

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Architectural Engineering -- Dissertations and  
Student Research

Architectural Engineering

---


12-2015

# SUBJECTIVE PERCEPTION OF VARYING REFLECTION DENSITIES IN ROOM IMPULSE RESPONSES

Hyun Hong

University of Nebraska - Lincoln, icarus1021@gmail.com

Follow this and additional works at: <http://digitalcommons.unl.edu/archengdiss>

 Part of the [Architectural Engineering Commons](#), [Construction Engineering and Management Commons](#), and the [Other Engineering Commons](#)

---

Hong, Hyun, "SUBJECTIVE PERCEPTION OF VARYING REFLECTION DENSITIES IN ROOM IMPULSE RESPONSES" (2015). *Architectural Engineering -- Dissertations and Student Research*. 40.  
<http://digitalcommons.unl.edu/archengdiss/40>

This Article is brought to you for free and open access by the Architectural Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Architectural Engineering -- Dissertations and Student Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

SUBJECTIVE PERCEPTION OF VARYING REFLECTION DENSITIES IN  
ROOM IMPULSE RESPONSES

by

Hyun Hong

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Architectural Engineering

Under the Supervision of Professor Lily M. Wang

Lincoln, Nebraska

December, 2015

SUBJECTIVE PERCEPTION OF VARYING REFLECTION DENSITIES IN  
ROOM IMPULSE RESPONSES

Hyun Hong, Ph.D.

University of Nebraska, 2015

Advisor: Lily M. Wang

Reverberation time (RT) is a metric commonly used to describe room acoustic conditions, but different rooms which have the same reverberation time can have different reflection densities. Much less is known about how humans perceive different reflection densities and how sensitive humans are to changes in reflection density. Previous investigations in the existing literature have studied the upper limit of distinguishable reflection density using artificial impulse responses, but not with more realistic impulse responses simulated in room acoustic software or measured from real rooms. The aim of this dissertation is to investigate methods for quantifying reflection density from measured impulse responses, and to understand human perception of reflection density more completely by determining the upper limit of distinguishable reflection density and just noticeable difference of reflection density.

This dissertation presents three studies on the perception of reflection densities. What is the upper limit of distinguishable reflection density when using artificial impulse responses convolved with a clapping signal, and how does this limit change with different reverberation times (Study 1)? What if the impulse responses are simulated from room acoustic software instead? Does the upper limit change if the source signal changes

from clapping to speech (Study 2)? And finally, how sensitive are humans to the change of reflection density (Study 3)? In each study, a number of listeners completed three-alternative forced-choice subjective tests using the one-up two-down adaptive testing method, comprised of different RTs, reflection densities, and source signals (clapping or speech).

The results showed relatively large variation among testing subjects, possibly due to other perceptual cues rather than reflection density. After grouping the results based on how well subjects demonstrated convergence, the upper limit of distinguishable reflection density and just noticeable difference of reflection density have been identified. These results are useful for understanding how reflection density may be applied as an additional room acoustic parameter.

# **Copyright**

Copyright 2015, Hyun Hong

## Acknowledgements

This dissertation covered a majority period of my doctoral training, which has been an extremely rewarding experience. I would like to take this opportunity to express my gratitude to those who have helped me through this journey.

First and foremost, I recognize my advisor and mentor, Dr. Lily M. Wang, for her guidance, enthusiasm, and encouragement have inspired me to achieve my potential. I have been always proud of having her as my advisor. She is a role model who I will forever look up to.

I am very grateful to have a dissertation committee. Thank you Dr. Roger Schwenke for providing me the main idea triggered this research. Thank you Dr. Erica Ryherd for becoming my committee member and a reader for my dissertation. Thank you Dr. Mahboub Baccouch and Dr. Moe Alahmad for their service and guidance as I developed and undertook this research.

I would also like to acknowledge Dr. Siu-Kit “Eddie” Lau previously at UNL for his help and guidance on research with him for two years.

Lastly, I would not have accomplished this dissertation without my extraordinary support system of friends and family. I want to thank my friends in Omaha, NE who shared this journey with me with much joy and laughter: Zhao (Ellen) Peng, Ph.D., Carl

Hart, Ph.D., Matthew Blevins, Joonhee Lee, and Christopher Ainley. I am deeply influenced and grateful to my parents: Thank you Mother and Father, for supporting me and teaching me passion in Engineering, Acoustics and music, these are the foundation of my research pursuit. I could not have done this alone without you!

## Table of Contents

<b>Chapter 1 – Introduction.....</b>	<b>1</b>
1.1 Reflection Density in Room Acoustics .....	1
1.2 Outline of Dissertation .....	4
<b>Chapter 2 – Previous Research .....</b>	<b>6</b>
2.1 Introduction to Reflection Density.....	6
2.2 The Maximum Audible Reflection Density .....	9
2.3 Just Noticeable Difference .....	13
2.4 Room Size Perception.....	14
2.5 Other Studies on Single and Early Versus Late Reflections .....	18
<b>Chapter 3 - Methodology.....</b>	<b>22</b>
3.1 Introduction.....	22
3.2 Quantification of Reflection Density .....	22
3.3 Testing Facilities and Equipment Setup.....	25
3.4 Signal Generation.....	26
3.4.1 Artificial Impulse Response .....	27



3.4.2 Simulated Impulse Responses from Odeon .....	31
3.4.3 Convolution and Loudness Normalization .....	32
3.5 Listener Testing Procedure.....	35
3.5.1 Initial Screen .....	35
3.5.2 Main Experiment .....	36
<b>Chapter 4 – Study 1: The Maximum Distinguishable Reflection Density with Artificial Impulse responses.....</b>	<b>45</b>
4.1 Introduction.....	45
4.2 Methodology.....	45
4.2.1 Test stimuli.....	45
4.2.2 Subjective Testing .....	51
4.2.3 Participants.....	53
4.3 Results and Analysis .....	53
4.3.1 RT of 0.3 sec case .....	53
4.3.2 RT of 1 sec case .....	57
4.3.3 RT of 2 sec case .....	61
4.3.4 Upper Limit of Distinguishable Reflection Density .....	65

4.4 Summary and Conclusions .....	67
<b>Chapter 5 - Study 2: The Maximum Distinguishable Reflection Density with Simulated Impulse Responses .....</b>	<b>68</b>
5.1 Introduction.....	68
5.2 Methodology .....	68
5.2.1 Test Stimuli .....	68
5.2.2 Three-Alternative Forced-Choice Method with 1Up-2Down Adaptive Method .....	75
5.2.3 Participants.....	77
5.3 Result and Analysis.....	77
5.3.1 Clapping Results .....	78
5.3.2 Speech Results .....	86
5.3.3 Upper Limit of Reflection Density .....	93
5.4 Conclusions.....	96
<b>Chapter 6 - Study 3: Just Noticeable Difference of Reflection Density .....</b>	<b>97</b>
6.1 Introduction.....	97
6.2 Methodology.....	97

<b>6.2.1</b> Test Stimuli .....	97
<b>6.2.2</b> Three-Alternative Forced-Choice Method with 1Up-2Down Adaptive Method .....	102
<b>6.2.3</b> Participants.....	105
<b>6.3</b> Result and Analysis.....	105
<b>6.3.1</b> Clapping Results .....	105
<b>6.3.2</b> Speech Results .....	123
<b>6.3.3</b> Just Noticeable Difference (JND) of Reflection Density.....	136
<b>6.4</b> Conclusions.....	140
<b>Chapter 7 - Conclusion</b> .....	<b>141</b>
<b>7.1</b> Summary of Findings and Conclusions .....	141
<b>7.2</b> Future Work .....	145
<b>References</b> .....	<b>147</b>
<b>Appendix A   Sound Booth</b> .....	<b>151</b>

## List of Figures

Figure 2.1 - Construction of an image source (Adapted from Kuttruff (2009)) .....	7
Figure 2.2 - Image sources in a rectangular room (Adapted from Kuttruff (2009)).....	8
Figure 3.1 - Sample impulse response with overlaying grid of different cut-off levels and time windows used to quantify reflection density.....	25
Figure 3.2 - Sample $Y_n$ with reflection density of 500 reflections/sec across a 0.5 sec window.....	28
Figure 3.3 - Sample $Y_n$ with reflection density of 500 reflections/sec across a 0.1 sec window.....	28
Figure 3.4 - Impulse responses with different reverberation times (RT) having reflection density of 500 reflection/sec: (a) RT = 0.3 sec, (b) RT = 1 sec, and (c) RT = 2 sec .....	29
Figure 3.5 - Combinations of different reflection density and reverberation time .....	30
Figure 3.6 - Example room used for Odeon simulations with location of source (P1) and receiver (1) shown.....	31
Figure 3.7 - Comparison between the new and the previous characteristics of equal- loudness-level contours. Figure from ISO 226:2003. ....	34
Figure 3.8 - Typical psychometric function. The threshold is often taken as the stimulus level halfway between complete lack of differentiation and complete differentiation.....	37
Figure 3.9 - Psychometric functions of a two-alternative, three-alternative, and four- alternative (2AFC, 3AFC, and 4 AFC, respectively) forced choice test. The threshold $\alpha$ corresponds to proportion correct of 0.75, 0.667 and 0.625,	

respectively (Adapted from Kindom and Prins (2010)).....	39
Figure 3.10 - The pairs of AFC and adaptive up-down method: (a) 2AFC with a staircase target of 71% and 3AFC with a staircase target of 71% and 79%	43
Figure 3.11 - Example sequence of a subject's one-up two-down results, showing location of reversals .....	44
Figure 4.1 - Sample reflection densities of 500, 1000, 5000, and 10000 reflections/sec across a 0.1 sec time window .....	47
Figure 4.2 - Impulse responses with different reverberation time (RT) having reflection density of 500 reflection/sec: (a) RT = 0.3 sec, (b) RT = 1 sec, and (c) RT = 2 sec.....	48
Figure 4.3 - Combination of different reflection density (RD) and reverberation time (RT) .....	50
Figure 4.4 - Frequency spectrum plot and time plot of a sample signal with RD of 1000 reflection/sec and RT of 1 sec .....	50
Figure 4.5 - Example of a testing screen.....	52
Figure 4.6 – Detailed results of individual trials from subjects demonstrating good convergence in RT 0.3 sec case.....	55
Figure 4.7 - Reversals of subjects demonstrating good convergence in RT 0.3 sec case	55
Figure 4.8 - Detailed results of individual trials from subjects demonstrating weak convergence in RT 0.3 sec case.....	56
Figure 4.9 - Reversals from subjects demonstrating weak convergence in RT 0.3 sec case.....	57
Figure 4.10 – Detailed results of individual trials from subjects demonstrating good	

convergence in RT 1 sec case.....	59
Figure 4.11 - Reversals from subjects demonstrating good convergence in RT 1 sec case.....	59
Figure 4.12 – Detailed results of individual trials from subjects demonstrating weak convergence in RT 1 sec case.....	60
Figure 4.13 - Reversals from subjects demonstrating weak convergence in RT 1 sec case.....	61
Figure 4.14 – Detailed results of individual trials from subjects demonstrating good convergence in RT 2 sec case.....	63
Figure 4.15 - Reversals from subjects demonstrating good convergence in RT 2 sec case.....	63
Figure 4.16 - Detailed results of individual trials from subject demonstrating weak convergence in RT 2 sec case.....	64
Figure 4.17 - Reversals from subjects demonstrating weak convergence in RT 2 sec case.....	65
Figure 4.18 - Upper limits of distinguishable reflection density with different RT conditions .....	67
Figure 5.1 - Relative locations of a source (red) and receiver (blue) in the largest and the smallest room used in the study .....	69
Figure 5.2 - Reflection density calculated from Odeon for each of the different room sizes .....	71
Figure 5.3 - Reflection density calculated by cS/4V from each of the different room sizes .....	72

Figure 5.4 - Sample impulse response with overlaying grid of different cut-off levels and time windows used to quantify reflection density.....	73
Figure 5.5 - Reflection density based on the cut-off level and time window shown in the legend .....	74
Figure 5.6 - Sample screenshot of the testing program.....	76
Figure 5.7 Detailed results of individual clapp.....	80
Figure 5.8 Reversals from clapping tests of subjects demonstrating good convergence	80
Figure 5.9 - Detailed results of individual clapping trials from subjects demonstrating convergence at later reversals among the weak convergence group.....	82
Figure 5.10 - Reversals from clapping tests of subjects demonstrating convergence at later reversals among the weak convergence group.....	83
Figure 5.11 - Detailed results of individual clapping trials from subjects excluded from final mean calculation .....	84
Figure 5.12 - Reversals from clapping tests of subjects excluded from the final mean calculation .....	84
Figure 5.13 - Detailed results of individual clapping trials from subjects demonstrating ability to distinguish all comparison RD.....	86
Figure 5.14 - Detailed results of individual speech trials from subjects demonstrating ability to distinguish most comparison samples .....	87
Figure 5.15 - Reversals from speech tests of subjects demonstrating ability to distinguish most comparison samples.....	88
Figure 5.16 - Detailed results of individual speech trials from subjects demonstrating inability to distinguish differences .....	90

Figure 5.17 - Reversals from speech tests of subjects demonstrating inability to distinguish differences .....	90
Figure 5.18 - Reversals from speech tests of subjects excluded from mean calculation	91
Figure 5.19 - Detailed results of individual speech trials from subjects among the weak or no convergence group, included in final mean calculation.....	92
Figure 5.20 - Reversals from speech tests of subjects among the weak or no convergence group, included in final mean calculation.....	93
Figure 5.21 - Median of upper limits of distinguishable reflection density for the clapping and speech signals .....	95
Figure 6.1 - Relative locations of a source (red) and receiver (blue) in the largest and the smallest room used in the study .....	98
Figure 6.2 - Percentage difference of comparison samples .....	99
Figure 6.3 - Reflection densities (Odeon provided) of sound samples for three testing groups.....	101
Figure 6.4 - Sample screenshot of the testing program.....	103
Figure 6.5 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and up direction. Median JND was 26%, and Subject 11 was an outlier. ....	109
Figure 6.6 – Reversals from subjects who had $SD < 10\%$ with 5 Up condition .....	110
Figure 6.7 – Reversals from subjects who had $10\% < SD < 15\%$ with 5 Up condition	110
Figure 6.8 - Reversals from subjects who had $15\% \leq SD$ with 5 Up condition .....	111
Figure 6.9 - Trials of clapping results with RD label 5 and down direction that show inability to distinguish samples.....	113



Figure 6.10 - Trials of results that marked reversals at the first comparison sample but was included for mean calculation.....	114
Figure 6.11 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and down direction. Median JND was 68%. .....	116
Figure 6.12 - Reversals from subjects who had $SD < 10\%$ with 5 Down condition .....	117
Figure 6.13 - Reversals from subjects who had $10\% < SD \leq 20\%$ with 5 Down condition.....	117
Figure 6.14 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 10 and up direction. Median JND was 32%, and Subject 1, 9 and 18 were outliers.....	121
Figure 6.15 - Reversals from subjects who had $SD < 10\%$ with 10 Up condition .....	122
Figure 6.16 - Reversals from subjects who had $10\% < SD \leq 15\%$ with 10 Up condition.....	122
Figure 6.17 - Reversals from subjects who had $15\% < SD < 20\%$ with 10 Up condition.....	123
Figure 6.18 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and up direction. Subject 1 and 12 were outlier and median JND was 13%. .....	126
Figure 6.19 - Reversals from subjects who had $SD < 5\%$ with 5 Up condition .....	126
Figure 6.20 - Reversals from subjects who had $5\% < SD < 20\%$ with 5 Up condition	127
Figure 6.21 – Detailed results of individual speech trials from subjects demonstrating inability to distinguish differences with RD label 5 and up direction.....	128

Figure 6.22 - Reversals from subjects who had $SD < 20\%$ with 5 Down condition ....	131
Figure 6.23 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 10 and up direction. Median JND was 26% .	134
Figure 6.26 - Reversals from subjects who had $SD < 15\%$ with 10 Up condition .....	135
Figure 6.27 - Reversals from subjects who had $15\% \leq SD < 20\%$ with 10 Up condition.....	135
Figure 6.28 - JND of reflection density (using Odeon provided values) from testing with (a) clapping signals, or (b) speech signals .....	137
Figure 6.29 - JND of reflection density (using $cS/4V$ ) from testing with (a) clapping signals, or (b) speech signals.....	138
Figure 6.30 - Reflection densities of tested comparison signals and upper limit of distinguishable reflection density.....	139

## List of Tables

Table 2.1 - Acoustics quantities grouped according to listeners aspects (ISO 3382-1:2009) .....	14
Table 3.1 - Summary of methodological similarities and differences between Study 1, Study 2, and Study 3 .....	23
Table 4.1 Testing signals listed by RT in three groups (0.3, 1, or 2 sec) with twenty reflection densities .....	49
Table 4.2 - Reversals and mean upper limit of distinguishable reflection densities with RT of 0.3 sec signals (All units in reflections/sec).....	54
Table 4.3 - Mean upper limit of distinguishable reflection densities with RT of 1 sec signals (All units in reflections/sec).....	58
Table 4.4 - Mean upper limit of distinguishable reflection densities with RT of 2 sec signals (All units in reflections/sec).....	62
Table 4.5 - Upper limit of distinguishable reflection densities (reflections/sec) for all three RT conditions (shaded values are excluded from mean calculation) ..	66
Table 5.1 - Reflection density (RD) of sound samples, based on Odeon-provided calculation .....	70
Table 5.2 - Reflection density (RD) of sound samples, based on the classic reflection density equation, $cS/4V$ , and the Odeon-provided calculation .....	71
Table 5.3 - Assorted clapping results with good convergence ( $RSD < 30\%$ ) of reversals.....	79
Table 5.4 - Assorted clapping results with a weak convergence ( $RSD \geq 30\%$ ) of reversals.....	81

Table 5.5 - Assorted speech results with a good convergence ( $RSD < 30\%$ ) of reversals.....	87
Table 5.6 - Assorted speech results with a weak convergence ( $RSD \geq 30\%$ ) of reversals.....	89
Table 5.7 - Upper limits of distinguishable reflection density of subjects.....	94
Table 6.1 - Reflection densities (RD) of sound samples, based on ODEON-provided calculation .....	100
Table 6.2 - Reflection density (RD) of sound samples, based on the classic reflection density equation, $cS/4V$ , and the Odeon-provided calculation .....	101
Table 6.3 - Comparison sample numbers at reversals of clapping results with RD label 5 and up direction condition.....	106
Table 6.4 - Percentage differences of reflection density of clapping comparison samples for RD label 5 and up direction case.....	107
Table 6.5 - Just noticeable difference (%) of reflection density for clapping signals with reference reflection density of RD label 5 and up direction. Subjects 2, 11, 16, 20 were excluded from the mean calculation. ( $N = 16$ ).....	108
Table 6.6 - Comparison sample numbers at reversals of clapping results with RD label 5 and down direction case.....	112
Table 6.7 - Percentage differences of reflection density of clapping comparison samples for RD label 5 and down direction case.....	112
Table 6.8 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and down direction .....	115
Table 6.9 - Comparison sample numbers at reversals of clapping results with RD label	

10 and up direction condition.....	118
Table 6.10 - Percentage differences of reflection density of clapping comparison samples for RD label 10 and up direction case .....	119
Table 6.11 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 10 and up direction .....	120
Table 6.12 - Comparison sample numbers at reversals of speech results with RD label 5 and up direction condition.....	124
Table 6.13 - Percentage differences of reflection density of speech comparison samples for RD label 5 and up direction case .....	124
Table 6.14 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and up direction .....	125
Table 6.15 - Comparison sample numbers at reversals of speech results with RD label 5 and down direction condition.....	128
Table 6.16 - Percentage differences of reflection density of speech comparison samples for RD label 5 and down direction case .....	129
Table 6.17 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and down direction .....	130
Table 6.18 - Comparison sample numbers at reversals of speech results with RD label 10 and up direction condition.....	132
Table 6.19 - Percentage differences of reflection density of speech comparison samples for RD label 10 and up direction case .....	132
Table 6.20 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 10 and up direction. Subject 10,	

12, 14, and 17 were excluded from the mean calculation (N=16). .....	133
Table 6.21 - JND of reflection densities (%) with clapping and speech signals using Odeon provided reflection densities.....	136
Table 6.22 - JND of reflection densities (%) with clapping and speech signals using reflection densities calculated by cS/4V .....	138
Table 7.1 - Upper limits of distinguishable reflection density (reflections/sec, generated by Matlab) under different RT conditions.....	142
Table 7.2 - Upper limit of distinguishable reflection density (reflections/ms, Odeon-provided) under a RT 1sec condition .....	143
Table 7.3 - JND (%) of reflection densities with clapping and speech signals using reflection densities calculated by Odeon.....	144
Table 7.4 – Perceptual limits of reflection density.....	144

## Chapter 1 – Introduction

### 1.1 Reflection Density in Room Acoustics

There are many parameters that have been developed for quantifying the quality of sound fields, such as reverberation time (RT), early decay time (EDT), clarity (C), definition (D), lateral energy fraction (LEF), interaural cross-correlation coefficient (IACC) and others. Among these, the reverberation time is one of the most common and widely used. The definition of RT is the time it takes for the sound level in a space to drop 60 dB, and it has been found to be proportional to the ratio of room volume over absorption area (Sabine, 1922). The absorption area is obtained by summing the areas of the absorbing surfaces multiplied by their absorption coefficients. In other words, RT depends on volume and absorption. This means that, in general, large rooms have longer reverberation times than smaller rooms. The key parameter investigated in this dissertation is the temporal reflection density within a room or the number of reflections per second. This reflection density is inversely proportional to the room volume, but not related to absorption. So one could surmise that, in general, large rooms have lower reflection densities than smaller rooms. When one considers two rooms of different sizes, one can make these two rooms have the same RT by adjusting absorption. In this case, there is no difference between these two rooms in terms of RT; however, the reflection density will be different because of the different room volumes. Consequently two different rooms having the same RT can have different reflection densities, resulting

in different subjective impressions. Because reflection density is inversely proportional to room volume, this difference of reflection density is expected to relate to perception of room size.

This dissertation studies two subjective aspects of reflection density. The first aim is to investigate the maximum audible reflection density, and the second is to investigate the just noticeable difference (JND) of reflection density. Ernst Weber (1795-1878) distinguished two types of thresholds in perception (Hyland, 2012): the absolute threshold, and the difference threshold. The absolute threshold is the minimum intensity of stimulation required to produce a detectable sensory experience. In this thesis, the maximum audible reflection density corresponds to the absolute threshold of interest, and the shortest time gap between reflections corresponds to the minimum intensity of stimulation of interest. The difference threshold is the minimum change in intensity required to produce a detectable change in sensory experience, also known as a just noticeable difference (JND). By studying these two targets, one can better understand how to control reflection density effectively to impact the spatial impression from that room.

When an acoustic impulse is emitted in a room, that impulse signal generates reflections from interacting with the room surfaces, and these reflections theoretically increase in number as a quadratic function of time (Bolt et al., 1950). After a certain time has passed, the reflection density is high, and the human auditory system can no



longer distinguish one reflection from another because of the limited human auditory time resolution (Jeong et al., 2010). Measurement of human temporal resolution can be categorized in several areas, and each area analyzes a different aspect of temporal processing: (a) detection of gaps in tones and noise, (b) discrimination of gap duration and signal duration, (c) detection of signals as a function of duration (i.e., temporal integration), (d) detection of tones in temporally modulated noise (Reed et al., 2009). The time resolution often referred to as being 100-200 ms is usually a threshold of temporal integration. These studies involve threshold measurement of tones (in quiet or in a background noise) as a function of signal duration. The thresholds decrease by roughly 3 dB/doubling of duration in the range from about 10 to 200 ms and remain constant above 200 ms (Plomp, 1959; Watson and Gengel, 1969). Different from temporal integration, studies on gap-detection ask subjects to discriminate between a reference signal which is continuous throughout the presentation intervals and a comparison signal containing a silent interval (gap). Gap-detection thresholds decrease from 25 ms at 20 dB SPL (stimulus levels near absolute threshold) to an asymptotic value of roughly 3 ms for levels in the range of 50 to 90 dB SPL (Florentine and Buus, 1984). Since reflections in impulse responses could be considered as a signal with many gaps (gap-detection) having different durations (temporal integration), the threshold related to reflection density could be different from values investigated in gap-detection studies or time integration studies. The durations mentioned above are in the range of 3 to 200 ms, which corresponds to 333 to 5 reflections/sec.

The maximum audible reflection densities reported in the literature vary between 1000 and 10,000 reflections per second (Schroeder, 1962; Jot et al., 1991; Michelsen and Rubak, 1997; Kuttruff, 2009). These studies asked participants to listen directly to artificially generated impulse responses and a 20 kHz bandwidth; however, one does not commonly hear impulse responses in the everyday environment. Recently, Krueger et al. found the maximum audible reflection density to be lower when testing with speech signals, around 300 reflections per second (Krueger et al., 2012). Since the maximum audible reflection density using speech appears to be different from testing with impulse responses, further assessment of the maximum audible reflection density using speech or music is needed.

The results from this dissertation will provide a deeper understanding of reflection density perception in room acoustics. This will be useful for future development of multimedia or virtual reality algorithms, particularly for generating impressions of different room sizes. One may be able to manage reflection density effectively to create a desired spatial impression without changing the actual room size.

## **1.2 Outline of Dissertation**

In Chapter 2, previous studies on the maximum audible reflection density, the just noticeable difference of reflection density, and room size perception are introduced. Chapter 3 summarizes methodologies used for this dissertation. It encompasses a

number of quantification methods for reflection density, as well as the testing facilities, equipment setup, signal generation methods, and listener testing procedure used in this dissertation. Chapter 4 presents a study on assessing the maximum distinguishable reflection density. Most prior studies on reflection density generated room impulse responses using artificial reverberators. In this phase, the goal was to investigate the maximum audible reflection density using impulse responses simulated by Matlab. Chapter 5 discusses a similar study on the upper limit of distinguishable reflection density, but it uses more realistic impulse responses generated from Odeon room acoustics software.

In Chapter 6, the focus switches from the maximum audible reflection density to the just noticeable difference (JND) of reflection density. In this investigation, the reference reflection density is set to be within the range of the maximum audible reflection density found in Chapter 5. The JND is then determined by the smallest difference between the reference and comparison reflection density that a human subject can distinguish.

## Chapter 2 – Previous Research

### 2.1 Introduction to Reflection Density

The reverberation time is proportional to the ratio of volume to absorption area (Sabine, 1922):

$$T_{60} = 0.161 \frac{V}{A} \quad (2.1)$$

where  $T_{60}$  is the time it takes for the sound energy to decay by 60 dB,  $V$  is a volume in [m of ft cubed], and  $A$  is the absorption area, obtained by summing the areas of the absorbing surfaces multiplied with their absorption coefficient. In general, this equation indicates that larger room have longer reverberation times than smaller rooms.

Reflection density is simply the number of reflections per second. Suppose there is a point source  $A$  and a point receiver  $B$  in front of a plane wall (

Figure 2.1).

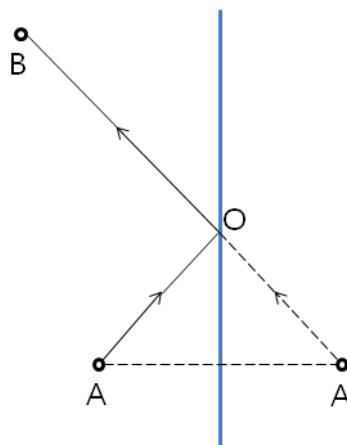


Figure 2.1 - Construction of an image source (Adapted from Kuttruff (2009))

Then the ray reflected from this wall can be considered as originating from a virtual sound source  $A'$  which is located behind the wall, on the line perpendicular to the wall, and at the same distance from it as the original source  $A$ . The line from image source  $A'$  to  $B$  is the imaginary path of ray, and  $A-O-B$  is the actual path of ray. This is called the image source method. When this image source method applies to a rectangular room, the result is the regular pattern of image rooms, as shown in Figure 2.2, each of them containing exactly one image source.

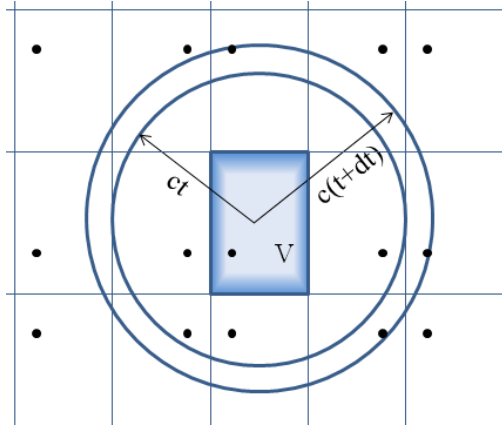


Figure 2.2 - Image sources in a rectangular room (Adapted from Kuttruff (2009))

When an impulse sound is emitted in a rectangular room, from time  $t$  to time  $(t+dt)$ , reflections arrive at center of room from sources that are  $(ct)$  to  $c(t+dt)$  away can be described (Kuttruff, 2009) by:

$$V_s = \frac{4}{3}\pi \left[ (c(t+dt))^3 - (ct)^3 \right] \quad (2.2)$$

with shell volume  $V_s$  in  $[m^3]$ ,  $c$  is the speed of sound  $[m/s]$ , and  $t$  is the time  $[s]$ .

Approximately, assuming  $dt$  is very small compared with  $ct$ , its volume can be described by the surface area of inner shell times thin thickness  $cdt$ , then

$$V_s \approx 4\pi(ct)^2 cdt = 4\pi c^3 t^2 dt \quad (2.3)$$

In this shell volume, the volume  $V$  of an image room is contained  $4\pi c^3 t^2 dt / V$  times. Each image room has one mirror source, so  $4\pi c^3 t^2 dt / V$  is also the number of

mirror sources  $N_r$  in the shell volume.

$$N_r = 4\pi c^3 t^2 dt / V \quad (2.4)$$

Hence, reflection density is the number of reflections within a time interval; it grows proportionally to the square of the elapsed time by dividing Eq. (2.4) by time interval  $dt$  to:

$$\frac{dN_r}{dt} = 4\pi \frac{c^3}{V} t^2 \quad (2.5)$$

While developed originally for a rectangular room, this expression can be extended to rooms with arbitrary shape (Kuttruff, 2009). This equation indicates that large rooms have lower reflection density than smaller rooms. From the two equations for RT and reflection density, one can see the possibility that two room with different volume can have the same RT yet different reflection densities. The perceptual differences between these cases are the focus of this dissertation.

## 2.2 The Maximum Audible Reflection Density

The maximum audible reflection density has been studied by some previous researchers, often using artificial reverberators that generate artificial impulse responses. Schroeder found that if reflection density is too low, this leads to a fluttering of the reverberated sound, especially short transients (Schroeder, 1962). He generated and tested artificial impulse responses and found that a minimum of approximately 1000

echoes (reflections) per second are required for a flutter-free reverberation. In this study, they called it as 'flutter-free' reverberation. This is not exactly the same as the maximum audible reflection density, at which one cannot perceive any more change due to the presence of more reflections, but rather a point at which the reverberation apparently sounded more realistic.

Griesinger noted that the reflection density should be larger when the bandwidth of the reverberator is increased (Griesinger, 1989). He suggested using a reflection density of up to 10,000 reflections per second or larger; however, there is no subjective data to support the use of this recommended reflection density. A later study also mentioned utilizing this 10,000 reflections per second based on Griesinger's report, but again no supported subjective testing data are provided (Jot et al., 1991). Jot et al. found that it needs to reduce the delay lengths with fewer filters, so that the time density remains same, but this causes the coloration increase, as the frequency density becomes lower.

Michelsen and Rubak studied the temporal distribution of reverberant energy and the fine structure of early reflections (Michelsen & Rubak, 1997). They found that distance perception is highly dependent on the temporal distribution of reverberant energy. In this study, a lower reflection density of on average 4410 reflections per second was used. They reported that the frequency spectrum was roughly uniform from 20 Hz to 20 kHz. It seems there was some coloration in the impulse response due to the



lower reflection density.

Krueger et al. investigated audible reflection density with four different transition times of 25, 50, 75, and 100 ms through a headphone listening test, where the transition time is defined as the time at which human ear cannot discriminate the change and growth of the density anymore (Krueger et al., 2012). Binaural impulse responses were simulated using Matlab in three rooms with different characteristics (an empty office, a lecture room, and an auditorium); these were also convolved with speech signals for use as additional test stimuli. Two types of room impulse responses were prepared. One signal was the reference signal that follows the theoretical reflection density increasing in the whole time. The simulated signal was the test signal, which follows an increasing reflection density up to a certain transition time and maintain a constant reflection density afterwards. Subjects were asked to increase or decrease the reflection density of the stimulus until they were not able to distinguish the test signal from for the different transition times in the three rooms. Results show that the maximum audible reflection density for impulse responses deviates with values below 1500 reflections per second. For speech signals, Danish sentences with the same long-term spectrum were used (Nielsen and Dau, 2009). The maximum audible reflection density for speech signals was found to be lower, around 300 reflections per second, regardless of the room and transition time. It is interesting that the maximum audible reflection density was much lower with speech stimuli than with impulse responses. Considering the fact that one hears speech signals more frequently than impulse responses in daily life, testing both

impulse responses and speech signals would be a good approach to further study reflection density.

The human hearing is generally much less sensitive to reflections in music, in contrast to speech. This is the general result of the study investigated by Schubert (1969), who measured the threshold with various music motifs. It may be of interest to test music signals as well, although the expectation may be that the maximum audible reflection density is even lower than 300 reflections per second.

So far all studies mentioned above have been primarily interested in generating artificial reverberation in rooms. However, the aim of this dissertation is not to add artificial reverberation, but to quantify reflection densities from more realistic room impulse responses and understand how humans perceive those reflection densities. It is trickier to quantify reflection density from measured room impulse responses because reflection density grows with time in real rooms, so reflection density could be changed by how to set the time window. Also, reflection density depends on how we define true reflections in real impulse responses. Recently Jeon et al. presented a method to quantify reflections in impulse responses for evaluation of scattered sounds in concert halls (Jeon et al., 2013). In this paper, they counted the number of local maxima as reflections under the assumption that they have enough time between them. However, these local maxima can overlap and build some arbitrary peak structure depending on the sampling frequency; unfortunately they do not provide information on the sampling

frequency used in their study. Also, this study set a -20 dB cut-off level for determining reflections; however, this number seems to be selected arbitrarily. One should be careful that the background noise level is at least 20 dB lower than the direct sound level then, or else such a method will count local maxima under the background noise level. More work on how to quantify reflection density from realistic impulse responses needs to be conducted, through determining appropriate time windows, cut-off levels, and how to deal with the reflection density changing over time.

### **2.3 Just Noticeable Difference**

Room acoustics quantities such as reverberation time, clarity, and definition can be controlled by acoustical treatments. However, if listeners cannot perceive different after installing acoustical treatments, those treatments are meaningless. In order to apply acoustical treatment effectively by checking whether an anticipated change of acoustics quantities, it is important to know the just noticeable difference (JND) of these acoustical parameters. Examples of JND of acoustical parameters are shown in Table 2.1.

Table 2.1 - Acoustics quantities grouped according to listeners aspects (ISO 3382-1:2009)

<b>Subjective listener aspect</b>	<b>Acoustic quantity</b>	<b>JND</b>
<b>Subjective level of sound</b>	Sound strength, $G$ , in decibels	1 dB
<b>Perceived reverberance</b>	Reverberation time ( $T_{60}$ ) in seconds Early decay time (EDT) in seconds	Rel. 5 %
<b>Perceived clarity of sound</b>	Clarity, $C_{80}$ , in decibels Definition, $D_{50}$ Centre time, $T_S$ , in milliseconds	1 dB 0.05 10 ms
<b>Apparent source width (ASW)</b>	Early lateral energy fraction, $J_{LF}$ or $J_{LFC}$	0.05
<b>Listener envelopment (LEV)</b>	Late lateral sound level, $L_J$	Not known
<b>Spatial impression</b>	Interaural cross correlation coefficient	Assumed 0.075

## 2.4 Room Size Perception

A number of investigators have looked into how humans perceive room size from auditory stimuli, some using stimuli taken from real rooms and others from simulated cases. Among those that have tested real rooms, McGrath et al. attempted to identify the ability to describe the properties and their location in two different rooms and the properties and locations of three different objects based on auditory cues generated by the participants themselves (McGrath et al., 1999). All 12 participants were male with 5 being totally blind and 7 participants having normal vision were blindfolded. In the room description test, the participant was led into the first room and seated somewhere in the

room. The participants then told to describe the room using the sound of his voice. After answering several questions about the room, the participants did the same test in another room after the half-hour break. The object perception test involved describing various objects that were placed in front of participants. There were three objects (a sheet of aluminum, a sheet of aeroboard, and leather football) with nine locations where the object could have been placed. The participant was allowed to make any sound he wanted with his voice to detect the object and asked to point out the direction. They found that both blind and sighted subjects were able to judge the size of rooms and their own location in real rooms from sound, such as that from their own speech and other sounds.

In another investigation of room size perception, longer reverberation time and greater source-receiver distance make subjects perceive rooms as being larger while background noise level has no significant effects on room size perception (Mershon et al., 1989). In this study, two levels of broadband background noise were used as a masking stimulus. Target sounds were presented at distances between 0.75 and 6.0 m, and they collected verbal reports of distance from 288 listeners in two separate experiments. Several other groups have used simulations of rooms to study room size perception. Sandvad performed three experiments on the perception of reverberant surroundings (Sandvad, 1999). In the first experiment, the subjects were presented binaural recordings of a speech signal and pictures of different rooms where the signal was recorded. Then subjects were asked to point out where each recording was made. Since

most listeners performed well, providing over 70% correct answers, they concluded that some features of the room impulse response act as cues for room size perception. In the second experiment, listeners were asked to estimate the room size after listening to the recordings, without any visual cues. They found that some listeners used the direct to reverberant energy ratio (D/R ratio) as a cue for room size, while others used the reverberation time. The third experiment was a comparison between artificial reverberation obtained by a geometric room simulation and signals generated directly from measurements. The results showed that the energy measures, such as early-to-late energy ratio (clarity index) and D/R ratio, are the most important for estimating the room dimensions.

Later Hameed et al. investigated the effects of reverberation time and the D/R ratio on room size perception using the method of paired comparisons (Hameed et al., 2004). The simulated impulse responses were convolved with speech signals. Three D/R ratios (-28, -25, and -23 dB) and three RTs (0.62, 0.73, and 0.83 sec) were used to simulate total 9 simulated impulse responses. The listening tests were conducted for ten subjects in an anechoic chamber with a 16-channel loudspeaker setup. The task was to answer which of two presented sound samples sounds like a larger room. This study concluded that reverberation time is the most dominant cue in room size perception. They also found that although D/R ratio did have some influence on the listeners, it is not a salient cue used in room size perception. This fact was commented previously by Sandvad (1999) that some subjects apply D/R ratio as a cue for room size while others

use the reverberation time. From this study, it is clear that reverberation time plays a big role in room size perception.

Pop and Cabrera also studied auditory room size perception by subjective experiments using the paired comparison method (Pop and Cabrera, 2005). This investigation consisted of three experiments using three types of rooms: (1) computer-modeled rooms, (2) a real room with variable reverberation, and (3) a concert auditorium. In Experiment 1, they used auralizations produced by a room acoustics software (Catt-Acoustics) to simulate three reverberant rooms. These rooms had the identical shape but large differences in room volume (31 m<sup>3</sup>, 249 m<sup>3</sup>, 1997 m<sup>3</sup>, 1:8:64), resulting in three different reverberation times of 0.5 seconds, 1 second, and 2 seconds, respectively. Various distances were also tested. Stimuli for the experiment were generated by a person in an anechoic chamber saying "I'm speaking from over here". The details of how these stimuli were generated are described in Cabrera and Gilfillan's study (Cabrera and Gilfillan, 2002). The recording was convolved with the three types of binaural impulse responses. They found that reverberation time and early-to-late energy ratio (known as clarity index) were good predictors of room size perception. The correlation between perceived room size and reverberation time was  $r=0.93$ , the correlation for  $C_{50}$  is  $r=-0.79$ , and for  $C_{80}$  is  $r=-0.84$ .

In Experiment 2, binaural recordings of the same speech were played from a loudspeaker and recorded by a KEMAR dummy head in a rectangular reverberation

room. These recordings were made at three distances from the source, across three reverberant conditions achieved in space by adding absorptive materials in the room. Note that actual room size did not change in this experiment. The subjective test was conducted with binaural recordings made in the above conditions, reproduced over headphones in an anechoic chamber. The conclusion of this experiment was that the reverberation time is the strongest cue for room size perception. Experiment 3 examined how perceived room size varies within the same auditorium. Binaural recordings of music were made at various seats in a large auditorium (the Michael Fowler Centre in New Zealand). In this experiment, it was found that a correlation exists between both C80 (clarity index) and sound pressure level at high frequencies with room size perception. In summary, this study indicates that auditory perception of room size may be related to actual room size, but that other room acoustical characteristics such as reverberation time and clarity can have stronger effects. In situations where the actual room size is held constant, while reverberation time and source-receiver distance varied, clarity index may be a good predictor of perceived room size.

## **2.5 Other Studies on Single and Early Versus Late Reflections**

A number of other studies have focused on just a few single reflections or early versus late reflections and how that can impact human perception of localization and source-receiver distance; a brief review of some of these papers is provided here,



although these studies do not relate as directly to the reflection density investigations pursued in this dissertation. Wallach et al. described and named the precedence effect (Wallach et al., 1949). They showed that when two identical sounds are presented with small timing difference, they will be heard as a single fused sound. In the experiments, fusion occurred when the time gap between the two sounds was in the range 1 to 5 ms for clicks, and up to 40 ms for more complex sounds such as orchestra or piano music. When the lag was longer, the second sound was heard as an echo. Wallach linked this precedence effect to the localization of sound (Wallach, 1940). Since the perceived location of the entity corresponds to the direction of the direct sounds, the source is correctly localized in space, even though reflections come from many different directions. This study did not systematically varied the intensities of the two sounds, although they cited a study by Langmuir (Langmuir, 1981) which suggested that if the second arriving sound is at least 15 dB louder than the first, the precedence effect breaks down (McGrath et al., 1999). Haas gave another interpretation of this precedence effect (Haas, 1972). The Haas effect mainly links the precedence effect to the perception of speech in a reverberant environment.

Shinn-Cunningham has studied sensitivity to reverberation patterns (Shinn-Cunningham, 2003). This study measured sensitivity to differences in reverberation time with changes in listener location in one particular room. Listeners were asked to tell their location from headphone simulations of sources located at different azimuths and distances relative to the head. Total four listener locations of the corner, center, left and

back were decided in the ordinary classroom (5 m x 9 m x 3.5 m). For each listener location, HRTFs were measured for nine source positions (all combinations of azimuths 0, 45, and 90° to the right and distances 0.15, 0.40, and 1m). Noise samples were convolved with the set of 36 HRTFs (9 source locations times 4 listener locations) to generate binaural stimuli. The results showed that there was a modest effect of azimuth; however, distance caused no statistically-significant effect. The author concluded that while listeners are sensitive to gross characteristics of the reverberation pattern reaching their ears, they are not particularly adept at discriminating between the exact timing and direction of the echoes reaching the ears.

Pop and Cabrera stated that early reflections strongly affect the sense of space and perhaps the size of the room, while slightly later reflections have been found to contribute more to a sense of envelopment (Pop and Cabrera, 2005). This acoustical term 'envelopment' can be affected by lateral reflections and reverberant sound (Long, 2006). Barron and Marshall developed LEF (lateral energy fraction) to quantify the spaciousness, which may be linked to room size (or at least room width) perception (Barron and Marshall, 1981). This parameter is the ratio of the laterally reflected sound energy in a room over sound energy arriving from all directions including the direct sound energy from the source.

Damaske and Ando developed the interaural cross-correlation (IACC) to measure spaciousness (Damaske and Ando, 1972). IACC is a measure of the similarity of sound

arriving at two points (the two ears of the listener). This study showed strong correlation between the IACC and room width, but it has weakness on the measurement of front-back spaciousness.

In summary, previous work indicates that reverberation time, source-receiver distance, and clarity are correlated to room size perception. Few previous studies have sought to control these other parameters and determine how reflection density in of itself impacts room size perception. This dissertation seeks to fill this knowledge gap.

## **Chapter 3 - Methodology**

### **3.1 Introduction**

This chapter discusses the general methodology used for the subjective perception tests, including quantification methods for reflection density, the testing facility and equipment, creation of test signals, and testing procedure. Study 1 investigated the upper limit of distinguishable reflection density (RD) with artificial impulse responses (IRs) generated by MATLAB; Study 2 investigated the upper limit of distinguishable RD with more realistic IRs simulated by ODEON; and Study 3 investigated the just noticeable difference (JND) of RD with more realistic IRs simulated by ODEON. The following table summarizes the similarities and differences in the methodologies between the three studies.

### **3.2 Quantification of Reflection Density**

Three methods of quantifying reflection density are investigated in this dissertation: theoretical mean reflection density, calculation from Odeon, and the combination of using a certain dB drop and time window.

The mean reflection density may be calculated theoretically based on global characteristics of an enclosure including volume and boundary surface area, as given in Equation (3.1):

$$\bar{n} = \frac{cS}{4V} \quad (3.1)$$

where  $\bar{n}$  is mean reflection density,  $c$  is speed of sound in [m/s],  $S$  is the surface area in [m<sup>2</sup>], and  $V$  is the volume of the room in [m<sup>3</sup>]. This equation assumes a diffuse field, which is not always easily achieved in real rooms.

Table 3.1 - Summary of methodological similarities and differences between Study 1, Study 2, and Study 3

Methodology	Study 1	Study 2	Study 3
	Upper limit with artificial IRs	Upper limit with simulated IRs	JND with simulated IRs
Acoustic Conditions	Background Noise Level (BNL): 42 dBA		
Testing Facility	Sound attenuated booth		
Source Signals	3 Claps in a row Speech	5 claps in a row Speech	
Impulse Response Generation	MATLAB	Odeon	
Auralization (Convolution)	MATLAB	Odeon	
Loudness Normalization	RMS	ITU-R algorithm	
Testing procedure	3 Alternative Forced Choice (3AFC) method 1Up-2Down adaptive method		
Repetition of samples	Manual	Automatic (3 repetitions)	Automatic (2 repetitions)

Reflection density calculated by Odeon is affected by the number of late rays and transition order. The number of late rays determines the density of reflections in the late part of the decay. Technically the reflection density in real situations will vary in time as shown in Equation (2.5) of Section 2.1, with reflections increasing exponentially. In

Odeon the late reflection process does not produce an exponentially growing number of reflections (with respect to the time) as the image source method would suggest, but it maintains the same reflection density allowing for reasonable calculation times (Christensen, 2011).

In this dissertation, the number of late rays within Odeon was set to be 10,000 and the transition order was 2. The weakness of this method is that the calculated reflection density can be changed by these settings, so that the calculated reflection density could be different for the same room and same receiver position in a room. For example, if the number of late rays is changed to 20,000, the calculated reflection density is much higher than that calculated with 10,000 late rays. If the detailed setting of Odeon do not match, it is impossible to reproduce the same reflection density with the same room and receiver location. Although this method has a limitation, its strength is that it is easy to control the reflection density of impulse responses generated by Odeon.

The last method is to count reflections from room impulse responses within a certain dB drop-off from the initial direct level and time window size. Since ISO 3382-1:2009 recommends a source level of 45 dB above the background level for T30 measurements, most measured impulse responses are expected to have a source level of at least 35 dB above the background noise level. For this reason, the cut-off level of -35 dB was selected and -50 dB was also selected for comparison. Time windows were tested in increasing 50 ms increments, out to a maximum of 1 second, and then

reflections above each cut-off level were counted until the limit of each time window was reached. An example of this method is depicted in Figure 3.1.

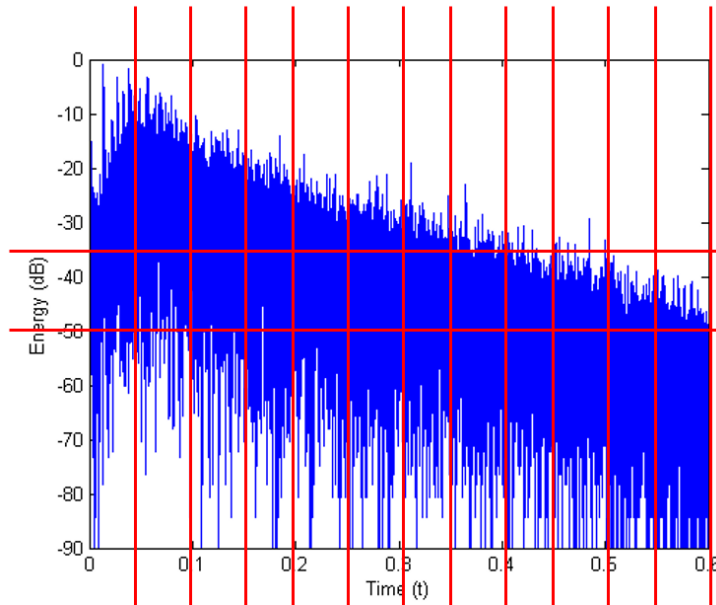


Figure 3.1 - Sample impulse response with overlaying grid of different cut-off levels and time windows used to quantify reflection density

### 3.3 Testing Facilities and Equipment Setup

The testing was conducted using headphones in a sound attenuated booth. The sound booth had a heavy metal enclosure with a floor area of 3.4 m<sup>2</sup> (35 ft<sup>2</sup>) and a height of 1.98 m (6 ft-6 in). It had low background noise levels measured at RC-30 RV (or 42 dBA). The detailed ambient background noise levels per 1/3 octave band frequency are included in Appendix A. During the subjective testing, some of the subjects reported that noise from a testing laptop was distracting their concentration, so background noise

levels were measured with the testing laptop presented in the booth. The background noise level with a testing laptop was a little bit higher than the original condition between 125 Hz and 315 Hz. However, the overall value remained the same at RC-30 RV (or 42 dBA).

The hardware used for the testing in the sound booth included an Alesis MultiMix8 multichannel USB audio interface, a Dell Inspiron 15 laptop computer and a Sennheiser HE 60 headphone. The software MATLAB and Adobe Audition CC were used for editing the testing materials. All equipment was adjusted to ensure a maximum sound pressure level of 65 dBA (re 20  $\mu$ Pa) at the subjects' ears, and this level was measured and adjusted using a Larson and Davis AEC 101 headphone coupler. The sampling frequency was set at 44.1 kHz with 16 bits resolution for all testing materials.

### **3.4 Signal Generation**

Testing signals were generated by convolving anechoically recorded source signals with simulated impulse responses. Two source signals were used: a series of claps and a short speech segment. Both the clapping and speech sound samples were cropped from signals that came from the Odeon version 11 database, and the length of the cropped samples was about 3 seconds each. The clapping signal included three claps for Study 1, and five claps for Study 2 and Study 3, while the speech signal stated 'When you are applying for a job, you need to have a good resume prepared' used only in Study



2 and Study 3. These source signals were convolved with impulse responses generated by MATLAB (Study 1) or Odeon (Study 2 and Study 3). To maintain the same loudness across signals, a RMS normalization technique was used for Study 1, while the normalization algorithm ITU-R in Adobe Audition was used for Study 2 and Study 3.

### **3.4.1 *Artificial Impulse Response***

For Study 1, the reflections in an artificial impulse response were generated within Matlab as a discrete stochastic process,  $Y[n] * w[n]$  with the following properties (Rubak and Johansen, 1998, 1999):

- a)  $Y[n]$  is given by -1, 0 and +1 with the probabilities  $p$ ,  $1-2p$  and  $p$ , respectively
- b)  $w[n]$  is an exponentially decaying weighting function
- c)  $p$  is selected to give an appropriate reflection density

$Y[n]$  consisted of randomly generated +1, -1, and 0, with +1 and -1 representing reflections. The sampling frequency was 44100 Hz. If the original method by Rubak and Johansen is used, the actual reflection density within a certain time period could be different from the desired reflection density, since reflections in  $Y[n]$  are controlled by the probability  $p$ . In order to have finer control of the reflection density, reflections (+1 and -1) were generated and placed randomly every 1/10 sec window. In this way, the actual reflection density is ensured to match the desired reflection density at least within a 1/10 sec window.

One of the  $Y[n]$  sample signals with a reflection density of 500 reflections/sec is depicted in Figure 3.2, and Figure 3.3 shows a magnified view of the 0.1 second window.

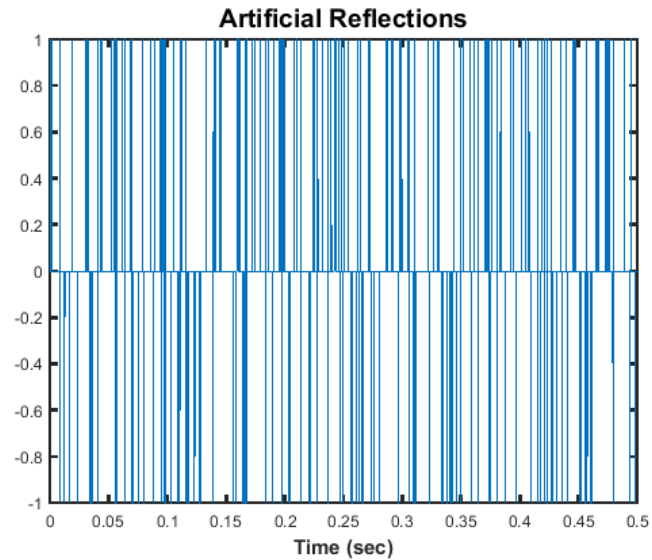


Figure 3.2 - Sample  $Y[n]$  with reflection density of 500 reflections/sec across a 0.5 sec window

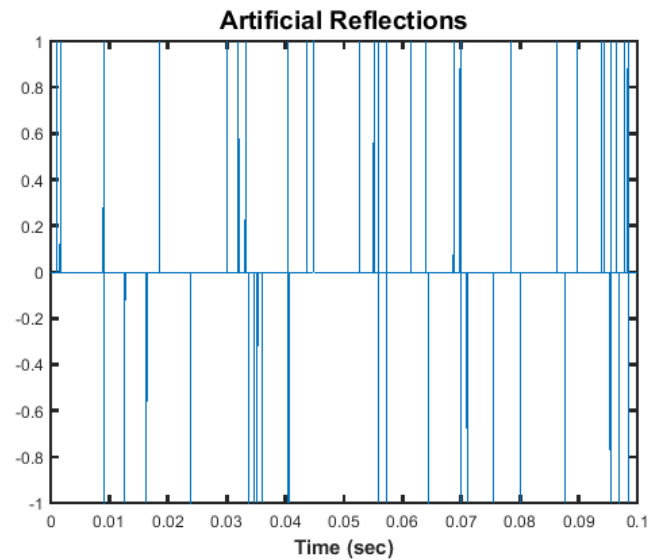


Figure 3.3 - Sample  $Y[n]$  with reflection density of 500 reflections/sec across a 0.1 sec window

An exponential decay weighting  $w[n]$  was applied on this signal to control reverberation time. The reverberation time ( $T_{60}$ ) was calculated based on Schroeder backward integration (Schroeder, 1965) within Matlab, and it was set to be 0.3, 1, or 2 sec.

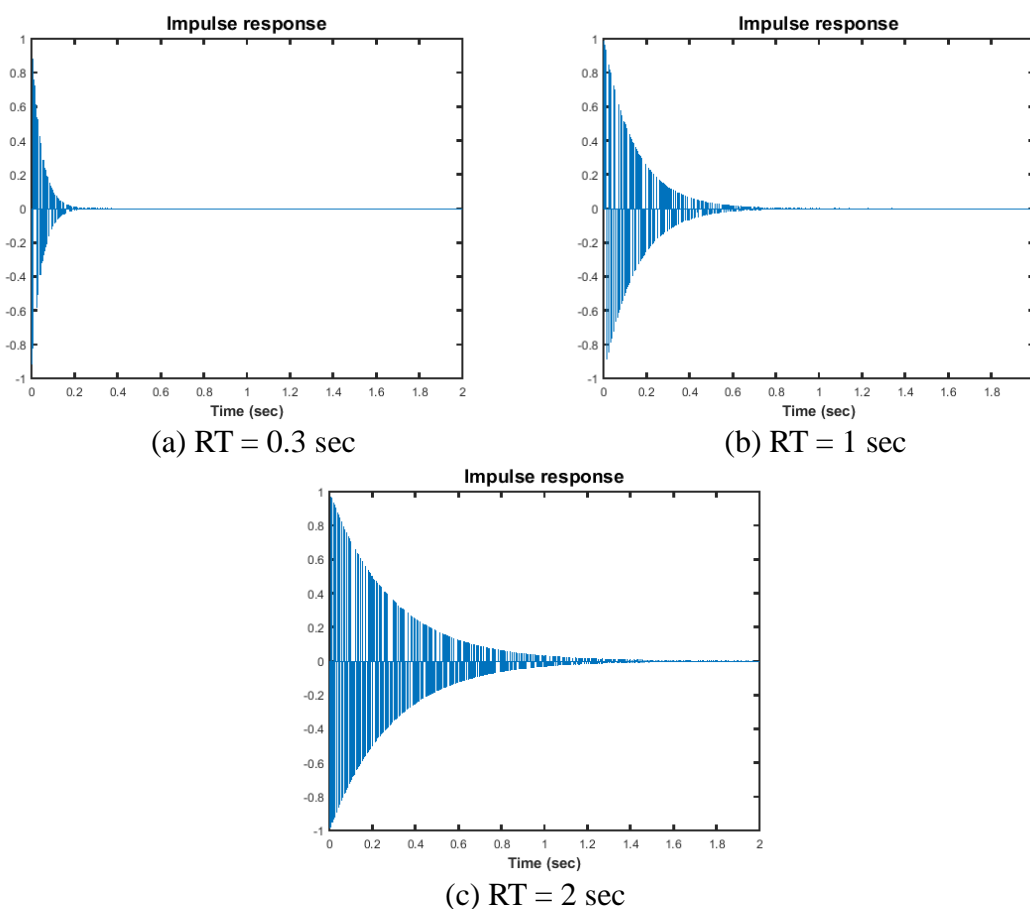


Figure 3.4 - Impulse responses with different reverberation times (RT) having reflection density of 500 reflection/sec: (a) RT = 0.3 sec, (b) RT = 1 sec, and (c) RT = 2 sec

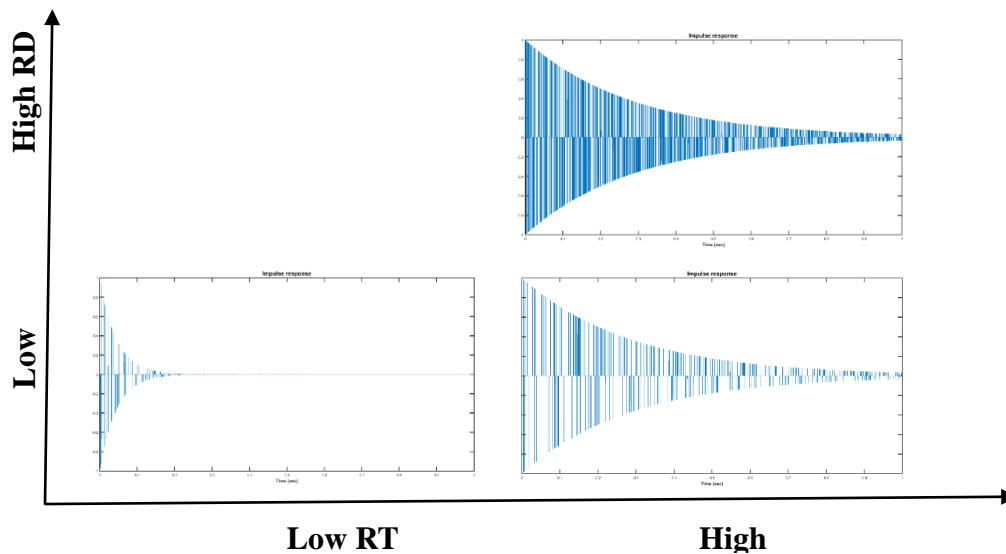


Figure 3.5 - Combinations of different reflection density and reverberation time

The sample combinations of different reflection density and reverberation time are shown in Figure 3.5. A total of 60 impulse responses, combination of three RT and twenty reflection densities for each RT, were studied, and Figure 3.5 depicts some of the more extreme cases.

### 3.4.2 Simulated Impulse Responses from Odeon

For Study 2 and Study 3, the room acoustic software, Odeon, has been used to generate impulse responses. Odeon was chosen because of its accuracy proven through round robin testing comparing modeled results to real room measurement ( Vorlander, 1995; Bork, 2000, 2005a, 2005b).

As shown by Equation (3.1) in Section 3.2, the theoretical mean reflection density is inversely proportional to the volume. In Study 2 and Study 3, different sizes of rooms were generated, and impulse responses were simulated by Odeon. The room model 'Example room' from Odeon was used, and the general room configuration and sound source and receiver locations are shown in Figure 3.6. The relative location of the source and receiver in each room was maintained across all different room sizes.

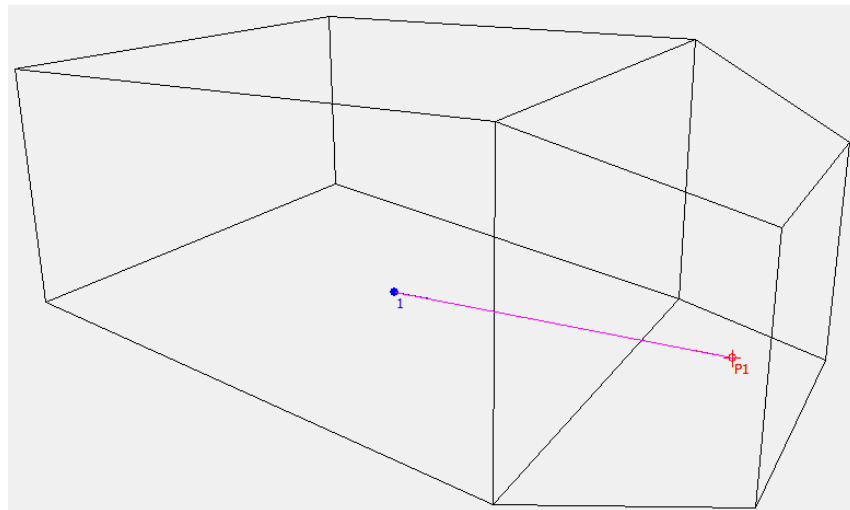


Figure 3.6 - Example room used for Odeon simulations with location of source (P1) and receiver (1) shown

In the setting, the number of late rays were set to be 10,000, and the transition order was 2. The air attenuation setting was on, and the condition was set to be 20°C and 50% relative humidity.

### ***3.4.3 Convolution and Loudness Normalization***

For Study 1, Matlab was used to convolve the anechoic source signals and impulse responses, while Odeon was used for Study 2 and Study 3. The resulting sound levels of the convolved test signals were different, which could act as an additional cue for differentiating the test signals rather than having subjects perceive differences based on reflection density as desired. For this reason, loudness normalization was applied to the test signals in each study. A review of various loudness normalization methods is provided here.

#### *3.4.3.1 Peak Volume Normalization*

A peak-volume normalization technique only considers how loud the peaks of the waveform are for deciding the overall volume of the file. This is the best technique to use if one wishes to make the audio signals as loud as possible. In digital audio the highest peak is at 0 dBFS (decibels relative to full scale), so normalizing to this value will create the loudest file.

#### *3.4.3.2 RMS Volume Normalization*

The RMS volume normalization technique considers the ‘overall’ loudness of a

file, taking into account not only the peaks, but also the remaining parts of the signal. Essentially all signals are set to have the same root-mean-square (RMS) value across the entire signal. This method was applied to signals in Study 1. The problem is that RMS volume normalization does not thoroughly account for human perception of loudness, particularly that humans perceive sound across different frequencies at different sensitivities.

#### 3.4.3.3 *ITU-R Volume Normalization*

The human auditory system can detect sounds from 20 Hz to a maximum of 20 kHz, although the upper frequency hearing limit decreases with age. Within this range, the human ear is most sensitive between 1000 Hz and 6000 Hz. Fletcher and Munson first measured equal-loudness contour using headphones (Fletcher and Munson, 1933). In 1956 Robinson and Dadson produced a new experimental determination (Robinson and Dadson, 1956), and it became the basis for a standard (ISO 226) that was considered definitive until 2003, when ISO revised the standard on the basis of recent assessment by research groups in various countries. The latest curves are those defined in the international standard ISO 226:2003, as shown below:

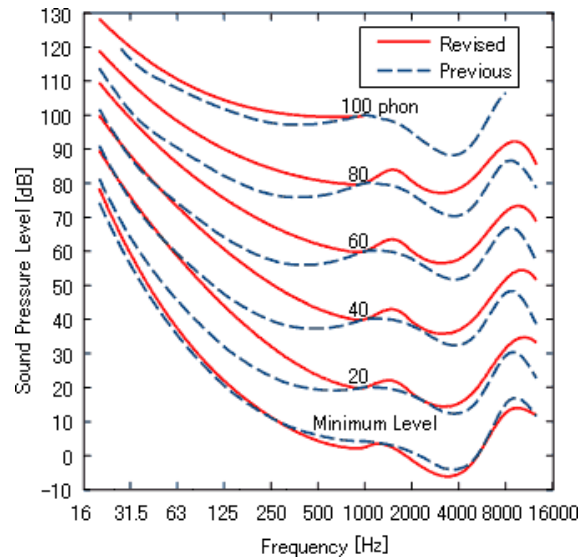


Figure 3.7 - Comparison between the new and the previous characteristics of equal-loudness-level contours. Figure from ISO 226:2003.

The ITU-R (International Telecommunication Union-Radiocommunication sector) volume detection technique assesses the loudness of a signal in a similar way to the RMS volume normalization method, but also considers human ear sensitivity across frequency. It listens to the volume intelligently and reflects how humans will hear it. It understands that humans hear sound between 1000 Hz and 6000 Hz as louder and takes that into account. RMS normalization does not take this into account. This method was applied to Study 2 and Study 3, within the Adobe Audition CC software.



## **3.5 Listener Testing Procedure**

Prepared test stimuli were presented to participants on a laptop using a custom JAVA GUI and MATLAB GUI. The tests were performed in a sound booth and presented over headphones (Sennheiser HE 60) to the subjects.

### **3.5.1 Initial Screen**

Subjective test participants were recruited by flyers posted on the University of Nebraska-Omaha campus. Once participants contacted the primary investigator, detailed information including the informed consent form, total testing hours, and compensation was provided.

At the beginning of the initial screen, the participants previewed the testing procedures in a PowerPoint orientation slideshow. They were then asked to read and sign the informed consent form. Participants were encouraged to ask questions during the screening process.

After the signed informed consent form had been collected, an audiometric screen was given in the sound booth using a Grason-Stadler GSI17 audiometer. Eligible participants needed to be able to listen to pure tones at 25 dB hearing level or lower from 125 Hz to 8000 Hz for both ears. If participants failed to meet the hearing screen requirements, they were given a \$5 Amazon gift card and asked not to participate further in the study.

### ***3.5.2 Main Experiment***

Subjects who met the minimum hearing requirements were placed in the sound booth and asked to use the testing software. Subjects listened to three sound samples in each trial and were asked to pick out the sound sample that was different from other two. One set of testing consisted of many trials, generally lasting between thirty minutes and forty-five minutes. The subjects stayed in a sound booth until the program notified them that the test had finished. Output files were then saved to a testing laptop for later analysis.

#### *3.5.2.1 Three Alternative Forced Choice*

In this dissertation, the three-alternative forced choice task was paired with a 1-up 2-down adaptive method. With the alternative forced choice method, the subject is asked to indicate which of the signals from multiple sequential presentations in each trial differs from the others. Because one of the three signals presented in each trial is different (the target stimulus) and the subject must select one of the signals as their response, this is called a forced-choice method. If there are two, three, or  $m$  choices, it is referred to as two-alternative forced choice (2AFC), three-alternative forced choice (3AFC), or  $m$ -alternative forced choice ( $m$ -AFC), respectively. Increasing the number of presented choices in a forced choice test decreases the chance of a “false correct”, where the subject can select the correct answer simply by chance, rather than through actual differentiation. The relationship between the number of alternatives and the probability of a by-chance

correct answer is given by Equation 3.2.

$$\gamma = \frac{1}{n} \quad (3.2)$$

where  $\gamma$  = probability of selecting a correct answer purely by chance (or the guessing rate), and  $n$  = number of available alternatives.

The guessing rate of 2AFC, 3AFC, and 4AFC are 0.5 (50%), 0.333 (33.3%), and 0.25 (25%), respectively, based on Equation (3.2).

Figure 3.8 shows a typical psychometric function. The difference threshold is conventionally taken to be the halfway point between complete differentiation and complete lack of differentiation.

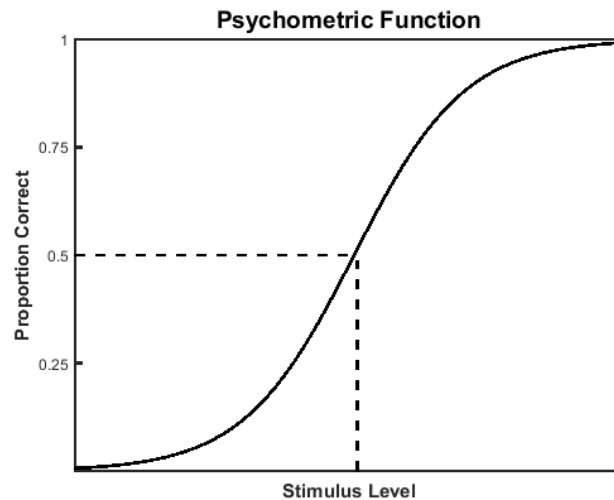


Figure 3.8 - Typical psychometric function. The threshold is often taken as the stimulus level halfway between complete lack of differentiation and complete differentiation.

In a forced choice test, a complete lack of differentiation is demonstrated when the subject performs at guessing rate. That means the subject performs no better than if he or she were randomly selecting answers. Complete differentiation is demonstrated when the subject performs at a 100% probability level (proportion correct). That means the subject can correctly pick out the different sample in every trial. Then the following equation shows how the threshold proportion correct value may be calculated:

$$\alpha = \gamma + \frac{1-\gamma}{2} \quad (3.3)$$

where  $\alpha$  is the threshold value for the proportion correct and  $\gamma$  is the guessing rate (Kingdom and Prins, 2010). The threshold values for 2AFC, 3AFC, and 4AFC are 0.75 (75%), 0.667 (66.7%), and 0.625 (62.5%), respectively, based on Equation (3.3).

As shown in Figure 3.9, a forced-choice procedure such as the 2AFC truncates the 0%-100% psychometric function, since the guessing rate is 50%. By using 3AFC or 4AFC, the usable range on the psychometric function increases and the error in difference threshold measurement created by an incorrectly measured probability level decreases. The accuracy of the three-alternative forced choice used in this dissertation is, therefore, higher than that of the two-alternative forced choice test. This was investigated in Schlauch and Rose's study, showing that 3AFC and 4AFC procedures were more satisfactory measurements of psychometric performance than 2AFC, even taking into account the greater experimental time necessary to present a greater number of choices (Schlauch and Rose, 1990). Although the 4AFC procedure has even better efficiency

and less bias, the additional time required to play four selections in each trial was not deemed worthwhile. For this reason, the 3AFC procedure was chosen for the studies in this dissertation.

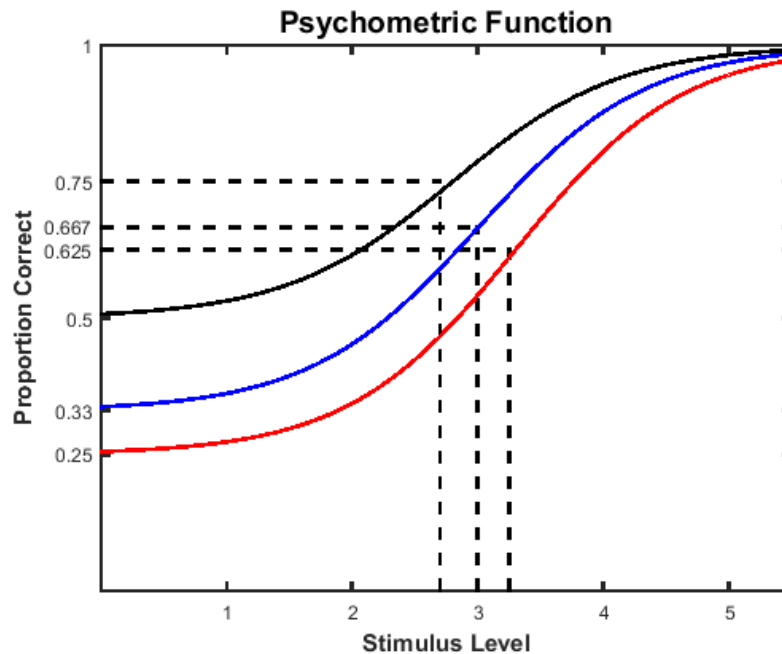


Figure 3.9 - Psychometric functions of a two-alternative, three-alternative, and four-alternative (2AFC, 3AFC, and 4 AFC, respectively) forced choice test. The threshold  $\alpha$  corresponds to proportion correct of 0.75, 0.667 and 0.625, respectively (Adapted from Kindom and Prins (2010)).

All three studies used the 3AFC procedure. The guessing rate was therefore defined by Equation (3.2) to be approximately 33.3%. The probability level indicating complete differentiation is 1 (100%). The proportional correct (probability level) is therefore taken to be 66.7 %.

### 3.5.2.2 *Up and Down Adaptive Method*

The up-and-down adaptive testing method used in this dissertation can be made to track many different probability levels by modifying its reactions to sequences of correct or incorrect responses. The ratio of consecutive correct responses (DOWN response) required to decrease the comparison difference to the number of incorrect responses (UP response) required to increase the comparison difference, denoted by  $R$ , governs the tracked probability level through the following equations (Levitt, 1971):

$$P(x)^R = 0.5 \quad (3.4)$$

where  $P(x)$  is tracked probability, and  $R$  is the ratio of sequential correct to sequential incorrect to effect change.

For the 1-up 2-down method, the  $R$  is 2, which results in  $P(x) = 0.707$ ; this is close to the desired proportional correct of 66.7% using 3 AFC. Other tracked probabilities corresponding to 1-up 1-down (also known as staircase method) and 1-up 3-down are 0.5 and 0.79, respectively.

Several studies have investigated the best combination of the AFC method and the up and down adaptive method. Kollmeier et al. showed that a 2AFC procedure used in a staircase targeting 71% (1-up 2-down) was the least efficient procedure of those tested whereas the 3AFC 79% staircase (1-up 3-down) was the most efficient (Kollmeier et al., 1988). Schlauch and Rose also reported that the 2AFC 71% target (1-up 2-down) was

more biased than the 4AFC 79% (1-up 3-down), and there was more bias for larger step sizes, especially with the 2AFC 71% procedure (Schlauch and Rose, 1990). McKee et al. provided a clear recommendation for the pairing of AFC and up-down adaptive methods (McKee et al. 1985). In general, points falling lower than the midpoint of the truncated functions generally have greater variability. McKee et al. suggested that it is more reliable if it is on the upper side of the midpoint of the psychometric function. The poorer psychometric performance of the 2AFC 71% combination (target lower than the 75% midpoint of the 2AFC function) could be explained by this suggestion. The examples are shown in Figure 3.10.

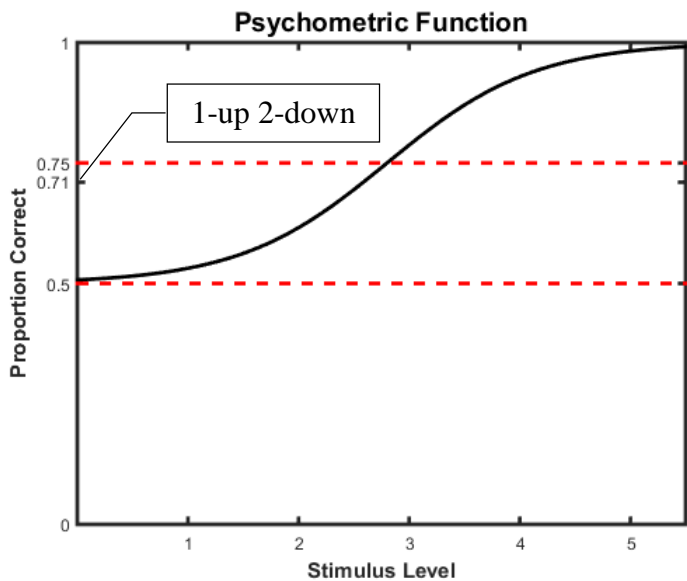
Both 3AFC tasks paired with 1-up 2-down and 1-up 3-down have good psychometric performance since corresponding targeting proportion corrects of 71% and 79% are higher than the 66.7% midpoint of the 3AFC function. This dissertation used a 3AFC paired with the 1-up 2-down adaptive method to reduce the testing time.

The number of times that the presented signals in each trial were played was controlled by each test subject in Study 1, so he or she could listen to the testing signals as many times as they want. This approach was chosen to provide test subjects enough exposure to the test signals to be certain about their choice. However, the range of testing times was very large, with some taking over 30 minutes longer than others, and it seemed that the testing results could be affected strongly by how many times each test subject listened to the test signals. Therefore, in the later studies (Study 2 and Study 3),

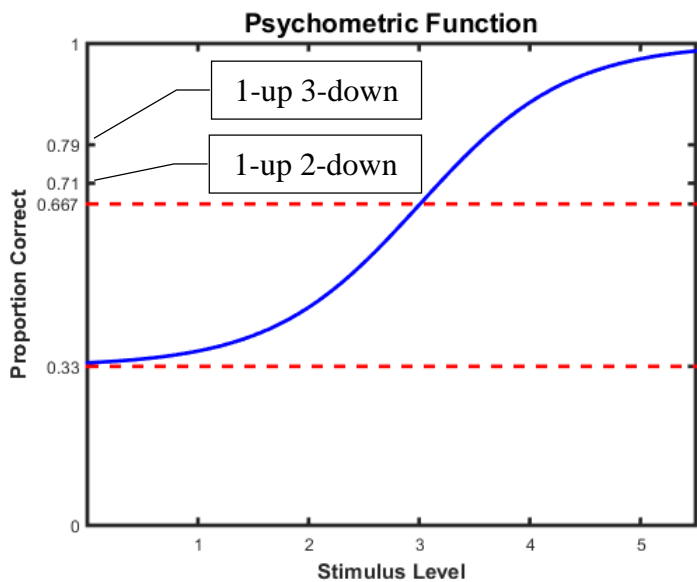
each signal was replayed a fixed number of times, controlled automatically by the testing software.

One test sequence was considered to be complete after five reversals in the 1-up 2-down procedure had been logged (Figure 3.11). This number of reversals was chosen to allow the subject to settle into a regular results pattern while keeping the overall length of testing to a reasonable length (less than one hour) to prevent subject fatigue. Increasing the number of reversals may improve the confidence of resulting averaged value. However, it also increases the testing time which can lead to subject fatigue. The threshold was calculated by averaging the reflection density of comparison stimuli at each logged reversal. As an alternative, terminating testing after a fixed number of trials was considered; however, it could force the subject to perform trials after the necessary information was gathered, or could terminate the test before the subject had a chance to provide enough information (Gescheider, 1997).





(a) 2AFC task paired with a staircase target of 71%: Poor performance



(b) 3AFC task paired with a staircase target of 71% (1 up 2 down) and 79% (1 up 3 down): Good performance

Figure 3.10 - The pairs of AFC and adaptive up-down method: (a) 2AFC with a staircase target of 71% and 3AFC with a staircase target of 71% and 79%

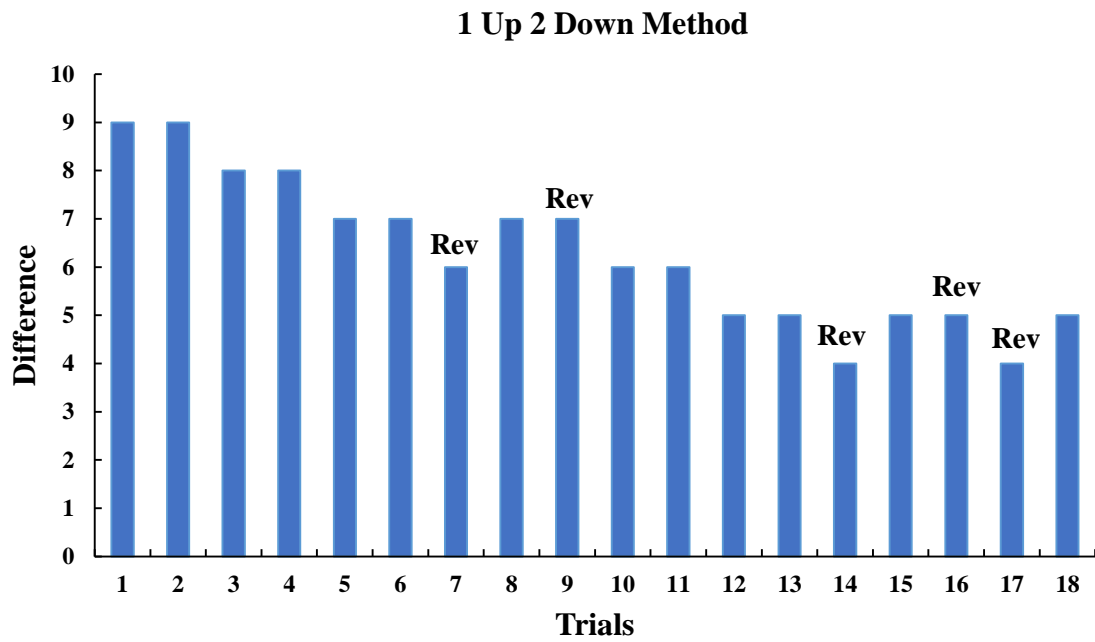


Figure 3.11 - Example sequence of a subject's one-up two-down results, showing location of reversals

## **Chapter 4 – Study 1: The Maximum Distinguishable Reflection Density with Artificial Impulse responses**

### **4.1 Introduction**

This chapter presents an investigation of the subjective perception of reflection densities from artificial impulse responses, generated as discrete stochastic processes in Matlab. While realistic room impulse responses do not demonstrate a fixed reflection density over time as used in this study, this first study helped the author to bridge prior studies that used artificial reverberators with fixed reflection densities to the other studies presented in this dissertation.

### **4.2 Methodology**

#### **4.2.1 *Test stimuli***

As presented in Section 3.4, the anechoic source signal convolved with impulse responses in this first study was a sequence of three claps edited from an original within the Odeon version 11 database. This clapping source signal was convolved in Matlab with the artificial impulse responses. Section 3.4.1 describes the method by which the artificial impulse responses were created for Study 1. The final signals were 2 sec in

length.

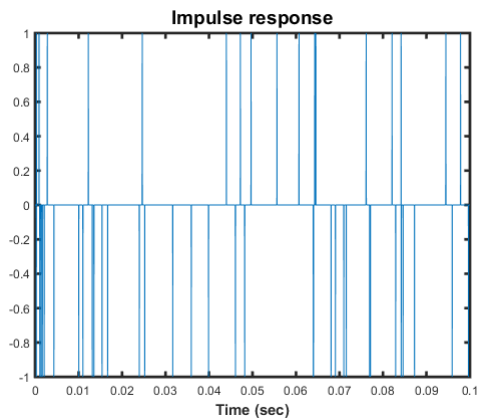
A total of twenty reflection densities were prepared, with comparison reflection densities ranging from 500 reflections/sec to 9500 reflections/sec in 500 reflection/sec increments, and the reference reflection density was taken to be the highest reflection density of 10,000 reflections/sec. The value of the highest reflection density (reference) was chosen to be 10,000 since this was the highest maximum distinguishable reflection density found among previous studies (Griesinger, 1989).

The sample signals  $Y[n]$  with reflection densities of 500, 1,000, 5,000, 10,000 reflections/sec are depicted in Figure 4.1.

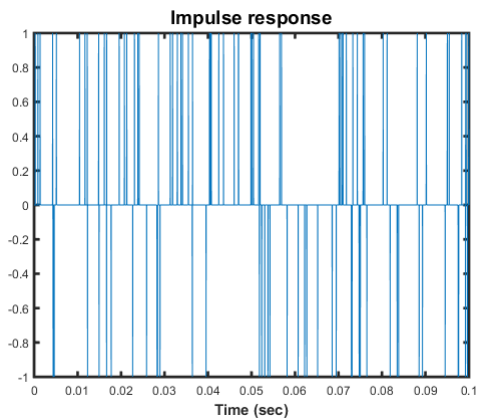
An exponential decay weighting function  $w[n]$  was multiplied with  $Y[n]$  to achieve the desired reverberation times:

$$IR[n] = Y[n] * w[n] = Y[n] * \exp(\textit{gain} * [n]) \quad (4.1)$$

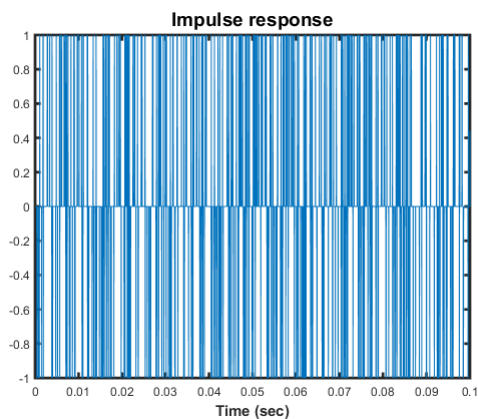
where the gain is a constant value that controls the degree of decay of the exponential curve. Three reverberation times ( $T_{60}$ ) were generated: 0.3, 1 and 2 sec, as calculated based on Schroeder backward integration (Schroeder, 1965) in Matlab. The proper gain corresponding to these three reverberation times were found to be -0.000514, -0.000156, and -0.0000785, respectively. Figure 4.2 plots the RT = 0.3, 1, and 2 sec signals with RD of 500 reflections/sec.



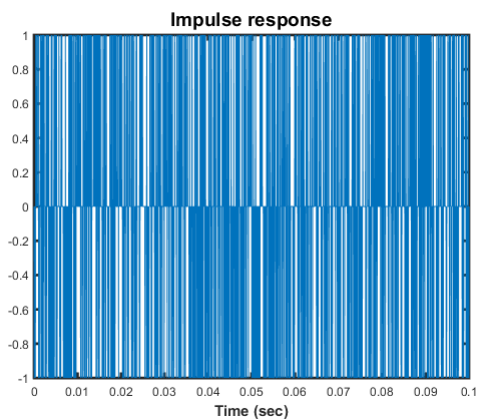
(a) 500 reflections/sec



(b) 1000 reflections/sec



(c) 5000 reflections/sec



(d) 10000 reflections/sec

Figure 4.1 - Sample reflection densities of 500, 1000, 5000, and 10000 reflections/sec across a 0.1 sec time window

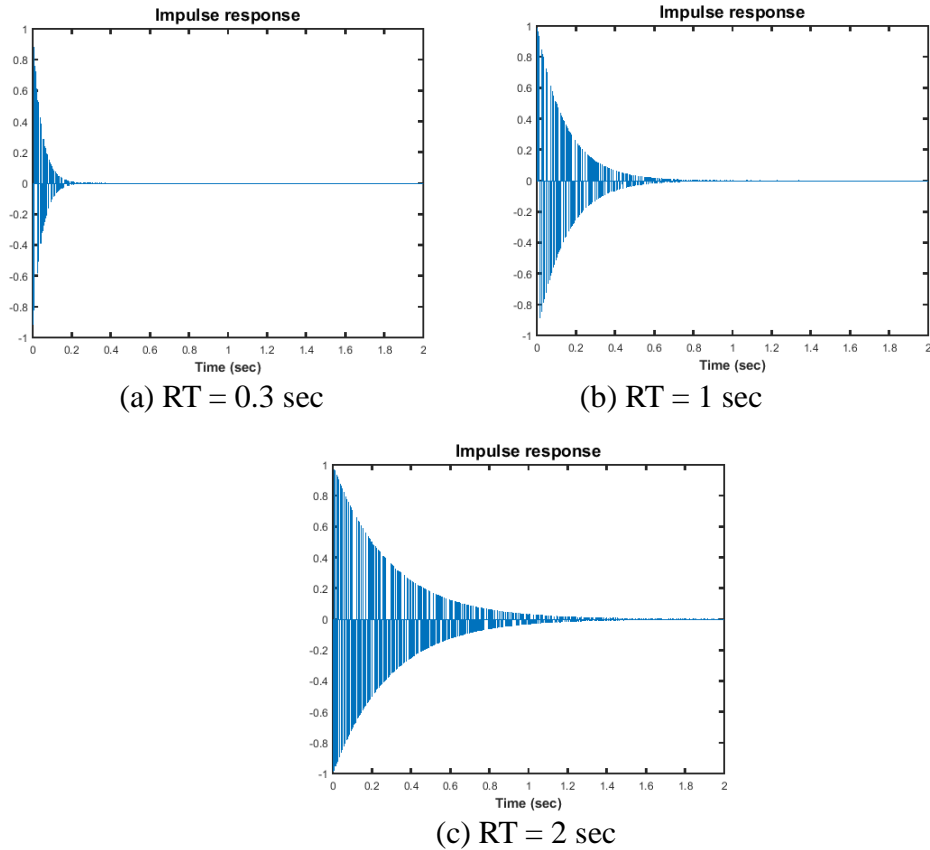


Figure 4.2 - Impulse responses with different reverberation time (RT) having reflection density of 500 reflection/sec: (a) RT = 0.3 sec, (b) RT = 1 sec, and (c) RT = 2 sec

In total, there were sixty signals created. Three reverberation time of 0.3, 1, and 2 sec were used; 19 comparison RD (from 500 reflections/sec to 9500 reflection/sec with 500 reflection density increment) and one reference RD of 10000 reflections/sec were paired with each RT. A list of testing signals is provided in Table 4.1.

Table 4.1 Testing signals listed by RT in three groups (0.3, 1, or 2 sec) with twenty reflection densities

	Group	A	B	C
	Reflection density (reflections/s)	RT 0.3 sec	RT 1 sec	RT 2 sec
Reference	10000	A20	B20	C20
Comparison	9500	A19	B19	C19
	9000	A18	B18	C18
	8500	A17	B17	C17
	8000	A16	B16	C16
	7500	A15	B15	C15
	7000	A14	B14	C14
	6500	A13	B13	C13
	6000	A12	B12	C12
	5500	A11	B11	C11
	5000	A10	B10	C10
	4500	A9	B9	C9
	4000	A8	B8	C8
	3500	A7	B7	C7
	3000	A6	B6	C6
	2500	A5	B5	C5
	2000	A4	B4	C4
	1500	A3	B3	C3
	1000	A2	B2	C2
	500	A1	B1	C1

The combinations of the most different reflection densities and reverberation times are shown in Figure 4.3.

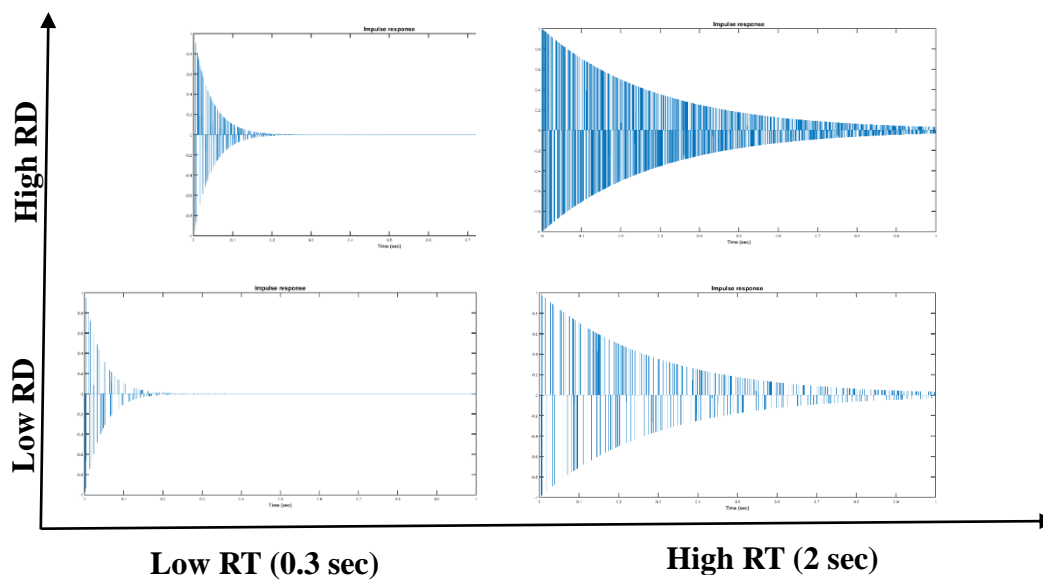


Figure 4.3 - Combination of different reflection density (RD) and reverberation time (RT)

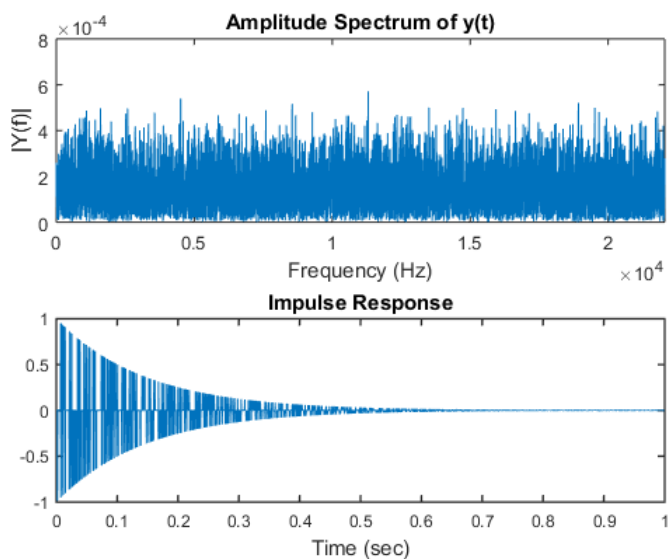


Figure 4.4 - Frequency spectrum plot and time plot of a sample signal with RD of 1000 reflection/sec and RT of 1 sec



Figure 4.4 shows a sample spectrum of one of the signals; all of the generated signals similarly showed no presence of strong tones.

The generated impulse responses were all monaural. The loudness of the convolved signals was normalized by RMS normalization, as described previously in Section 3.4.3. All testing signals were calibrated to playback at 65 dBA over the headphones.

#### ***4.2.2 Subjective Testing***

As described in Section 3.5.2, the 3AFC method combined with a 1-up 2-down adaptive method was used to determine the upper limit of distinguishable reflection density with artificial impulse responses in this first study. The entire test was expected to last around 30 minutes for each subject. In this first study, participants were allowed to play each audio sample as many times as they wanted within a trial; Figure 4.5 shows a screen shot of the testing graphical user interface. The test was finished once the testing program achieved enough data, and then a pop-up notification asked the subjects to exit the testing chamber.

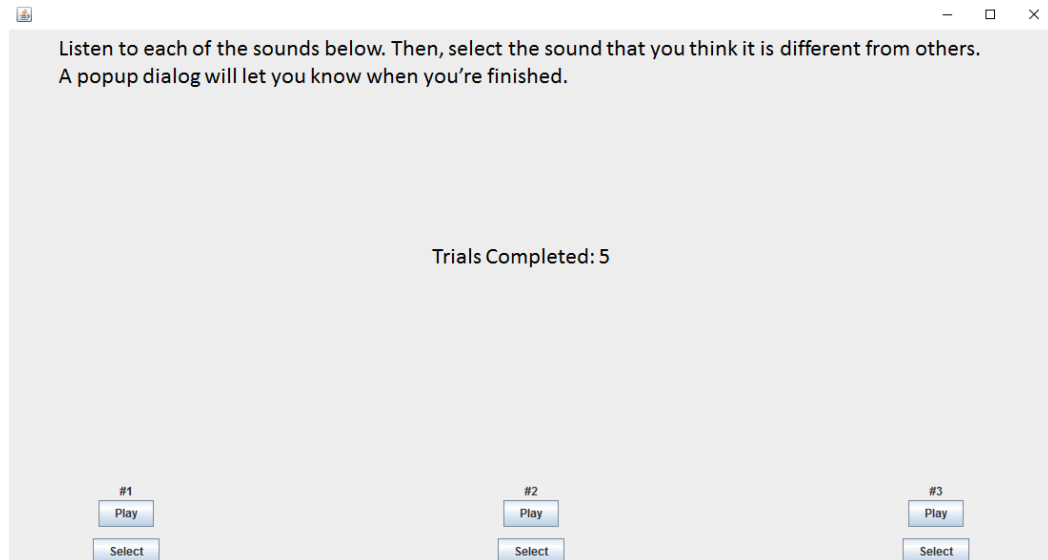


Figure 4.5 - Example of a testing screen

The 3AFC testing method presented two identical reflection density (RD) samples and one comparison RD sample for each and every comparison. The reference RD was the highest reflection density of 10,000 reflections/sec. The comparison RD approached the reference value from below when test subjects made correct selections. With the 1-up 2-down adaptive method, two consecutive correct answers decreased (down) the difference between the reference and comparison RD, while one incorrect answer increased (up) the difference. As shown previously in Chapter 3.5.2, there were three 1-up 2-down testing sequences (A, B, and C), each corresponding to a set RT of 0.3, 1, or 2 sec, respectively. The three testing sequences were mixed together randomly so that participants could not easily tell how well they were doing on any one sequence. Reversals within each testing sequence were recorded, and once five reversals in one

testing sequence had been reached, that sequence was not included in the rest of the test session. After five reversals from all three testing sequences had been recorded, the averaged comparison RD at the five reversals of each testing sequence were reported as the upper limit of distinguishable reflection density.

### **4.2.3 *Participants***

A total of ten listeners (3 female) were recruited on the University of Nebraska at Omaha campus to participate in the study. All listeners were older than 19 years old, and their average testing time was 25 min (Max: 34, Min: 10, SD = 6.7 min). All listeners had at least three years of musical training or experience. As described in Section 3.5.1, all listeners met minimum hearing thresholds and years of musical training or experience. Listeners were given a \$15 Amazon gift card for participating in the study.

## **4.3 Results and Analysis**

### **4.3.1 *RT of 0.3 sec case***

In this first study, the five reversals from a number of test subjects in the RT of 0.3 sec test sequence converged as expected, but a few did not, specifically Subjects 4, 6, 8, and 9. Subjects 4, 6, 8, and 9 recorded more than 30% relative standard deviation (RSD) between their reversal points, so these data were excluded from the calculation of the mean value for maximum distinguishable reflection density (Table 4.2). Cases where the

RSD is larger than 30% are shaded in the table. The mean upper limit of distinguishable reflection density from the remaining data was found to be 3000 reflections/sec (RSD = 23%, N = 6); when all subjects are included, this value drops to 2640 reflections/sec (RSD = 32%, N = 10).

Table 4.2 - Reversals and mean upper limit of distinguishable reflection densities with RT of 0.3 sec signals (All units in reflections/sec)

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
1	5500	3000	5000	4000	4500	4400	860	20%
2	2500	2000	3000	2000	3000	2500	447	18%
3	3000	2500	3000	2500	3500	2900	374	13%
4	1000	500	1500	1000	2000	1200	510	42%
5	3500	3000	3500	3000	3500	3300	245	7%
6	2500	2000	2500	2000	5000	2800	1122	40%
7	3000	2000	2500	2000	3500	2600	583	22%
8	3000	2000	3000	2000	4500	2900	917	32%
9	2500	1500	2000	500	1000	1500	707	47%
10	2500	2000	2500	2000	2500	2300	245	11%
Mean (All, N = 10)						2640	849	32%
Mean (RSD<30%, N = 6)						3000	702	23%

The following figures group the reversal results by subjects who demonstrated good convergence and those who demonstrated weak convergence. In Figure 4.6 and Figure 4.7, results from Subjects 2, 3, 5, 7, and 10 converged to around 3000 reflections/sec. Although Subject 1 showed relatively larger standard deviation compared to other subjects, convergence can be observed.

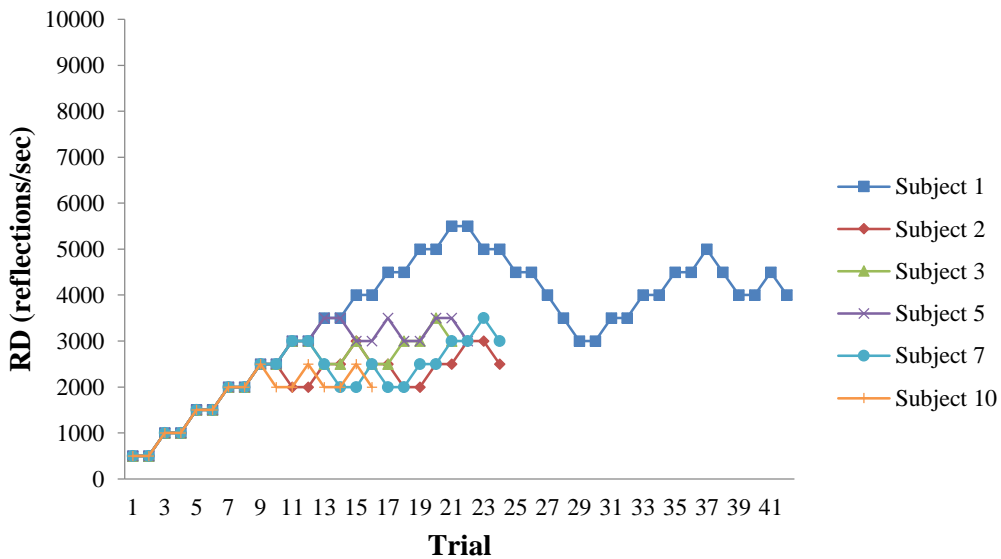


Figure 4.6 – Detailed results of individual trials from subjects demonstrating good convergence in RT 0.3 sec case

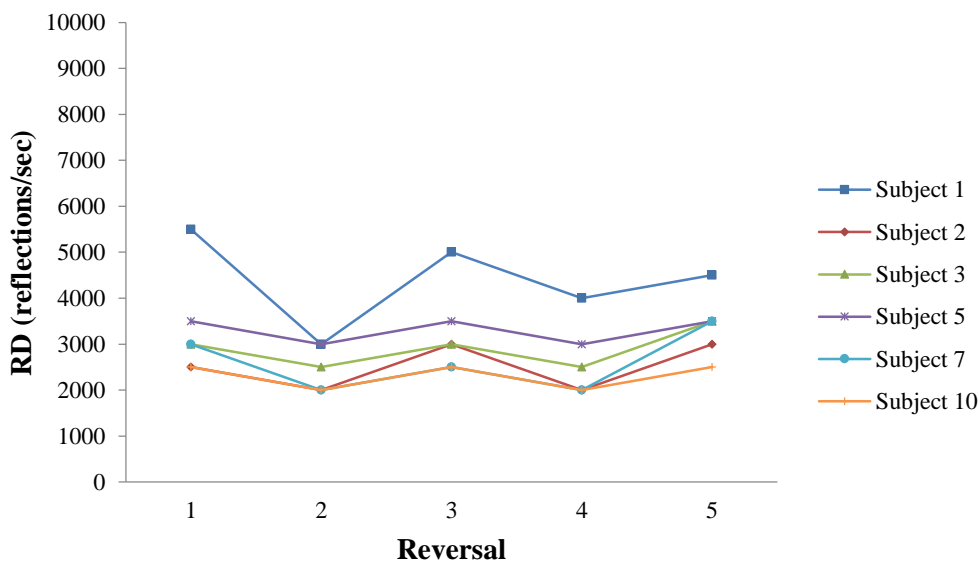


Figure 4.7 - Reversals of subjects demonstrating good convergence in RT 0.3 sec case

Figure 4.8 and Figure 4.9 depicts results from subjects who demonstrated a weak

convergence of reversals. Subjects 6 and 8 showed a good convergence to around 3000 reflections/sec from the 1st reversal to 4th reversal, but their 5th reversals increased by around 3000 reflection/sec.

It seems that Subjects 6 and 8 could catch extra cues in the testing signals after the 4th reversal rather than basing results on a perceived difference of reflection density, or learned how to differentiate samples better. Subjects 4 and 9 showed gradual increase or decrease of reversals, which may reflect that subjects either did not know how to differentiate samples (Subject 4) or lost their concentration (Subject 9).

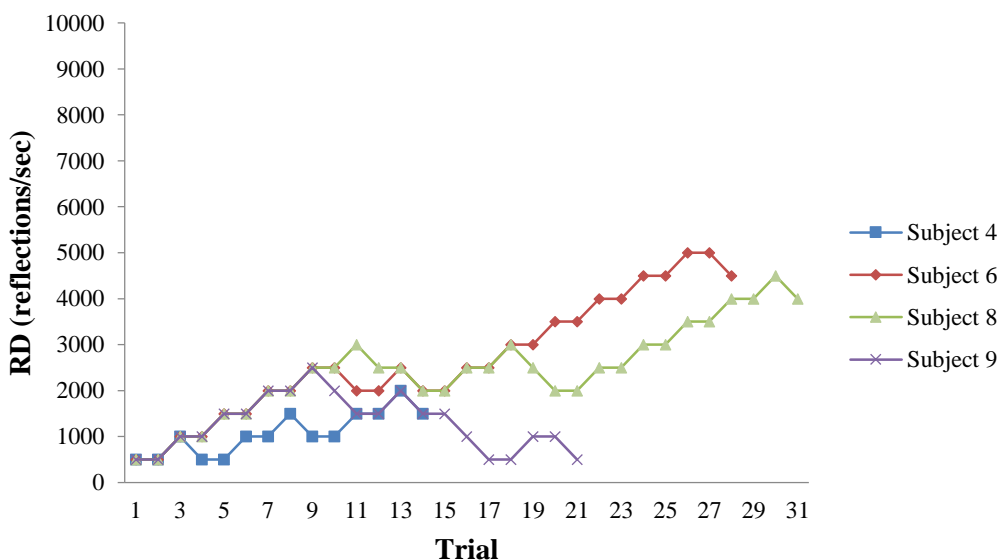


Figure 4.8 - Detailed results of individual trials from subjects demonstrating weak convergence in RT 0.3 sec case

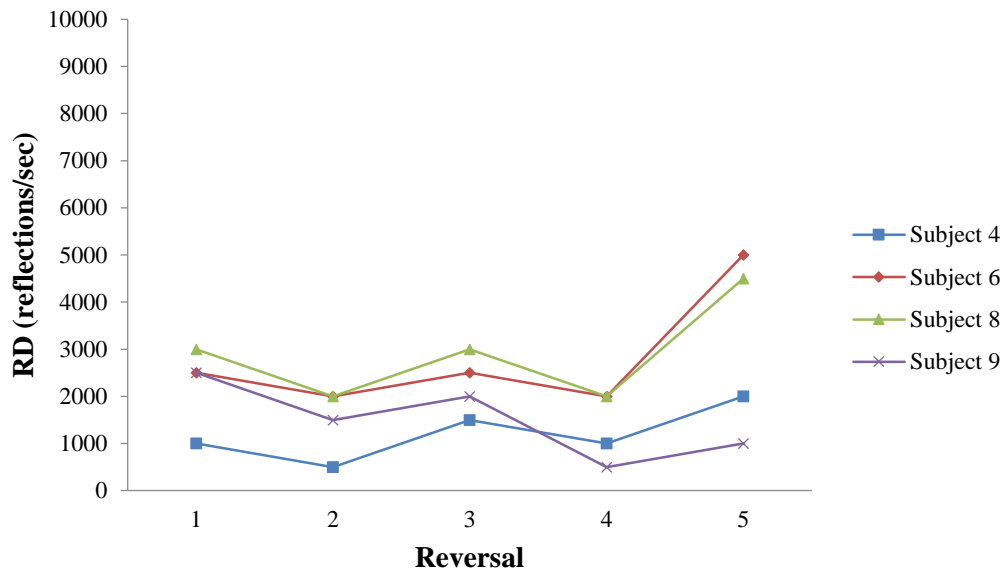


Figure 4.9 - Reversals from subjects demonstrating weak convergence in RT 0.3 sec case

### 4.3.2 *RT of 1 sec case*

For the testing sequence related to the RT of 1 sec condition, Subjects 2, 7, 8, and 10 had relatively large RSD (shaded in the table), which were equal or more than 30%. These subjects were excluded from the mean calculation. The mean upper limit of distinguishable reflection density for the remaining subjects was found to be 4075 reflections/sec (RSD = 60%, N = 6); this value becomes 4060 reflections/sec (RSD = 54%, N = 10) when results from all subjects are included (Table 4.3).

Table 4.3 - Mean upper limit of distinguishable reflection densities with RT of 1 sec signals (All units in reflections/sec)

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
1	7000	6500	8000	7500	8500	7500	707	9%
2	2000	1500	2000	1500	3500	2100	735	35%
3	8000	7500	8500	7500	8000	7900	374	5%
4	1500	1000	1500	1000	1500	1300	245	19%
5	5500	5000	5500	4000	5500	5100	583	11%
6	3000	2500	3500	2500	3000	2900	374	13%
7	2500	2000	4500	4000	5500	3700	1288	35%
8	5500	5000	5500	2500	3000	4300	1288	30%
9	4000	3500	4500	4000	5500	4300	678	16%
10	2000	1500	2000	500	1500	1500	548	37%
Mean (All, N = 10)						4060	2174	54%
Mean (RSD<30%, N = 6)						4075	2426	60%

Subject 1, 3, 4, 5, 6, 9 had good convergences (RSD < 30%) and their reversals are shown in Figure 4.11 and subjects' responses by trials are depicted in Figure 4.10s. A noticeable difference from the RT of 0.3 sec results is that subjects recorded distinctly different mean reversals among subjects. This observation can also be verified by the fact that the RSD value of assorted subjects' reversals in RT 1 sec case (54%) is much larger than the one in RT 0.3 sec case (23%).



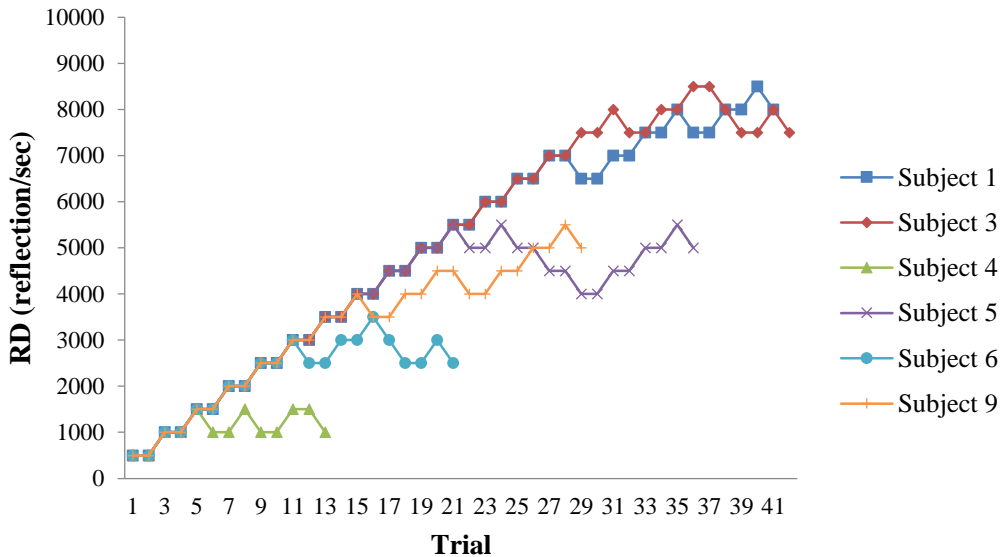


Figure 4.10 – Detailed results of individual trials from subjects demonstrating good convergence in RT 1 sec case

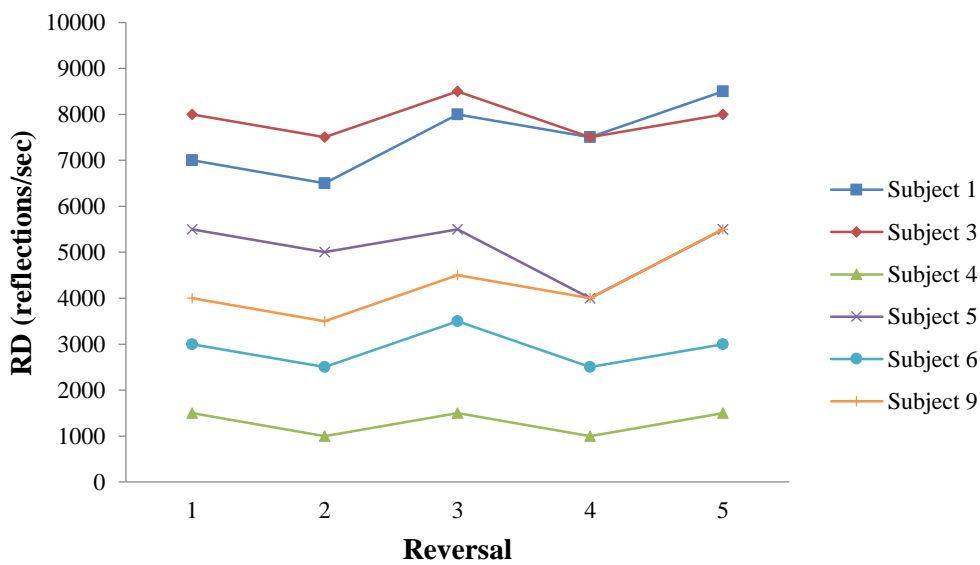


Figure 4.11 - Reversals from subjects demonstrating good convergence in RT 1 sec case

Figure 4.12 and Figure 4.13 show the results from subjects demonstrating weak convergences with individual RSD of 30% or more. Subjects 7 and 8 recorded steep increases or decreases in reversals with RSDs of 35% and 30%, respectively. Subject 10 showed a gradual decrease, while Subject 2 maintained a good convergence except at the 5th reversal which increased by 2000 reflections/sec (Figure 4.13). As shown in Figure 4.12, Subject 8 could not select two consecutive correct answers between the 26th and 35th trials, and Subject 7 did not show any convergence. Subject 10 reached the minimum comparison reflection density at the 14th trial after several incorrect answers, and the minimum reflection density was taken as one of five reversals. It seems Subjects 8 and 10 lost their concentration after a certain trial and could not recover.

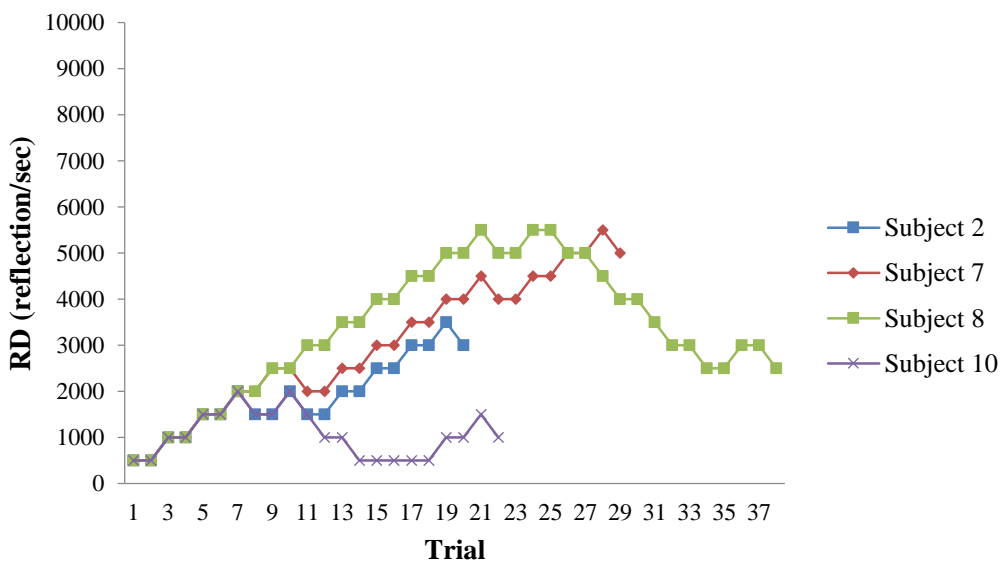


Figure 4.12 – Detailed results of individual trials from subjects demonstrating weak convergence in RT 1 sec case

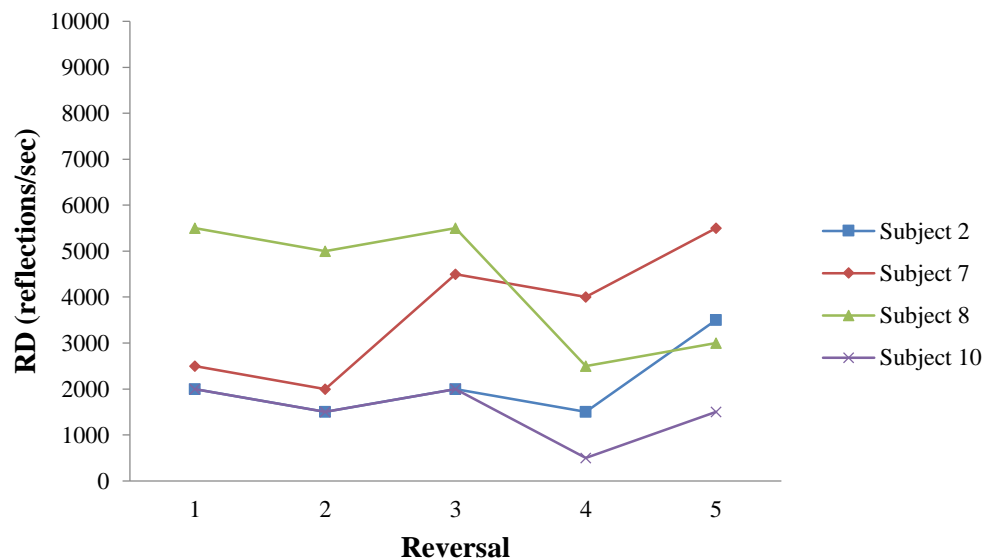


Figure 4.13 - Reversals from subjects demonstrating weak convergence in RT 1 sec case

### 4.3.3 *RT of 2 sec case*

Compared to RT 0.3 and 1 sec case, RT 2 sec had the largest number of excluded subjects from the mean calculation, due to large RSD and an exceptional subject who distinguished all comparison reflection densities (Table 4.4). Subject 5 reached the reference reflection density (10000 reflections/sec), and it could be either the upper limit of Subject 5 could be higher than the chosen reference reflection density or an extra cue rather than reflection density has been used for differentiation. A mean upper limit of distinguishable reflection density for assorted subjects was found to be 3875 reflections/sec (RSD = 18%, N = 4), whereas 3700 (RSD = 51%, N = 10) for all subjects.

Table 4.4 - Mean upper limit of distinguishable reflection densities with RT of 2 sec signals (All units in reflections/sec)

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
1	3000	2500	5000	4000	4500	3800	927	24%
2	4000	3500	4500	4000	7500	4700	1435	31%
3	5500	4500	5000	4500	5000	4900	374	8%
4	1000	500	4000	3000	3500	2400	1393	58%
5	7500	7000	8000	7500	10000	8000	1049	13%
6	3000	2000	3500	2500	3500	2900	583	20%
7	4000	3500	4000	3500	4500	3900	374	10%
8	1500	1000	2000	1500	3000	1800	678	38%
9	4000	2500	3000	2500	7000	3800	1691	45%
10	1000	500	1000	500	1000	800	245	31%
Mean (All)						3700	1880	51%
Mean (RSD<30%)						3875	708	18%

Assorted subjects had small RSD of upper limits (18%) between subjects. Figure 4.14 and Figure 4.15 shows reversals and responses by trials.

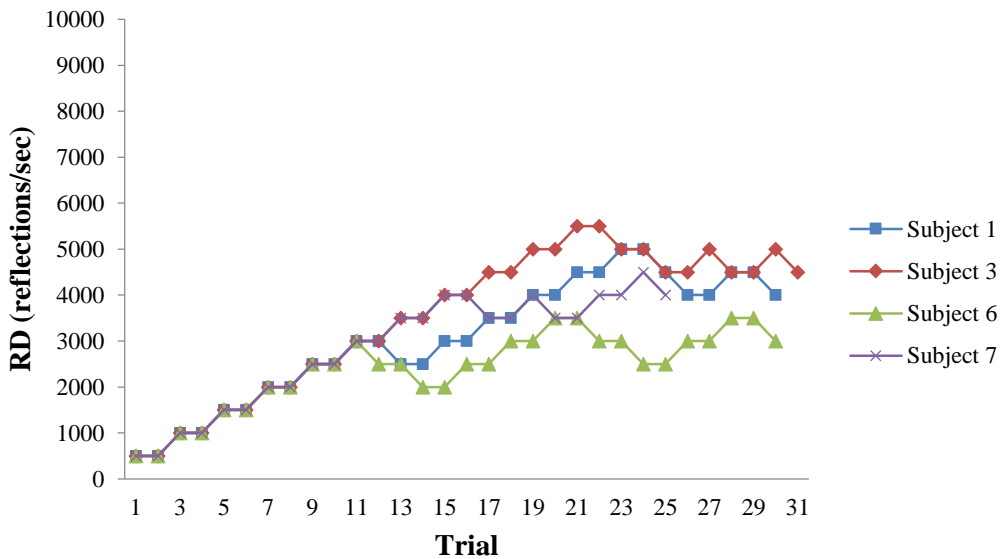


Figure 4.14 – Detailed results of individual trials from subjects demonstrating good convergence in RT 2 sec case

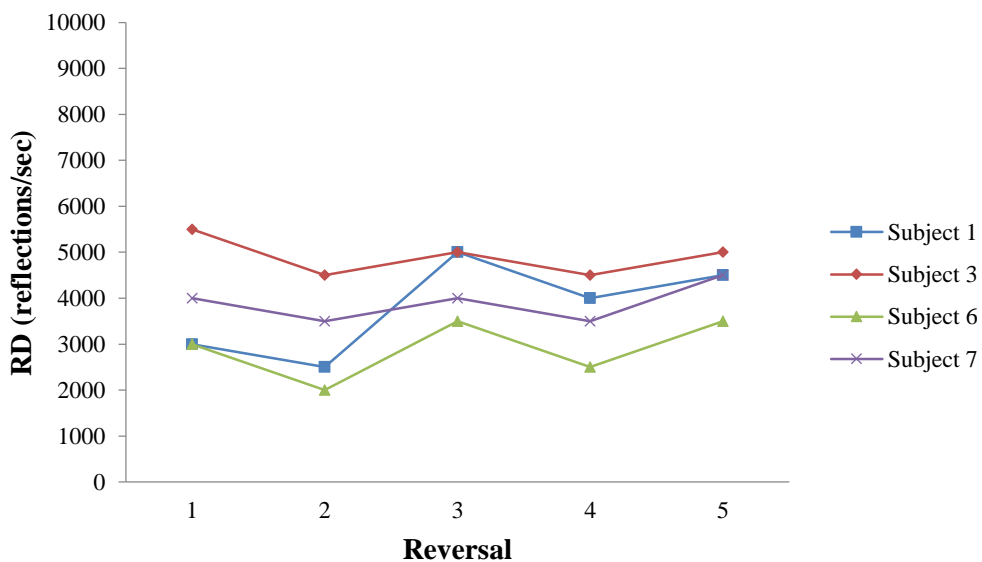


Figure 4.15 - Reversals from subjects demonstrating good convergence in RT 2 sec case

Results of subjects with the low convergence of reversals are shown in Figure 4.16 and Figure 4.17. Subject 2, 5, and 9 showed a steep increase from 4th reversal to 5th reversal. Subject 10 could distinguish only first comparison reflection density. Three subjects (Subject 2, 4, and 9) seemed they had a kind of an upper limit between around 2500 and 4000 reflections/sec, since most of their reversals located around that range. However, these responses could not be included in the mean calculation, because the rest of their reversals caused a large deviation.

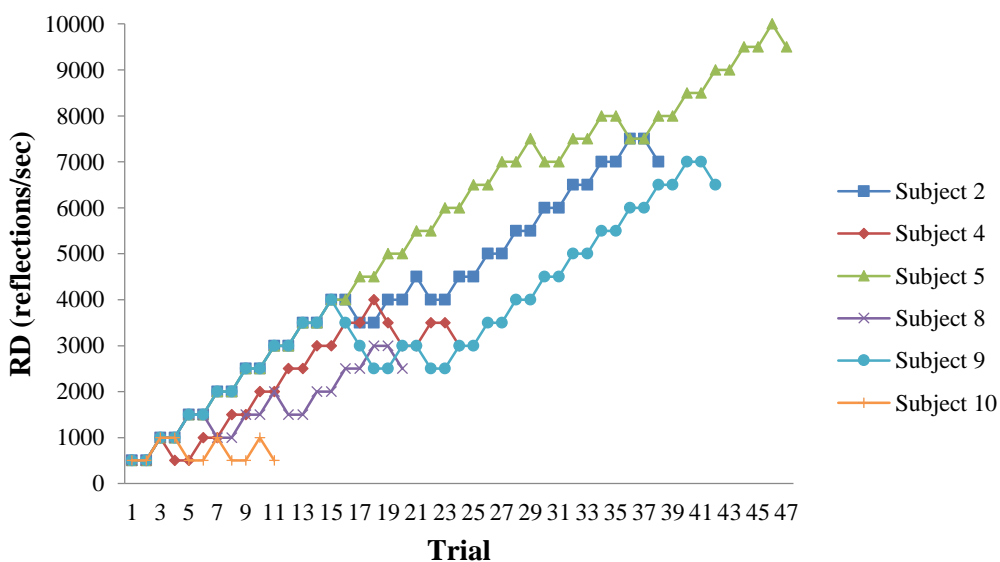


Figure 4.16 - Detailed results of individual trials from subject demonstrating weak convergence in RT 2 sec case

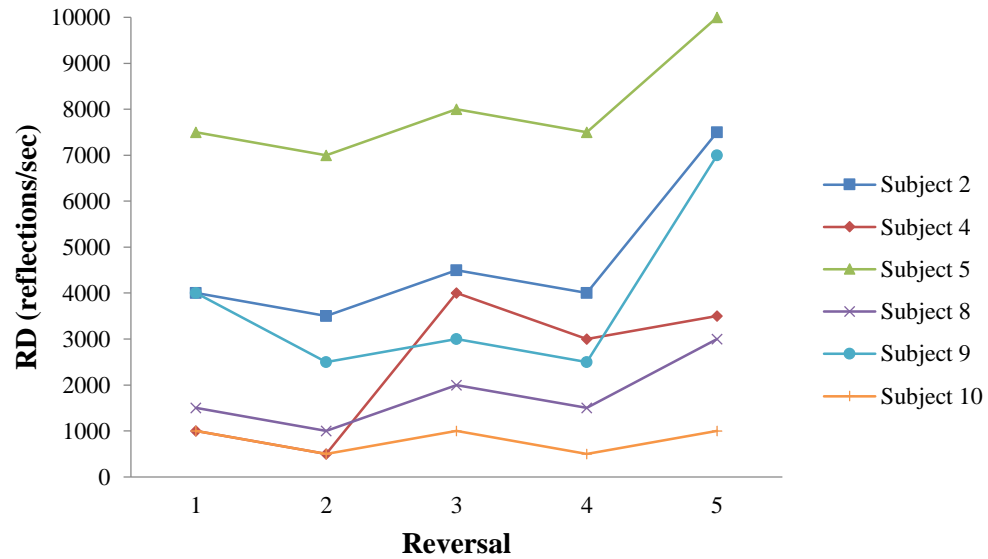


Figure 4.17 - Reversals from subjects demonstrating weak convergence in RT 2 sec case

#### 4.3.4 *Upper Limit of Distinguishable Reflection Density*

The upper limits of distinguishable reflection density are averaged only for those who demonstrated a good convergence of five reversals. A total of four to six subjects out of the ten were included for the mean calculation. Excluded data are shaded in Table 4.5.

Table 4.5 - Upper limit of distinguishable reflection densities (reflections/sec) for all three RT conditions (shaded values are excluded from mean calculation)

RT (sec)	Subject										Mean	SD
	1	2	3	4	5	6	7	8	9	10		
0.3	4400	2500	2900	1200	3300	2800	2600	2900	1500	2300	3000 (N = 6)	702
1	7500	2100	7900	1300	5100	2900	3700	4300	4300	1500	4075 (N = 4)	2426
2	3800	4700	4900	2400	8000	2900	3900	1800	3800	800	3875 (N = 4)	708

Figure 4.18 shows the mean upper limits of distinguishable reflection density for the three different RTs (0.3, 1, and 2 sec). The upper limit for the RT 0.3 sec condition was found to be 3000 reflections/sec (N = 6), 4075 reflections/sec (N = 4) for the RT 1 sec condition, and 3875 reflections/sec (N = 4) for the RT 2 sec condition. The standard deviations for the RT 0.3 and 2 sec case were similar and smaller than for the RT 1 sec case. The mean upper limit of distinguishable reflection densities for these three RT conditions were found to be similar as shown in Figure 4.18. The average of the upper limits from the three RT conditions was around 3700 reflections/sec.



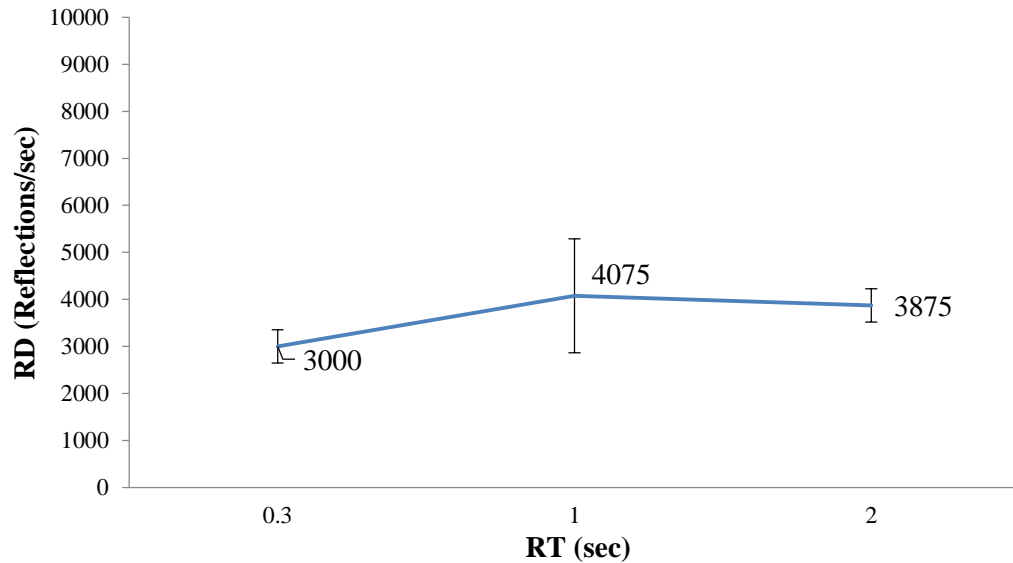


Figure 4.18 - Upper limits of distinguishable reflection density with different RT conditions

#### 4.4 Summary and Conclusions

The upper limit of distinguishable reflection densities using a clapping signal under different RT conditions was investigated using artificially generated impulse responses with constant reflection density, and tested across RTs: 0.3, 1, and 2 sec. Some results were excluded due to large relative standard deviations (more than 30%). The mean upper limits were found to be 3000, 4075, and 3875 reflections/sec for the RT of 0.3, 1, and 2 sec conditions, respectively. The final value averaged across the RT conditions was around 3700 reflections/sec.

## **Chapter 5 - Study 2: The Maximum Distinguishable Reflection Density with Simulated Impulse Responses**

### **5.1 Introduction**

This chapter focused on the testing the subjective perception of reflection densities from different room impulse responses. The room impulse responses were generated from simulated rooms by Odeon, and reflection densities were controlled by variation of room size. All rooms had reverberation time ( $T_{30}$ ) of 1 sec from 63 Hz to 8000 Hz by adjusting magnitude of absorption of all surfaces uniformly.

### **5.2 Methodology**

#### **5.2.1 *Test Stimuli***

##### *5.2.1.1 Source Signals*

As presented in Section 3.4, two anechoic source signals were used: clapping and speech. Both the clapping and speech sound samples were cropped from signals that came with Odeon version 11. The clapping signal included five claps while the speech sentence stated: ‘When you are applying for a job, you need to have a good resume prepared.’ These source signals were convolved in Odeon with the binaural room

impulse responses.

### 5.2.1.2 Impulse Responses

The room model 'Example room' from Odeon was used to simulate room impulse responses. A total of twenty different sizes of rooms were used, and the relative location of the sound source in each room was maintained across all twenty room sizes as illustrated in Figure 5.1.

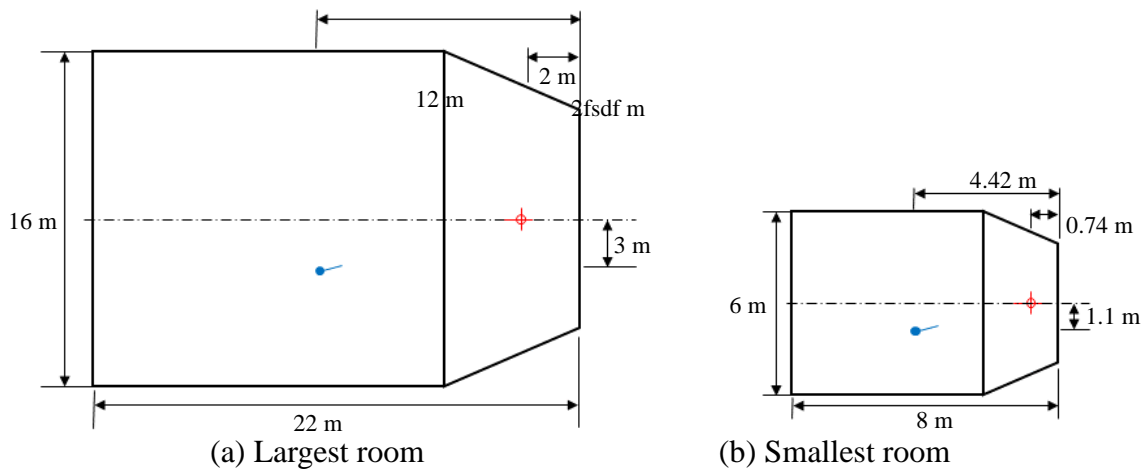


Figure 5.1 - Relative locations of a source (red) and receiver (blue) in the largest and the smallest room used in the study

The largest room volume was about  $2923 \text{ m}^3$ , and the smallest room was about  $146 \text{ m}^3$ . The source was located at the front center, and the receiver was at the middle of the room slightly to the right side. Twenty room sizes were selected to have a linear variation of reflection density, as calculated from Odeon and given as a single value

representing late reflection density. The reflection density of twenty rooms in this study is shown in Figure 5.2. All rooms had a reverberation time ( $T_{30}$ ) of 1 sec from 63 Hz to 8000 Hz by adjusting magnitude of absorption of all surfaces uniformly. The reflection density of twenty rooms in this study is shown in Table 5.1.

Table 5.1 - Reflection density (RD) of sound samples, based on Odeon-provided calculation

RD label	Volume (m <sup>3</sup> )	RD (/ms)	RD label	Volume (m <sup>3</sup> )	RD (/ms)
1	2923	357	11	423	705
2	2261	391	12	372	740
3	1769	427	13	325	774
4	1429	460	14	285	809
5	1154	496	15	253	844
6	961	530	16	224	878
7	812	564	17	199	912
8	671	601	18	180	948
9	575	634	19	161	983
10	490	670	20 (Reference)	146	1017

Reflection densities were calculated by the theoretical reflection density equation,  $cS/4V$ , for all rooms (Table 5.2). The volume ( $V$ ) and surface area ( $S$ ) values were taken from Odeon, and  $c$  was taken to be 343 m/s. The calculated reflection densities were also approximately linear across the 20 rooms (Figure 5.2).

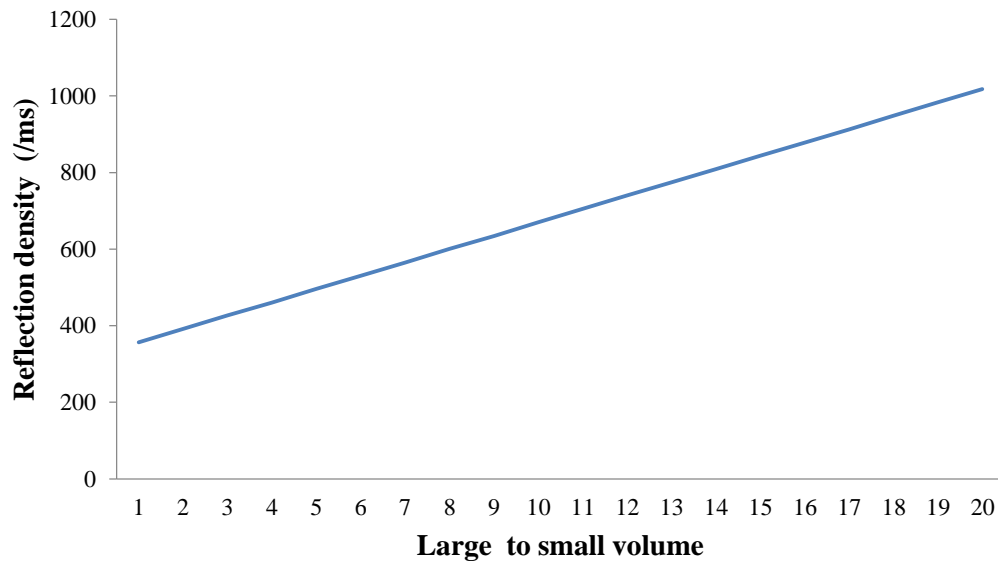


Figure 5.2 - Reflection density calculated from Odeon for each of the different room sizes

Table 5.2 - Reflection density (RD) of sound samples, based on the classic reflection density equation,  $cS/4V$ , and the Odeon-provided calculation

RD label	Odeon (/ms)	$cS/4V$ (/s)	Room label	Odeon (/ms)	$cS/4V$ (/s)
1	357	37	11	705	71
2	391	40	12	740	74
3	427	44	13	774	77
4	460	47	14	809	81
5	496	51	15	844	84
6	530	54	16	878	88
7	564	57	17	912	91
8	601	61	18	948	94
9	634	64	19	983	98
10	670	68	20	1017	101

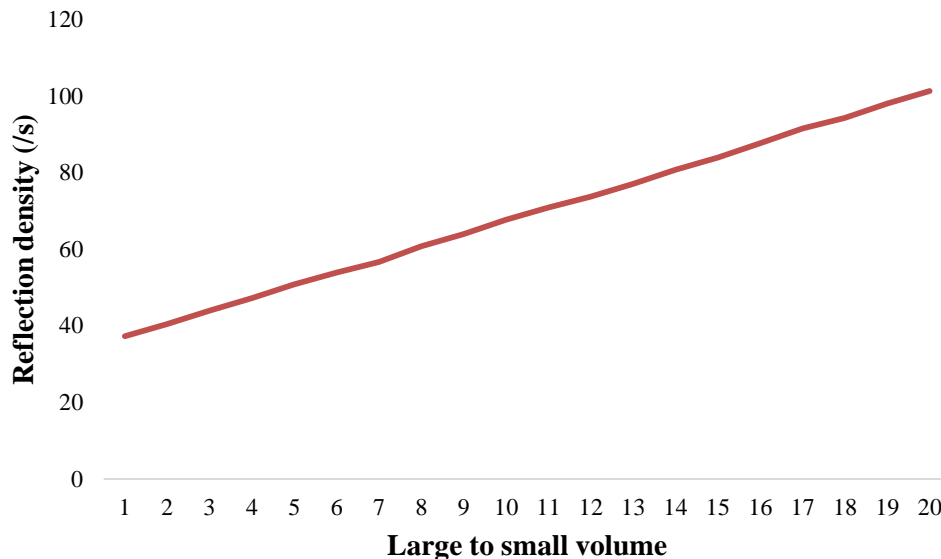


Figure 5.3 - Reflection density calculated by  $cS/4V$  from each of the different room sizes

As a third method, a cut-off level and a time window were used to quantify reflection density. Since ISO 3382-1:2009 recommends a source level of 45 dB above the background level for  $T_{30}$  reverberation time measurements, most measured impulse responses are expected to have a source level of minimum 35 dB above the background noise level. For this reason, the cut-off level of -35 dB from a direct sound was selected and -50 dB was also examined for comparison. Since binaural impulse responses may be different between the left and right channels, monaural impulse responses were used for this quantification method. Time windows were tested in increasing 50 ms increments, out to a maximum of 1 sec, and then reflections above each cut-off level were counted until the limit of each time window was reached. An example of this method is

depicted in Figure 5.4.

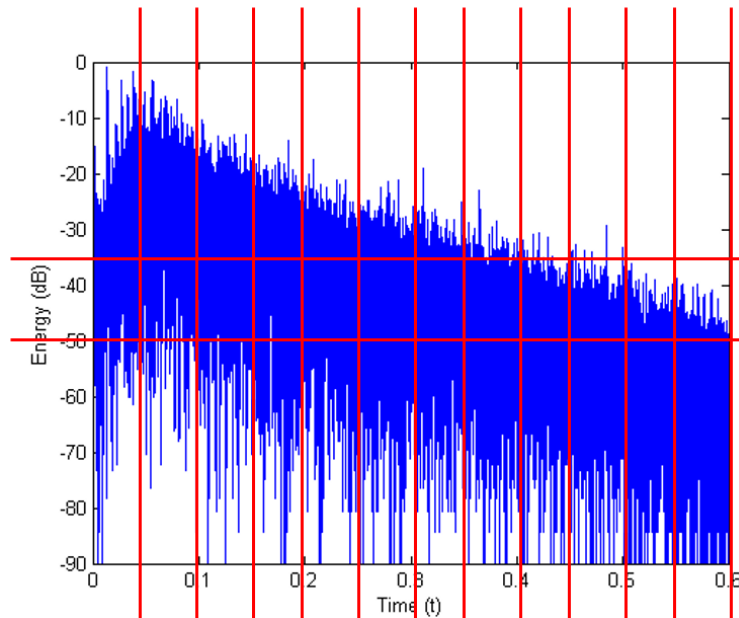


Figure 5.4 - Sample impulse response with overlaying grid of different cut-off levels and time windows used to quantify reflection density

A Matlab code was applied to the Odeon simulated impulse responses to determine the number of reflections. The sampling frequency of the impulse responses was 44,100 Hz. Figure 5.5 shows the total number of reflections from the different combinations of cut-off level and time windows for each of the twenty rooms. The number of reflections was expected to have a linear or gradual change by room size variation, based on the theoretical reflection density equation; however, the results showed many fluctuations.

There were three quantification methods of the reflection density considered in

this dissertation (Chapter 3.2):  $cS/4V$ , Odeon-provided value, and dB cut-off. Among these methods, the dB cut-off method was not appropriate to use since it showed many fluctuations with room size variation which is different from expectation as described above.  $cS/4V$  is a theoretical mean value appropriate for diffuse fields and also could not be used, as  $cS/4V$  does not have meaningful interpretation with respect to these simulated impulse responses (values are shown in Table 5.2, though). The remaining Odeon-provided reflection density was used for quantification of reflection densities in impulse responses simulated by Odeon.

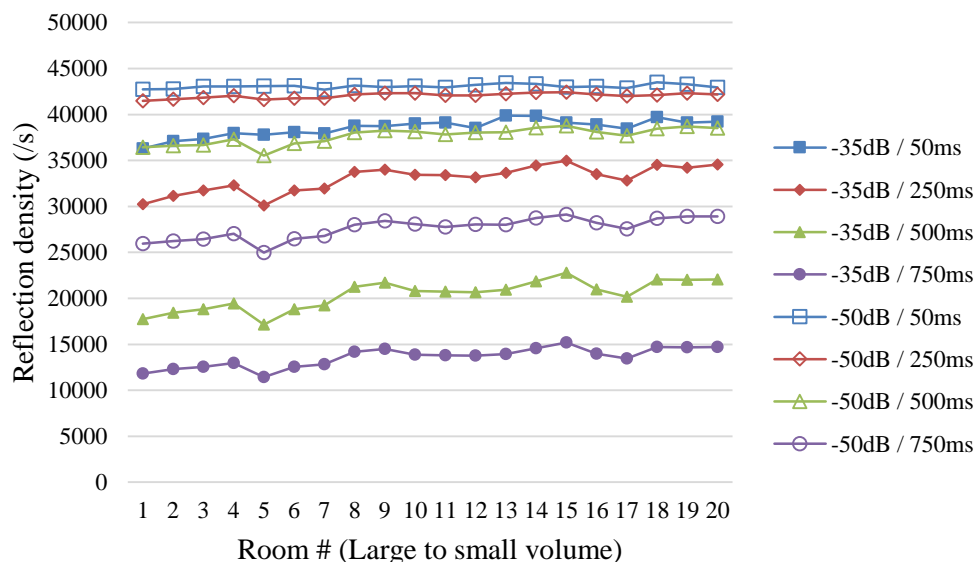


Figure 5.5 - Reflection density based on the cut-off level and time window shown in the legend



### 5.2.1.3 *Convolution and Loudness Normalization*

Source signals were convolved with binaural impulse responses from Odeon. Due to the difference in room volumes and distances between source and receiver, the loudness of the convolved signals from each room varied. To maintain the same loudness, algorithm ITU-R in the Adobe Audition software was used. The loudness was normalized at -23 LUFS (Loudness Units relative to Full Scale). Prepared test stimuli were presented to participants on a laptop using a custom Matlab GUI. The tests were conducted in a sound booth and presented over headphones to the subjects. The loudness from headphone was maintained around 65 dBA (re 20  $\mu$ Pa).

### 5.2.2 *Three-Alternative Forced-Choice Method with 1Up-2Down Adaptive Method*

Test subjects participated in two separate sessions, each using different source signals: clapping and speech. Each session lasted last 30 minutes to 45 minutes. In each trial, the subject was asked to select which of three samples sounds was different from others. The subject heard each of the sound samples three times and then was asked to answer the question after the signal presentations were completed.

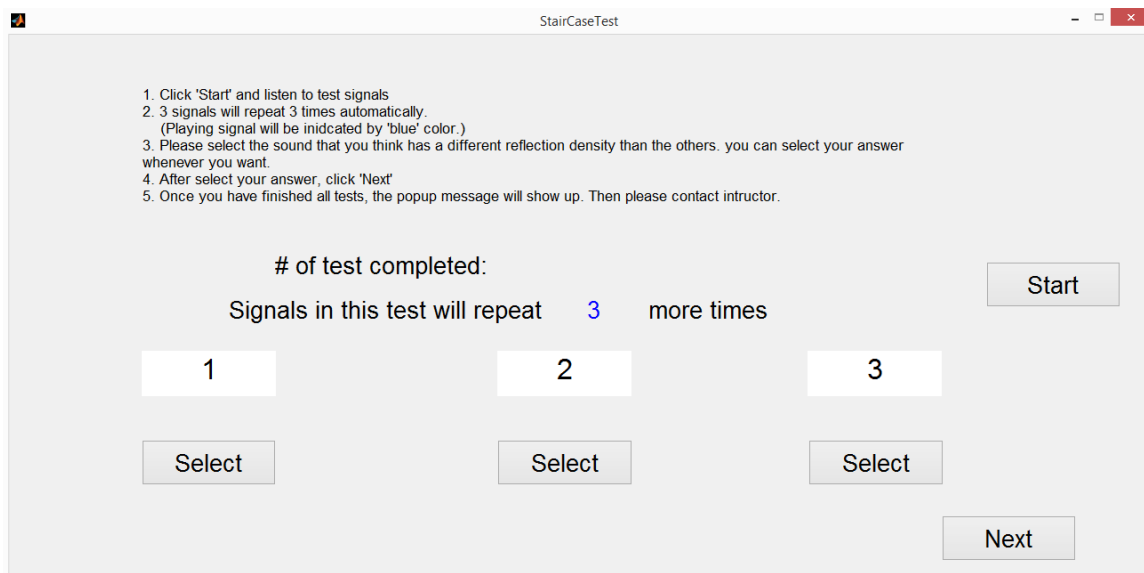


Figure 5.6 - Sample screenshot of the testing program

In total, these sessions including subject screening lasted no more than 90 minutes. A three-alternative forced-choice (3AFC) method combined with a 1-up 2-down adaptive method was used to determine the upper limit of distinguishable reflection density produced from the twenty simulated rooms.

The three-alternative force-choice testing method presents two identical reference reflection density (RD) samples and one comparison RD sample in each trial. The reference RD is designated reflection density that performs as a point of comparison; in this study, it was the highest RD obtained from the smallest room among twenty. The reference RD was approached from below (lower RD). With a 1-up 2-down method, two consecutive correct answers decreased (down) a difference between the reference and

comparison RD, and one incorrect answer increased (up) the difference. In other words, two consecutive correct answers increased comparison RD, and one incorrect answer decreased comparison RD. A reversal was recorded when a participant made changes of direction from up to down or down to up. After five reversals had been made, an average comparison RD at those reversals was reported as the participant's upper limit of distinguishable reflection density.

### **5.2.3 *Participants***

A total of twenty-one listeners (11 male, 10 female) were recruited on the University of Nebraska at Omaha campus, and twenty subjects (10 male, 10 female) finished all tests. One subject (Subject# 6) failed the initial screening, so he was excluded from the study. A total of ten subjects were native English speakers among participated twenty subjects. The average age was 25.5 years (SD = 6.5 years). All listeners had pure tone thresholds below 25 dB hearing level between 250 and 8000 Hz. Listeners provided informed consent for their participation in the study and were paid \$15 Amazon gift card for their time.

## **5.3 Result and Analysis**

The results showed a relatively large variation among subjects. For this reason, this subchapter will analyze grouped results to show detailed interpretation of subject

responses.

### **5.3.1 *Clapping Results***

Results of tests with clapping signal could be categorized by three groups: i) good convergence, ii) weak convergence and iii) ability to distinguish all.

#### **5.3.1.1 *Group Showing Good Convergence***

If participants performed the test well by differentiating comparison RDs and the reference RD accurately, then the reversals are expected to converge well on a certain RD with small variance. Table 5.3 shows results from assorted subjects who achieved good convergence with relative standard deviation (RSD) equal or less than 30%. The last trial of each subject was a final reversal. An asterisk mark (\*) indicates subjects who distinguished all comparison samples, who reached the reference reflection density (RD label of 20). These subjects are grouped separately later in this chapter and excluded from the mean calculation, since the upper limit could be beyond the reference reflection density. There were errors in the testing for Subject 13, 14, and 18, who had to stop the testing after the 4<sup>th</sup> reversal. Their mean upper limits of distinguishable reflection density are consequently averaged only four reversals.

Table 5.3 - Assorted clapping results with good convergence (RSD < 30%) of reversals

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
4*	13	12	20*	19	20*	16.8	3.5	21%
5	12	11	17	16	17	14.6	2.6	18%
7	5	4	8	6	8	6.2	1.6	26%
10*	19	18	20*	19	20*	19.2	0.7	4%
12	11	10	11	10	11	10.6	0.5	5%
13	8	7	10	8	·	8.3	1.1	13%
14	7	5	6	5	·	5.8	0.8	14%
17	4	3	6	4	6	4.6	1.2	26%
18	14	10	11	10	·	11.3	1.6	15%
19	15	14	15	14	17	15.0	1.1	7%
Mean (RSD<30%)						11.2	4.8	43%

Figure 5.7 and Figure 5.8 depict responses of eight subjects who demonstrated good convergence as listed in Table 5.3. The Y-axis is marked by RD label listed in Table 5.2, and RD label 20 is the reference RD.

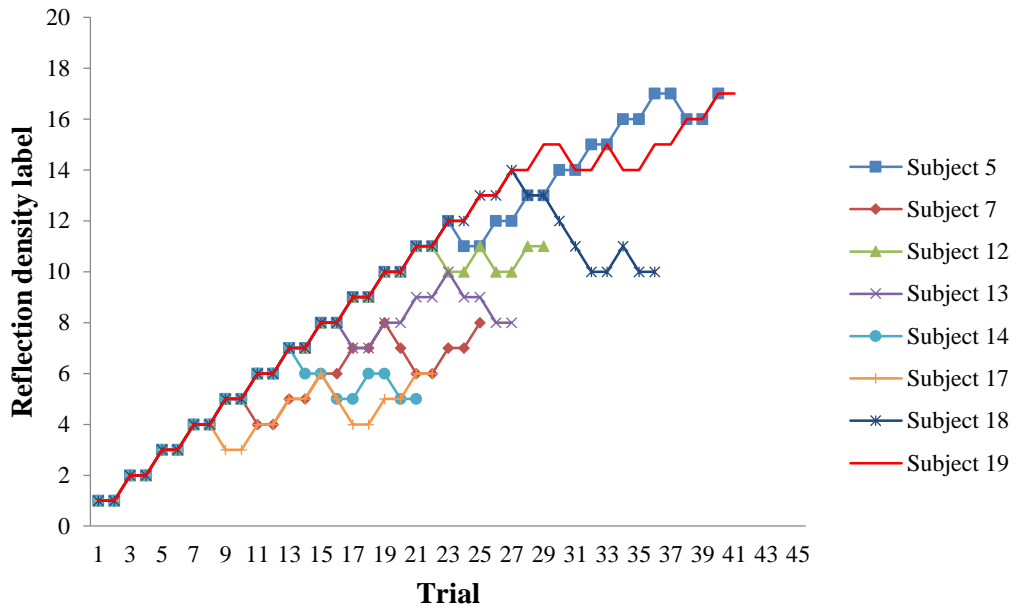


Figure 5.7 Detailed results of individual clapping trials from subjects demonstrating good convergence

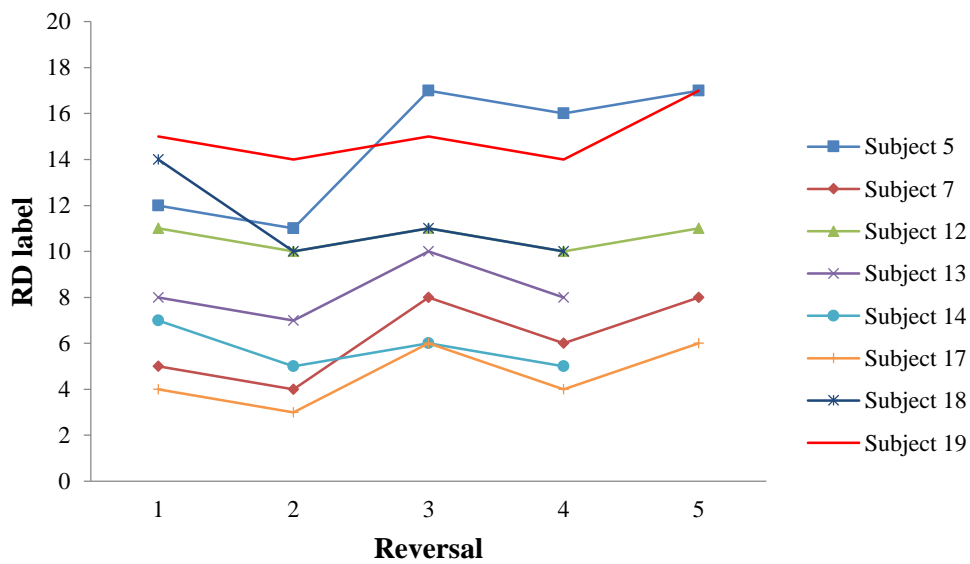


Figure 5.8 Reversals from clapping tests of subjects demonstrating good convergence

Reversals of Subject 12 converged within a range of 1 RD label difference, while Subjects 7, 13, 14, 17, 18, and 19 converged within a range of 4 RD label difference. Although reversals of Subject 5 converged within a range of 5 RD label difference, the relative standard deviation between these was 18%, so it was included in the good convergence group. These results are considered as ‘well converged’ compared to other results that will be described below.

### 5.3.1.2 Group with Weak Convergence

Some subjects showed some convergence in their result, however, the convergence was not strong. Table 5.4 lists subjects who demonstrated weak convergence with relative standard deviation (RSD) equal or more than 30%.

Table 5.4 - Assorted clapping results with a weak convergence ( $RSD \geq 30\%$ ) of reversals

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
1	3	2	4	3	5	3.4	1.0	30%
2	2	1	2	1	6	2.4	1.9	77%
3	2	1	5	4	5	3.4	1.6	48%
8	2	1	13	12	16	8.8	6.1	69%
9	1	6	5	6	5	4.6	1.9	40%
11	2	1	4	3	7	3.4	2.1	61%
15	4	3	6	5	15	6.6	4.3	65%
16	2	1	6	5	6	4.0	2.1	52%
20	1	2	1	5	3	2.4	1.5	62%
21	1	5	3	11	8	5.6	3.6	63%
Mean (RSD>30%)						4.5	1.9	43%

Subjects 3 and 16 had similar reversals from 3rd to 5th reversals, and Subject 9 had similar reversals from 2nd to 5th reversals, as shown in Figure 5.9 and Figure 5.10. Subject 16 could not choose two consecutively correct answers from the 5th to 9th trials, so RD label 1 was recorded as one of the reversals for averaging.

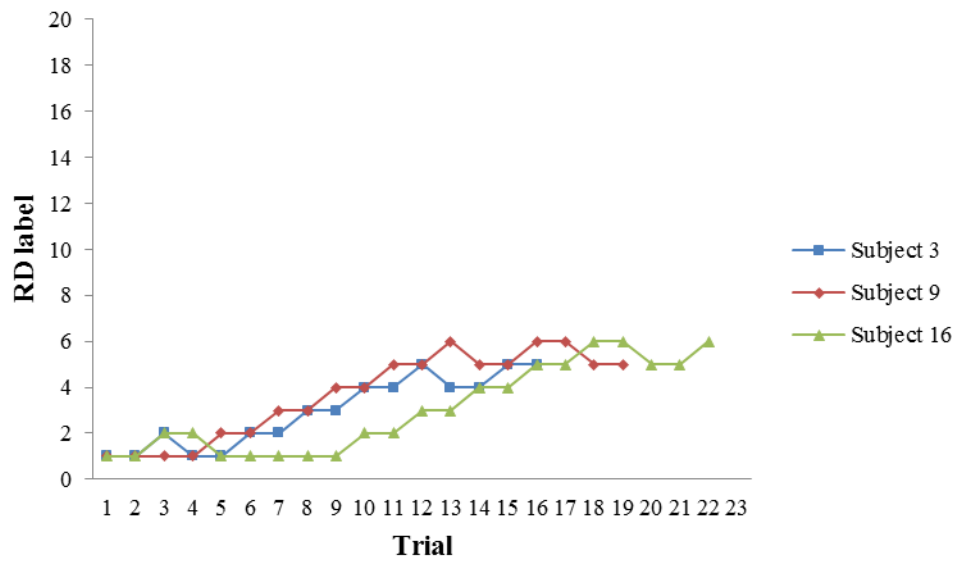


Figure 5.9 - Detailed results of individual clapping trials from subjects demonstrating convergence at later reversals among the weak convergence group



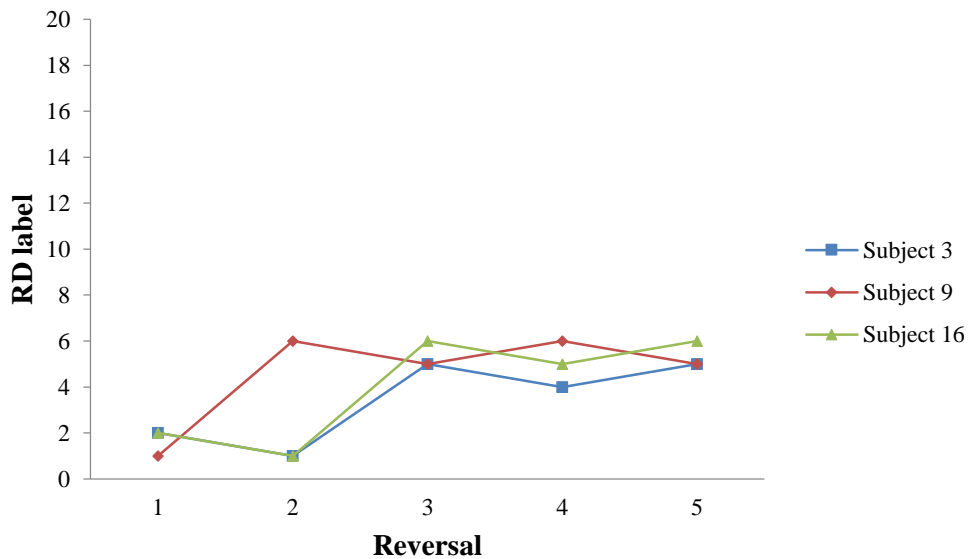


Figure 5.10 - Reversals from clapping tests of subjects demonstrating convergence at later reversals among the weak convergence group

The remaining subjects in the weak convergence group also have been excluded from the final mean calculation presented later in this chapter; their results are depicted in Figure 5.11 and Figure 5.12.

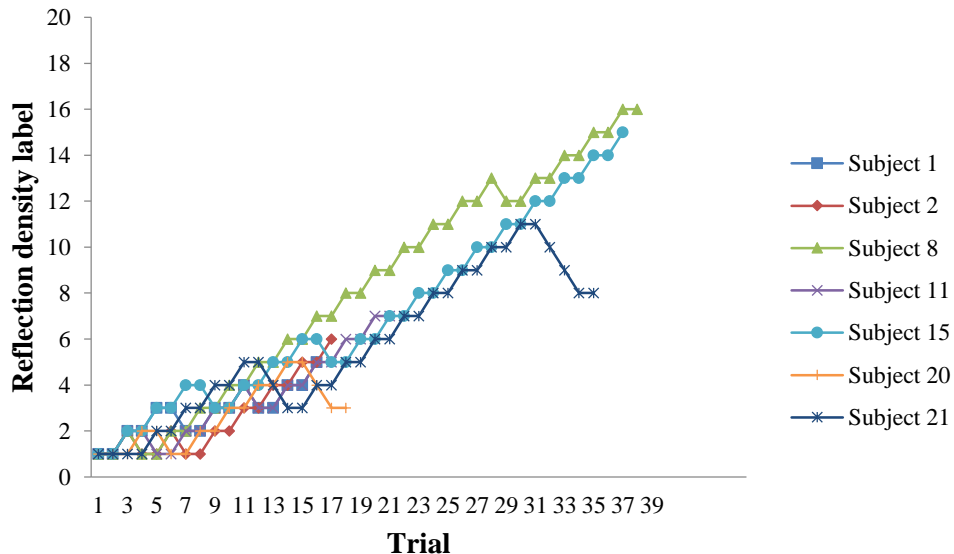


Figure 5.11 - Detailed results of individual clapping trials from subjects excluded from final mean calculation

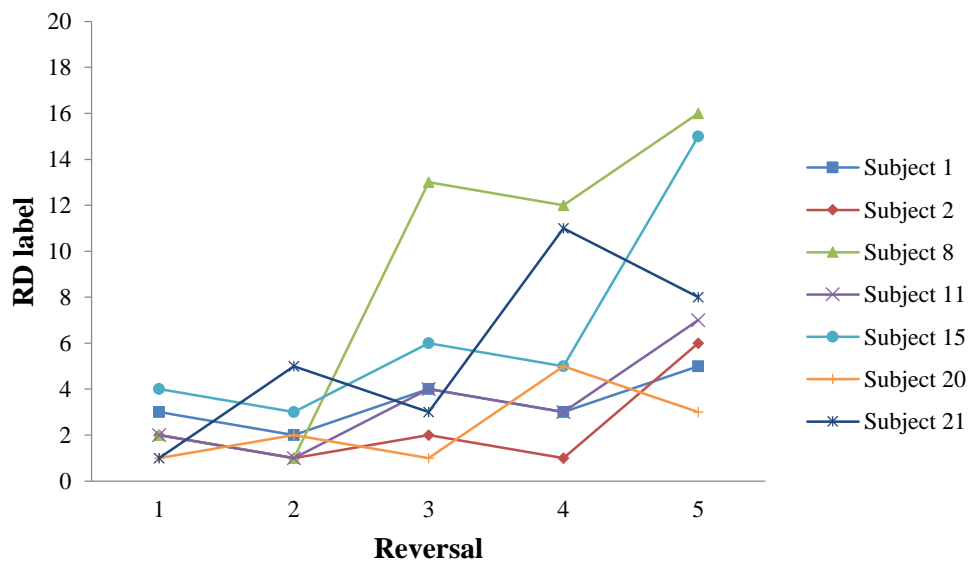


Figure 5.12 - Reversals from clapping tests of subjects excluded from the final mean calculation

Subjects 1 and 11 showed a gradual increase of reversals and Subjects 2, 8, and 15 showed steep increases. The last reversals of Subjects 20 and 21 decreased, but the overall tendency was increasing. Subject 21 seemed to get tired and lost their concentration. This can be observed by consecutive incorrect answers (downward movement) after 26 trials. For these reasons, these subjects were excluded.

#### 5.3.1.3 *Group with Ability to Distinguish All*

Subject 4 and Subject 10 were able to distinguish all of the comparison RD up to the reference RD (Figure 5.13). Both subjects made some wrong answers before reaching the reference reflection density (RD label 20), but eventually differentiated all comparison reflection density. This could indicate that there was another cue in the test samples, other than reflection density, that some could use to distinguish samples. The additional cues could be frequency spectrum differences, small loudness differences (even though loudness normalization was applied), or differences in another room acoustic parameter.

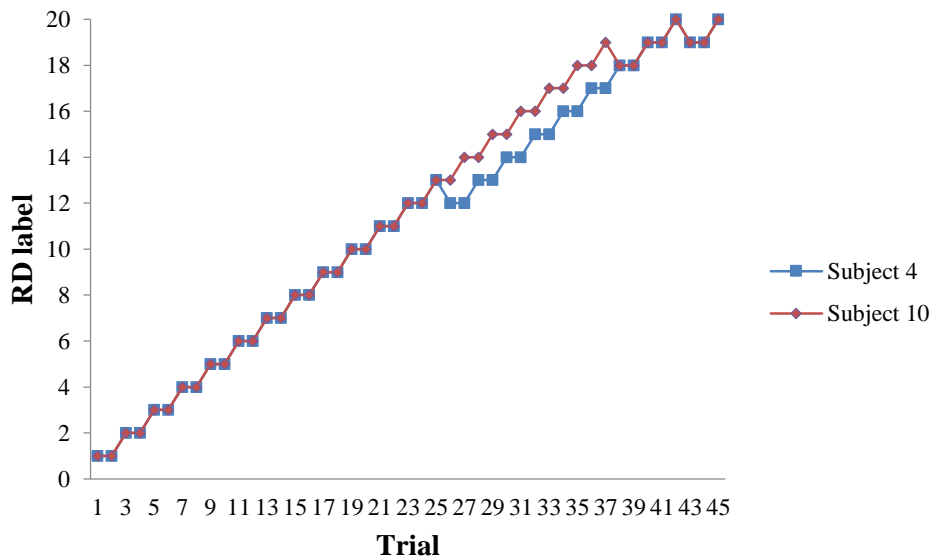


Figure 5.13 - Detailed results of individual clapping trials from subjects demonstrating ability to distinguish all comparison RD

### 5.3.2 *Speech Results*

Results of tests with speech signal can be categorized in two groups: those showing i) ability to distinguish most and ii) weak convergence. The overall standard deviation was much larger than from clapping signals.

#### 5.3.2.1 *Group with Ability to Distinguish Most*

As with the clapping signal, some subjects could distinguish all comparison RD for the speech signal, and the portion of these subjects out of the total participants was bigger than for the clapping signal case. The clapping signal had only five discontinuous claps whereas the speech signal was continuous, so it appears that test subjects were able

to differentiate comparison RD signals more with the speech signal than with clapping.

Four subjects could distinguished most comparison samples, and two subjects (Subject 9 and 10, marked by asterisk mark (\*) in Table 5.5) distinguished all. Results from Subjects 9 and 10 were excluded from the final mean calculation.

Table 5.5 - Assorted speech results with a good convergence (RSD < 30%) of reversals

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
4	18	17	19	18	19	18.2	0.7	4%
9*	20*	19	20*	19	20*	19.6	0.5	2%
10*	20*	·	·	·	·	20.0	0.0	0%
18	17	16	17	16	17	16.6	0.5	3%
Mean (RSD<30%)						17.4	0.8	5%

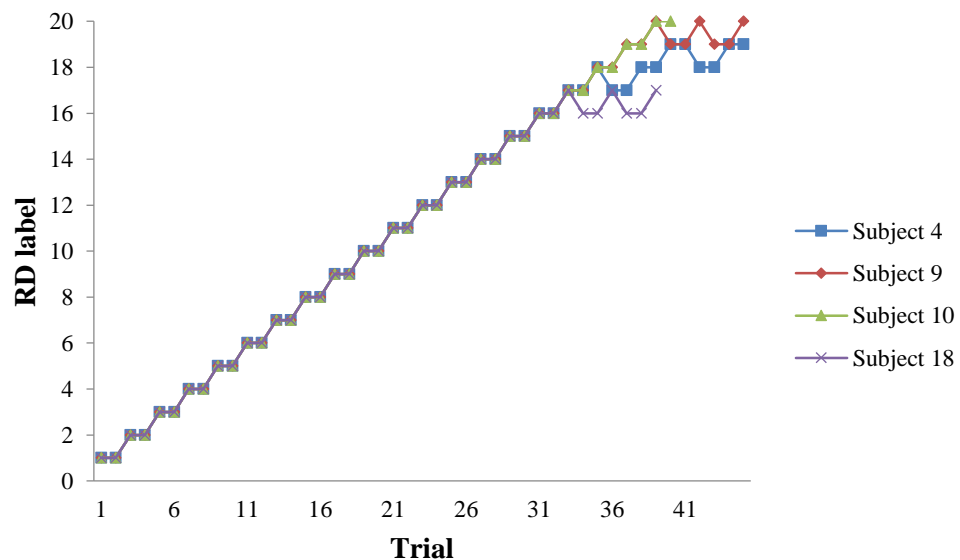


Figure 5.14 - Detailed results of individual speech trials from subjects demonstrating ability to distinguish most comparison samples

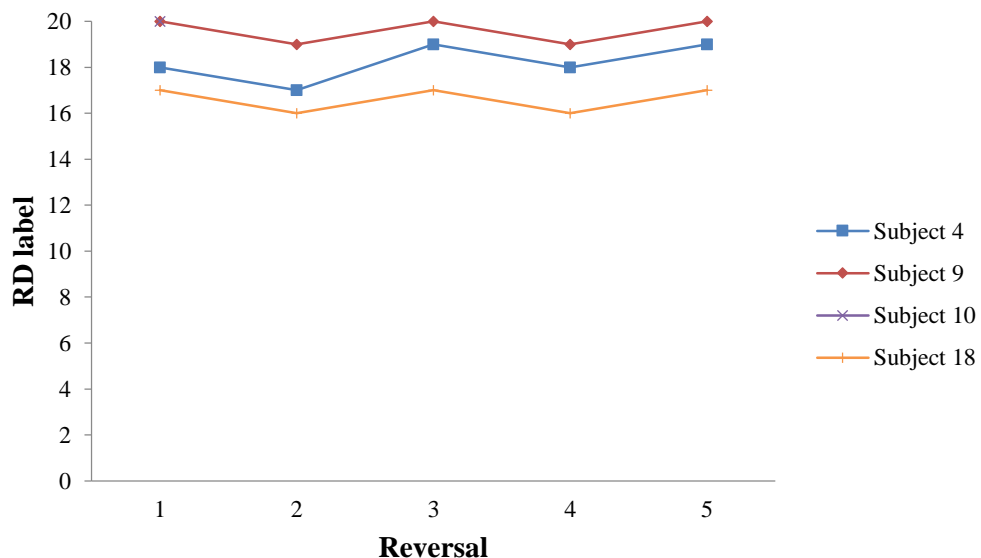


Figure 5.15 - Reversals from speech tests of subjects demonstrating ability to distinguish most comparison samples

In Figure 5.14 and Figure 5.15, Subjects 9 and 10 differentiated all comparison RD without any wrong answers. Subject 10 was asked to stop after the investigator checked that Subject 10 reached the reference RD, so that case has only one reversal.

### 5.3.2.2 *Weak or No Convergence*

Subjects shown in

Table 5.6 had weak convergence with RSD of more than 30%. Subjects 2, 13, and 16 were unable to complete a full test due to a malfunction in the Matlab testing program.

Table 5.6 - Assorted speech results with a weak convergence ( $RSD \geq 30\%$ ) of reversals

Subject	Reversal					Mean	SD	RSD (%)
	1	2	3	4	5			
1	1	2	1	2	1	1.4	0.5	35%
2	1	6	·	·	·	3.5	2.5	71%
3	2	1	2	1	2	1.6	0.5	31%
5	1	3	2	3	1	2.0	0.9	45%
7	1	2	1	2	1	1.4	0.5	35%
8	1	16	15	16	15	12.6	5.8	46%
11	2	1	2	1	7	2.6	2.2	86%
12	1	2	1	2	1	1.4	0.5	35%
13	1	5	2	7	·	3.8	2.4	64%
14	1	2	1	5	4	2.6	1.6	62%
15	2	1	6	5	7	4.2	2.3	55%
16	1	19	12	·	·	10.7	7.4	69%
17	1	2	1	7	6	3.4	2.6	76%
19	6	5	13	11	14	9.8	3.7	37%
20	1	6	5	15	14	8.2	5.4	66%
21	4	3	6	5	13	6.2	3.5	57%

There were also a few subjects who did not reach above RD label 3, so they apparently could not distinguish differences at all, or did not understand the testing instructions (Figure 5.17). These subjects did demonstrate five reversals; however, many of those reversals were at RD label 1. For this reason, these subjects were excluded from the final mean calculation.

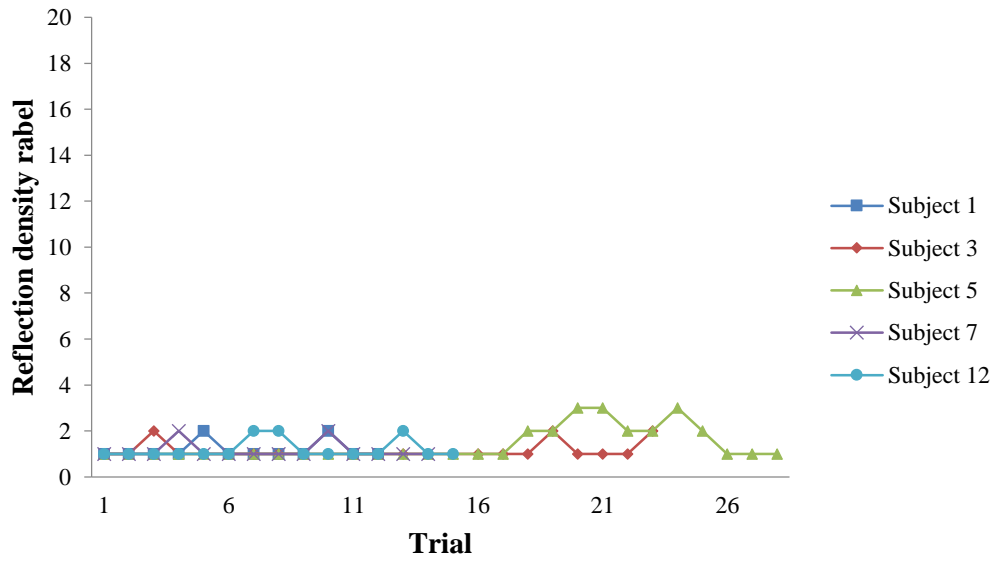


Figure 5.16 - Detailed results of individual speech trials from subjects demonstrating inability to distinguish differences

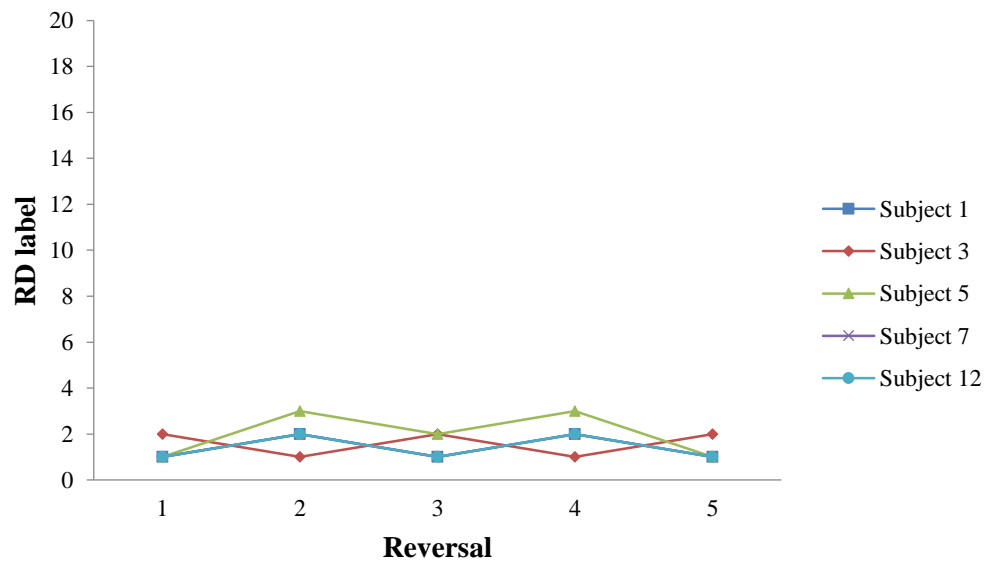


Figure 5.17 - Reversals from speech tests of subjects demonstrating inability to distinguish differences



Six subjects (Subject 2, 11, 16, 19, 20, and 21) were excluded due to the weak or no convergence of reversals. The 5th reversals of Subjects 11 and 21 increased steeply, Subjects 2 and 19 showed an increasing tendency of reversals. Subject 16 made big drop after the 2nd reversal and had to stop due to the malfunction of the testing program.

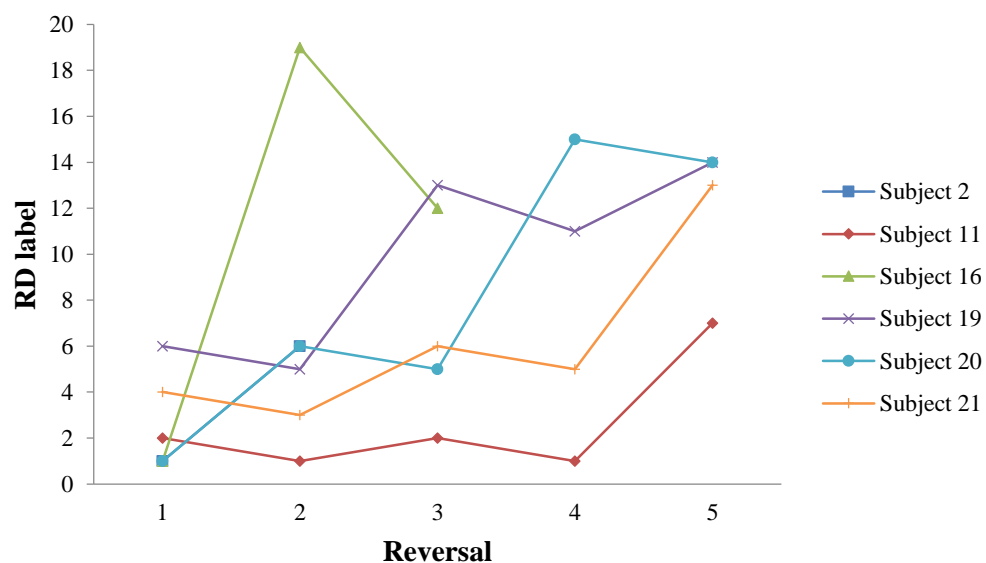


Figure 5.18 - Reversals from speech tests of subjects excluded from mean calculation

The remaining subjects in the ‘weak or no convergence’ group for speech signals were also excluded from the final mean calculation, even though these results demonstrated some convergence in latter trials. These included Subjects 8, 13, 14, 15, and 17 (Figure 5.20). Subject 8 reached RD label 16, while the reversals of the others

remained below RD label 7. Although Subjects 13, 14, 15, and 17 made a reversal at the minimum reflection density as subjects in Figure 5.17 did, they made other reversals higher than RD label 3.

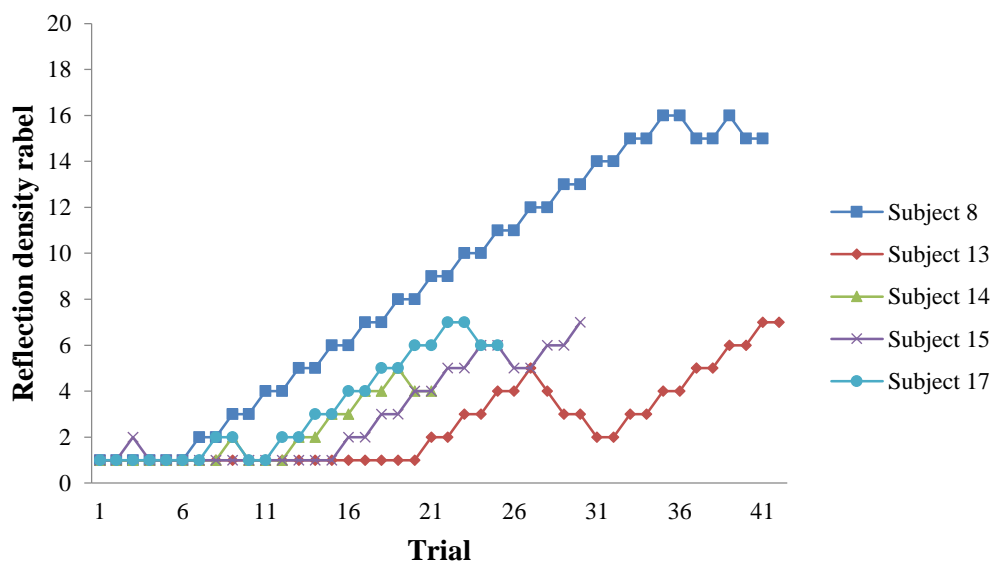


Figure 5.19 - Detailed results of individual speech trials from subjects among the weak or no convergence group, included in final mean calculation

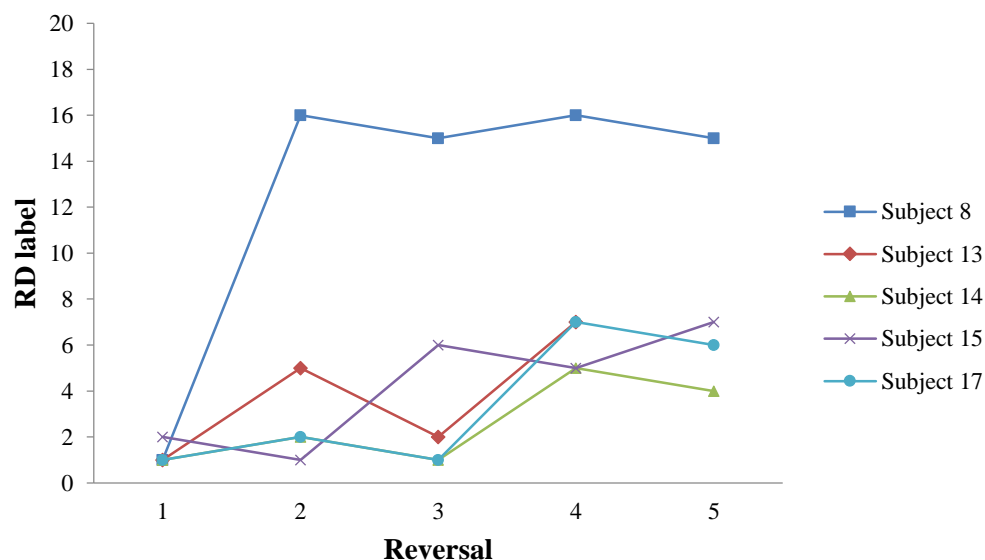


Figure 5.20 - Reversals from speech tests of subjects among the weak or no convergence group, included in final mean calculation

### 5.3.3 Upper Limit of Reflection Density

The results of both the clapping and speech signals showed large variance among subjects, so only those results with reasonable convergence of reversals were included to determine the mean of the upper limit of distinguishable reflection density for clapping and speech signals as depicted in Figure 5.21. The calculated upper limit for each subject was converted from the RD label to the Odeon calculated values (reflections/ms) (Table 5.7). Shaded cells are excluded from mean calculation, due to reasons presented in the sections above.

Table 5.7 - Upper limits of distinguishable reflection density of subjects

Subject	Clapping	Speech	Subject	Clapping	Speech
1	440	371	12	691	371
2	405	444	13	611	453
3	440	377	14	523	413
4	905	955	15	550	467
5	830	391	16	460	695
7	537	371	17	482	440
8	627	760	18	716	898
9	482	1003	19	844	663
10	990	1017	20	405	608
11	440	413	21	516	537
			Mean (RSD<30%)	654 (N=8)	858 (N=2)
			SD	130	98

The median RD of the upper limit of reflection density for clapping signals was found to be 684 /ms, and 858 /ms for speech signals. The mean RDs for clapping and speech signals were found to be 654 reflections/ms (SD = 130, N = 8) and 858 reflections/ms (SD = 98, N = 2), respectively (Figure 5.21). Since only two subjects were included for the mean calculation of the upper limit of reflection density of speech signals, it is hard to conclude that these two subjects represent a broad group of listeners under the speech signals.

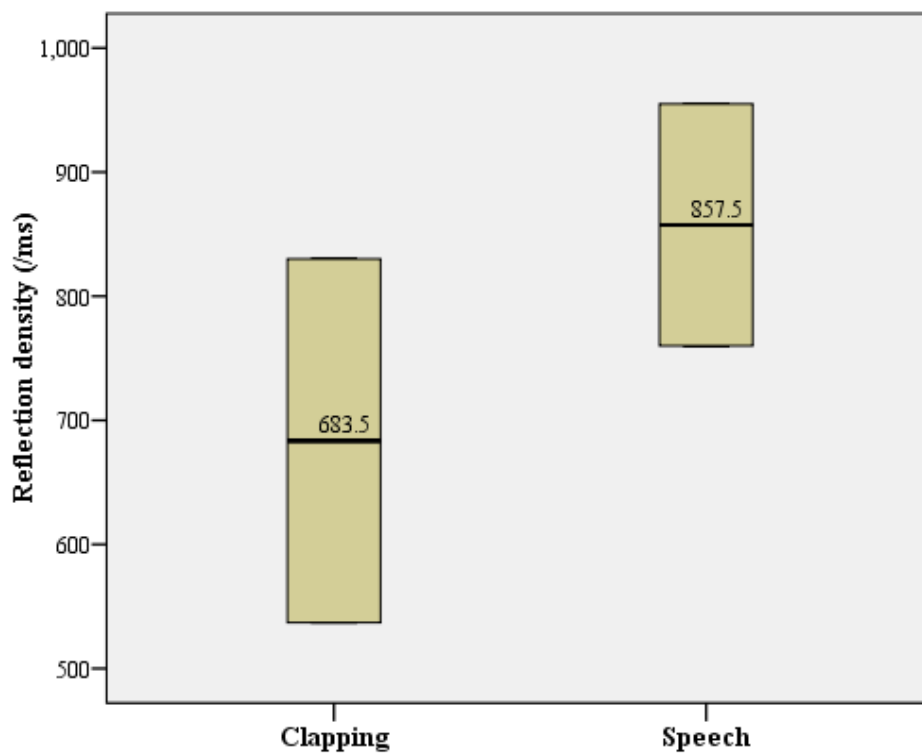


Figure 5.21 - Median of upper limits of distinguishable reflection density for the clapping and speech signals

## 5.4 Conclusions

Three methods for quantifying reflection density were investigated in this study. One of these is based on a set time window and cut-off level, but it requires further development as it did not produce linear results with increasing simulated room volume size as originally expected, likely due to random variances.

The upper limit of audible reflection density using both clapping and speech signals was examined through subjective testing. While there was a wide range in subject responses, the mean values for the upper limit of audible reflection density of clapping and speech signals, based on Odeon-provided calculations, were found to be 654 reflections/ms (SD = 130, N = 8), and 858 reflections/ms (SD = 98, N = 2), respectively.

As mentioned in section 2.1 the Odeon-provided reflection density has a limitation, since this value can be changed by the particular setup in Odeon. For simulations that use the same settings as presented in this study, the upper limit found can be useful to model and diagnose room acoustics. Now that an upper limit for distinguishable reflection density has been investigated, future work may focus on determining the just noticeable difference of reflection density.

## **Chapter 6 - Study 3: Just Noticeable Difference of Reflection Density**

### **6.1 Introduction**

This chapter investigates the just noticeable difference (JND) of reflection densities of impulse responses. The impulse responses were generated from simulated rooms in Odeon, and reflection density was controlled by variation of room size as used in Chapter 5. All rooms had reverberation times ( $T_{30}$ ) of 1 sec from 63 Hz to 8000 Hz by adjusting the magnitude of absorption of all surfaces uniformly. Unlike the method in Chapter 5, the variation of reflection density among comparison samples was not linear. Instead, larger differences were applied at lower reflection density samples, and the difference decreased as comparison samples approached the reference reflection density.

### **6.2 Methodology**

#### **6.2.1 *Test Stimuli***

##### *6.2.1.1 Source Signals*

Two kinds of source signals were used: clapping and speech. Both the clapping and speech sound samples were cropped from signals that came from Odeon version 11

database. The clapping signal included five claps while the speech sentence stated ‘When you are applying for a job, you need to have a good resume prepared.’ These source signals were convolved in Odeon with the binaural room impulse responses.

### 6.2.1.2 Impulse Responses

The room model ‘Example room’ from Odeon was used to simulate room impulse responses. A total of thirty different sized rooms were used, and the relative location of the sound source in each room was maintained across all thirty room sizes as shown in Figure 5.1.

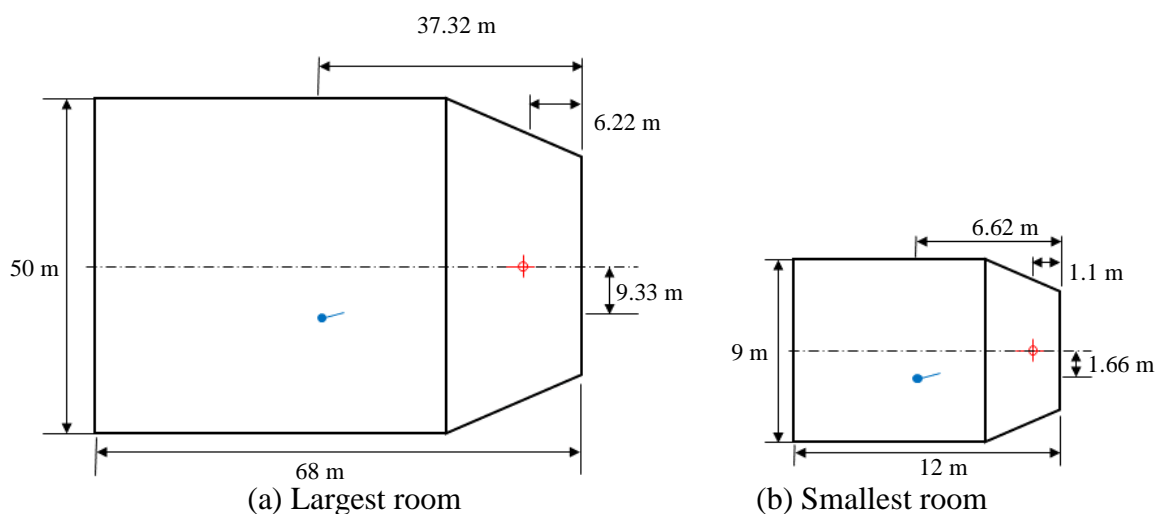


Figure 6.1 - Relative locations of a source (red) and receiver (blue) in the largest and the smallest room used in the study

The largest room volume was about  $88438 \text{ m}^3$ , and the smallest room was about



490 m<sup>3</sup>. The source was located at the frontal center, and the receiver was at the middle of the room slightly to the right side.

The percentage differences of comparison reflection density were a combination of linear and log scale variation, and it was 80, 60, 40, 20, 10, 5.33, 3, 0.95, and 0.53% of a reference reflection density. A linear decrease was applied from the 1st comparison sample to the 4th comparison sample (80% to 20%), and then a log scale decrease was used from the 5th comparison level to the 9th comparison level (10% to 0.52%) (Figure 6.2).

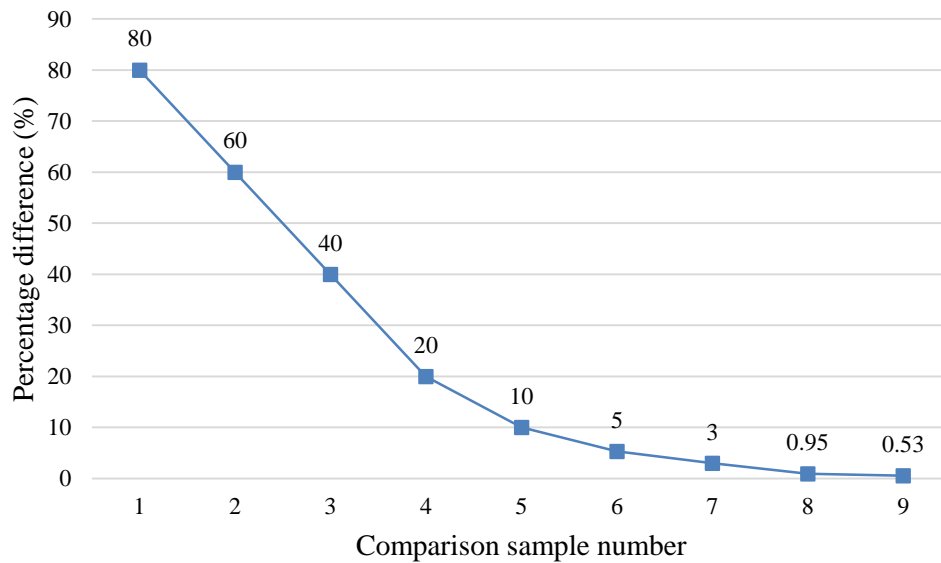


Figure 6.2 - Percentage difference of comparison samples

There were three groups of comparison samples (5 Up, 5 Down, and 10 Up), and each group has one reference and nine comparison samples (Table 6.1). As explained in

Section 5.2.1.2, the Odeon provided reflection densities were used due to the limitation of other quantification methods. This reflection density is a single value representing the late reflection density. Two reference reflection densities, corresponding to RD label 5 and RD label 10 in Chapter 5, were used. The Up and Down directions refer to the approaching directions of comparison samples to the corresponding reference, and it will be explained in detail later.

Table 6.1 - Reflection densities (RD) of sound samples, based on ODEON-provided calculation

Percentage difference (%)	5 Up			5 Down			10 Up		
	Sample number	Volume (m <sup>3</sup> )	RD (/ms)	Sample number	Volume (m <sup>3</sup> )	RD (/ms)	Sample number	Volume (m <sup>3</sup> )	RD (/ms)
80	1	88438	99.43	1	214	893.50	1	423	705
60	2	14500	198.75	2	301	794.44	2	372	740
40	3	4786	297.67	3	442	694.95	3	325	774
20	4	2158	397.36	4	680	595.73	4	285	809
10	5	1541	446.75	5	872	545.90	5	253	844
5.33	6	1342	470.11	6	975	522.74	6	224	878
3	7	1253	481.66	7	1074	511.55	7	199	912
0.95	8	1189	491.67	8	1126	501.39	8	180	948
0.53	9	1179	493.71	9	1137	498.90	9	161	983
	10 (Reference)	1153	496.42	10 (Reference)	1153	496.42	10 (Reference)	146	1017

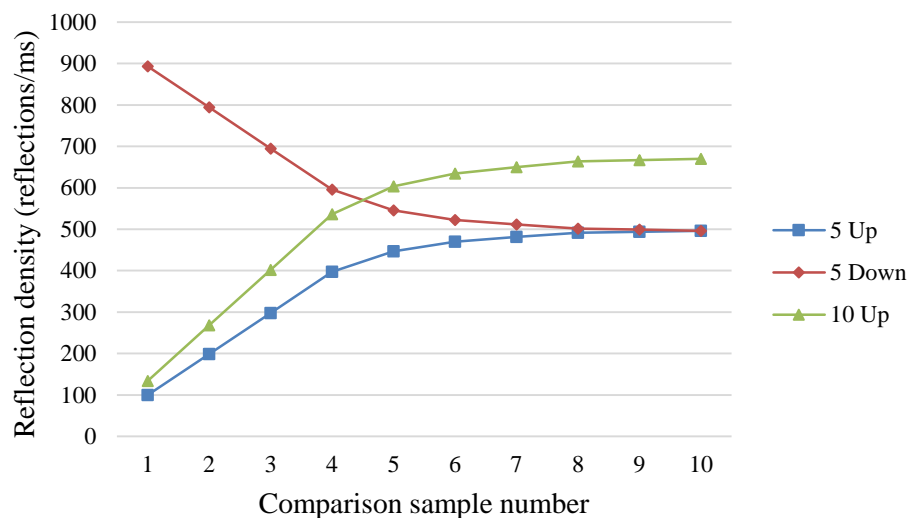


Figure 6.3 - Reflection densities (Odeon provided) of sound samples for three testing groups

Mean reflection densities were also calculated by the theoretical reflection density equation,  $cS/4V$ , for all rooms (Table 6.2).

Table 6.2 - Reflection density (RD) of sound samples, based on the classic reflection density equation,  $cS/4V$ , and the Odeon-provided calculation

Percentage difference (%)	5 Up			5 Down			10 Up		
	Sample number	$cS/4V$ (/s)	Odeon (/ms)	Sample number	$cS/4V$ (/s)	Odeon (/ms)	Sample number	$cS/4V$ (/s)	Odeon (/ms)
80	1	11.89	99.43	1	88.57	893.50	1	15.89	133.98
60	2	21.80	198.75	2	79.25	794.44	2	28.76	268.28
40	3	31.85	297.67	3	69.88	694.95	3	41.64	401.94
20	4	41.19	397.36	4	60.84	595.73	4	55.24	536.23
10	5	46.37	446.75	5	56.16	545.90	5	61.63	603.44
5.33	6	48.44	470.11	6	54.50	522.74	6	64.01	634.62
3	7	49.53	481.66	7	51.61	511.55	7	65.61	650.28
0.95	8	50.19	491.67	8	51.19	501.39	8	67.55	663.79
0.53	9	50.22	493.71	9	51.16	498.90	9	67.63	667.03
	10 (Reference)	50.83	496.42	10 (Reference)	50.83	496.42	10 (Reference)	67.63	670.24

The values in Table 6.1 are also shown in Figure 6.3 above. All rooms had a reverberation time ( $T_{30}$ ) of 1 sec from 63 Hz to 8000 Hz by adjusting the magnitude of absorption of all surfaces uniformly.

#### 6.2.1.3 *Convolution and Loudness Normalization*

Source signals were convolved with binaural impulse responses from thirty rooms in Odeon. Due to differences in the room volumes and distances between source and receiver, the loudness of the convolved signals from each room was different. To maintain the same loudness, the algorithm ITU-R in Adobe Audition software was used. The loudness was normalized at -23 LUFS (Loudness Units relative to Full Scale). Prepared test stimuli were presented to participants on a laptop using a custom Matlab GUI. The tests were conducted in a sound booth and presented over headphones to the subjects. The loudness from headphone was maintained around 65 dBA (re 20  $\mu$ Pa).

#### **6.2.2 *Three-Alternative Forced-Choice Method with 1Up-2Down Adaptive Method***

Test subjects participated in two separate sessions that were using different source signals: clapping and speech. Each session consisted of testing pairs, and it was expected to last 30 minutes to 45 minutes. The subject was asked to select which of three samples has a different reflection density than the others. The subjects had an opportunity to listen to maximum and minimum reflection density samples of clapping

and speech signals, so they got used to testing signals. The subject heard each of the sound samples two times and then was asked to answer the question after the signal presentations were completed.

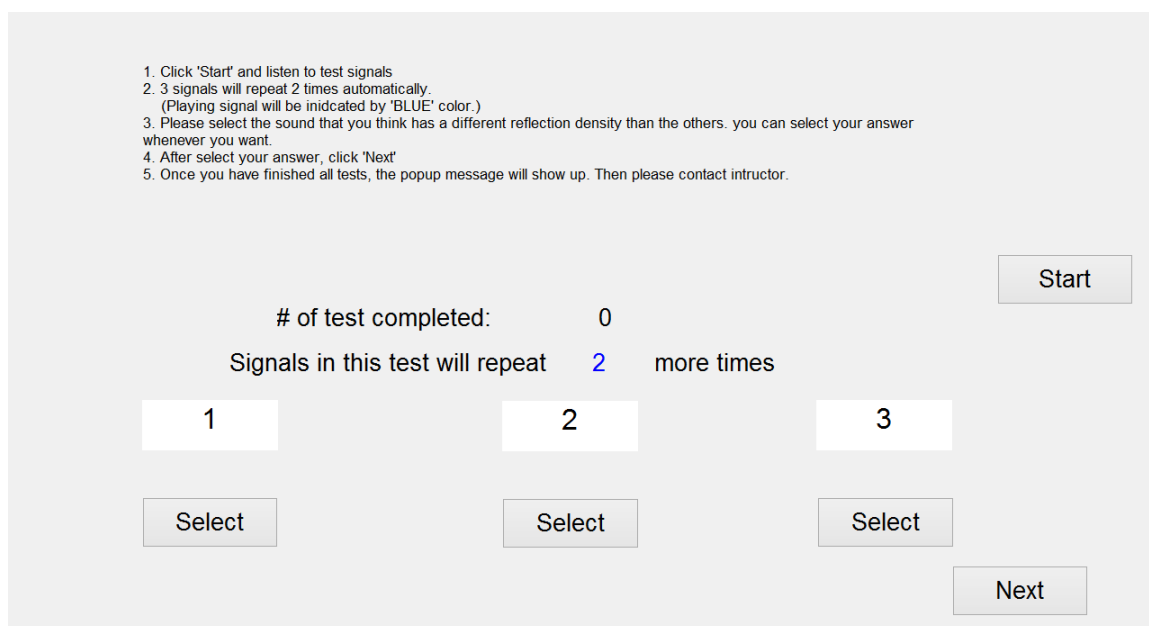


Figure 6.4 - Sample screenshot of the testing program

In total, these sessions including screening were expected to last no more than 90 minutes. A three-alternative forced-choice (3AFC) method combined with a 1-up 2-down adaptive method was used to determine the upper limit of distinguishable reflection density with impulse responses from the thirty rooms.

The chosen three-alternative force-choice testing method presents two identical reference reflection density (RD) samples and one comparison RD sample for each and

every comparison. The reference RD is designated reflection density that performs as a point of comparison. The reference RD was approached from below (lower RD) or above (higher RD) when test subjects made correct answers. With a 1-up 2-down method, two consecutive correct answers decreased (down) a difference between the reference and a comparison RD, and one incorrect answer increased (up) a difference. A reversal was recorded when test subject made changes of direction from up to down or down to up. After five reversals had been made, an averaged comparison RD at those reversals was reported as an upper limit of distinguishable reflection density. There were three groups of testing signals, and these groups' signals were mixed as a whole testing group. The presenting order of each group's signal was randomly chosen. The 3AFC mechanism applied to each group independently, so five reversals were counted separately for each group.

The three groups of testing samples were 5 Up, 5 Down, and 10 Up cases. Numbers 5 and 10 indicate the reference reflection density: RD label 5 or RD label 10 from Chapter 5. RD label 5 is 496 reflections/ms, and RD label 10 is 670 reflections/ms. Up and Down refer to the approaching direction of comparison reflection density. For example, 5 Up means the reference reflection density is RD label 5 (670 reflections/ms) and comparison reflection densities are lower than 670 reflections/ms. The comparison reflection density increases (Up), becoming closer to the reference reflection density as a subject selects correct answers.

### **6.2.3 *Participants***

A total of twenty subjects (11 male, 9 female) finished all tests, and 11 subjects were native English speakers. The average age was 25.1 years (SD = 7.9 years). All listeners had pure tone thresholds below 25 dB hearing level between 250 and 8000 Hz and had at least three years of musical training or experience. Listeners provided informed consent for their participation in the study and were paid \$15 Amazon gift card for their time.

## **6.3 Result and Analysis**

### **6.3.1 *Clapping Results***

Results of tests with clapping signal were categorized by reference reflection density and approaching direction: i) RD label 5 and up direction, ii) RD label 5 and down direction, and iii) RD label 10 and up direction.

#### **6.3.1.1 *RD Label 5 and Up Direction***

Subjects 2, 16, and 20 could differentiate all comparison reflection densities and reach the reference reflection density (sample #10). These subjects were excluded from the final mean calculation. Table 6.3 shows reversals of these subjects and shaded cells indicate reference samples (sample #10).

The reflection densities of comparison samples were achieved by variation of

room volume. However, it was not perfectly matched to the targeted reflection density. Reflection densities by  $cS/4V$  and percentage differences were also calculated for all rooms (Table 6.4). Since sizes of simulated rooms were adjusted based on the Odeon provided reflection density, Odeon provided reflection density will be used to describe results hereafter.

Table 6.3 - Comparison sample numbers at reversals of clapping results with RD label 5 and up direction condition

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	3	2	4	3	4	11	3	2	3	2	3
2	5	4	10	9	10	12	4	3	5	4	5
3	5	3	5	3	4	13	4	3	5	3	4
4	5	4	5	3	5	14	3	2	4	2	3
5	4	3	4	2	5	15	4	3	6	5	6
6	4	3	4	3	5	16	4	3	10	9	10
7	4	3	5	4	5	17	5	4	6	4	7
8	5	4	5	4	5	18	5	3	4	3	4
9	3	2	4	3	6	19	5	4	5	4	5
10	4	3	4	3	4	20	4	3	5	4	10



Table 6.4 - Percentage differences of reflection density of clapping comparison samples for RD label 5 and up direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (/ms)	Actual % difference	RD (/s)	Actual % difference
1	80	99.43	79.97	11.89	76.60
2	60	198.75	59.96	21.80	57.11
3	40	297.67	40.04	31.85	37.34
4	20	397.36	19.95	41.19	18.97
5	10	446.75	10.01	46.37	8.79
6	5.33	470.11	5.30	48.44	4.71
7	3	481.66	2.97	49.53	2.55
8	0.95	491.67	0.96	50.19	1.26
9	0.53	493.71	0.55	50.22	1.21
10	0	496.42	0	50.83	0

Percentage differences corresponding to reversals recorded in Table 6.3 are shown in Table 6.4. The percentage difference in this table used Odeon provided reflection densities. Mean reversal values of Subjects 2, 6, and 20 were excluded from the total mean calculation and marked as · in the table. From a boxplot of subjects' JNDs, Subject 11 was identified as an outlier, so it was excluded from the mean calculation (Figure 6.5). Subject 11 is marked with asterisk mark (\*) in Table 6.5. Since the difference between samples are much larger (maximum of 20%) than in Study 2, the RSD of 30% may not be a good criteria to sort the data. In order to sort results with better convergence among subjects, this maximum difference between samples of 20% was used as a limit. For RD label 5 and up direction, all subjects had SD of reversals less

than 20%. The mean JND of reflection density with RD label 5 and up direction was found to be about 24.17% with standard deviation of 8.56%.

Table 6.5 - Just noticeable difference (%) of reflection density for clapping signals with reference reflection density of RD label 5 and up direction. Subjects 2, 11, 16, 20 were excluded from the mean calculation. (N = 16)

Subject	Reversal						
	1st	2nd	3rd	4 <sup>th</sup>	5th	Mean	SD
1	40.04	59.96	19.95	40.04	19.95	35.99	15.0
2	10.01	19.95	0.00	0.55	0.00	·	·
3	10.01	40.04	10.01	40.04	19.95	24.01	13.6
4	10.01	19.95	10.01	40.04	10.01	18.00	11.7
5	19.95	40.04	19.95	59.96	10.01	29.98	17.9
6	19.95	40.04	19.95	40.04	10.01	26.00	12.0
7	19.95	40.04	10.01	19.95	10.01	19.99	11.0
8	10.01	19.95	10.01	19.95	10.01	13.99	4.9
9	40.04	59.96	19.95	40.04	5.30	33.06	18.8
10	19.95	40.04	19.95	40.04	19.95	27.99	9.8
11	40.04	59.96	40.04	59.96	40.04	48.01*	9.8
12	19.95	40.04	10.01	19.95	10.01	19.99	11.0
13	19.95	40.04	10.01	40.04	19.95	26.00	12.0
14	40.04	59.96	19.95	59.96	40.04	43.99	15.0
15	19.95	40.04	5.30	10.01	5.30	16.12	13.1
16	19.95	40.04	0.00	0.55	0.00	·	·
17	10.01	19.95	5.30	19.95	2.97	11.64	7.2
18	10.01	40.04	19.95	40.04	19.95	26.00	12.0
19	10.01	19.95	10.01	19.95	10.01	13.99	4.9
20	19.95	40.04	10.01	19.95	0.00	·	·
Mean (SD < 20%, N = 16)					24.17 % (SD = 8.56)		

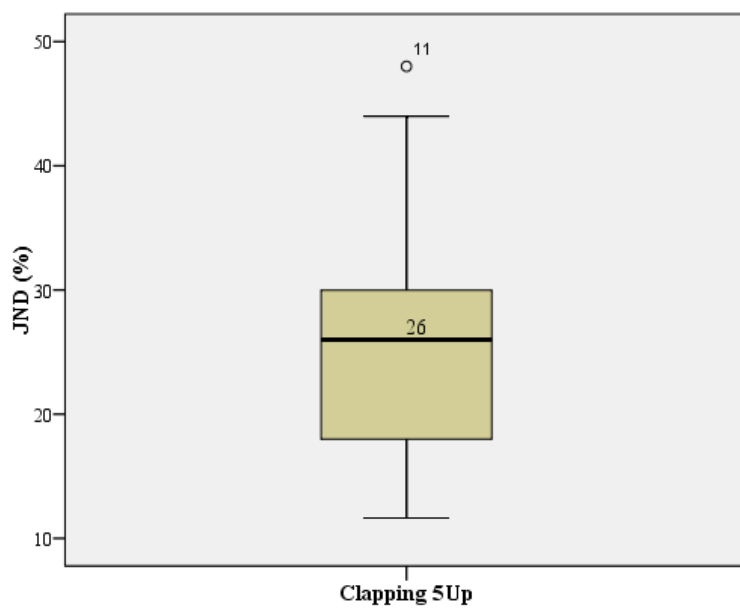


Figure 6.5 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and up direction. Median JND was 26%, and Subject 11 was an outlier.

The reversals from subjects are grouped and shown in Figure 6.6, Figure 6.7, and Figure 6.8.

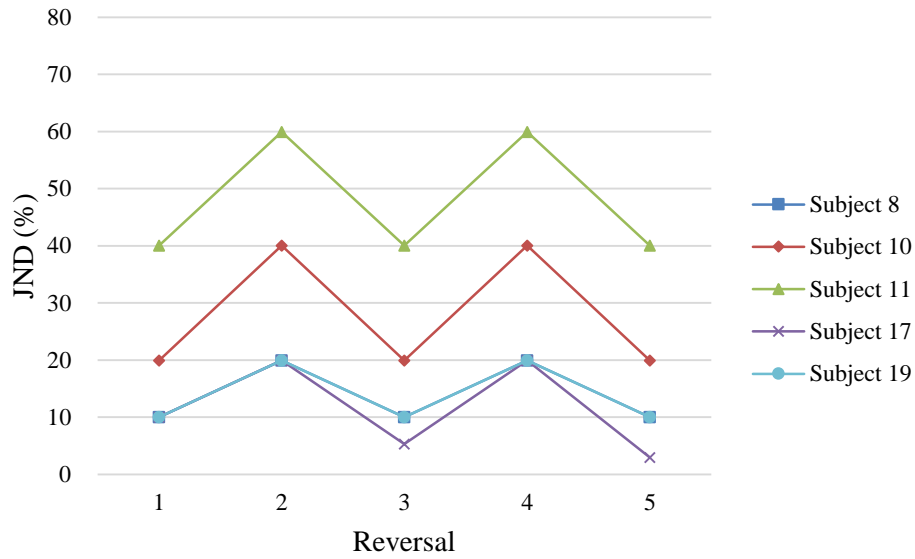


Figure 6.6 – Reversals from subjects who had  $SD < 10\%$  with 5 Up condition

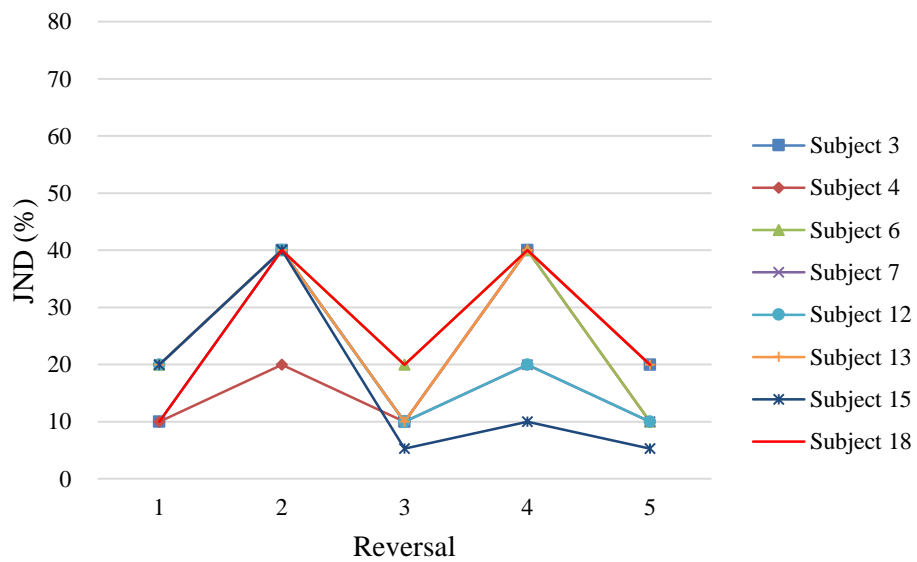


Figure 6.7 – Reversals from subjects who had  $10\% < SD < 15\%$  with 5 Up condition

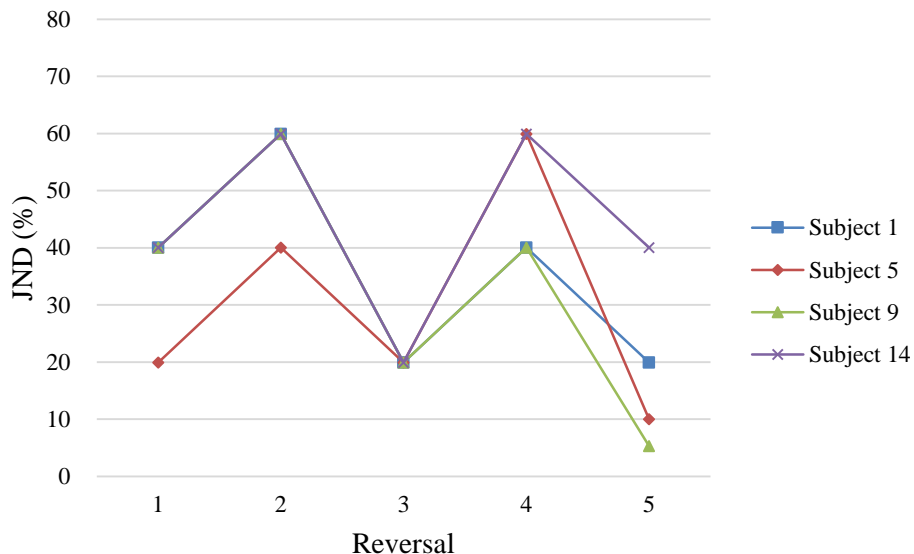


Figure 6.8 - Reversals from subjects who had  $15\% \leq SD$  with 5 Up condition

### 6.3.1.2 RD Label 5 and Down Direction

In results with a reference reflection density of RD label 5 and down direction case, Subject 2 distinguished all comparison reflection densities and reached the reference reflection density (sample #10). For this reason, Subject 2 was excluded from the final mean calculation. Table 6.6 shows reversals of subjects, and a shaded cell indicates reference samples (Sample #10).

Reflection densities used for RD label 5 and down direction are shown in Table 6.7. It also includes values of reflection densities calculated by  $cS/4V$ .

Table 6.6 - Comparison sample numbers at reversals of clapping results with RD label 5 and down direction case

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	3	1	4	3	4	11	1	3	1	2	1
2	1	6	5	10	9	12	1	2	1	2	1
3	1	2	1	4	1	13	1	2	1	2	1
4	4	3	6	2	4	14	1	3	2	4	2
5	1	2	1	2	1	15	1	3	1	5	4
6	1	2	1	4	2	16	1	3	1	2	1
7	1	2	1	2	1	17	1	2	1	3	1
8	1	2	1	3	1	18	1	2	1	2	1
9	1	2	1	2	1	19	2	1	2	1	4
10	1	3	2	3	1	20	6	5	6	5	7

Table 6.7 - Percentage differences of reflection density of clapping comparison samples for RD label 5 and down direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (Odeon, /ms)	Actual % difference	RD (cS/4V/s)	Actual % difference
1	80	893.50	79.99	88.57	74.24
2	60	794.44	60.03	79.25	55.90
3	40	694.95	39.99	69.88	37.46
4	20	595.73	20.01	60.84	19.70
5	10	545.90	9.97	56.16	10.48
6	5.33	522.74	5.30	54.50	7.22
7	3	511.55	3.05	51.61	1.54
8	0.95	501.39	1	51.19	0.69
9	0.53	498.90	0.5	51.16	0.65
10	0	496.42	0	50.83	0

There were more subjects who marked reversals at the first comparison sample (Sample #1, 80% designed difference) than the case of RD label 5 and up direction, so these subjects were analyzed more by responses at trials. Subjects 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19 were included for this analysis. Subjects 5, 7, 9, 18 could not distinguish the second comparison sample (60.03%), and differentiated only the first comparison sample (79.99%). Figure 6.9 depicts these subjects, and the comparison sample #10 is the reference reflection density. In this situation, it is possible that just noticeable differences could be much larger than 80%, so these subjects were excluded from the mean calculation. Subjects 3, 6, 8, 10, 11, 12, 13, 15, 17, 19 also made at least one reversal at the first comparison sample. However, these subjects made reversals at higher reflection densities (smaller differences).

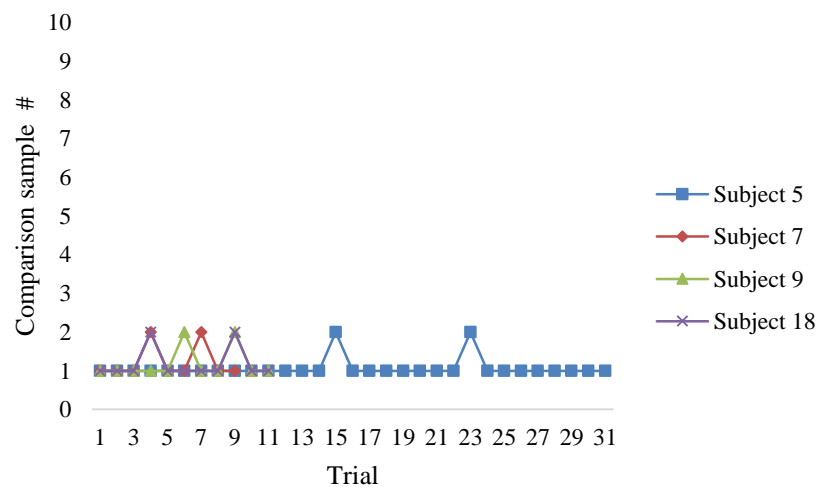


Figure 6.9 - Trials of clapping results with RD label 5 and down direction that show inability to distinguish samples

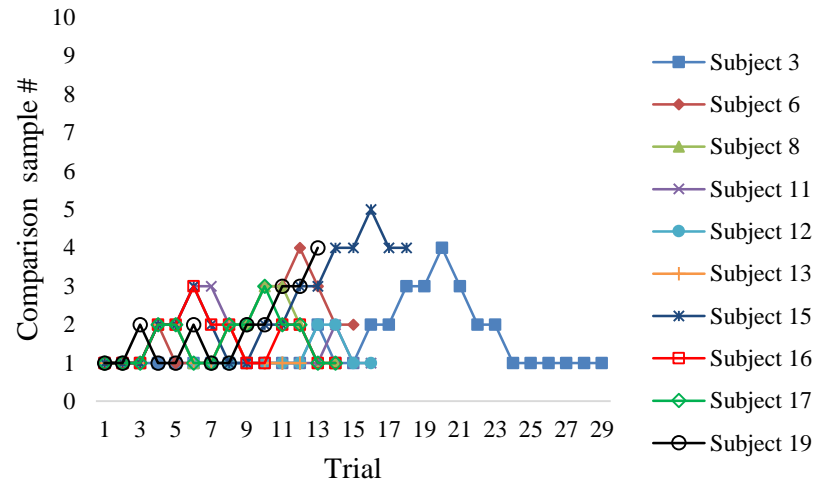


Figure 6.10 - Trials of results that marked reversals at the first comparison sample but was included for mean calculation

Percentage differences corresponding to reversals recorded in Table 6.6 are shown in Table 6.8. Subjects 1, 3, 6, 14, 15, and 19 exceeded 20% of standard deviation and excluded from the final mean calculation. The box plot for JNDs of assorted results are shown in Figure 6.11.



Table 6.8 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and down direction

Subject	Reversal						
	1st	2nd	3rd	4th	5th	Mean	SD
1	39.99	79.99	20.01	39.99	20.01	40.00	21.9
2	79.99	5.30	9.97	0.00	0.50	.	.
3	79.99	60.03	79.99	20.01	79.99	64.00	23.3
4	20.01	39.99	5.30	60.03	20.01	29.07	19.0
5	79.99	60.03	79.99	60.03	79.99	.	.
6	79.99	60.03	79.99	20.01	60.03	60.01	21.9
7	79.99	60.03	79.99	60.03	79.99	.	.
8	79.99	60.03	79.99	39.99	79.99	68.00	16.0
9	79.99	60.03	79.99	60.03	79.99	.	.
10	79.99	39.99	60.03	39.99	79.99	60.00	17.9
11	79.99	39.99	79.99	60.03	79.99	68.00	16.0
12	79.99	60.03	79.99	60.03	79.99	72.01	9.8
13	79.99	60.03	79.99	60.03	79.99	72.01	9.8
14	79.99	39.99	60.03	20.01	60.03	52.01	20.4
15	79.99	39.99	79.99	9.97	20.01	45.99	29.4
16	79.99	39.99	79.99	60.03	79.99	68.00	16.0
17	79.99	60.03	79.99	39.99	79.99	68.00	16.0
18	79.99	60.03	79.99	60.03	79.99	.	.
19	60.03	79.99	60.03	79.99	20.01	60.01	21.9
20	5.30	9.97	5.30	9.97	3.05	6.72*	2.8
				All (N = 15)		55.59 % (SD = 17.75)	
				Assorted (SD < 20%, N = 9)		56.87 % (SD = 21.73)	

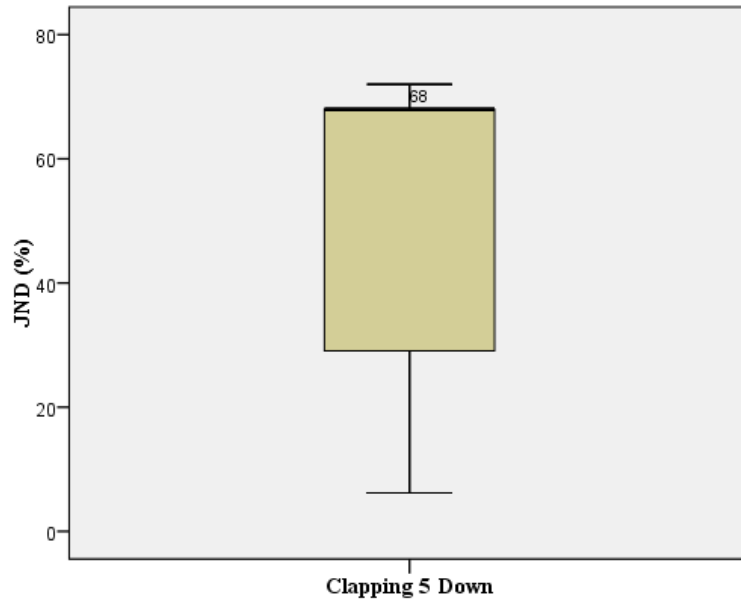


Figure 6.11 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 5 and down direction. Median JND was 68%.

By excluding only subjects exceeding 20% of standard deviation or distinguished all samples, Subjects 4, 8, 10, 11, 12, 13, 16, 17, 20 are included for the final mean calculation, and the mean JND of reflection density with RD label 5 and down direction was found to be about 56.87% (N = 9) with a standard deviation of 21.72% (Table 6.8). The JND of clapping signals with RD label 5 and down direction was much higher than up direction by over 30%. The reversals from subjects are grouped and shown in Figure 6.12 and Figure 6.13.

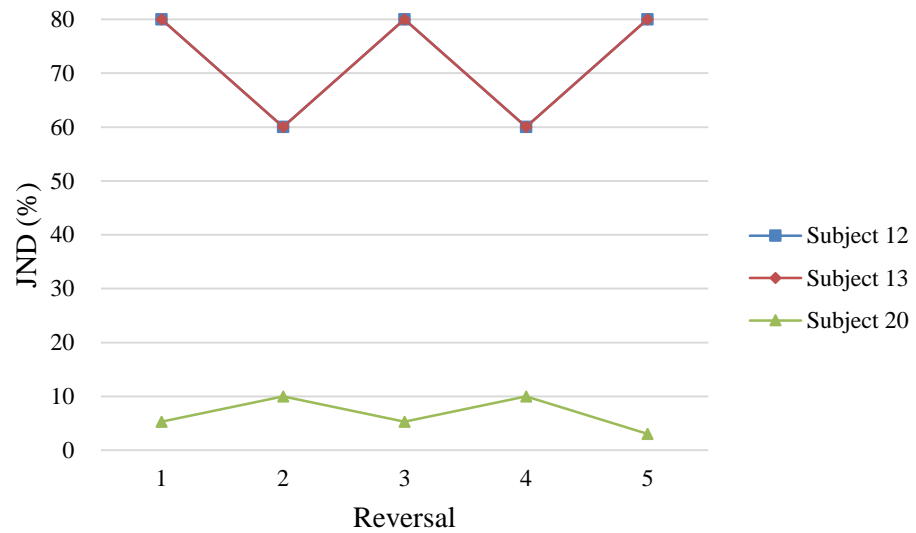


Figure 6.12 - Reversals from subjects who had  $SD < 10\%$  with 5 Down condition

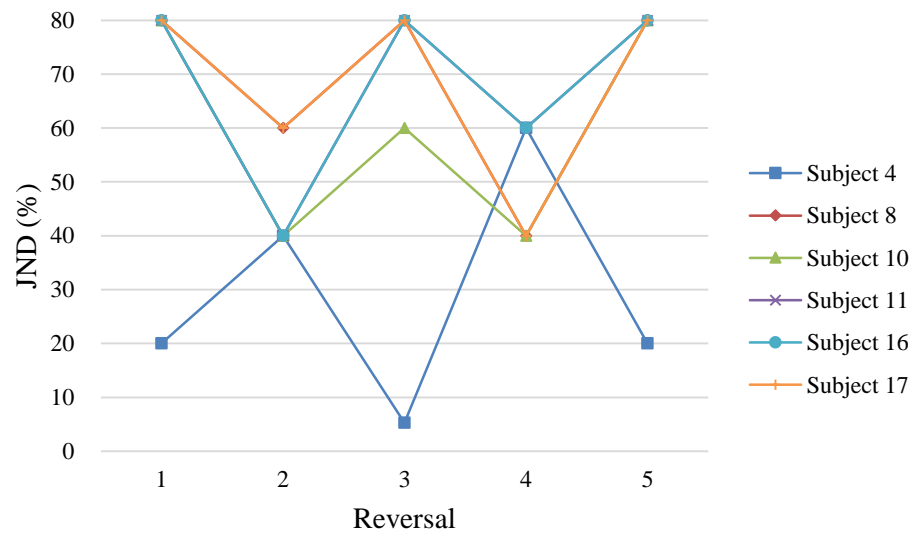


Figure 6.13 - Reversals from subjects who had  $10\% < SD \leq 20\%$  with 5 Down condition

### 6.3.1.3 RD Label 10 and Up Direction

Subjects 2 and 20 were excluded since these subjects distinguished all comparison reflection density. Table 6.9 shows reversals of subjects and shaded cells indicate reference samples (Sample #10).

Table 6.9 - Comparison sample numbers at reversals of clapping results with RD label 10 and up direction condition

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	2	1	3	2	3	11	3	2	3	2	4
2	4	3	4	3	10	12	4	3	4	3	5
3	4	1	3	2	3	13	4	3	5	3	4
4	4	3	4	2	3	14	4	3	4	3	4
5	4	3	4	3	4	15	3	2	4	3	5
6	4	2	4	3	4	16	4	3	4	2	4
7	4	2	4	3	4	17	3	2	4	3	5
8	2	1	3	2	5	18	7	6	7	6	7
9	3	2	3	2	3	19	4	3	5	3	7
10	5	3	4	2	4	20	10	9	10	9	·

Actual reflection densities used for testing are shown in Table 6.10.

Table 6.10 - Percentage differences of reflection density of clapping comparison samples for RD label 10 and up direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (Odeon, /ms)	Actual % difference	RD (cS/4V,/s)	Actual % difference
1	80	133.98	80.01	15.89	76.50
2	60	26828	59.97	28.76	57.47
3	40	401.94	40.03	41.64	38.43
4	20	536.23	19.99	55.24	18.32
5	10	603.44	9.97	61.63	8.88
6	5.33	634.62	5.31	64.01	5.36
7	3	650.28	2.96	65.61	2.98
8	0.95	663.79	0.96	67.55	0.12
9	0.53	667.03	0.48	67.63	0
10	0	670.24	0	67.63	0

The mean just noticeable difference of reflection density was calculated with eighteen subjects. Subjects 2 and 20 were excluded as they distinguished all comparison samples. Subjects 3 and 8 were also excluded since their SD exceeded 20% (Table 6.11).

Table 6.11 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 10 and up direction

Subject	Reversal						
	1st	2nd	3rd	4th	5th	Mean	SD
1	59.97	80.01	40.03	59.97	40.03	56.00*	15.0
2	19.99	40.03	19.99	40.03	0.00	.	.
3	19.99	80.01	40.03	59.97	40.03	48.01	20.4
4	19.99	40.03	19.99	59.97	40.03	36.00	15.0
5	19.99	40.03	19.99	40.03	19.99	28.01	9.8
6	19.99	59.97	19.99	40.03	19.99	32.00	16.0
7	19.99	59.97	19.99	40.03	19.99	32.00	16.0
8	59.97	80.01	40.03	59.97	9.97	49.99	23.7
9	40.03	59.97	40.03	59.97	40.03	48.01*	9.8
10	9.97	40.03	19.99	59.97	19.99	29.99	17.9
11	40.03	59.97	40.03	59.97	19.99	44.00	15.0
12	19.99	40.03	19.99	40.03	9.97	26.00	12.0
13	19.99	40.03	9.97	40.03	19.99	26.00	12.0
14	19.99	40.03	19.99	40.03	19.99	28.01	9.8
15	40.03	59.97	19.99	40.03	9.97	34.00	17.4
16	19.99	40.03	19.99	59.97	19.99	32.00	16.0
17	40.03	59.97	19.99	40.03	9.97	34.00	17.4
18	2.98	5.31	2.98	5.31	2.98	3.91*	1.1
19	19.99	40.03	9.97	40.03	2.98	22.60	15.2
20	0.00	0.48	0.00	0.48	.	.	.
All (N = 18)					32.59 % (SD = 11.47)		
Assorted (SD < 20%, N = 13)					31.12 % (SD = 5.22)		

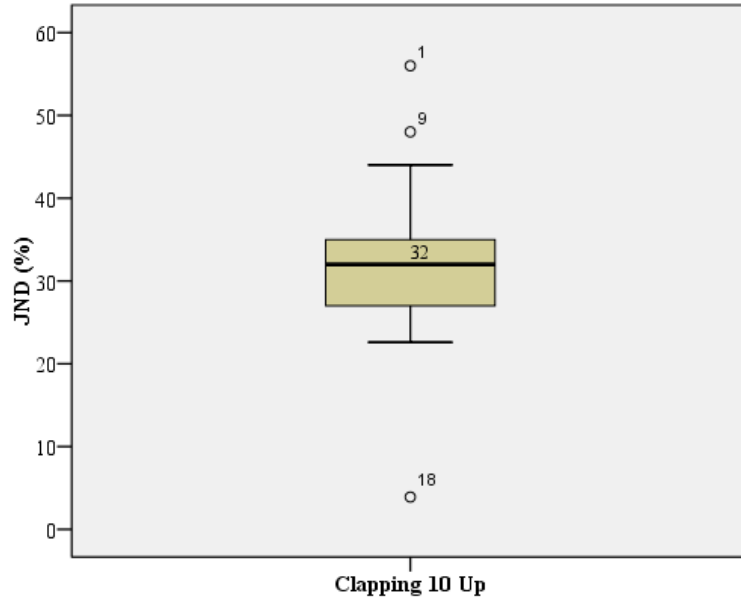


Figure 6.14 - Just noticeable difference of reflection density for clapping signals with reference reflection density of RD label 10 and up direction. Median JND was 32%, and Subject 1, 9 and 18 were outliers

Subjects 1, 9 and 18 were found to be outliers in the boxplot and excluded from the final mean calculation (Figure 6.14). Finally, Subjects 1, 2, 3, 8, 9, 18, 20 were excluded from the mean calculation, and the mean JND of the reflection density with reference of RD label 10 and up direction was found to be 31.12% (N = 13) with a standard deviation of 5.22%.

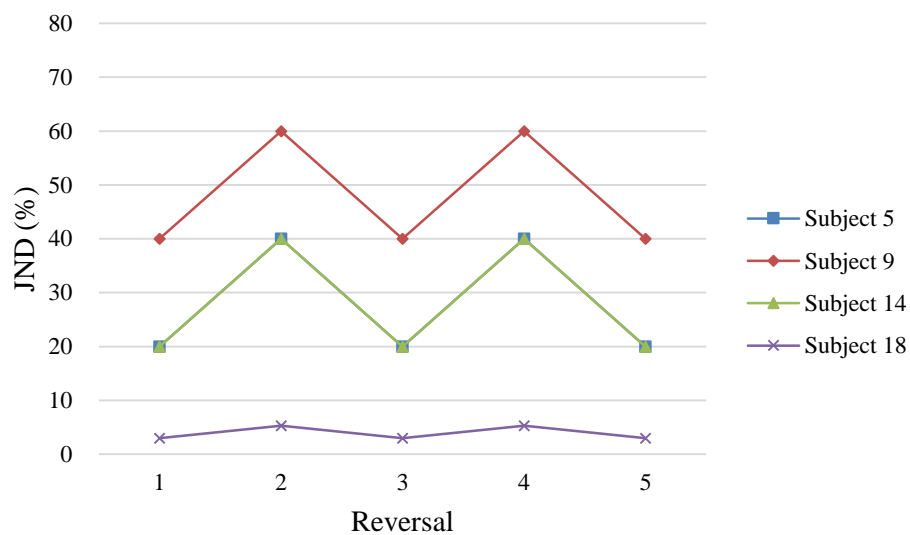


Figure 6.15 - Reversals from subjects who had  $SD < 10\%$  with 10 Up condition

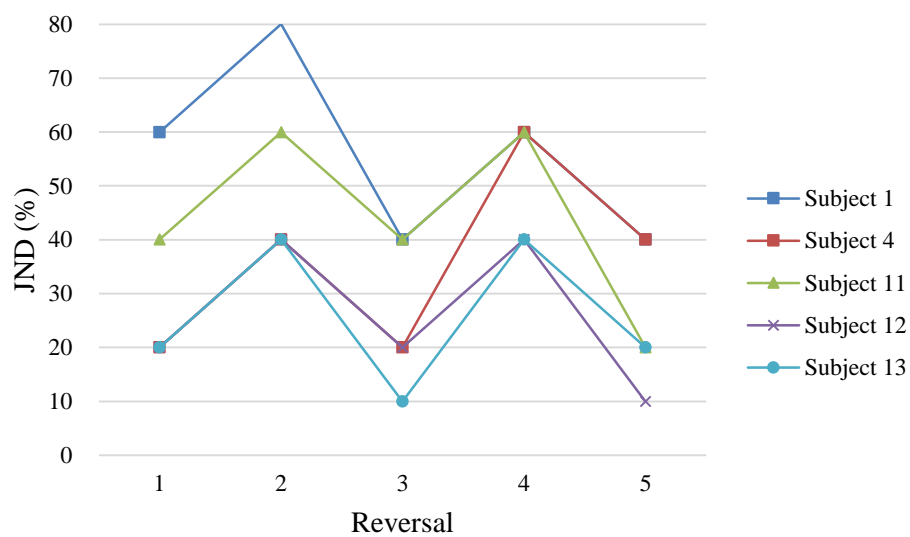


Figure 6.16 - Reversals from subjects who had  $10\% < SD \leq 15\%$  with 10 Up condition



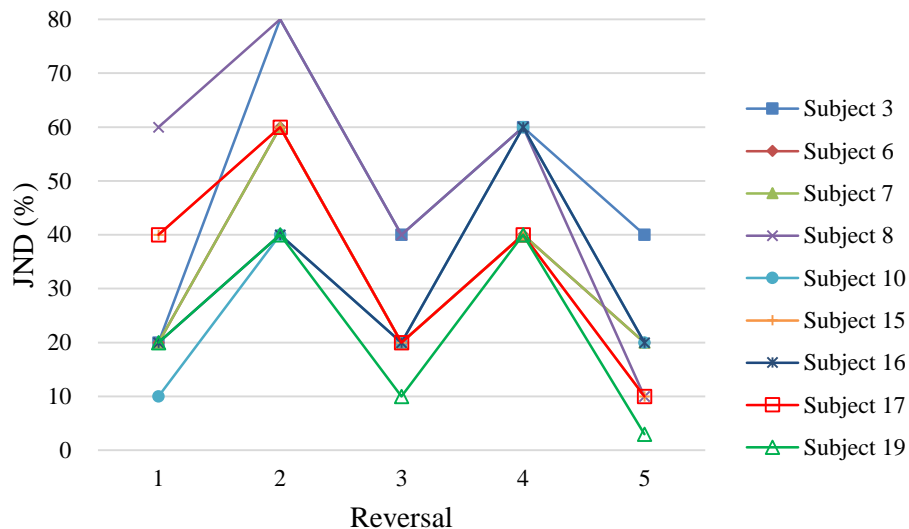


Figure 6.17 - Reversals from subjects who had  $15\% < SD < 20\%$  with 10 Up condition

### 6.3.2 Speech Results

Results of tests with a speech signal were also categorized by reference reflection density and approaching direction as a clapping signal: three groups: i) RD label 5 and up direction, ii) RD label 5 and down direction, and iii) RD label 10 and up direction.

#### 6.3.2.1 RD Label 5 and Up Direction

In the tests with a reference of RD label 5 and up direction, five subjects showed an ability to distinguish all comparison samples, and those are Subjects 8, 9, 14, 19, and 20 shown with shaded cells in Table 6.12. These subjects were excluded from the mean calculation. The percentage differences used for the RD label 5 with up direction case is shown in Table 6.13.

Table 6.12 - Comparison sample numbers at reversals of speech results with RD label 5 and up direction condition

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	2	1	3	2	3	11	2	1	3	2	6
2	6	5	6	5	6	12	3	2	3	2	3
3	6	5	7	6	7	13	2	1	4	3	5
4	3	1	4	3	6	14	10	9	10	9	10
5	5	4	5	4	7	15	6	4	6	5	6
6	5	4	5	4	5	16	5	4	6	5	6
7	4	3	6	4	5	17	5	4	6	4	5
8	4	3	10	9	10	18	5	4	5	4	5
9	3	2	10	9	10	19	5	4	10	9	10
10	6	4	5	4	7	20	6	5	10	·	·

Table 6.13 - Percentage differences of reflection density of speech comparison samples for RD label 5 and up direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (/ms)	Actual % difference	RD (/s)	Actual % difference
1	80	99.43	79.97	11.89	76.60
2	60	198.75	59.96	21.80	57.11
3	40	297.67	40.04	31.85	37.34
4	20	397.36	19.95	41.19	18.97
5	10	446.75	10.01	46.37	8.79
6	5.33	470.11	5.30	48.44	4.71
7	3	481.66	2.97	49.53	2.55
8	0.95	491.67	0.96	50.19	1.26
9	0.53	493.71	0.55	50.22	1.21
10	0	496.42	0	50.83	0

Subjects 4, 11 and 13 were excluded since their SD exceeded 20% (Table 6.14).

Subjects 1 and 12 were found to be outliers in the boxplot (Table 6.19) and also excluded.

The mean JND was found to be 18.35% (N = 12) with a standard deviation of 15.51%

(Table 6.14)

Table 6.14 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and up direction

Subject	Reversal						
	1st	2nd	3rd	4th	5th	Mean	SD
1	59.96	79.97	40.04	59.96	40.04	55.99	14.9
2	5.30	10.01	5.30	10.01	5.30	7.18	2.3
3	5.30	10.01	2.97	5.30	2.97	5.31	2.6
4	40.04	79.97	19.95	40.04	5.30	37.06	25.1
5	10.01	19.95	10.01	19.95	2.97	12.58	6.5
6	10.01	19.95	10.01	19.95	10.01	13.99	4.9
7	19.95	40.04	5.30	19.95	10.01	19.05	11.9
8	19.95	40.04	0.00	0.55	0.00	.	.
9	40.04	59.96	0.00	0.55	0.00	.	.
10	5.30	19.95	10.01	19.95	2.97	11.64	7.2
11	59.96	79.97	40.04	59.96	5.30	49.05	25.3
12	40.04	59.96	40.04	59.96	40.04	48.01	9.8
13	59.96	79.97	19.95	40.04	10.01	41.99	25.6
14	0.00	0.55	0.00	0.55	0.00	.	.
15	5.30	19.95	5.30	10.01	5.30	9.17	5.7
16	10.01	19.95	5.30	10.01	5.30	10.11	5.4
17	10.01	19.95	5.30	19.95	10.01	13.04	5.9
18	10.01	19.95	10.01	19.95	10.01	13.99	4.9
19	10.01	19.95	0.00	0.55	0.00	.	.
20	5.30	10.01	0.00	.	.	.	.
All (N = 15)					23.22 % (SD = 17.10)		
Assorted (SD<20%, N = 12)					18.35 % (SD = 15.51)		

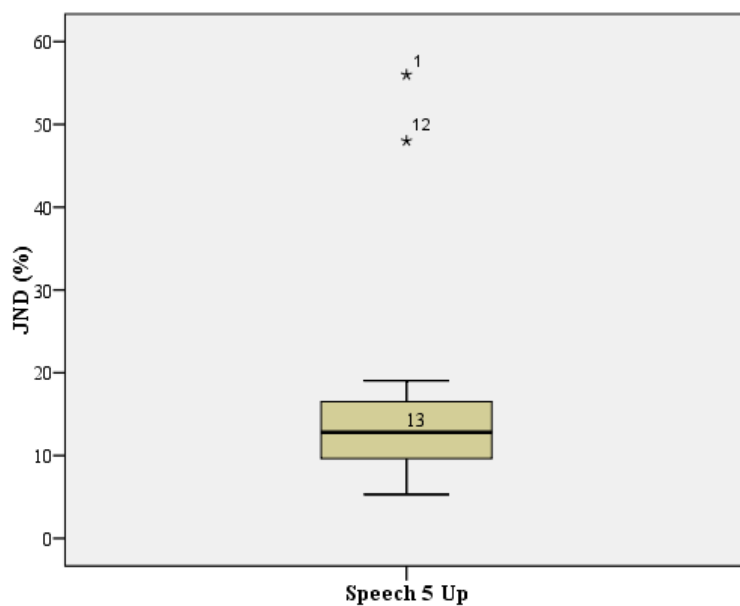


Figure 6.18 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and up direction. Subject 1 and 12 were outlier and median JND was 13%.

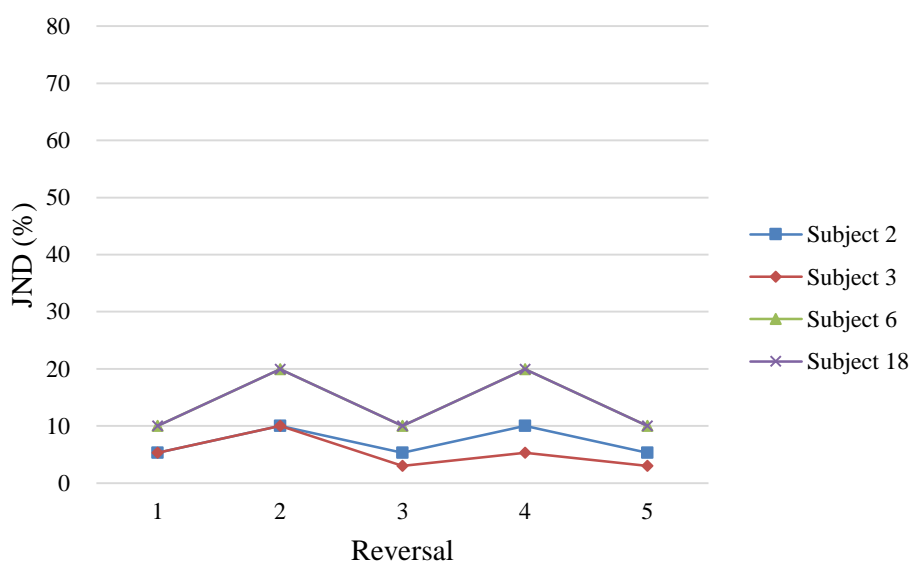


Figure 6.19 - Reversals from subjects who had SD < 5% with 5 Up condition

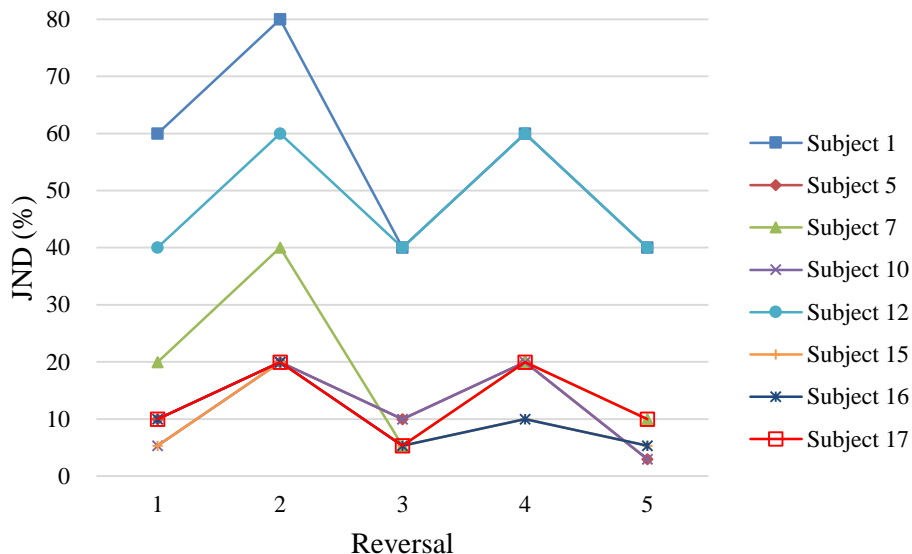


Figure 6.20 - Reversals from subjects who had  $5\% < SD < 20\%$  with 5 Up condition

### 6.3.2.2 RD Label 5 and Down Direction

Similar to clapping signals with a reference of RD label 5 and down direction, many subjects could not distinguish well between the first two comparison samples, Samples #1 and #2 (Table 6.15). Subjects 2, 4, 10, 15 could not distinguish the comparison sample with a reflection density of more than the second comparison sample. Although these subjects made a correct answer at comparison sample #2, they could not make two consecutively correct answers (Figure 6.21). These four subjects were excluded from the mean calculation.

Table 6.15 - Comparison sample numbers at reversals of speech results with RD label 5 and down direction condition

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	7	4	5	4	5	11	2	1	3	2	5
2	1	2	1	2	1	12	1	2	1	4	1
3	1	3	1	2	1	13	2	1	2	1	3
4	2	1	2	1	2	14	3	1	8	6	7
5	2	1	4	3	4	15	1	2	1	2	1
6	1	6	5	6	3	16	2	1	4	3	4
7	1	4	3	6	4	17	1	4	3	4	3
8	3	2	3	1	4	18	1	7	4	5	4
9	1	6	5	9	8	19	7	5	6	5	6
10	2	1	2	1	2	20	2	1	4	3	6

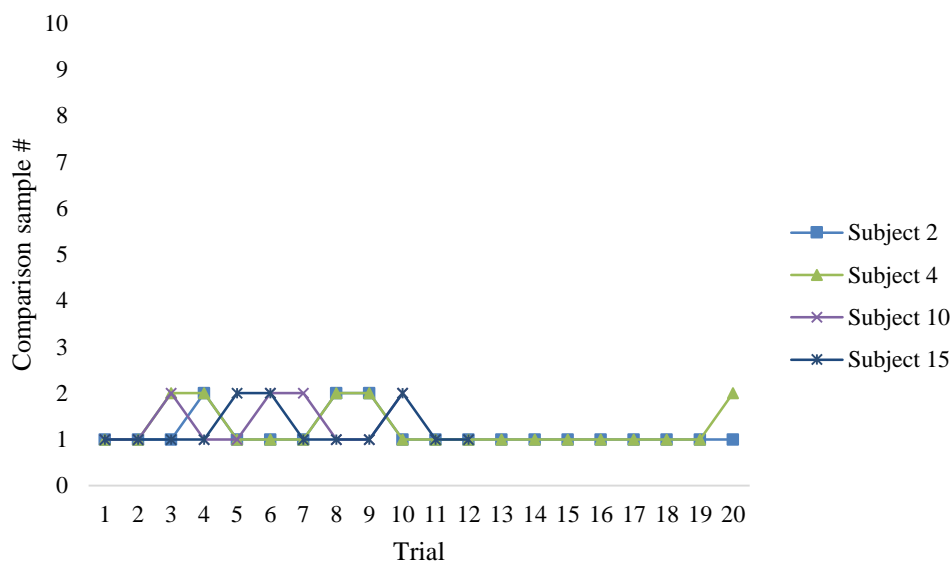


Figure 6.21 – Detailed results of individual speech trials from subjects demonstrating inability to distinguish differences with RD label 5 and up direction

The percentage differences used for the RD label 5 with down direction case are shown in Table 6.16.

Table 6.16 - Percentage differences of reflection density of speech comparison samples for RD label 5 and down direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (Odeon, /ms)	Actual % difference	RD (cS/4V/s)	Actual % difference
1	80	893.50	79.99	88.57	74.24
2	60	794.44	60.03	79.25	55.90
3	40	694.95	39.99	69.88	37.46
4	20	595.73	20.01	60.84	19.70
5	10	545.90	9.97	56.16	10.48
6	5.33	522.74	5.30	54.50	7.22
7	3	511.55	3.05	51.61	1.54
8	0.95	501.39	1.00	51.19	0.69
9	0.53	498.90	0.50	51.16	0.65
10	0	496.42	0	50.83	0

Only Subjects 1, 3, 13, 19 demonstrated SD less than 20%, so these subjects were included for the final mean calculation. The mean JND was found to be 37.83 % (N = 4) with a standard deviation of 5.58% (Table 6.17).

Table 6.17 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 5 and down direction

Subject	Reversal						
	1st	2nd	3rd	4th	5th	Mean	SD
1	3.05	20.01	9.97	20.01	9.97	12.60	6.6
2	79.99	60.03	79.99	60.03	79.99	·	·
3	79.99	39.99	79.99	60.03	79.99	68.00	16.0
4	60.03	79.99	60.03	79.99	60.03	·	·
5	60.03	79.99	20.01	39.99	20.01	44.00	23.3
6	79.99	5.30	9.97	5.30	39.99	28.13	29.0
7	79.99	20.01	39.99	5.30	20.01	33.07	25.9
8	39.99	60.03	39.99	79.99	20.01	48.00	20.4
9	79.99	5.30	9.97	0.50	1.00	19.36	30.5
10	60.03	79.99	60.03	79.99	60.03	·	·
11	60.03	79.99	39.99	60.03	9.97	50.00	23.7
12	79.99	60.03	79.99	20.01	79.99	64.00	23.3
13	60.03	79.99	60.03	79.99	39.99	64.01	15.0
14	39.99	79.99	1.00	5.30	3.05	25.87	30.6
15	79.99	60.03	79.99	60.03	79.99	·	·
16	60.03	79.99	20.01	39.99	20.01	44.01	23.3
17	79.99	20.01	39.99	20.01	39.99	40.00	21.9
18	79.99	3.05	20.01	9.97	20.01	26.60	27.5
19	3.05	9.97	5.30	9.97	5.30	6.72	2.8
20	60.03	79.99	20.01	39.99	5.30	41.06	26.8
				All (N = 16)		38.46% (SD = 17.59)	
				Mean (SD<20%, N = 4)		37.83 % (SD = 28.28)	



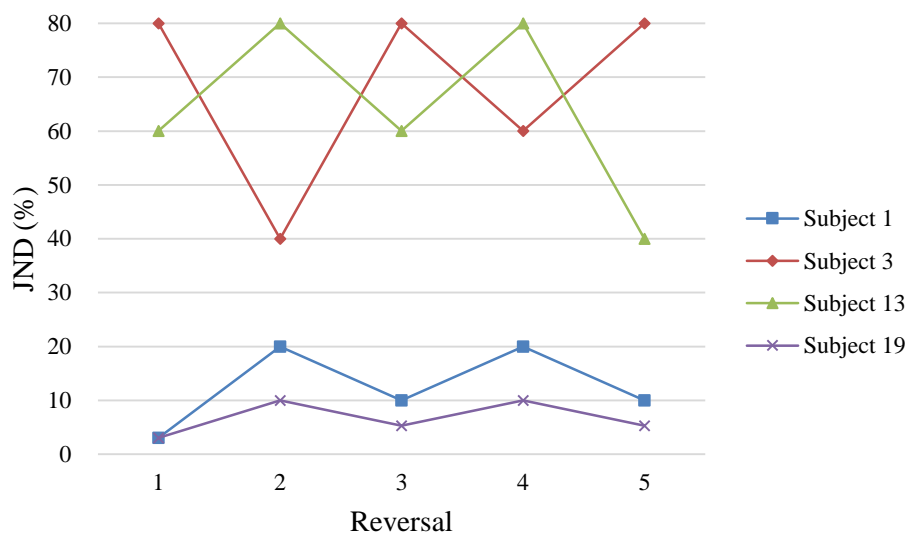


Figure 6.22 - Reversals from subjects who had SD < 20% with 5 Down condition

### 6.3.2.3 RD Label 10 and Up Direction

For the test with a reflection density of RD label 10 and up direction, some subjects made reversals at comparison sample #1 (Table 6.18). Although these subjects made incorrect answers at comparison sample #1, they could differentiate smaller differences than comparison sample #2, so they were all included for mean calculation. The percentage differences used for the RD label 10 with up direction case are shown in Table 6.19.

Table 6.18 - Comparison sample numbers at reversals of speech results with RD label 10 and up direction condition

Subject	Reversal					Subject	Reversal				
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th
1	3	1	2	1	2	11	4	3	4	3	6
2	2	1	3	2	3	12	2	1	2	1	4
3	4	3	4	3	4	13	4	3	5	3	5
4	2	1	2	1	3	14	2	1	3	2	5
5	5	3	6	4	5	15	5	4	6	5	6
6	3	2	5	3	5	16	5	4	5	3	5
7	4	3	4	3	5	17	2	1	5	4	5
8	5	4	6	4	5	18	2	1	2	1	3
9	3	2	3	2	4	19	7	4	8	6	8
10	2	1	4	3	6	20	5	3	5	4	5

Table 6.19 - Percentage differences of reflection density of speech comparison samples for RD label 10 and up direction case

Sample #	Designed % difference	Odeon		cS/4V	
		RD (Odeon, /ms)	Actual % difference	RD (cS/4V,/s)	Actual % difference
1	80	133.98	80.01	15.89	15.89
2	60	268.28	59.97	28.76	28.76
3	40	401.94	40.03	41.64	41.64
4	20	536.23	19.99	55.24	55.24
5	10	603.44	9.97	61.63	61.63
6	5.33	634.62	5.31	64.01	64.01
7	3	650.28	2.98	65.61	65.61
8	0.95	663.79	0.96	67.55	67.55
9	0.53	667.03	0.48	67.63	67.63
10	0	670.24	0	67.63	67.63

Table 6.20 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 10 and up direction. Subject 10, 12, 14, and 17 were excluded from the mean calculation (N=16).

Subject	Reversal						
	1st	2nd	3rd	4th	5th	Mean	SD
1	40.03	80.01	59.97	80.01	59.97	64.00	15.0
2	59.97	80.01	40.03	59.97	40.03	56.00	15.0
3	19.99	40.03	19.99	40.03	19.99	28.00	9.8
4	59.97	80.01	59.97	80.01	40.03	64.00	15.0
5	9.97	40.03	5.31	19.99	9.97	17.07	12.4
6	40.03	59.97	9.97	40.03	9.97	32.00	19.4
7	19.99	40.03	19.99	40.03	9.97	26.00	12.0
8	9.97	19.99	5.31	19.99	9.97	13.07	5.9
9	40.03	59.97	40.03	59.97	19.99	44.00	15.0
10	59.97	80.01	19.99	40.03	5.31	41.07	26.8
11	19.99	40.03	19.99	40.03	5.31	25.07	13.3
12	59.97	80.01	59.97	80.01	19.99	59.99	21.9
13	19.99	40.03	9.97	40.03	9.97	24.00	13.6
14	59.97	80.01	40.03	59.97	9.97	49.99	23.7
15	9.97	19.99	5.31	9.97	5.31	10.11	5.4
16	9.97	19.99	9.97	40.03	9.97	17.98	11.7
17	59.97	80.01	9.97	19.99	9.97	35.98	28.7
18	59.97	80.01	59.97	80.01	40.03	64.00	15.0
19	2.98	19.99	0.96	5.31	0.96	6.04	7.2
20	9.97	40.03	9.97	19.99	9.97	17.98	11.7
All (N = 20)					34.82% (SD = 18.88)		
Assorted (SD<20%, N = 16)					31.83% (SD = 19.50)		

Subject 10, 12, 14, and 17 exceeded SD of 20%, so these subjects were excluded from the final mean calculation. The mean JND was found to be 31.83 % (N = 16) with a standard deviation of 19.50% (Figure 6.20). The box plot is shown in Figure 6.23.

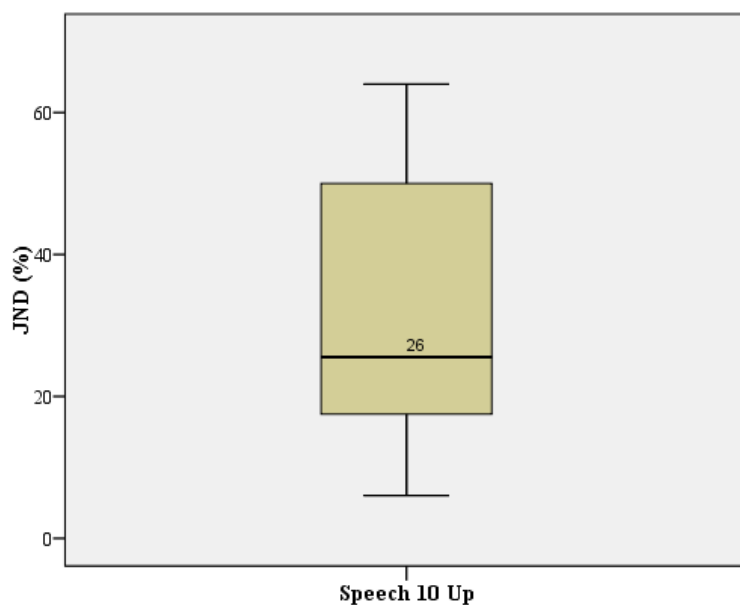


Figure 6.23 - Just noticeable difference of reflection density for speech signals with reference reflection density of RD label 10 and up direction. Median JND was 26%.

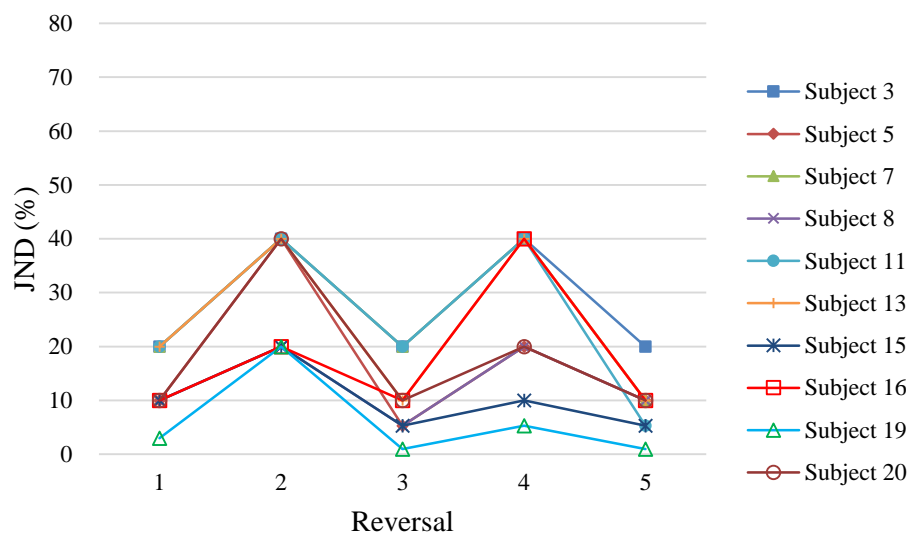


Figure 6.24 - Reversals from subjects who had  $SD < 15\%$  with 10 Up condition

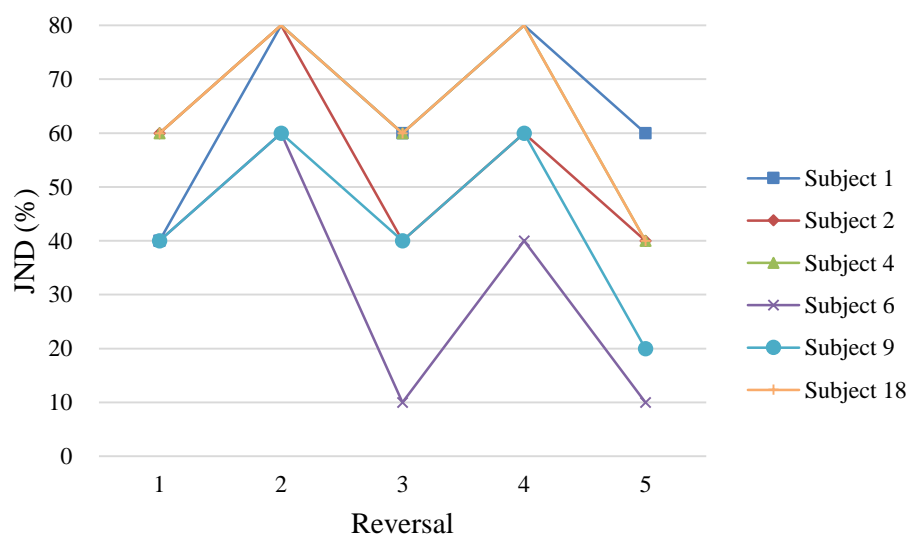


Figure 6.25 - Reversals from subjects who had  $15\% \leq SD < 20\%$  with 10 Up condition

### 6.3.3 *Just Noticeable Difference (JND) of Reflection Density*

The summarized just noticeable differences are shown in Table 6.21. ‘All’ group includes all subjects except who distinguished all comparison samples, and ‘Assorted’ group includes only subjects with SD less than 20%. The assorted JNDs using clapping and speech signals were 37.39% and 29.34%, respectively, and it was higher with 5 Down (RD label 5 and down direction) and 10 Up (RD label 10 and up direction) than with 5 Up (RD label 5 and Up direction) case.

Table 6.21 - JND of reflection densities (%) with clapping and speech signals using Odeon provided reflection densities

		5 Up	5 Down	10 Up	Mean
Clapping	All	24.17 (N = 16)	55.59 (N = 15)	32.59 (N = 18)	37.45
	Assorted		56.87 (N = 9)	31.12 (N = 13)	37.39
Speech	All	23.22 (N = 15)	38.46 (N = 16)	34.82 (N = 20)	32.17
	Assorted	18.35 (N = 12)	37.83 (N = 4)	31.83 (N = 16)	29.34

These results are depicted graphically in Figure 6.26.

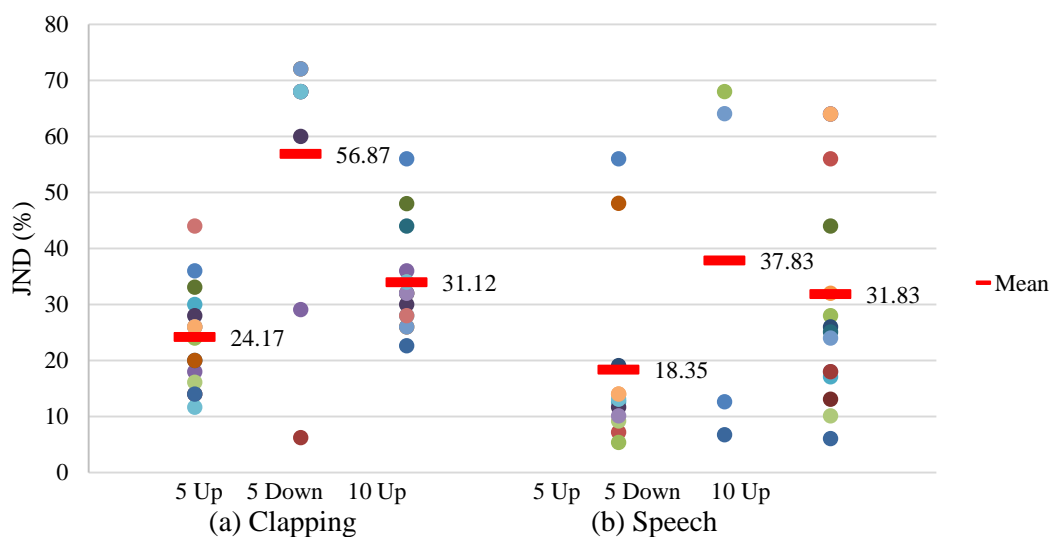


Figure 6.26 - JND of reflection density (using Odeon provided values) from testing with (a) clapping signals, or (b) speech signals

Just noticeable differences were located between 20% and 40% except for the 5 down case using clapping signals (56.81%). One observation is that the 5 Up case has lower JND (around 24%) than the 5 Down or 10 Up cases. Also, speech signals resulted in larger deviations than clapping signals.

The JND was also checked with reflection densities calculated by cS/4V. The values of reflection densities used are shown in Table 6.4, Table 6.7, Table 6.10, Table 6.13, Table 6.16, and Table 6.19. Only subjects included for the mean calculation when using Odeon provided reflection densities were included.

Table 6.22 - JND of reflection densities (%) with clapping and speech signals using reflection densities calculated by cS/4V

	5 Up	5 Down	10 Up	Mean
Clapping	22.59	53.10	32.16	35.95
Speech	17.07	35.63	30.20	27.63

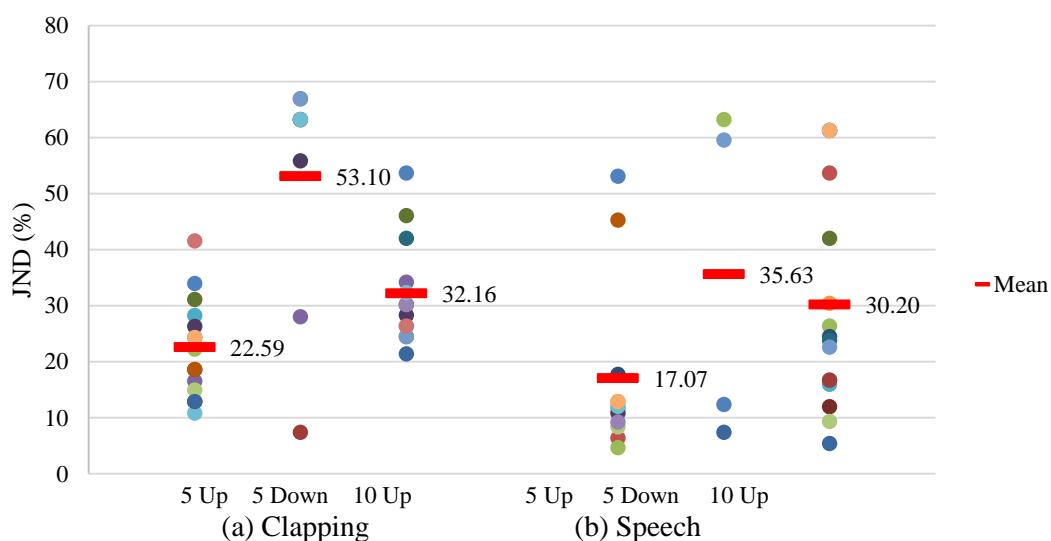


Figure 6.27 - JND of reflection density (using cS/4V) from testing with (a) clapping signals, or (b) speech signals

The JND results were generally similar to those calculated with the Odeon provided reflection densities. For the 5 Up case, JNDs were smaller than when using Odeon provided reflection density.

For both clapping and speech signals, 5 Up cases had lower JNDs than 5 Down or 10 Up. This was analyzed further with the relationship between upper limits of distinguishable reflection density studied in Chapter 5.



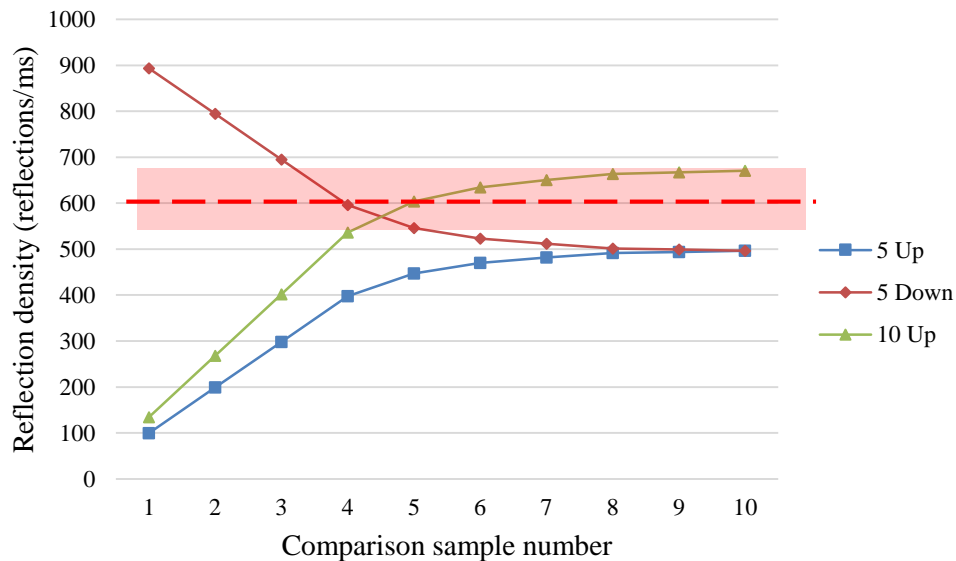


Figure 6.28 - Reflection densities of tested comparison signals and upper limit of distinguishable reflection density

The upper limits found in Chapter 5 were around 600 reflections/sec, indicated by a dotted line in Figure 6.28. The shaded range is an imaginary standard deviation. For the 5 Up case, all comparison reflection density samples (sample number 1 to 9) and the reference reflection density (sample number 10) are below the upper limit of distinguishable reflection density. For the 5 Down case, many comparison samples are above the upper limit of distinguishable reflection density, and this could be the reason that 5 Down case has higher JND than others. Also for the 10 Up case, later comparison samples might be above the upper limit of differentiation, so it possibly increased the JND. For this reason, the JND for the 5 Up case may be more valid.

## 6.4 Conclusions

The just noticeable difference (JND) of reflection densities using clapping and speech signals were investigated with two reference reflection densities (RD label 5 and RD label 10) and two approaching directions (up and down). The JNDs for 5 Up (RD label 5 and up direction) were 23% and 17% for clapping and speech signals, respectively, and it was smaller than in the 5 Down or 10 Up cases. Some of the comparison reflection densities in the 5 Down and 10 Up cases were above the upper limit of distinguishable reflection density determined in Chapter 5, which may account for the higher JNDs in those cases. For this reason, the JNDs of 24% and 18% found for the 5 Up case using the clapping and speech signals, respectively, may be more trustworthy results.

## Chapter 7 - Conclusion

### 7.1 Summary of Findings and Conclusions

In this dissertation, the perception of varying reflection density in impulse responses has been studied extensively. To investigate the sensitivity of the human auditory system to reflection density, subjective tests to study the upper limit of distinguishable reflection density and the just noticeable difference (JND) of reflection density have been performed. The room impulse responses (RIR) were generated by Matlab code and Odeon software, and then those RIRs were convolved with either a clapping or a speech signal. Reflection densities were controlled by Matlab code or through variation of room volume in Odeon. In order to keep reflection density as a main cue to differentiate testing samples, reverberation time (RT) was maintained to be the same in all testing groups while reflection densities changed between testing samples. The three-alternative forced-choice (3AFC) method paired with 1-up 2-down adaptive method was used. In the subjective test, three sound samples were presented; two samples were the same fixed reference samples and one was the comparison sample with different reflection density. Subjects were asked to choose the one which was different from the others, and the next testing reflection density was controlled by a subject's answer. With the 1-up 2-down method, one incorrect answer increased the difference between the comparison and reference sample, and two consecutive correct answers decreased the difference. When a subject's answers were changed from incorrect to

correct or correct to incorrect, these trials were marked as reversals, and reflection densities attained at the first five reversals were averaged and recorded to calculate the upper limit or JND.

In Study 1, RIRs were generated as a discrete stochastic process by Matlab code. Reflection densities from 500 reflection/sec to 10000 reflection/sec under three RT conditions (0.3, 1, and 2 sec) were tested. The source signal was a clapping signal with three claps in a row. The test subject could play testing samples as many times as they wanted. Results show that listeners performed similarly under the three different RT conditions. The upper limit of distinguishable reflection density was found to be around 3700 reflections/sec (Table 7.1)

Table 7.1 - Upper limits of distinguishable reflection density (reflections/sec, generated by Matlab) under different RT conditions

RT (sec)	0.1	1	2	Mean
All	2640 (N = 10)	4060 (N = 10)	3700 (N = 10)	3467
Assorted (RSD<30%)	3000 (N = 6)	4300 (N = 4)	3875 (N = 4)	3725

In Study 2, RIRs were generated from different sized rooms simulated in Odeon software. Absorption of all surfaces were uniformly managed to maintain 1 sec of RT ( $T_{30}$ ), and twenty reflection densities were used. Reflection densities of comparison samples had linear variation between samples. Unlike in Study 1, test subjects had to listen to a testing pair only three times. The source signals were clapping (five claps) and

speech, and upper limits were found to be 654 reflections/ms (N = 8) and 858 reflections/ms (N = 2), respectively (Table 7.2). Speech signals showed much larger deviation than clapping signals.

Table 7.2 - Upper limit of distinguishable reflection density (reflections/ms, Odeon-provided) under a RT 1sec condition

Source	Clapping	Speech	Mean
Upper limit (reflections/ms)	654 (N = 8)	858 (N = 2)	756

In Study 3, the just noticeable difference (JND) was investigated, and RIRs were generated by the same method used for Study 2. Reflection densities of comparison samples did not vary linearly, but had larger differences farther away from the reference signal. Two source signals were again tested: clapping (three claps) and speech signals. Similar to Study 2, test subjects could listen to each testing pair for a fixed repetition, reduced to two times for Study 3. Two reference reflection density were chosen from Study 2, which were RD label 5 (496 reflections/ms) and RD label 10 (670 reflections/ms). JNDs were found to be larger in the 5 Down (RD label with down direction) and 10 Up (RD label 10 and up direction) cases than in the 5 Up (RD label 5 with up direction), as shown in Table 7.3.

Some comparison reflection densities in the 5 Down and 10 Up groups were higher than the upper limit found in Study 2, which could be a reason that the JNDs were higher in the 5 Down and 10 Up cases. For this reason, the JND in the 5 Up condition

could be more valid than from the 5 Down or 10 Up conditions, and it was around 21%.

Table 7.3 - JND (%) of reflection densities with clapping and speech signals using reflection densities calculated by Odeon

		5 Up	5 Down	10 Up	Mean
Clapping	All	24.17 (N = 16)	55.59 (N = 15)	32.59 (N = 18)	37.45
	Assorted		56.87 (N = 9)	31.12 (N = 13)	37.39
Speech	All	23.22 (N = 15)	38.46 (N = 16)	34.82 (N = 20)	32.17
	Assorted	18.35 (N = 12)	37.83 (N = 4)	31.83 (N = 16)	29.34
Overall	All	23.70	47.03	33.71	34.81
	Assorted	21.26	47.35	31.48	33.36

In conclusion, it was verified that different reflection densities under the same RT are possible, and there was a limitation in perceiving the highest reflection density and differentiating change of reflection density. From the findings in this dissertation, the perceptual limits of reflection density are summarized in Table 7.4.

Table 7.4 – Perceptual limits of reflection density

	Clapping	Speech
Upper limit of distinguishable reflection density	654 reflections/ms (N = 8)	858 reflections/ms (N = 2)
JND	24.17 % (N = 16)	18.35% (N = 12)

These values apply to Odeon simulated impulse responses using a transition order

of 2 and late rays of 10,000 in the Odeon setting, taken as a single number for late reflection density provided by Odeon. However, as reviewed in this thesis, reflection density should increase with time. With the Odeon simulated impulse responses, it was not possible to use  $cS/4V$ , and the dB cut-off and time window method did not vary as expected; however, as proposed in the next section, future work using impulse responses from real rooms is encouraged. In real rooms, these two other metrics may perform more robustly and be more generally applicable.

## **7.2 Future Work**

This dissertation identified perceptual limits of reflection density to understand better reflection density as an additional acoustical parameter for room acoustics. However, there was a limitation in that impulse responses for subjective tests were artificially generated by software (Matlab and Odeon), so it was less realistic than measured impulse responses from real rooms. This dissertation also had a limited number of subjects, and many of subjects were excluded due to the large deviation in their responses. One of the testing signals was an English sentence, but this study did not control for native or non-native English listeners which may have affected the results. Also the current method of calculating reflection density may not work well for spaces with obstacles or major surfaces in the enclosure such as classroom desks, hospital equipment, partial office partitions, etc. Consequently, three directions for future work

are summarized below.

*1) How do subjects perform with measured room impulse responses?*

This dissertation used simulated room impulse responses by Matlab or Odeon, so it was less realistic than using measured impulse responses. Future work using measured impulse responses is suggested, perhaps gathered from a fixed volume room with active acoustics. Such a space could be used to generate physically a number of impulse responses that have the same RT but different reflection densities that could be used in subjective testing, as conducted in this thesis.

*2) Better method to quantify reflection density from measured room impulse responses*

This dissertation discussed three methods to quantify the reflection density:  $cS/4V$ , Odeon-provided values, and dB cut-off and time window. In particular, the theoretical mean reflection density and the dB cut-off and time window methods could be explored further using impulse responses from real rooms.

*3) How does reflection density affect perception of room acoustics?*

The reflection density is generally inversely proportional to the volume, so it may relate to room size perception. With upper limit of distinguishable reflection density and JND studied in this dissertation, the relationship between reflection density and room size perception could next be studied effectively. Also, the degree of reflection density that people prefer for speech or music could be another area to study.



## References

- International Organization for Standardization (2003). “ISO 226:2003 Acoustics – Normal Equal-Loudness-Level Contours
- International Organization for Standardization (2009). “ISO 3382-1:2009 Acoustics – Measurement of Room Acoustic Parameters – Part 1: Performance Spaces
- Barron, M., & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibration*, 77(2), 211–232.
- Bolt, R. H., Doak, P. E., & Westervelt, P. J. (1950). Pulse Statistics Analysis of Room Acoustics, 22(3), 328–340.
- Bork, I. (2000). A comparison of room simulation software - The 2nd Round Robin on Room Acoustical Computer Simulation. *Acustica*, 86(6), 943–956.
- Bork, I. (2005a). Report on the 3rd round robin on room acoustical computer simulation - Part I: Measurements. *Acta Acustica United with Acustica*, 91(4), 740–752.
- Bork, I. (2005b). Report on the 3rd round robin on room acoustical computer simulation - Part II: Calculations. *Acta Acustica United with Acustica*, 91(4), 753–763.
- Cabrera, D., & Gilfillan, D. (2002). Auditory distance perception of speech in the presence of noise. In *Int. Conf. on Auditory Display, Kyoto* (pp. 1–9).
- Christensen, C. L. (2011). *User Manual, Odeon Room Acoustics Software Version 11* (2nd ed.).
- Damaske, P., & Ando, Y. (1972). Interaural Crosscorrelation for Multichannel Loudspeaker Reproduction. *Acustica: Spatial Sound Techniques: Part 2*, 27(4), 232–238.
- Fletcher, H., & Munson, W. A. (1933). Loudness, its Definition, Measurement and Calculation. *Journal of the Acoustical Society of America*, 5, 82–108.
- Gescheider, G. A. (1997). *Psychophysics: The Fundamentals*. scan psychology (Vol. 435).
- Griesinger, D. (1989). Practical processors and programs for digital reverberation. *Audio*

- Engineering Society Conference: 7th International Conference: Audio in Digital Times*, 5, 187–195.
- Haas, H. (1972). The Influence of a Single Echo on the Audibility of Speech. *Journal of the Audio Engineering Society*, 20(2), 146–159.
- Hameed, S., Pakarinen, J., Valde, K., & Pulkki, V. (2004). Psychoacoustic Cues in Room Size Perception. *Area*, 1–7.
- Hyland, M. C. C. M. E. (2012). *History and philosophy of psychology*. Hoboken : John Wiley & Sons.
- Jeon, J. Y., Jang, H. S., & Kim, Y. H. (2013). Subjective and objective evaluations of scattered sounds in concert halls. *Proceedings of International Symposium on Room Acoustics*, 1–7.
- Jeong, C.-H., Brunskog, J., & Jacobsen, F. (2010). Room acoustic transition time based on reflection overlap. *The Journal of the Acoustical Society of America*, 127(5), 2733–2736.
- Jot, J.-M., Jot, J.-M., Chaigne, A., & Chaigne, A. (1991). Digital Delay Networks for Designing Artificial Reverberators. *AES Convention*, 3030, preprint no. 3030.
- Kingdom, F., & Prins, N. (2010). *Psychophysics: a practical introduction*.
- Kollmeier, B., Gilkey, R. H., & Sieben, U. K. (1988). Adaptive staircase techniques in psychoacoustics: a comparison of human data and a mathematical model. *The Journal of the Acoustical Society of America*, 83(5), 1852–1862.
- Krueger, D., Jeong, C., Brunskog, J., & Buchholz, J. (2012). Audible reflection density for different late reflection criteria in rooms. *Internoise*.
- Kuttruff, H. (2009). *Room Acoustics*. CRC Press.
- Langmuir, I. (1981). *A study of binaural perception of the direction of a sound source*. [Springfield Va.]: [NTIS].
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2)
- Long, M. (2006). *Architectural Acoustics*. Elsevier/Academic Press.
- McGrath, R., Waldmann, T., & Fernström, M. (1999). Listening to rooms and objects.

- AES 16th International Conference on Spatial Sound Reproduction*, 1–11.
- McKee, S. P., Klein, S. A., & Teller, D. Y. (1985). Statistical properties of forced-choice psychometric functions: implications of probit analysis. *Perception & Psychophysics*, 37(4), 286–298.
- Mershon, D. H., Ballenger, W. L., Little, a D., McMurtry, P. L., & Buchanan, J. L. (1989). Effects of room reflectance and background noise on perceived auditory distance. *Perception*, 18(3), 403–16.
- Michelsen, J., & Rubak, P. (1997). Parameters of distance perception in stereo loudspeaker scenario. *Methods*.
- Nielsen, J. B., & Dau, T. (2009). *Assessment of speech intelligibility in background noise and reverberation*. Technical University of Denmark.
- Pop, C. B., & Cabrera, D. (2005). Auditory Room Size Perception for Real Rooms. In *Proc. of Acoust., 2005* (pp. 115–121).
- Robinson, D. W., & Dadson, R. S. (1956). A re-determination of the equal loudness relation for pure tones. *British Journal of Applied Physics*, 7, 166–181.
- Rubak, Per; Johansen, L. G. (1998). Artificial Reverberation Based on a Pseudo-Random Impulse Response I (Vol. 4725).
- Rubak, Per; Johansen, L. G. (1999). Artificial Reverberation Based on a Pseudo-Random Impulse Response II, 900.
- Sabine, W. C. (1922). *Collected papers on acoustics*. Cambridge : Harvard University Press.
- Sandvad, J. (1999). Auditory Perception of Reverberant Surroundings. *The J. of the Acoustical Soc. of America*, 105(2), 1193.
- Schlauch, R. S., & Rose, R. M. (1990). Two-, three-, and four-interval forced-choice staircase procedures: estimator bias and efficiency. *The Journal of the Acoustical Society of America*, 88(2), 732–740.
- Schroeder, M. R. (1962). Natural Sounding Artificial Reverberation. *J. Audio Eng. Soc*, 10(3), 219–223.
- Schroeder, M. R. (1965). New Method of Measuring Reverberation Time. *The Journal of*

- the Acoustical Society of America*, 37(6), 1187.
- Schwenke, R., & Long, B. (2011). Physiological and Psycho-Acoustic basis for Multi-Resolution Frequency Response Analysis. *SMPTE Conf. Proc*, 1100.
- Shinn-Cunningham, B. G. (2003). Acoustics and perception of sound in everyday environments. *Proc. 3rd Int. Workshop on Spatial Media*, 1–9.
- Vorlander, M. (1995). International Round Robin on Room Acoustical Computer Simulations. *15th International Congress on Acoustics*.
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, 27(4), 339–368.
- Wallach, H., Newman, E. B., & Resenzweig, M. R. (1949). The Precedence Effect in Sound Localization. *The American Journal of Psychology*, 62(3), 315–336.
- Watson, C. S., & Gengel, R. W. (1969). Signal duration and signal frequency in relation to auditory sensitivity. *The Journal of the Acoustical Society of America*, 46(4), 989–997.

## Appendix A Sound Booth

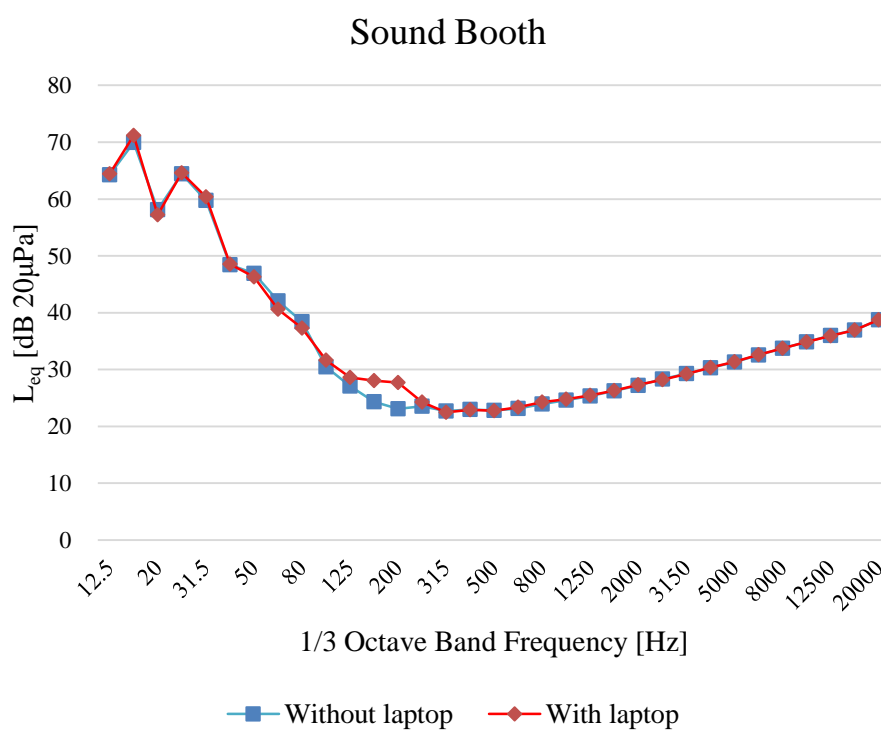


Figure A. 1 - Ambient background noise level in sound booth with/without laptop