2.20	-4	17	22	-12	15	6.57	0.68	1
7	23	-15	10	7.90	-6.12	-9	17	23	-15	14	8.32	2.20	3
7	24	-10	15	6.77	8.48	10	17	24	-16	9	6.79	-8.60	-12
8	1	-11	10	6.36	-1.20	-2	18	1	-13	15	7.98	2.92	4
8	2	-12	11	6.88	0.68	2	18	2	-13	14	6.92	3.56	5
8	3	-6	15	5.87	4.68	5	18	3	-6	15	6.78	5.08	5
8	4	-10	12	5.77	2.80	3	18	4	-11	14	7.30	4.12	5
8	5	-12	7	4.95	-0.08	1	18	5	-4	12	3.65	5.48	5
8	6	-10	13	5.44	2.12	2	18	6	-16	15	7.28	4.88	4
8	7	-11	12	6.74	-1.20	-2	18	7	-7	15	5.89	4.68	4
8	8	-10	10	5.25	-1.24	-1	18	8	-6	15	5.81	3.92	4
8	9	-13	11	5.94	-2.80	-3	18	9	-13	16	7.41	2.96	3
8	10	-15	5	5.60	-3.08	-3	18	10	-9	13	5.79	2.84	3
8	11	-13	14	7.96	1.00	1	18	11	-15	14	6.94	-6.96	-9
8	12	-16	13	7.15	2.40	3	18	12	-16	12	7.50	-6.48	-9
8	13	1	16	4.46	8.48	10	18	13	-16	9	8.24	-4.80	-7
8	14	-16	15	7.15	5.28	5	18	14	-16	16	7.90	-6.48	-9
8	15	-4	16	5.42	9.80	11	18	15	-16	16	8.23	-9.36	-13
8	16	-12	14	5.96	2.79	3	18	16	-16	4	5.81	-7.94	-9
8	17	-14	15	6.22	2.56	3	18	17	-16	10	7.39	-6.44	-8
8	18	-8	16	5.86	4.04	4	18	18	-16	12	8.26	-7.20	-10
8	19	-10	14	5.41	2.96	3	18	19	-16	12	6.71	1.68	2
8	20	3	16	4.32	10.24	11	18	20	-13	16	7.61	6.76	10
8	21	-9	13	6.02	2.20	3	18	21	-4	16	4.55	5.08	4
8	22	-7	13	5.05	3.96	4	18	22	-10	12	6.05	3.28	4
8	23	-14	6	6.05	-3.12	-5	18	23	-14	13	7.73	3.16	5
8	24	-1	15	4.04	7.92	9	18	24	-16	10	8.43	-5.76	-9
9	1	-15	15	8.74	0.32	-1	19	1	-8	15	6.13	5.24	6
9	2	-15	15	9.31	1.16	3	19	2	-12	14	6.61	3.84	4
9	3	-15	15	9.52	3.12	6	19	3	-14	16	6.22	6.08	6
9	4	-15	13	8.46	1.44	3	19	4	-10	14	7.12	4.40	4
9	5	-12	13	7.85	2.12	3	19	5	-10	14	6.00	5.80	6
9	6	-7	15	6.50	3.52	4	19	6	-11	15	6.35	7.40	9
9	7	-15	12	8.39	0.32	2	19	7	-7	16	6.50	5.88	7
9	8	-15	12	7.66	-0.36	-1	19	8	-7	15	6.59	5.92	6
9	9	-15	11	7.19	-2.68	-4	19	9	-13	16	8.67	2.76	4
9	10	-14	11	7.58	-1.52	-2	19	10	-11	15	6.57	4.28	5
9	11	-14	15	9.03	0.64	1	19	11	-12	14	7.46	-1.88	-3
9	12	-14	16	7.32	0.72	0	19	12	-16	16	8.07	-1.64	0
9	13	-13	14	7.57	6.12	7	19	13	-14	12	7.15	-3.36	-4
9	14	-13	16	7.94	5.75	6.5	19	14	-16	16	7.93	-3.32	-2
9	15	-10	16	6.95	7.80	10	19	15	-16	13	6.46	-8.24	-10

Table A.2 – continued from previous page

Scale	Signal	Min	Max	Stdev	Mean	Median	Scale	Signal	Min	Max	Stdev	Mean	Median
9	16	-16	12	8.65	-0.56	2	19	16	-15	10	8.20	-2.77	-2
9	17	-14	15	8.27	1.72	3	19	17	-16	13	7.83	-2.24	-3
9	18	-15	15	9.47	1.68	2	19	18	-15	12	7.69	-2.04	-2
9	19	-9	16	6.95	4.52	5	19	19	-4	16	4.51	6.04	5.5
9	20	-16	16	10.44	-5.68	-9	19	20	-13	15	8.08	5.68	6
9	21	-14	15	8.47	1.60	3	19	21	-10	14	5.69	7.44	8
9	22	-10	14	6.11	5.04	5	19	22	-11	13	5.87	5.56	6
9	23	-16	16	8.49	-0.92	-3	19	23	-6	15	6.07	6.72	7
9	24	-7	16	5.95	7.40	9	19	24	-15	12	7.49	-3.28	-3
10	1	-10	10	5.92	-0.48	1	20	1	-16	13	7.04	5.12	6
10	2	-12	15	7.12	3.96	5	20	2	-11	14	6.54	3.12	4
10	3	-14	15	7.01	5.00	6	20	3	-9	15	5.91	5.00	4
10	4	-12	15	7.40	3.00	3	20	4	-10	15	6.41	4.56	5
10	5	-12	12	5.72	1.20	1	20	5	-4	15	5.61	7.68	9
10	6	-6	13	4.93	1.56	3	20	6	-2	14	4.14	6.68	7
10	7	-14	13	7.22	-1.16	1	20	7	-6	15	5.30	6.04	6
10	8	-9	9	4.88	-0.56	-1	20	8	-5	14	5.87	7.20	9
10	9	-15	6	6.04	-3.44	-3	20	9	-12	16	8.33	4.36	6
10	10	-14	6	4.87	-3.32	-3	20	10	-12	16	6.90	6.52	7
10	11	-15	12	6.35	-1.68	-3	20	11	-15	16	9.44	-0.16	3
10	12	-16	7	6.03	-1.32	0	20	12	-13	14	6.05	0.68	2
10	13	-10	13	5.35	5.20	6	20	13	-15	11	7.08	-5.68	-6
10	14	-5	13	3.96	4.33	4	20	14	-14	15	7.89	-2.52	-4
10	15	-11	16	6.58	7.64	9	20	15	-15	11	6.96	-6.48	-9
10	16	-15	9	6.20	-1.31	1	20	16	-12	15	7.83	-2.08	-4
10	17	-16	10	6.32	0.76	2	20	17	-11	14	7.94	0.88	2
10	18	-11	13	5.79	0.44	0	20	18	-13	12	8.21	-0.44	-1
10	19	-10	13	5.76	1.44	3	20	19	-10	15	6.29	5.00	4
10	20	-6	16	5.41	10.00	11	20	20	-13	16	7.94	0.04	-2
10	21	-9	7	4.65	1.72	3	20	21	-9	15	6.19	5.28	5
10	22	-15	12	5.90	1.32	2	20	22	-10	14	7.00	3.12	2
10	23	-10	9	4.97	-2.00	-3	20	23	-16	15	7.14	6.00	7
10	24	-9	15	5.22	8.24	9	20	24	-15	11	6.73	-3.32	-3

Calculated values for the metrics analyzed in Chapter 2 are in Tables A.3 and A.4.

Table A.3 Metric values for signals using in Earphone (1st number) and Simulator (2nd number) semantic differential experiments.

Signals	PL	Lamax	Lcmax	SELA	SELC	Mm	Zm
1	72.58 81.44	82.1 61.68	94.6 82.61	57.01 67.47	81.14 88.26	9.73 18.57	11.69 17.79
2	82.26 85.93	93.21 66.18	101.8 88.37	68.27 70.15	88.02 92.12	25.27 33.10	26.95 31.79
3	83.71 87.01	95.26 67.54	101.52 91.24	69.11 71.63	90.44 96.08	30.73 37.26	29.76 35.70
4	83.91 86.79	94.70 66.17	102.25 91.98	68.67 70.67	92.28 96.78	27.68 33.77	27.36 33.20
5	84.40 87.77	91.83 66.52	103.96 90.48	69.61 70.87	91.33 95.17	23.58 30.82	30.09 34.32
6	87.46 90.73	92.74 69.72	105.23 92.04	72.97 73.98	93.59 97.77	28.93 38.28	37.68 44.42
7	77.05 82.39	87.83 62.00	100.09 85.95	63.17 65.42	85.94 89.33	14.56 22.56	17.75 22.99
8	75.84 81.29	86.22 60.09	99.65 88.58	61.01 64.69	87.7 93.73	12.88 19.62	16.30 20.69
9	72.24 79.61	85.59 57.46	98.83 84.59	57.65 63.91	83.97 90.12	9.51 16.75	13.22 17.86
10	76.03 81.83	88.72 59.95	99.91 87.17	61.16 64.34	85.71 92.60	12.55 20.24	16.48 21.75
11	64.68 83.67	83.51 55.34	94.32 77.07	51.30 65.96	73.48 91.56	7.17 23.90	7.94 26.15
12	79.54 83.91	90.63 56.60	100.16 78.82	65.22 67.88	85.94 93.57	19.38 26.29	22.01 27.00
13	78.47 79.45	93.02 63.54	92.27 64.58	68.84 71.25	68.13 71.07	35.99 33.78	28.58 27.60
14	71.23 76.60	88.66 55.87	90.24 59.49	58.55 62.42	61.79 67.65	21.69 23.66	16.34 18.18
15	77.44 79.27	92.67 56.71	92.78 62.47	63.57 65.72	68.24 72.70	36.42 37.21	25.00 25.35
16	69.53 77.87	86.22 55.94	96.26 65.98	56.56 60.63	78.09 79.54	10.22 19.05	11.21 16.36
17	75.76 80.76	89.22 56.96	99.27 73.10	62.58 64.77	84.11 85.73	14.69 27.16	17.53 27.04
18	79.54 83.23	90.99 57.55	101.03 77.01	66.10 68.22	87.64 89.27	21.40 24.04	22.72 22.27
19	84.26 88.53	95.03 71.04	102.67 85.74	72.92 75.98	86.91 88.94	28.70 37.16	35.26 40.09
20	92.50 90.60	98.18 76.96	97.21 77.51	82.90 84.82	82.90 84.92	60.10 55.82	53.62 49.92
21	87.46 90.73	92.74 69.72	105.23 92.04	72.97 73.98	93.59 97.77	28.93 38.28	37.68 44.42
22	84.26 77.33	95.03 55.97	102.67 90.81	72.92 62.59	86.91 97.96	28.70 12.22	35.26 11.82
23	72.58 81.44	82.10 61.68	94.6 82.61	57.01 67.47	81.14 88.26	9.73 18.57	11.69 17.79
24	78.47 77.15	93.02 55.96	92.27 90.78	68.84 62.56	68.13 97.93	35.99 12.26	28.58 11.60

Table A.4 Metrics used in models of two or more metrics in earphone (1st number) and simulator (2nd number) semantic differential experiments

Signals	MRT	ZRT	MdLmax	ZdLmax	Smax	Smean	Smin
1	0.10 0.09	0.12 -0.13	256.24 248.53	197.42 215.59	0.93 2.04	0.62 1.34	0.34 1.09
2	0.06 0.09	0.05 -0.10	1278.01 942.62	1265.42 776.33	1.35 2.20	0.53 1.14	0.35 0.80
3	0.05 0.09	0.07 -0.23	1583.01 1143.24	1915.29 1241.60	1.15 1.66	0.56 1.02	0.36 0.80
4	0.08 0.09	0.11 -0.24	1410.11 1054.17	1347.25 848.77	1.36 1.84	0.55 1.15	0.30 0.73
5	0.06 0.10	0.06 -0.12	1008.08 772.87	814.15 723.37	0.88 1.52	0.48 1.05	0.30 0.71
6	0.06 0.08	0.06 -0.12	1397.50 1148.96	1097.88 767.86	0.88 1.84	0.52 0.91	0.29 0.62
7	0.05 0.08	0.06 -0.10	648.41 495.45	660.88 574.38	1.18 2.04	0.53 1.27	0.32 0.93
8	0.06 0.09	0.08 -0.23	437.95 341.34	318.54 257.03	1.14 1.79	0.59 1.14	0.31 0.96
9	0.05 0.09	0.09 -0.13	339.21 251.21	324.43 271.29	1.26 2.07	0.60 1.43	0.33 1.08
10	0.06 0.08	0.10 -0.08	515.27 440.52	460.63 358.08	1.20 2.10	0.59 1.34	0.31 0.92
11	0.08 0.09	0.12 -0.10	325.08 558.59	236.40 526.61	0.59 2.39	0.23 1.37	0.10 0.85
12	0.08 0.08	0.09 -0.09	1071.12 631.84	945.86 642.22	0.59 2.15	0.22 1.25	0.02 0.81
13	0.13 0.10	0.11 -0.08	1897.51 1143.74	2030.27 1509.28	1.95 2.41	1.80 1.91	0.80 1.59
14	0.20 0.19	0.19 -0.20	991.03 497.57	995.49 548.88	1.43 2.49	1.19 1.90	0.87 1.36
15	0.07 0.09	0.03 -0.08	2251.97 1613.49	2287.91 1407.14	1.63 2.51	1.32 1.80	0.86 1.25
16	0.13 0.11	0.17 -0.14	460.57 354.58	418.08 294.26	0.27 2.49	0.19 1.93	0.11 1.20
17	0.17 0.11	0.17 -0.12	645.97 595.68	612.08 555.01	0.28 2.47	0.20 1.55	0.11 0.83
18	0.09 0.11	0.17 -0.10	795.14 530.44	756.51 449.60	0.29 2.50	0.20 1.56	0.10 0.92
19	0.05 0.22	0.05 -0.23	1204.83 1035.26	1188.67 1196.99	0.97 1.86	0.68 1.08	0.44 0.71
20	0.47 0.16	0.45 -0.18	759.41 1397.40	900.05 675.92	1.65 1.45	1.32 1.39	1.08 1.15
21	0.06 0.08	0.06 -0.12	1397.50 1148.96	1097.88 767.86	0.88 1.84	0.52 0.91	0.29 0.62
22	0.05 0.08	0.05 -0.30	1204.83 185.12	1188.67 118.54	0.97 2.39	0.68 1.87	0.44 1.51
23	0.10 0.09	0.12 -0.13	256.24 248.53	197.42 215.59	0.93 2.04	0.62 1.34	0.34 1.09
24	0.13 0.09	0.11 -0.36	1897.51 157.33	2030.27 155.80	1.95 2.47	1.80 1.91	0.80 1.51

Appendix B: Initial Startle Experiment

The test instructions for the experiment described in Chapter 3 are given to each subject are below.

In this experiment, you will complete an arithmetic and memory task on the computer. First, you will see a math problem on your screen. Solve the problem, then click "ok". Then a possible answer will be displayed, if it is correct click "yes", if not click "no". After this you will be given a letter to remember. This process will repeat itself until a screen will be displayed asking you to recall the letters in the order they were given. Please place the order number (e.g. 1 for the first letter, 2 for the second) by the letter you remember. Then the process will repeat itself with different letters. During the experiment, you will hear sounds through your earphones. Please ignore them and continue with the task.

The Letter correct and time per each letter block in the arithmetic and memory task are in Tables B.1 and B.2.

Table B.1 Letters correct in arithmetic and memory task for each subject.

	Signal Number											
Subject	1	2	3	4	5	6	7	8	9	10	11	12
1	3	6	5	6	4	6	4	4	6	6	6	5
2	3	5	5	6	6	6	6	4	4	5	4	6
3	4	3	6	6	5	5	6	5	4	4	4	3
4	6	6	6	5	3	5	6	6	4	4	4	4
5	4	4	6	4	4	5	6	4	5	6	6	3
6	3	6	4	6	3	6	5	5	5	4	4	5
7	4	6	6	6	5	3	6	5	6	4	6	5
8	5	4	6	6	6	4	5	5	3	6	6	4
9	3	6	4	6	5	6	4	5	6	5	5	4
10	6	4	5	6	6	4	6	4	5	3	5	4

Table B.2 Duration of each letter block for each subject. Signal playbacks are display in Figure 3.3.

	Signal Number									
Letter Block	1	2	3	4	5	6	7	8	9	10
1	0.0	61.7	140.0	78.7	102.3	110.3	76.5	131.9	78.4	93.3
2	0.0	60.5	102.2	68.1	86.2	100.8	65.8	100.2	76.8	76.5
3	0.0	55.1	109.0	59.8	68.1	116.9	45.0	96.9	56.0	61.0
4	0.0	73.6	120.3	64.5	76.2	102.4	58.6	63.7	56.9	52.8
5	0.0	54.7	94.6	65.5	74.0	85.2	46.1	68.6	61.8	83.8
6	0.0	62.4	108.8	57.2	77.3	89.8	46.6	61.0	75.8	51.0
7	0.0	59.3	103.0	60.4	66.3	77.3	50.8	60.5	47.6	48.6
8	0.0	61.1	105.5	59.3	71.5	94.3	44.6	84.1	44.9	45.9
9	0.0	59.0	86.1	60.0	55.2	69.5	44.6	56.0	51.5	51.3
10	0.0	52.2	115.2	65.0	53.5	75.2	51.5	53.5	47.2	46.8
11	0.0	57.6	85.0	64.8	61.9	82.4	46.9	57.4	48.5	46.2
12	0.0	59.6	98.5	51.0	60.2	95.9	45.0	84.2	47.5	49.7
13	0.0	58.9	103.0	55.8	70.2	71.9	51.2	73.9	43.5	49.1
14	0.0	45.6	94.6	58.8	65.7	65.4	46.1	100.9	51.9	49.3
15	0.0	65.3	94.7	50.9	74.5	79.3	46.2	59.3	49.0	46.0
16	0.0	50.3	97.5	54.9	58.0	64.6	43.2	56.3	40.9	42.0
17	0.0	47.2	108.0	54.6	72.0	95.9	47.6	62.4	51.8	56.6
18	0.0	45.7	95.5	54.3	59.4	65.2	49.5	50.7	49.2	43.5
19	0.0	59.4	91.8	59.2	70.7	57.0	54.6	61.6	56.5	63.4
20	0.0	53.6	87.8	49.9	58.1	59.9	52.4	60.7	55.0	49.5
21	0.0	59.1	85.0	50.5	58.6	69.7	40.8	64.7	47.1	45.0
22	0.0	50.2	123.4	60.4	54.5	68.5	56.0	71.1	45.3	50.5

Appendix C: Startle Repeatability Study

The instructions for this experiment are below:

In this experiment you will look at pictures while you are seated and using the chin-rest. During the experiment you may hear sounds through your earphone, please ignore them. Please keep you left hand relaxed on the desk and avoid movement. (after picture viewing) Now you will make judgments about the sounds you previously heard. In this test, you will hear pairs of sounds, one played after another. Please judge which of the two sounds is more startling. After making your selection, you will hear another pair and so on. Remember there are no right or wrong answers. We are interested in your opinion.

The probability matrices from the paired comparison experiments for day 1 and day 2 experiments are in Tables C.1 and C.2, respectively.

Table C.1 Probabilities of first sound (row) being judged more startling than second sound (column) on day 1.

	Sound 1	Sound 2	Sound 3	Sound 4	Sound 5
Sound 1	0.50	0.84	0.67	0.14	0.18
Sound 2	0.25	0.50	0.33	0.12	0.19
Sound 3	0.39	0.65	0.50	0.44	0.47
Sound 4	0.82	0.67	0.58	0.50	0.47
Sound 5	0.91	0.79	0.53	0.28	0.50

Table C.2 Probabilities of first sound (row) being judged more startling than second sound (column) on day 2.

	Sound 1	Sound 2	Sound 3	Sound 4	Sound 5
Sound 1	0.50	0.86	0.77	0.14	0.16
Sound 2	0.12	0.50	0.35	0.12	0.21
Sound 3	0.30	0.67	0.50	0.25	0.37
Sound 4	0.84	0.68	0.65	0.50	0.40
Sound 5	0.95	0.81	0.61	0.14	0.50

Appendix D: Startle Modeling Experiment

The instructions for this experiment are below:

The purpose of this experiment is to determine the characteristics of a set of transient sounds. In this experiment, you will here a variety of transient sounds from various sources. Upon hearing a sound, please mark on the scale provided how startling you would find it if you were outdoors in a park or garden and heard this sound intermittently throughout the day. Some of the sounds had noticeable background noise. Please consider only the main impulsive sound event when making on the scale. When finished marking all of the scales for one sound, click "next sound" and the next sound will play. Remember there are no right or wrong answers, we are interested in your opinion.

The mean startle ratings and metric values for the stimuli in this experiment are displayed in Table .

Table D.1: Mean ratings, standard deviation of the estimated mean and metric values for the startle modeling experiment stimuli. N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones) and Dur is duration (ms).

Signal	Startle	Stdev	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur
1	18.10	5.46	33.26	107.52	16.37	1.11	0.70	254.00
2	22.12	6.56	33.48	657.20	16.77	1.25	0.78	348.00
3	24.75	6.81	33.35	915.76	18.10	1.23	0.63	366.00
4	24.94	6.15	33.63	1395.69	17.65	1.24	0.73	350.00
5	26.55	7.36	33.33	2053.00	16.54	1.26	0.63	368.00
6	25.50	6.66	33.69	2818.86	19.76	1.25	0.68	396.00
7	48.42	6.72	32.95	183.22	15.30	2.34	1.45	219.00
8	50.80	7.28	33.06	176.76	15.33	2.33	1.50	223.00
9	44.75	6.52	32.94	672.20	15.85	2.16	1.36	319.00
10	50.98	6.87	33.31	700.84	16.19	2.43	1.50	323.00
11	46.24	6.81	32.92	963.50	17.03	2.28	1.45	362.00
12	48.70	6.58	33.10	970.11	17.19	2.33	1.55	365.00
13	49.25	7.35	33.08	1435.45	16.65	2.20	1.41	347.00
14	53.80	7.47	33.75	1448.94	17.06	2.46	1.56	357.00
15	48.17	6.54	32.93	2105.47	15.63	2.28	1.48	349.00
16	52.42	7.01	33.73	2099.02	16.05	2.50	1.64	353.00
17	58.90	7.10	33.00	2863.74	18.75	2.18	1.42	391.00

Table D.1 – continued from previous page

Signal	Startle	Stdev	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur
18	61.06	6.99	33.16	2854.35	18.82	2.26	1.54	390.00
19	28.61	5.86	32.77	146.13	15.29	1.39	1.14	201.00
20	50.68	7.13	32.84	179.66	15.27	2.04	1.31	200.00
21	30.21	5.57	33.30	663.29	15.99	1.36	1.11	318.00
22	41.64	6.81	32.89	676.96	15.81	1.83	1.26	318.00
23	26.08	5.95	32.93	945.21	17.00	1.36	1.07	359.00
24	35.96	6.36	32.81	934.82	16.94	1.75	1.25	357.00
25	31.69	6.13	33.18	1397.53	16.77	1.37	1.07	349.00
26	49.44	7.10	33.05	1432.17	16.64	1.92	1.26	346.00
27	31.80	6.35	32.85	2096.06	15.59	1.34	1.08	348.00
28	44.12	7.14	32.88	2104.45	15.61	1.95	1.36	349.00
29	57.51	7.07	33.12	2864.69	18.71	1.71	1.22	390.00
30	53.15	6.82	33.06	2869.24	18.66	1.85	1.27	389.00
31	16.63	5.85	25.38	93.94	13.55	1.12	0.74	233.00
32	19.57	6.14	25.80	356.56	14.14	1.11	0.78	309.00
33	25.80	7.06	25.24	576.65	15.52	1.17	0.64	357.00
34	22.36	6.60	25.70	903.61	14.48	1.21	0.76	315.00
35	26.76	7.05	25.30	1383.21	13.94	1.23	0.63	352.00
36	27.31	7.03	25.78	1950.90	16.18	1.30	0.76	372.00
37	35.49	6.40	25.05	135.50	12.62	2.21	1.41	182.00
38	36.88	6.57	25.14	385.34	13.11	2.16	1.40	287.00
39	51.85	7.87	26.63	436.42	14.17	2.98	1.64	312.00
40	35.38	7.32	24.67	602.92	14.34	2.27	1.53	342.00
41	38.82	7.47	25.03	596.98	14.55	2.42	1.60	348.00
42	37.45	7.03	24.94	929.78	13.51	2.18	1.39	312.00
43	44.60	7.35	25.62	955.18	13.89	2.34	1.52	319.00
44	55.07	8.23	26.79	996.49	14.53	2.94	1.68	330.00
45	38.70	8.08	25.00	1418.26	13.02	2.25	1.51	320.00
46	44.24	7.70	25.24	1402.21	13.23	2.33	1.56	325.00
47	47.17	7.58	25.13	1984.46	15.14	2.13	1.39	353.00
48	48.06	7.39	25.60	1987.82	15.52	2.39	1.52	357.00
49	62.21	7.89	26.56	2030.86	16.18	2.94	1.69	366.00
50	25.64	6.14	24.85	129.80	12.58	1.37	1.15	181.00
51	30.88	6.91	24.86	162.59	12.56	1.74	1.26	181.00
52	37.71	6.54	24.94	166.51	12.57	1.98	1.31	181.00
53	23.56	5.82	25.10	374.69	13.04	1.34	1.15	282.00
54	35.49	7.43	25.08	380.33	13.05	1.84	1.30	285.00
55	20.74	5.81	24.80	605.88	14.37	1.30	1.12	339.00
56	34.54	6.47	24.62	596.91	14.31	2.00	1.41	342.00
57	37.19	6.94	25.01	927.15	13.50	1.88	1.28	311.00
58	25.14	5.96	25.29	1429.00	13.11	1.31	1.11	320.00
59	32.15	6.20	24.73	1414.10	12.97	1.95	1.38	320.00

Table D.1 – continued from previous page

Signal	Startle	Stdev	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur
60	39.01	6.38	25.18	1986.96	15.08	1.64	1.23	351.00
61	15.51	5.45	17.79	142.48	11.23	1.15	0.74	222.00
62	18.29	6.25	17.76	370.74	11.59	1.12	0.70	275.00
63	20.39	6.44	17.38	904.78	11.00	1.20	0.72	284.00
64	20.21	6.42	17.74	1375.09	10.77	1.24	0.80	266.00
65	28.89	7.28	18.04	2030.14	11.90	1.34	0.86	302.00
66	30.46	5.71	17.15	195.20	10.25	2.11	1.42	174.00
67	33.12	6.59	17.11	391.03	10.59	2.12	1.44	225.00
68	34.25	6.80	17.79	427.15	10.95	2.37	1.52	247.00
69	28.76	5.86	16.91	560.27	10.79	2.21	1.44	268.00
70	37.24	7.52	18.57	621.07	11.85	2.92	1.62	293.00
71	32.81	7.27	17.16	939.42	10.19	2.14	1.51	264.00
72	32.43	7.09	17.37	947.48	10.41	2.19	1.54	271.00
73	34.23	6.81	18.08	971.91	10.73	2.36	1.59	279.00
74	35.10	6.93	17.27	1394.58	10.13	2.05	1.45	255.00
75	36.73	7.39	18.23	1432.25	10.50	2.24	1.52	265.00
76	45.27	7.67	19.04	1451.44	10.87	2.70	1.68	279.00
77	30.46	5.57	18.06	206.13	10.68	2.29	1.49	186.00
78	37.43	7.14	17.93	2043.29	11.22	2.05	1.44	287.00
79	41.30	7.82	18.21	2036.49	11.47	2.14	1.46	293.00
80	54.69	8.56	19.04	2080.63	12.00	2.65	1.62	306.00
81	22.96	7.14	17.52	536.79	11.80	1.25	0.73	294.00
82	17.67	4.78	16.97	188.20	10.13	1.33	1.16	151.00
83	22.70	6.18	17.10	184.22	10.16	1.64	1.28	164.00
84	18.83	5.06	16.99	406.17	10.52	1.32	1.19	218.00
85	24.46	5.49	16.99	372.17	10.47	1.63	1.27	218.00
86	19.01	5.12	17.02	569.28	10.77	1.32	1.18	266.00
87	25.39	6.03	16.86	562.13	10.73	1.71	1.29	265.00
88	24.70	5.70	16.96	924.52	10.24	1.67	1.32	263.00
89	25.10	5.60	17.22	1409.54	10.12	1.58	1.28	253.00
90	38.98	7.55	17.78	2023.20	11.09	1.76	1.30	286.00
91	17.95	5.87	9.75	220.91	9.47	1.08	0.78	301.00
92	17.10	6.79	9.74	316.05	8.96	1.11	0.82	253.00
93	21.75	6.35	9.62	543.41	9.46	1.18	0.80	256.00
94	15.36	5.71	9.76	117.09	8.56	1.10	0.85	227.00
95	15.20	6.05	10.00	996.63	8.21	1.32	1.00	183.00
96	18.57	6.03	10.24	1378.00	8.88	1.37	1.06	207.00
97	26.81	6.58	9.49	268.16	8.61	2.01	1.57	284.00
98	25.80	6.30	9.55	360.03	8.20	1.94	1.56	225.00
99	27.86	6.77	9.51	595.74	8.67	1.99	1.61	230.00
100	23.42	6.23	9.75	150.03	7.88	1.97	1.53	147.00
101	23.23	6.73	9.96	1015.75	7.74	1.77	1.34	156.00

Table D.1 – continued from previous page

Signal	Startle	Stdev	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur
102	31.10	7.15	10.25	1007.72	7.96	1.97	1.42	177.00
103	27.77	7.34	10.97	1055.80	8.33	2.30	1.52	199.00
104	25.48	6.36	10.13	1405.86	8.35	1.91	1.38	198.00
105	36.42	7.41	10.89	1455.34	9.03	2.41	1.55	229.00
106	16.18	4.65	9.31	271.06	8.34	1.28	1.21	271.00
107	21.79	5.65	9.51	256.54	8.44	1.55	1.36	278.00
108	25.79	6.28	9.49	264.46	8.55	1.76	1.46	281.00
109	16.98	4.95	9.18	367.68	8.09	1.29	1.21	223.00
110	20.24	5.88	9.46	364.26	8.14	1.55	1.34	225.00
111	25.56	6.08	9.50	354.38	8.18	1.77	1.44	226.00
112	19.19	5.92	9.24	569.53	8.53	1.28	1.19	227.00
113	24.05	5.90	9.52	578.19	8.57	1.57	1.38	226.00
114	26.77	6.56	9.43	569.62	8.51	1.73	1.48	226.00
115	14.21	4.52	9.31	146.50	7.86	1.31	1.23	136.00
116	19.06	5.80	9.49	154.45	7.66	1.52	1.32	114.00
117	18.13	5.19	9.56	152.23	7.90	1.74	1.42	140.00
118	18.08	5.53	9.94	1009.58	7.79	1.56	1.28	161.00
119	28.06	6.08	33.26	107.52	16.37	1.11	0.70	254.00
120	28.19	6.23	33.48	657.20	16.77	1.25	0.78	348.00
121	35.54	7.17	33.35	915.75	18.10	1.23	0.62	366.00
122	31.95	6.68	33.62	1395.68	17.64	1.23	0.72	350.00
123	34.27	7.01	33.33	2053.00	16.54	1.26	0.63	368.00
124	33.75	6.92	33.69	2818.86	19.76	1.25	0.68	396.00
125	20.14	5.80	25.38	93.94	13.55	1.12	0.74	233.00
126	21.48	6.09	25.80	356.56	14.14	1.11	0.78	309.00
127	26.95	6.91	25.24	576.65	15.52	1.17	0.64	357.00
128	27.5	6.44	25.70	903.61	14.48	1.21	0.76	315.00
129	27.49	6.75	25.30	1383.22	13.94	1.23	0.63	352.00
130	29.46	6.11	25.78	1950.90	16.18	1.30	0.76	372.00
131	14.44	4.65	17.79	142.48	11.23	1.15	0.74	222.00
132	22.33	6.21	17.76	370.74	11.59	1.12	0.70	275.00
133	22.01	5.67	17.38	904.78	11.00	1.20	0.72	284.00
134	20.64	5.89	17.74	1375.09	10.77	1.24	0.80	266.00
135	25.70	6.45	18.04	2030.14	11.90	1.34	0.86	302.00
136	15.77	4.65	9.75	220.91	9.47	1.08	0.78	301.00
137	14.25	4.48	9.74	316.05	8.96	1.11	0.82	253.00
138	16.65	6.12	9.62	543.41	9.46	1.18	0.80	256.00
139	13.31	4.52	9.76	117.08	8.56	1.10	0.85	227.00
140	13.86	5.34	10.00	996.63	8.21	1.32	0.99	183.00
141	17.33	5.28	10.23	1377.99	8.88	1.37	1.06	207.00
142	28.92	7.41	14.73	527.73	11.52	1.10	0.64	305.00
143	24.57	6.23	12.04	392.13	10.26	1.12	0.68	287.00

Table D.1 – continued from previous page

Signal	Startle	Stdev	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur
144	19.88	6.20	10.06	287.21	9.51	1.13	0.69	275.00

Appendix E: Verification Experiment

The instructions for this experiment (described in Chapter 6) are below:

The purpose of this experiment is to determine the characteristics of a set of transient sounds. In this experiment, you will here a variety of transient sounds from various sources. Upon hearing a sound, please mark on the scale provided how startling you would find it if you were outdoors in a park or garden and heard this sound intermittently throughout the day. Some of the sounds had noticeable background noise. Please consider only the main impulsive sound event when making on the scale. When finished marking all of the scales for one sound, click "next sound" and the next sound will play. Remember there are no right or wrong answers, we are interested in your opinion. (After startle section) In this section, upon hearing a sound please judge how annoying you would find the sound if you were outdoors in a part or garden and heard this sound intermittently throughout the day. (After annoyance section) In this section, upon hearing a sound please judge if the sound would be acceptable to you if you were outdoors in a part or garden and heard this sound intermittently throughout the day.

Startle ratings and corresponding metric values are in Table . Annoyance ratings, acceptability ratings and corresponding metric values are in Table .

Table E.1: Mean ratings, standard deviation of the estimated mean and metric values for the startle ratings in Chapter 6. N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones), Dur is duration (ms), PL is Steven's Mk7 PL (dB) and Z_{max} is maximum Zwicker loudness (sones).

Signal	Startle	StdevM	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur	PL	Z_{max}
1	18.88	6.58	33.23	2714.97	14.14	1.08	0.57	204.00	82.74	26.24
2	20.30	6.63	33.20	920.05	16.12	1.26	0.53	392.00	84.88	34.65
3	20.93	7.16	33.25	2056.67	14.40	1.27	0.51	357.00	84.93	33.50
4	16.50	6.58	25.11	578.28	13.55	1.24	0.55	376.00	78.77	21.05
5	16.68	6.86	25.26	1384.89	11.79	1.26	0.52	336.00	79.23	21.95
6	17.33	6.35	25.74	1944.92	14.17	1.24	0.63	442.00	79.80	20.88
7	13.77	5.34	17.59	367.02	9.42	0.88	0.52	344.00	74.94	15.27
8	13.18	4.96	17.39	907.23	8.82	1.25	0.58	306.00	72.39	12.82

Table E.1 – continued from previous page

Signal	Startle	StdevM	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur	PL	Z_{max}
9	15.83	6.00	17.99	2015.45	9.84	1.28	0.79	405.00	73.51	14.87
10	12.50	4.68	9.71	545.97	7.35	1.25	0.68	250.00	66.91	8.41
11	10.15	4.51	10.19	1366.48	6.78	1.32	1.06	357.00	68.02	10.38
12	25.47	8.01	35.12	2148.39	20.24	1.27	0.48	404.00	88.79	40.74
13	40.27	8.41	32.98	2734.36	13.90	2.34	1.96	193.00	77.93	29.71
14	33.95	7.88	33.02	1422.83	15.17	2.21	1.51	386.00	79.02	30.83
15	40.78	7.98	32.71	2088.55	13.94	2.34	1.59	355.00	77.48	27.52
16	27.22	6.95	25.40	368.34	11.77	2.11	1.48	420.00	74.26	22.21
17	38.52	8.09	26.89	995.53	12.94	2.95	1.87	375.00	80.48	27.01
18	32.03	7.12	25.39	1971.84	13.73	2.12	1.46	432.00	76.28	23.01
19	21.97	6.70	36.58	451.86	16.84	1.20	0.56	273.00	85.37	36.59
20	21.18	6.27	17.11	391.19	9.08	2.15	1.88	338.00	68.57	14.08
21	20.62	5.64	16.93	940.23	8.42	2.19	1.56	313.00	67.38	12.47
22	24.48	6.77	17.12	1399.19	8.17	2.10	1.48	317.00	69.57	14.41
23	24.17	7.73	35.12	2148.39	20.24	1.27	0.48	404.00	88.79	40.74
24	20.65	6.79	9.30	442.43	7.13	1.97	1.69	308.00	63.69	7.14
25	38.85	7.98	18.96	2082.76	10.31	2.67	1.69	402.00	77.26	20.19
26	17.07	4.86	10.11	1382.21	6.68	1.84	1.33	358.00	67.85	10.77
27	47.82	6.46	36.52	1771.63	26.27	1.90	1.68	437.00	78.95	30.70
28	31.28	6.33	17.75	847.04	12.88	1.75	1.61	383.00	68.66	13.81
29	40.40	7.71	26.75	1295.94	19.28	1.85	1.66	414.00	74.33	21.63
30	22.27	5.84	9.92	425.91	7.44	1.56	1.41	310.00	62.15	7.48
31	17.58	6.41	14.71	558.43	9.22	1.16	0.52	291.00	74.11	14.50
32	20.03	6.18	12.01	395.21	7.93	1.16	0.55	274.00	71.47	11.66
33	15.75	5.85	10.28	302.82	7.15	1.17	0.56	266.00	70.10	10.60
34	13.93	4.45	16.41	1320.77	8.92	1.71	1.59	161.00	68.81	13.62
35	26.40	6.57	24.69	2073.55	12.32	1.72	1.53	197.00	73.65	20.92
36	19.23	6.33	19.11	117.36	9.92	1.53	0.98	237.00	68.40	14.33
37	35.38	6.70	21.86	847.95	16.81	1.12	1.08	1064.00	72.15	16.87
38	51.33	7.64	37.08	2087.72	22.72	1.12	1.08	1063.00	77.80	26.91
39	49.15	6.97	36.55	1704.99	26.47	1.12	1.08	1065.00	79.39	30.80
40	11.17	5.20	9.86	260.69	7.48	1.12	1.09	394.20	70.20	9.82
41	22.10	7.31	21.30	967.53	13.25	1.12	1.07	635.38	80.34	23.23
42	64.98	8.03	32.55	2530.71	21.31	1.39	0.98	2530.00	80.79	26.47
43	56.00	8.65	23.80	1182.14	14.92	1.31	0.95	2547.00	75.94	18.73
44	22.85	6.40	37.95	1305.98	17.78	1.27	0.56	398.00	86.60	35.91
45	68.08	6.98	48.75	2211.16	36.11	1.15	0.57	951.00	92.36	55.10
46	39.17	7.24	16.57	252.24	14.50	1.20	1.04	616.00	69.47	14.00
47	56.00	7.80	20.31	285.89	17.54	1.37	1.19	3500.00	71.62	15.67
48	44.65	7.69	19.82	521.64	15.44	1.22	0.98	3110.00	72.00	14.74
49	29.25	7.19	33.48	663.41	14.53	1.31	1.09	443.00	78.76	31.90
50	29.65	6.63	33.19	1405.35	15.22	1.38	1.07	386.00	78.16	29.70

Table E.1 – continued from previous page

Signal	Startle	StdevM	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur	PL	Z_{max}
51	21.52	5.06	25.22	371.84	11.68	1.79	1.33	421.00	74.13	22.91
52	28.48	6.47	24.95	919.39	11.94	1.88	1.36	370.00	74.46	23.07
53	28.48	5.98	25.30	1976.44	13.70	1.62	1.23	430.00	75.92	24.13
54	18.82	5.95	16.95	1290.70	8.61	1.66	1.46	159.00	67.98	14.32
55	18.90	5.96	17.11	1402.91	8.16	1.61	1.29	316.00	68.80	14.89
56	25.25	6.29	17.83	2038.12	9.54	1.76	1.30	394.00	72.80	17.90
57	14.90	4.88	9.34	463.30	7.06	1.57	1.41	305.00	63.51	7.45
58	13.72	4.86	9.94	1011.59	5.96	1.58	1.24	337.00	66.09	9.82
59	53.55	7.58	37.10	1252.30	28.28	1.82	1.47	534.00	79.92	32.95

Table E.2: Mean annoyance ratings, standard deviation of the estimated mean, acceptability fractions, confidence intervals and associated metric values for the sounds in Chapter 6. Ann is Annoyance, Acc is acceptability, N_{max} is maximum of Moore's time-varying loudness (sones), dL_{max} is maximum loudness derivative (sones/s), S is sharpness (acum), LN_{max} is maximum of long-term loudness (sones), Dur is duration (ms), PL is Steven's Mk7 PL (dB) and Z_{max} is maximum Zwicker loudness (sones).

Signal	Ann	Stdev	Acc	CI	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur	PL	Z_{max}
1	20.73	6.58	0.90	0.05	33.23	2714.97	14.14	1.08	0.57	204.00	82.74	26.24
2	29.62	6.63	0.67	0.08	33.20	920.05	16.12	1.26	0.53	392.00	84.88	34.65
3	27.48	7.16	0.70	0.08	33.25	2056.67	14.40	1.27	0.51	357.00	84.93	33.50
4	22.05	6.58	0.90	0.05	25.11	578.28	13.55	1.24	0.55	376.00	78.77	21.05
5	26.62	6.86	0.73	0.08	25.26	1384.89	11.79	1.26	0.52	336.00	79.23	21.95
6	25.47	6.35	0.83	0.07	25.74	1944.92	14.17	1.24	0.63	442.00	79.80	20.88
7	18.65	5.34	0.85	0.06	17.59	367.02	9.42	0.88	0.52	344.00	74.94	15.27
8	21.00	4.96	0.80	0.07	17.39	907.23	8.82	1.25	0.58	306.00	72.39	12.82
9	23.78	6.00	0.80	0.07	17.99	2015.45	9.84	1.28	0.79	405.00	73.51	14.87
10	22.63	4.68	0.85	0.06	9.71	545.97	7.35	1.25	0.68	250.00	66.91	8.41
11	15.03	4.51	0.93	0.04	10.19	1366.48	6.78	1.32	1.06	357.00	68.02	10.38
12	32.15	8.01	0.68	0.08	35.12	2148.39	20.24	1.27	0.48	404.00	88.79	40.74
13	48.18	8.41	0.43	0.09	32.98	2734.36	13.90	2.34	1.96	193.00	77.93	29.71
14	48.42	7.88	0.23	0.07	33.02	1422.83	15.17	2.21	1.51	386.00	79.02	30.83
15	52.32	7.98	0.28	0.08	32.71	2088.55	13.94	2.34	1.59	355.00	77.48	27.52
16	37.87	6.95	0.58	0.09	25.40	368.34	11.77	2.11	1.48	420.00	74.26	22.21
17	51.43	8.09	0.33	0.08	26.89	995.53	12.94	2.95	1.87	375.00	80.48	27.01
18	40.27	7.12	0.43	0.09	25.39	1971.84	13.73	2.12	1.46	432.00	76.28	23.01
19	23.32	6.70	0.63	0.08	36.58	451.86	16.84	1.20	0.56	273.00	85.37	36.59
20	26.10	6.27	0.83	0.07	17.11	391.19	9.08	2.15	1.88	338.00	68.57	14.08
21	24.02	5.64	0.83	0.07	16.93	940.23	8.42	2.19	1.56	313.00	67.38	12.47
22	28.03	6.77	0.83	0.07	17.12	1399.19	8.17	2.10	1.48	317.00	69.57	14.41
23	37.20	7.73	0.60	0.09	35.12	2148.39	20.24	1.27	0.48	404.00	88.79	40.74
24	23.98	6.79	0.65	0.08	9.30	442.43	7.13	1.97	1.69	308.00	63.69	7.14
25	46.28	7.98	0.30	0.08	18.96	2082.76	10.31	2.67	1.69	402.00	77.26	20.19
26	21.78	4.86	0.83	0.07	10.11	1382.21	6.68	1.84	1.33	358.00	67.85	10.77
27	56.73	6.46	0.20	0.07	36.52	1771.63	26.27	1.90	1.68	437.00	78.95	30.70
28	39.15	6.33	0.65	0.08	17.75	847.04	12.88	1.75	1.61	383.00	68.66	13.81
29	47.68	7.71	0.43	0.09	26.75	1295.94	19.28	1.85	1.66	414.00	74.33	21.63
30	27.07	5.84	0.85	0.06	9.92	425.91	7.44	1.56	1.41	310.00	62.15	7.48
31	24.15	6.41	0.75	0.08	14.71	558.43	9.22	1.16	0.52	291.00	74.11	14.50
32	22.92	6.18	0.80	0.07	12.01	395.21	7.93	1.16	0.55	274.00	71.47	11.66

Table E.2 – continued from previous page

Signal	Ann	Stdev	Acc	CI	N_{max}	dL_{max}	LN_{max}	S_{max}	S_{mean}	Dur	PL	Z_{max}
33	17.62	5.85	0.80	0.07	10.28	302.82	7.15	1.17	0.56	266.00	70.10	10.60
34	21.93	4.45	0.83	0.07	16.41	1320.77	8.92	1.71	1.59	161.00	68.81	13.62
35	34.72	6.57	0.77	0.07	24.69	2073.55	12.32	1.72	1.53	197.00	73.65	20.92
36	19.22	6.33	0.80	0.07	19.11	117.36	9.92	1.53	0.98	237.00	68.40	14.33
37	39.75	6.70	0.53	0.09	21.86	847.95	16.81	1.12	1.08	1064.00	72.15	16.87
38	53.75	7.64	0.17	0.07	37.08	2087.72	22.72	1.12	1.08	1063.00	77.80	26.91
39	56.60	6.97	0.37	0.08	36.55	1704.99	26.47	1.12	1.08	1065.00	79.39	30.80
40	12.85	5.20	0.87	0.06	9.86	260.69	7.48	1.12	1.09	394.20	70.20	9.82
41	25.20	7.31	0.68	0.08	21.30	967.53	13.25	1.12	1.07	635.38	80.34	23.23
42	64.03	8.03	0.13	0.06	32.55	2530.71	21.31	1.39	0.98	2530.00	80.79	26.47
43	53.58	8.65	0.23	0.07	23.80	1182.14	14.92	1.31	0.95	2547.00	75.94	18.73
44	28.50	6.40	0.77	0.07	37.95	1305.98	17.78	1.27	0.56	398.00	86.60	35.91
45	67.02	6.98	0.10	0.05	48.75	2211.16	36.11	1.15	0.57	951.00	92.36	55.10
46	44.60	7.24	0.48	0.09	16.57	252.24	14.50	1.20	1.04	616.00	69.47	14.00
47	55.70	7.80	0.23	0.07	20.31	285.89	17.54	1.37	1.19	3500.00	71.62	15.67
48	46.82	7.69	0.30	0.08	19.82	521.64	15.44	1.22	0.98	3110.00	72.00	14.74
49	36.68	7.19	0.60	0.09	37.25	2712.78	20.63	1.08	0.54	289.00	85.73	38.01
50	36.12	6.63	0.53	0.09	35.33	2055.94	20.26	1.27	0.51	462.00	87.95	41.94
51	35.05	5.06	0.65	0.08	19.65	365.39	12.66	1.24	0.58	424.00	77.74	20.05
52	38.83	6.47	0.53	0.09	20.11	2015.19	12.77	1.28	0.70	483.00	76.00	19.10
53	40.40	5.98	0.53	0.09	33.48	663.41	14.53	1.31	1.09	443.00	78.76	31.90
54	24.30	5.95	0.82	0.07	33.19	1405.35	15.22	1.38	1.07	386.00	78.16	29.70
55	22.63	5.96	0.83	0.07	25.22	371.84	11.68	1.79	1.33	421.00	74.13	22.91
56	34.72	6.29	0.77	0.07	24.95	919.39	11.94	1.88	1.36	370.00	74.46	23.07
57	20.67	4.88	0.83	0.07	25.30	1976.44	13.70	1.62	1.23	430.00	75.92	24.13
58	20.43	4.86	0.90	0.05	16.95	1290.70	8.61	1.66	1.46	159.00	67.98	14.32
59	56.92	7.58	0.17	0.07	17.11	1402.91	8.16	1.61	1.29	316.00	68.80	14.89
60	33.60	6.41	0.65	0.08	17.83	2038.12	9.54	1.76	1.30	394.00	72.80	17.90
61	40.15	6.81	0.47	0.09	9.34	463.30	7.06	1.57	1.41	305.00	63.51	7.45
62	29.48	6.95	0.80	0.07	9.94	1011.59	5.96	1.58	1.24	337.00	66.09	9.82
63	31.82	6.11	0.80	0.07	37.10	1252.30	28.28	1.82	1.47	534.00	79.92	32.95

Appendix F: Miscellaneous Computer Programs

In this section, the source code of two computer programs that were used extensively in the research described in this document are listed. The first program is an implementation of Moore and Glasberg's time-varying loudness (2002). It is written in C++. The second series of program, written in the shell scripting language and matlab and utilizing the time-varying loudness program, were used to generate the boom-like noise signals used as stimuli in the experiments described in Chapter 5 and 6. The program creates amplitude modulated colored noise that is designed to have a loudness time-history close to that of the input.

F.1 Moore and Glasberg's Time-varying Loudness Program

This program calculating Moore and Glasberg's time-varying loudness for a given wave file or batch of wave files. It has been tested on windows systems (XP and later) and Unix/Linux systems.

To run a single sound:

```
MooreTV4.exe -wav $filename$ -calb $calbrationfactor$
```

Where the two arguments are the file name to be read (a WAV file) and the calibration factor. for the example:

```
MooreTV4.exe -wav boom.wav -calb 12.77
```

The output of the command is a .ML with the same name as the input wav file. It is an ascii file with each column corresponding to the instantaneous loudness, short-term loudness and long-term loudness. In figure F.1, the pressure time-history and the output of the program run with the above command are plotted. The signal is one of the high-pass filtered boom signals used in the experiment described in Chapter 5.

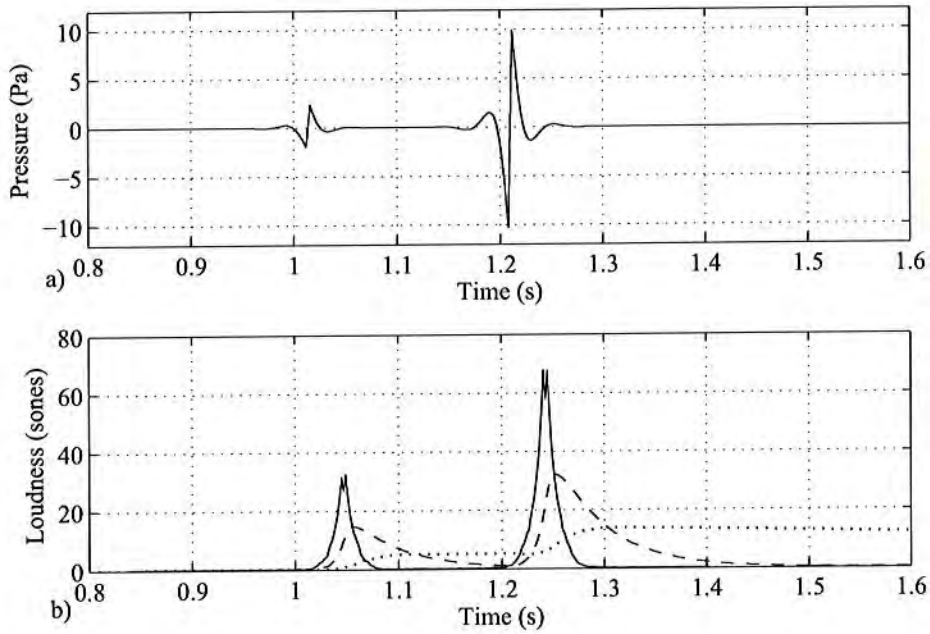

Figure F.1. a) Time-history and b) loudness time-history of a sonic boom. Solid line is instantaneous loudness, long dashes are short-term loudness and short dashes are long-term loudness.

```
/*
Glasberg and Moore's Time varying loudness program (V1.4)
written by Andrew Marshall

mixed radix fft uses FFTw3.0:
Matteo Frigo and Steven G. Johnson, "The Design and Implementation of FFTW3,"
Proceedings of the IEEE 93 (2), 216 231 (2005). Invited paper, Special Issue on Program
    Generation,
Optimization, and Platform Adaptation.

Based on the paper:
Glasberg and Moore, A Model of Loudness Applicable to Time-Varying Sounds,
J. Audio Eng. Soc., Vol. 50, No. 5, 2002 May

Work: unsensitive

Warning: Use at own risk, not guaranteed to produce accurate results.
Copyright 2009 PATRICIA DAVIES AND ANDREW MARSHALL, Ray W. Herrick Laboratories
and Purdue University. This program is distributed WITHOUT ANY WARRANTY; without
even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE

Dependencies: fftw3 library (www.fftw.org)

Changelog:
V1.4: added endian correction to fix cross-platform issues
V1.3: removed memory leak in wav-file reader
*/


#include <cstdlib>
#include <iostream>
#include <math.h>
#include <fstream.h>
#include <vector.h>
#include <string>
#include <time.h>
#include <inttypes.h>
#include <fftw3.h>

/*
TODO:
-more stress testings
*/
```

```
/*
design notes
-using half-complex form of FFT (faster than real-to-complex fft)
-using vectors for simplicity for latter sections where speed is less important
-main() is getting a bit unweildy=>split into different sections
-merged window function with multispectra calculation (eliminated several loops)
*/

//constants
const double pi=3.14592653589793238462;
//for get_S
const double c1=24.673;
const double c2=4.368;
const double c3=21.3655;

//=====prototypes==============================================================

//===Functions for operating modes
void SingleSound(char* Name,bool wav,double calb,bool ER2,char* Inpath, char* Outpath);//
    add in stuff for Eparams
void SameLoudness(char* Name,char* proc_Name, double M_value);

//===Functions for Moore's Loudness
//outer/inner ear filtering
//via time domain
void FIR_ear(double x[],double out[],double fs,int N,bool ER2);//time-domain convolution
//via frequency domain
void FIR_earF(double x[],double out[],double fs,int N,bool ER2,int delay);
//My radix-2 fft (used in FIR filter design for time domain convolution)
void FFT(double data[],int N,int isign,bool ps);

//function to calculate optimal fft size (alpha version)
int calc_FFT_size(int fs, int bound_t[]);
//calculates multispectrum (short-term spectrum)
void multi(double signal[],double out[],int center_i,int bound_f[],int bound_t[],int N,int
     fft_size,fftw_plan plan_fft,double* temp);
//calculates excitation pattern, give a short-term spectrum
void get_S(double post_FFT[],double intensity[],double post_FFT_freq[],int N,int start,
    double Estart, int Esize, double dE);
//transforms excitation pattern to intant loudness
double L_transform(double intensity[],int N,double dE, double G[],double A[],double alpha
    [],double Excitation_thrsq[]);
//generates loudness transform paramters for above function
```

```
void Setup_L_transform(double Estart, double dE,int N, double Excitation_thrsq[],double G
    [],double A[], double alpha[]);
//quadratic interpolation function
void interpolate(double x[],double y[],double xi[],int N,int P);
//linear interpolation function
void interpL(double x[],double y[],double xi[],int N,int P);
//temporal filter
vector<double> Tempor(vector<double> invec,double attack, double release);

//---misc helper functions---
double abs(double x){if (x < 0)return -x;else return x;}
//wavfile reader
bool wavin(char* Name,double calb,ifstream &InFile,int32_t &N,int32_t &fs,int32_t &Bits);
void load_samples(ifstream &inFile,double X[],int start,int Num,int &phold,double calb,
    int32_t N,int32_t fs,int32_t Bits,bool swap);
//overloaded function for asci files
void load_samples(ifstream &inFile,double X[],int start,int Num,int &phold,double calb,int
     N,int fs);
//endian swapping code (for cross-platform)
short ShortSwap(short s);
int LongSwap(int i);
//======Implementation=========================================================

int main(int argc, char *argv[])
{
    //check for arguments
    //test all
    bool batch=false;
    bool Same =false;
    int i,j;
    int N;
    bool wav=false;
    bool ER2=false;
    double calb = 1.0;
    char* Name="none       ";
    char* Inpath=" ";
    char* Outpath=" ";

    //command-line option parsing =>convert to function?
    if (argc >= 2)
    {
       for (i=1;i<argc;i++)
       {
           cout<<argv[i]<<" ";
```

```
if (!strcmp(argv[i],"-batch"))
{
       //case "-batch":
            cout<<"batch mode"<<endl;
            batch = true;
            i++;//exclude batchfile name
            Name = argv[i];

}else if (!strcmp(argv[i],"-Eparams"))//not implemented yet
{
       //case "-Eparams":
            cout<<"-Eparams"<<endl;
            i++;//allows a file for different E parameters
}else if (!strcmp(argv[i],"-SameLoudness"))
{
       //case "-SameLoudness":
            cout<<"-SameLoudness"<<endl;//not implemented yet
            Same = true;
            i++;
            double M_goal = atof(argv[i]);
}else if (!strcmp(argv[i],"-Inpath"))
{
      i++;
      Inpath = argv[i];
      }
else if (!strcmp(argv[i],"-Outpath"))
{     i++;
     Outpath = argv[i];
}
else if (!strcmp(argv[i],"-wav"))
{
     wav = true;
}
else if (!strcmp(argv[i],"-calb"))
{
     i++;
     calb = atof(argv[i]);
}
else if (!strcmp(argv[i],"-ER2"))
{
     ER2 =true;
}
else{
       //default:
```

```
                    cout<<"acsi-file"<<endl;
                    Name = argv[i];
            }
        }
    }
    //warn user of possible missed calibration
    if (wav == true && calb == 1.0)
    {
       cout<<"Warning: calibration factor for wav file is 1.0"<<endl;
       cout<<"did you remember to set the -calb option?";
    }

    if (!batch && !Same)
    {//single mode
       cout<<"processing single sound:"<<Name<<endl;
       SingleSound(Name,wav,calb,ER2,Inpath,Outpath);
    }else if (!Same)
    {
         ifstream batch_file;
         batch_file.open(Name);
         batch_file>>N>>wav>>calb>>ER2;//get number of files
         cout<<"Batch Mode:"<<Name<<" Files:"<<N<<" calb:"<<calb<< "ER?"<<ER2<<endl;
         for (i=0; i<N; i++)
         {   batch_file>>Name;
             cout<<"Sound:"<<Name<<endl;
             SingleSound(Name,wav,calb,ER2,Inpath,Outpath);
         }
         batch_file.close();
    }else
    {    //for loudness matching
         cout<<"Error: Not Implemented option";
    }
    cout<<"Finished Successfully";
    //clean up
    //system("PAUSE");
    return EXIT_SUCCESS;
}

//------------------------------------------------------------------------
//functions for operating modes
void SingleSound(char* Name,bool wav,double calb,bool ER2,char* Inpath, char* Outpath)//
    add in stuff for Eparams
{
    //check for arguments
```

```
//test all
ifstream inFile;
double * signal;
double * out;
bool fint=true;
int i;
int32_t fs;
int32_t N,Bits;
//char* Name="none";

//generate Loundness transform variables
double Estart = 1.8;
double dE = 0.1;
int Esize = 371;
double Excitation_thrsq[Esize];
double G[Esize];
double A[Esize];
double alpha[Esize];
Setup_L_transform(Estart,dE,Esize,Excitation_thrsq,G,A,alpha);


if (!wav)
{
 //prepare load_samples for asci files
 if (Name == "none")
    Name = "test.txt";
 inFile.open(Name);
 inFile>>N>>fs>>ER2;//get length, sampling rate and ER2 option
 cout<<N<<" "<<fs<<" "<<ER2<<endl;
 fint =true;
}
else
{
    //prepare load_samples for wavfiles
    fint=wavin(Name,calb,inFile,N,fs,Bits);
    cout<<N<<" "<<fs<<" "<<Bits<<endl;
    //out =new double[N+4096];
    //for (i=0;i<(N+4096);i++)
}

if (fint)
{
 cout<<fs<<endl;
 //standard time/frequency bounds for MooreTV
```

```
int bound_f[7]={15000,4050,2540,1250,500,80,10};
int bound_t[6]={2,4,8,16,32,64};
int bound_N = 6;

//calculate FFT size: now another function!
int fft_size;//=(fs*64)/1000;
fft_size = 2*calc_FFT_size(fs,bound_t);

//generate frequency vector
cout<<"fft_size="<<fft_size<<endl;
double freq_vector[fft_size];
for (i=0; i<fft_size; i++){
    freq_vector[i]=i*fs/fft_size;}
//generate frequency and time bounds
for (i=0; i<bound_N+1; i++)
{
    bound_f[i]=(fft_size*bound_f[i])/(fs)+1;
    cout<<bound_f[i]<<" ";
    if (i < bound_N){
       bound_t[i]=bound_t[i]*fs/1000;}
}
//correct top bound
if (bound_f[0] > (fft_size/2))
   {bound_f[0] = fft_size/2-1;
    cout<<"correcting:"<<bound_f[0]<<endl;}
//set time-step to 1ms
int ms=fs/1000;


//prep vars for calculation
signal = new double[4096];//signal
out = new double[8192];//likewise
for (i=0;i<8192;i++)//+fft_size/2;
{
    out[i] = 0.0;
    if (i < 4096)
       signal[i]=0.0;
}
double * intensity = new double[Esize];
vector<double> IL,SL,LL;

//prepare fftw library for multispectrum calculation
fftw_plan plan_fft;
double * fft_temp = new double[fft_size];
```

```
plan_fft = fftw_plan_r2r_1d(fft_size,fft_temp,fft_temp,FFTW_R2HC,FFTW_MEASURE);

int jj,hold,phold,delay,stt;
hold =0;
phold =0;
//load initial signal (zeros?)
jj=0;
delay = 0;//fft_size/2;
//endian test
unsigned char EndianTest[2]={1,0};
short test;
test = *(short *) EndianTest;
bool Endian;
if (test == 255)
  Endian = true;
else
  Endian = false;

ofstream tempO; //filtered signal temp file
tempO.open("TEMP.MTV");
//Read input signal in chunks and apply Outer Ear filter
while (phold < N+4096)
{
      if (wav)
         load_samples(inFile,signal,hold,4096,phold,calb,N,fs,Bits,Endian);
      else
         load_samples(inFile,signal,hold,4096,phold,calb,N,fs);
      FIR_earF(signal,out,fs,4096,ER2,0);
      for (jj=0;jj<4096;jj++)
      {
       tempO<<out[jj]<<" ";
       out[jj]=out[jj+4096];
       out[jj+4096]=0.0;
      }
      //phold = phold+4096;
}
cout<<"done filtering";
delete [] signal;
delete [] out;
inFile.close();
tempO.close();

ifstream tempI;
tempI.open("TEMP.MTV");
```

```
signal = new double[fft_size];
out = new double[bound_t[5]];
stt = 0;
int check=0;
for (i=0;i<phold-bound_t[5]/2;i=i+ms)
{
        for (jj=stt;jj<bound_t[5];jj++)
            tempI>>out[jj];

        //calculate multispectrum
        multi(out,signal,bound_t[5]/2,bound_f,bound_t,bound_N,fft_size,plan_fft,
            fft_temp);
        //calculate excitation pattern from multispectrum
        get_S(signal,intensity,freq_vector,bound_f[0],bound_f[6],Estart,Esize,dE);
        //calculate instantaneous loudness from excitation
        IL.push_back(2.0*L_transform(intensity,Esize,dE,G,A,alpha,Excitation_thrsq));
        stt = bound_t[5]-ms;
        for (jj=ms;jj<bound_t[5];jj++)
            out[jj-ms]=out[jj];
}
tempI.close();
//touch filter file, overwrite contents
tempO.open("TEMP.MTV");tempO<<"a";tempO.close();

//clean up
fftw_destroy_plan(plan_fft);
fftw_free(fft_temp);
delete [] intensity;
delete [] signal;
delete [] out;

//Apply temporal filtering
SL=Tempor(IL,0.045,0.02);//short term
LL=Tempor(SL,0.01,0.0005);//long term

//output instant, short-term and long-term loudness to file
ofstream myfile2;
Name[strlen(Name)-4]='x';
Name[strlen(Name)-3]='.';
Name[strlen(Name)-2]='M';
Name[strlen(Name)-1]='L';
myfile2.open(Name);
cout<<"starting write";
cout<<IL.size()<<" "<<SL.size()<<" "<<LL.size();
```

```
for (i=0;i<IL.size();i++)
    myfile2<<IL[i]<<" "<<SL[i]<<" "<<LL[i]<<endl;
myfile2.close();

}
else
{
//error message
cout<<Name<<" FAILED! File did not load/is of wrong format"<<endl;
}

}



//Moore's loudness functions
//---------------------------------------------------------------------------

//Constructions and applies outer/middle ear FIR filter in frequency domain
//
void FIR_earF(double x[],double out[],double fs,int N,bool ER2,int delay)
{/*
 x-input signal
  out-ouput array (adjusted by filter length)
  fs-sampling rate
  N-signal length (x)
  ER2-if outer-ear filter is to be used
 */
   //generate FIR filter
   double filter[8192];
   double temp_in[8192];
   fftw_complex * temp_out;

   int i,j;
   int Q;
   double Y;

   //insert filter gen code here or code to load in filter?
   double filter_f
       [40]={0,20,25,31.5,40,50,63,80,100,125,160,200,250,315,400,500,630,750,800,1000
       ,1250,1500,1600,2000,2500,3000,3150,4000,5000,6000,6300,8000,
       9000,10000,11200,12500,14000,15000,16000,20000};
   double inner_ear
       [40]={-50,-39.6,-32,-25.85,-21.4,-18.5,-15.9,-14.1,-12.4,-11,-9.6,-8.3,-7.4,
```

```
-6.2,-4.8,-3.8,-3.3,-2.9,-2.6,-2.6,-4.5,-5.4,-6.1,-8.5,-10.4,-7.3,-7,-6.6,-7,
    -9.2,-10.2,-12.2,-10.8,-10.1,-12.7,-15,-18.2,-23.8,-32.3,-50.0};
double outer_ear[40]={0,0,0,0,0,0,0,0,0,0.1,0.3,0.5,0.9,1.4,1.6,1.7,2.5,2.7,
    2.6,2.6,3.2,5.2,6.6,12 ,16.8,15.3,15.2,14.2,10.7,7.1,6.4,1.8,-0.9,
    -1.6,1.9,4.9,2,-2,2.5,0.0};


//generate frequency vector
for (i=0;i<4096;i++)
{
    if (i < 40 && ER2 == false)
       inner_ear[i]=pow(10.0,(inner_ear[i]+outer_ear[i])/20.0);
    else if (i < 40)
       inner_ear[i]=pow(10.0,inner_ear[i]/20.0);
    filter[i]=(fs*i)/8192.0;
}

interpolate(filter_f,inner_ear,filter,40,4096);//half of frequency response

//apply symmetry
for (i=0;i<4096;i++)
    {
    filter[8192-i]=filter[i];
    }

filter[4096]=filter[4095];
//allocated filter output
temp_out = (fftw_complex*) fftw_malloc(sizeof(fftw_complex) * 8192);
//generate FFTW plans for filter
fftw_plan in_FFT=fftw_plan_dft_r2c_1d(8192,temp_in,temp_out,FFTW_ESTIMATE);
fftw_plan out_FFT=fftw_plan_dft_c2r_1d(8192,temp_out,temp_in,FFTW_ESTIMATE);

//prepare input vector
i=0;
for (j=0;j<2048;j++)
    {temp_in[j] = 0.0; temp_in[8191-j]=0.0;}

for (j=0;j<4096;j++)
    {temp_in[j+2048]=x[i+j];}

//fft
fftw_execute(in_FFT);

//apply filter in frequency domain
```

```
    for (j=0;j<8192;j++)
        {temp_out[j][0]=temp_out[j][0]*filter[j];
        temp_out[j][1]=temp_out[j][1]*filter[j];
        }
    //ifft
    fftw_execute(out_FFT);

    i=delay;
    //scale output
    for (j=0;j<8192;j++)
        out[i+j] =out[i+j]+temp_in[j]/8192;

    //clean up
    fftw_destroy_plan(in_FFT);fftw_destroy_plan(out_FFT);
    fftw_free(temp_out);


}

//Outer/middle ear via time-domain convolution
void FIR_ear(double x[],double out[],double fs,int N,bool ER2)
{
/* x-input signal
   out-ouput array (adjusted by filter length)
   fs-sampling rate
   N-signal length (x)
   ER2-if outer-ear filter is to be used
 */
    //generate FIR filter
    double filter[4096];
    int i,j;
    int Q;
    double Y;

    //insert filter gen code here or code to load in filter?
    double filter_f
        [40]={0,20,25,31.5,40,50,63,80,100,125,160,200,250,315,400,500,630,750,800,1000,
        1250,1500,1600,2000,2500,3000,3150,4000,5000,6000,6300,8000,9000,
        10000,11200,12500,14000,15000,16000,20000};
    double inner_ear
        [40]={-50,-39.6,-32,-25.85,-21.4,-18.5,-15.9,-14.1,-12.4,-11,-9.6,-8.3,-7.4,
    -6.2,-4.8,-3.8, -3.3,-2.9,-2.6,-2.6,-4.5,-5.4,-6.1,-8.5,-10.4,-7.3,-7,-6.6,-7,
        -9.2,-10.2,-12.2,-10.8,-10.1,-12.7,-15, -18.2,-23.8,-32.3,-50.0};
```

```
double outer_ear[40]={0,0,0,0,0,0,0,0,0,0.1,0.3,0.5,0.9,1.4,
    1.6,1.7,2.5,2.7,2.6,2.6,3.2,5.2,6.6,12,16.8,15.3,
    15.2,14.2,10.7,7.1,6.4,1.8,-0.9,-1.6,1.9,4.9,2,-2,2.5,0.0};

//generate frequency vector
for (i=0;i<2048;i++)
{
    if (i < 40 && ER2 == false)
       inner_ear[i]=pow(10.0,(inner_ear[i]+outer_ear[i])/20.0);
    else if (i < 40)
       inner_ear[i]=pow(10.0,inner_ear[i]/20.0);
    filter[i]=(fs*i)/4096.0;
}

interpolate(filter_f,inner_ear,filter,40,2048);//half of frequency response
//interpL(filter_f,inner_ear,filter,40,2048);
//add symmetry
for (i=0;i<2048;i++)
    filter[4095-i]=filter[i];

//inverse DFT of FR to get filter
FFT(filter,4096,-1,false);

//shift filter (must be better way to do this...)
for (i=0; i<2048;i++)
{
    Y = filter[i]/4096;
    filter[i]=filter[i+2048]/4096;
    filter[i+2048]=Y;
}

//convolve with signal
for (i=0; i<(N+4096); i++)
{
    out[i] = 0.0;
    if (i < 4096)
        {for (j=i;j>=0;j--)
            out[i]=out[i]+x[j]*filter[i-j];}
    else{
        if (i < N)
           {for (j=i;j>=(i-4096);j--)
                out[i]=out[i]+x[j]*filter[i-j];}
        else
           {for (j=N;j>=(i-4096);j--)
```

```
                out[i]=out[i]+x[j]*filter[i-j];}

            }
    }
    //cout<<"post-convolve";

}

//determines length of multispectrum FFT
int calc_FFT_size(int fs, int bound_t[])
{
   //add-in O() analysis
   //int fft_size;
   if (fs > 40000 && fs < 48000)
      fs = 48000;
   if (fs > 64000)
      cout<<"WARNING:This sampling rate is really high";

   return (64*fs)/1000;
}

//inspired by numerical recipies in fortran (CITE)
//output is power spectra if ps is true
void FFT(double Data[], int NN,int isign,bool ps)
{
   //Data - input/output array
   //NN - length of Data
   //isign - fft=1, ifft = -1
   double wr,wi,wpi,wtemp,theta,wpr;
   int N = 2*NN;
   //data needs to be real,image state
   double* data = new double[N];
   int i;
   int j = 1;
   int M,ISTEP;
   double tempr,tempi;
   //setup TD
   for (i=0;i<NN;i++)
   {
       data[2*i]=Data[i];
       data[2*i+1]=0.0;
   }
   for (i=1;i<N;i=i+2)
   {
```

```
    if (j > i)
        {
        tempr=data[j-1];
        tempi=data[j];
        data[j-1]=data[i-1];
        data[j]=data[i];
        data[i-1]=tempr;
        data[i]=tempi;
    }
    M=NN;
    while ((M >=2) && (j > M))
    {
         j=j-M;
         M=M/2;
    }
    j=j+M;

}
int Mmax=2;

while (N > Mmax)
{
  ISTEP=2*Mmax;
  theta=6.28318530717959/(isign*Mmax);
  wpr = -2.0*pow(sin(0.5*theta),2);
  wpi = sin(theta);
  wr =1.0;wi=0.0;
  for (M=1; M<Mmax;M=M+2)
  {
     for (i=M;i<N;i=i+ISTEP)
     {
        j = i+Mmax;
        tempr=wr*data[j-1]-wi*data[j];
        tempi=wr*data[j]+wi*data[j-1];
        data[j-1]=data[i-1]-tempr;
        data[j]=data[i]-tempi;
        data[i-1]=data[i-1]+tempr;
        data[i]=data[i]+tempi;
     }
     wtemp = wr;
     wr=wr*wpr-wi*wpi+wr;
     wi=wi*wpr+wtemp*wpi+wi;
  }
  Mmax=ISTEP;
```

```
    }

    //now is done do sum squaring to convert to PS
    for(i=0;i<NN;i++)
    {
     if (ps == true)
       Data[i]=2.0*(pow(data[2*i],2)+pow(data[2*i+1],2))/NN;
     else
       Data[i]=data[2*i];

    }
    //cout<<"done reformatting";
    delete [] data;
    //cout<<"done deleting";
}

//calculate multispectrum
void multi(double signal[],double out[],int center_i,int bound_f[],int bound_t[],int N,int
     fft_size,fftw_plan plan_fft,double * temp)
{
/*
  signal -input signal
  center_i - index of window center
  bound_f - array of frequency indeces
  bound_t - array of window sizes
  N -length bound_t/bound_f-1
*/

    int i,j,k;
    double w,E;
    for (i = 0; i<N;i++)
    {
        E=0.0;
        for (j=0;j<fft_size;j++)
        {
         if (j<bound_t[i])
         {
            w = 0.5*(1.0-cos(2.0*pi*j/(bound_t[i]-1.0)));//hann windowing
            temp[j]=signal[center_i-bound_t[i]/2+j]*w;
            E = E+w*w;
         }
         else
            temp[j]=0.0; //zero padding
        }
```

```
        fftw_execute(plan_fft);

        //extract frequencies of interest, apply correction
        for (k=bound_f[i+1];k<=bound_f[i];k++)
            { out[k] = 2.0*(pow(temp[k],2)+pow(temp[fft_size-k-1],2) )/(E*fft_size); }

    }

}

//check p generation...
void get_S(double post_FFT[],double intensity[],double post_FFT_freq[],int N, int start,
    double Estart, int Esize, double dE)
{
  /*
  post_FFT -short term spectrum
  intensity - excitation pattern
  post_FFT_freq - frequency vector
  N - end frequency index
  start - start frequency index
  Estart,Esize,dE - Excitation parameters
  */

  double E;
  double * ERB_dB = new double[N];
  double df = post_FFT_freq[1]-post_FFT_freq[0];
  int k,j,cnt,lead;
  double signal_freq_E;
  double lower_freq,upper_freq;
  int lower_index,upper_index;
  double p,g,intensityf;
  int pu_level = 0;
  double p51;
  double p51_1k=4.0*1000.0/(c1*(c2+1.0));
  double filter_freq_Hz;
  double temp;


  ERB_dB[0] = 0.0;
  p51_1k= 4.0*1000.0/(24.673*(4.368+1));

  //calculat ERB_dB, e.g. estimate excitation
  //cout<<clock()<<" ";
  for (k=start; k<N; k++)
```

```
{
  post_FFT[k]=post_FFT[k]/pow(20e-6,2);
  signal_freq_E=21.3655*log10(c2*post_FFT_freq[k]/1000.0+1.0);
  //lower and upper levels
  //lower and upper freqs
  lower_freq=1000.0*(pow(10.0,((signal_freq_E-0.5)/21.3655))-1.0)/4.368;
  upper_freq=1000.0*(pow(10.0,((signal_freq_E+0.5)/21.3655))-1.0)/4.368;
  //lower and upper indeces
  lower_index=lower_freq/df;//+1;
  upper_index=upper_freq/df;//+1;

  //fix above
  if (lower_index<start)
    lower_index = start;
  if (upper_index > N-1)
    upper_index = N-1;
  p=4.0/24.673*(4.368/1000.0+1.0);
  //cout<<clock()<<" ";
  if (k -start> 1)
   {
   g = 0.0;
   intensityf = 0.0;
   for (j=lower_index; j<= upper_index; j++)
   {
      g = abs((post_FFT_freq[k]-post_FFT_freq[j])/post_FFT_freq[k]);
      intensityf=intensityf+(1.0+p*g)*exp(-p*g)*post_FFT[j];
   }
   if (intensityf != 0.0)
      ERB_dB[k-start]=10.0*log10(intensityf);
   else
      ERB_dB[k-start]=0.0;
    }
    //cout<<clock()<<" "<<endl;
    //system("PAUSE");
  }

ERB_dB[0] = 0.0;
//second step (after calculating ERB, get actual excitation)
//cout<<clock()<<" ";
for (cnt=0; cnt<Esize; cnt++)
{
     intensity[cnt] = 0.0;
     E=Estart+dE*cnt;//this is where modifications to E need to be done
     filter_freq_Hz = 1000.0*(pow(10.0,(E/21.3655))-1.0)/4.368;
```

```
        p51=4.0*filter_freq_Hz/(0.1078*filter_freq_Hz+24.674);
        //reduction in interation based upon g here (-1 to 4)

         for (lead=start;lead<N; lead++)
         {
             g = (post_FFT_freq[lead]-filter_freq_Hz)/filter_freq_Hz;

             if (g>3.0)  // in Moore Fortran code
                 break;
             else if (g<0.0)
             {
                 g = -g;
                 p=p51-0.35*(p51/p51_1k)*(ERB_dB[lead]-51.0);
             }
             else if (g>=0.0)
             {
                 p=p51;
                 if (pu_level==1)
                     p=p51+0.118*(p51/p51_1k)*(ERB_dB[lead]-51.0);
             }
             intensity[cnt]=intensity[cnt]+(1.0+p*g)*exp(-p*g)*post_FFT[lead];
         }
         if (intensity[cnt]<1e-10)
             intensity[cnt]=1e-10;
    }
  delete [] ERB_dB;
  //cout<<clock()
}

double L_transform(double intensity[],int N,double dE,double G[],double A[],double alpha
    [],double Excitation_thrsq[])
{
 /*
 intensity - Excitation
 N - Length of excitation
 dE - Excitation resolution
 A,G,alpha,Exciation_thrsq,C - parameters of loudness transform, see moore's paper
 */

 double C = 0.046871;
 int i;
 double S = 0.0;
 double * Ns = new double[N];
```

```
for (i =0; i<N; i++)
{
    if (intensity[i] <= pow(10.0,10.0) && intensity[i] >=pow(10.0,(Excitation_thrsq[i
        ]/10.0)))
        Ns[i] = C*( pow(G[i]*intensity[i]+A[i],alpha[i]) - pow(A[i],alpha[i]));
    else if (intensity[i] < pow(10.0,Excitation_thrsq[i]/10.0))
        Ns[i] = C*(pow(G[i]*intensity[i]+A[i],alpha[i]) - pow(A[i],alpha[i]))*pow(2.0*
            intensity[i]/(intensity[i]+pow(10.0,Excitation_thrsq[i]/10.0)),1.5);
    else
        Ns[i] = C*sqrt(intensity[i]/1.04e6);

    S = S+Ns[i]*dE;

}
delete [] Ns;
return S;
}

void Setup_L_transform(double Estart, double dE, int N, double Excitation_thrsq[],double G
    [],double A[], double alpha[])
{
  /*
  Estart,dE - Minimum E values and E resolution
  N - length of E vector
  Excitation_thrsq,G,A,alpha - parameter of loudness transform (see Moore's paper)
  */
  int i;
  //lookup tables
  //double f_elc[16]={50,63,80,100,125,160,200,250,315,400,500,630,750,800,1000,20000};//
      update
  double f_elc[16]={1.83294390180259,2.25561683142248,2.78079629927702,3.36285106050045,
      4.04255534967076,
      4.91754124507044,5.82587783110374,6.84896057821133,8.02983460819669,
      9.37711023069572,
  10.7462708934667, 12.2690167901693, 13.4824359053557, 13.9446576733465,
      15.5928091217470, 41.5827263121687};
  double Excitation_thrsqx[16] =
      {28.18,23.9,19.2,15.68,12.67,10.09,8.08,6.3,5.3,4.5,3.73,3.73, 3.73,3.73,3.73,3.73};
  double alpha_fixed[6]={0.2669,0.25016,0.23679,0.22228,0.21055,0.2};
  double G_dB[6]={-25.0,-20.0,-15.0,-10.0,-5.0,0.0};
  double A_fixed[6]={8.85200,7.8585,6.83957,5.94640,5.27688,4.72096};
  //generate E vector
  for (i=0;i<N;i++)
      Excitation_thrsq[i] = Estart +i*dE;
```

```
  //get 1st
  interpolate(f_elc,Excitation_thrsqx,Excitation_thrsq, 16,N);
  //generate G from Excitation_thrsq
  for (i=0;i<N;i++)
  {
      G[i] =pow(10.0,(Excitation_thrsq[N-1]-Excitation_thrsq[i])/10.0);
      A[i] =10*log10(G[i]);
      alpha[i]=A[i];

  }
  //interpolate rest of vars
  interpolate(G_dB,A_fixed,A,6,N);
  interpolate(G_dB,alpha_fixed,alpha,6,N);//some trouble with quadratic

}

vector<double> Tempor(vector<double> invec,double attack, double release)
{
     /*
     invec - incoming loudness through time
     attack - attack constant
     release - release constant
     out - returned integrated loudness
     */
     int i;
     vector<double> out;
     out.push_back(0.0);
     for (i=1;i<invec.size();i++)
     {
         if (invec[i-1] > out[i-1])
            out.push_back(attack*invec[i-1]+(1.0-attack)*out[i-1]);
         else
            out.push_back(release*invec[i-1]+(1.0-release)*out[i-1]);
      }
      return out;
}

void interpolate(double x[],double y[],double xi[],int N,int P)
{
   /*
   x-independant var
   y-dependant var
   xi-values to be interpolated
```

```
  N-length of x
  P-length of xi
  */
  int i;
  int q=0;//place holder
  int x1,x2,x3;
  for (i=0;i<P;i++)//walk through xi's
  {
      //set up of the problem
      while (x[q] < xi[i] && q <N)
            q++;
      if (q == 0) //begining of lookup
      {          x1=q;x2=q+1;x3=q+2;        }
      else if (q >= N-2)//end of lookup extrapolation
      {          x1=N-3;x2=N-2;x3=N-1;                }
      else //normal operation
      {
         if ((xi[i]-x[q] < x[q+1]-xi[i]) || (q+2>=N-1))
            {x1 = q-1;x2=q;x3=q+1;}
         else
            {x1 = q;x2=q+1;x3=q+2;}

      }
      //do actual interpolation-overwrite xi
      xi[i]= (xi[i]-x[x2])*(xi[i]-x[x3])/(x[x1]-x[x2])/(x[x1]-x[x3])*y[x1]+(xi[i]-x[x1])
          *(xi[i]-x[x3])/(x[x2]-x[x1])/(x[x2]-x[x3])*y[x2]+(xi[i]-x[x1])*(xi[i]-x[x2])/(x
          [x3]-x[x1])/(x[x3]-x[x2])*y[x3];


  }

}

//for problematic interpolations
void interpL(double x[],double y[],double xi[],int N,int P)
{
   int i;
   int q=0;//place holder
   int x1,x2;
   for (i=0;i<P;i++)//walk through xi's
   {
       while (x[q] < xi[i] && q <N)
             q++;
       if (q == 0) //begining of lookup
```

```
        {           x1=q;x2=q+1;        }
        else if (q == N)//end of lookup extrapolation
        {           x1=N-2;x2=N-1;              }
        else //normal operation
        {
            x1 = q;
            x2 = q+1;
        }
        //do actual interpolation-overwrite xi
        xi[i]= (xi[i]-x[x2])/(x[x1]-x[x2])*y[x1]+(xi[i]-x[x1])/(x[x2]-x[x1])*y[x2];
    }
}


//.wav file reader
//reads header and prepares streams for input samples
bool wavin(char* Name,double calb,ifstream &inFile,int32_t &N,int32_t &fs,int32_t &Bits)
{
    //endian test
    unsigned char EndianTest[2]={1,0};
    short test;
    test = *(short *) EndianTest;
    cout<<"Test"<<test<<endl;

    //ifstream inFile;
    inFile.open(Name,ios::binary);

    char* temp;
    char chunkID[4];
    int32_t chunkSize;
    char* Format;
    char name[4];
    int16_t nChannels;//short nChannels;
    //int Fs;
    int32_t blockrate;
    //int Bits;
    int32_t xtra;
    //int N;
    int i;
    int16_t format;//short format;
    //cout<<sizeof(int)<<" "<<sizeof(int16_t)<<" "<<sizeof(int32_t)<<" "<<sizeof(short)<<"
         "<<sizeof(char *);
    inFile.read((char*)&chunkID,4);//chunkID
    //cout<<chunkID<<endl;
```

```
//if (strcmp(chunkID,"RIFF"))
//    {cout<<"FAIL:not riff format";system("PAUSE");return 0;}
//delete temp;

inFile.read((char*) (&chunkSize),4);//chunkSize
if (test ==256)
    chunkSize = LongSwap(chunkSize);
cout<<chunkSize<<endl;
//delete temp;
inFile.read(chunkID,4);//Format
//cout<<chunkID;
//if (strcmp(chunkID,"WAVE"))
//    {cout<<"FAIL:not wave file";system("PAUSE");return 0;}

//cout<<chunkID<<endl;
//delete temp;
//system("PAUSE");

inFile.read(chunkID,4);//subchunk1ID
//cout<<chunkID<<endl;
//if (strcmp(chunkID,"fmt "))
//    {cout<<"FAIL:no format chunk";system("PAUSE");return 0;}
inFile.read((char *)(&xtra),4);//should be 16 unless funky business
inFile.read((char *) (&format),2);//shoulde be 1
inFile.read((char *) (&nChannels),2);
inFile.read((char *) (&fs),4);
inFile.read((char *) (&Bits),4);
if (test ==256)
    {
    xtra=LongSwap(xtra);
    format=LongSwap(format);
    nChannels=LongSwap(nChannels);
    fs=LongSwap(fs);
    Bits=LongSwap(Bits);
    }
Bits = Bits/fs*8;
inFile.read((char *) (&blockrate),2);//block align
inFile.read((char *) (&Bits),2);
if (test == 256)
{
    blockrate =LongSwap(blockrate);
    Bits = LongSwap(Bits);
}
cout<<xtra<<" "<<nChannels<<" "<<format<<" "<<fs<<" "<<Bits<<endl;
```

```
    if (xtra != 16)
       {cout<<"Huh";inFile.read(temp,xtra-16); delete temp;};

    //data reading time
    inFile.read(chunkID,4);
    //cout<<chunkID<<endl;
    inFile.read((char *)(&N),4);
    if (test == 256)
        N = LongSwap(N);
    N = N/(Bits/8);
    cout<<"Length:"<<N<<endl;
    return 1;
}

//load wavefile prepared by wavin()
void load_samples(ifstream &inFile,double X[],int start,int Num,int &phold,double calb,
    int32_t N,int32_t fs,int32_t Bits,bool swap)
{
       int j,i;
       unsigned int place;
       char* buffer=0;
           int32_t temp=0;
       //buffer = new char[Bits/8];
       //char temp[Bits/8];

       for (i=start;i<Num;i++)
       {
          phold++;
          if (phold < N)
          {
          inFile.read((char *) (&temp),Bits/8);

          switch (Bits)
             {
                case 8:
                     X[i]=(temp/256.0-128.0)*calb;
                     break;
                case 16:
                                         //if (swap)
                                         //      temp = LongSwap(temp);
                                         if (temp > 32768)
                                                X[i]=(temp-32768*2.0)/32768.0*calb;
                                         else
```

```
                                        X[i]=(temp)/32768.0*calb;
                        break;
                    default:
                        place = temp;
                        if (place > pow(2,23))
                           X[i]=(place-pow(2,24))/pow(2,23)*calb;
                        else
                           X[i]=(place)/pow(2,23)*calb;

                        break;
                }
            }
            else
                X[i] = 0.0;
        }

        //ofstream bork;
        //bork.open("mal.txt");
        //for (i=0;i<Num;i++)
        //    bork<<X[i]<<" ";
        //bork.close();
        //cout<<"test";
}

//load asci file
void load_samples(ifstream &inFile,double X[],int start,int Num,int &phold,double calb,int
     N,int fs)
{
     /*
     inFile - Stream to be read
     X - output
     start - starting index to start adding data to X[]
     Num - number of points desired to be read
     phold - running total of points read so far
     calb - calibration factor
     N - number of points in file
     fs - sampling rate
     Bits - number of bits per sample
     */
       int j,i;
       unsigned int place;
       char * buffer;
       for (i=start;i<Num;i++)
       {
```

```
            phold++;
            if (phold < N)
                inFile>>X[i];
            else
                X[i] = 0.0;
        }
}

//endian swapping madness
short ShortSwap(short s)
{
 unsigned char b1,b2;
 b1 = s& 255;
 b2 = (s>>8)&255;
 return (b1 <<8)+b2;
}

int LongSwap (int i)
{
  unsigned char b1,b2,b3,b4;
  b1 = i & 255;
  b2 = (i>>8)&255;
  b3 = (i>>16)&255;
  b4 = (i>>24)&255;
  return ((int)b1<<24)+((int)b2 <<16)+((int)b3 <<8)+b4;
}
```

F.2 Boom Surrogate Generator Program

This program was used to generate the boom-like noises used as stimuli in Chapters 5 and 6. In consists of two parts. A matlab program that adjusts the window of the input noise based upon the deviation between the desired loudness time-history and the current iteration of the noise signal. Due to a bug in Matlab, the fast fourier transfrom library used by the compile Moore and Glasberg's time-varying loudness program causes matlab to crash if it is called from Matlabs unix prompt. Thus, a shell script "test.sh" was written to run the loudness program in between iterations of the surrogate program.

In the following example, a boom-like white noise with a loudness time-history matching that for the signal described in Section F.1 is generated. To run the program the following commands are used:

```
From unix prompt:
>>sh test.sh&
From matlab prompt:
>>SurrGatUnix(Desired_short_term_loudness,Cal,sampling_rate,
   cut-off,hl,Name)
```

"Cal" is to set the calibration factor for the wave-file generated by the program. "cut-off" and "hl" are used if if colored noise is used. Highpass is set to 1 for low-pass filtering,-1 for high-pass filtering and 0 if no filtering is used. For the example described here the following command is used:

```
From unix prompt:
>>sh test.sh&
From matlab prompt:
>>SurrGatUnix(Desired_short_term_loudness,12.77,32000,0,0,'
   Boomlike')
```

The pressure and loudness time-histories of the original signal and the boom-like noise generated from this program are plotted in Figure F.2.

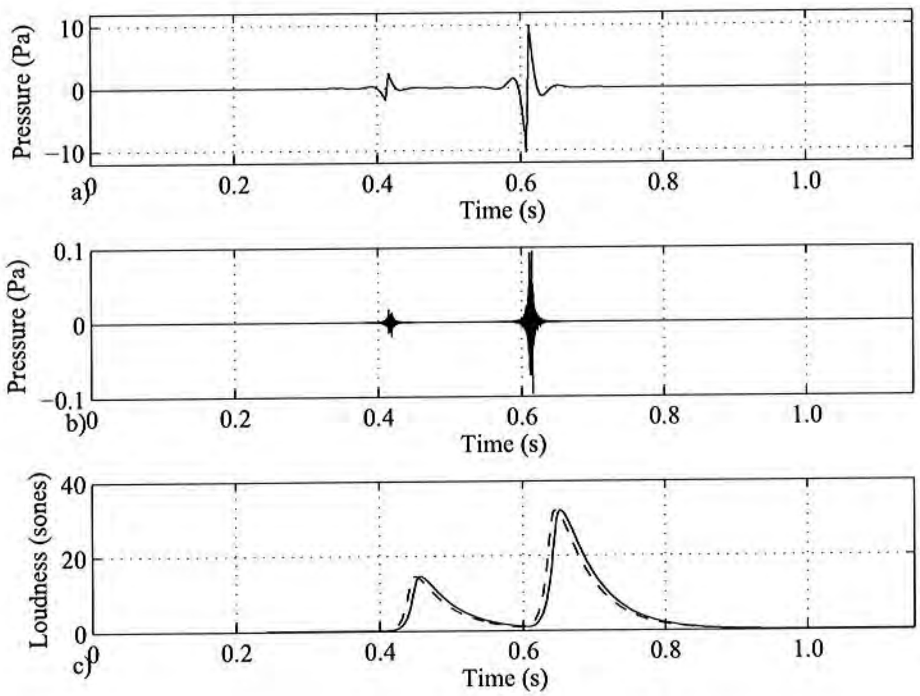

Figure F.2. a) Pressure time-history of original signal, b) pressure time-history of boom-like noise and c) loudness time-history of boom-like noise (dash) and original signal (solid). The loudness time-history of the boom-like noise is time-shifted to be more noticeable.

```
SurrGatUnix.m
%Creates windowed noise with desire loudness time-history
function main(X,fs,Cal,fc,hl,Name)
%schma tables for guesses
Level =[1 10 15 20 25 30 35 40 45 50 55 60 63 65 68 70 73 75 78 80 83 85 88 90 93 95 98
    100];
ILev  =[0.0034 0.0819 0.1993 0.4436 0.9814 1.7698 2.8367 4.7231 6.7367 9.6083 13.3226 ...
18.1693 22.2077 24.8403 28.0146 32.4268 39.2475 40.9676 48.8882 53.4784 ...
67.8263 72.2271 84.0927 91.7572 106.4155 119.1172 136.9933 151.2848];

guess=interp1(ILev*.7,10.^(Level/20)*20e-6,X,'spline');

%for all desired cut-off frequencies
for i = 1:length(fc)
    %use old correct response as initial guess
    [guess2,N]=runner(X,fs,Cal,fc(i),hl(i),strcat(Name,'-',num2str(i)),guess)
    unix('echo i > progress')
end

%-----------------------------------------------------------
function [MBW,N]=runner(X,fs,Cal,fc,hl,Name,MBW)
%X-instant loudness to match, bounded around event
%fs-sampling rate of noise surrogate
%calibration factor to be used

%generate filtered noise (fc cut-off), hl =1 highpass, hl=-1 lowpass)
N = wgn(32000*length(X)/1000,1,0);
N = N/max(abs(N));
if hl == 1
    [B,A]=butter(3,fc/(fs/2),'high');
    N = filtfilt(B,A,N);
elseif hl == -1
   [B,A]=butter(3,fc/(fs/2));
   N = filtfilt(B,A,N);
end
N = N/max(abs(N));
LBW=0;UBW=0;

Done =0;
Iter = 1;
runM(LBW,UBW,MBW,N,fs,Cal)

while Done == 0 && Iter < 1000
    [ErrM]=getErr(X);
```

```
title(strcat('Iter',num2str(Iter)))
    %drop current iteration number into a file
unix(['echo ' num2str(Iter) ' > Progress2']);

adjustLo=find(abs(ErrM) > .3 & ErrM < 0);
adjustHi=find(abs(ErrM) > .3 & ErrM > 0);
length(adjustLo)
length(adjustHi)

    if ~isempty(adjustHi)
        T=find(ErrM(adjustHi) > 10);
        MBW(adjustHi)=MBW(adjustHi)*.99;
        MBW(adjustHi(T))=MBW(adjustHi(T))*.8;

        P1=0;P2=0;
    else
        P1=1;P2=1;
    end
    if ~isempty(adjustLo)
        T=find(ErrM(adjustLo) < -10);
        MBW(adjustLo)=MBW(adjustLo)*1.01;
        MBW(adjustLo(T))=MBW(adjustLo(T))*1.2;
        P3=0;P4=0;
    else
        P3=1;P4=1;
    end
    P1
    P2
    P3
    P4
    Done = P1*P2*P3*P4;
    %regen MBW
    if Done == 0
        MBW=runM(LBW,UBW,MBW,N,fs,Cal)
        save(Name,'LBW','UBW','MBW','N','ErrM')
        Iter = Iter+1;
    else
        disp('finished')
        %save Place LBW UBW MBW ErrL ErrM ErrU
        save(Name,'LBW','UBW','MBW','N','ErrM')
    end

end
```

```
%---------------------------------------
%helper functions
function M=runM(L,U,M,N,fs,cal)
%generate test signal from Noise (N) and window (M)
Q =N.*resample(M,fs,1000)/cal;
if max(abs(Q)) > 1
    M = M/max(abs(Q));
end
wavwrite(Q,fs,'temp.wav');

%tell shell script to run
unix('echo 0 > bork')

%---------------------------------------
function [ErrM]=getErr(X)

%determine if Loudness code is finished then determine error
sent=0;
disp('waiting')
while sent ==0
   try
   temp=load('bork');
   catch
   temp = 0;
   end
   if temp == 1
       sent = 1;
       try
           X2=load('tempx.ML');
       catch
           sent =0;
       end
   end
end
disp('done waiting')
X2 = load('tempx.ML');


%error to be returned
ErrM =X2((1:length(X))+32,1)-X;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
test.sh script
```

```
#sh

echo $1
echo limit

var=3
echo $var
#get value of bork file to determine whether to start MooreTV
var='awk '{print }' bork'
echo $var
#keeping looking at book until value equals 5
while  [ "$var" -ne "5" ]
do
   var='awk '{print $1}' bork'
   #echo $var
   echo $var | grep "[^0-9]" > /dev/null 2>&1
   if [ -n "$var" ]#if not defined then try again
   then
   if [ "$var" -eq "0" ]
   then
      echo "yes" #run Moore
      ./MooreTV4 -calb 12.77 -wav temp.wav
      #rm nohup.out
      var = 1
      echo 1 > bork
   fi
   else
      var=0
      echo "strange"
   fi
done
#nohup ./MooreTV4 -calb 12.77 -wav $1
```