Interaction Between Loudspeakers and Room Acoustics Influences Loudspeaker Preferences in Multichannel Audio Reproduction

Sean E. Olive

Degree of Doctor of Philosophy

Schulich School of Music

McGill University

Montreal, Quebec, Canada

October 2007

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-50971-5
Our file Notre référence
ISBN: 978-0-494-50971-5

ABSTRACT

This dissertation examined the influence of room acoustic variations on listener loudspeaker preferences in multichannel audio reproduction, and the extent to which listeners adapt to the room acoustics. Listener preferences among 4 different 5-channel loudspeaker systems were observed in 4 different virtual listening rooms via a binaural room scanning (BRS) measurement and playback system. To study room acoustic adaptation, 2 groups of listeners evaluated identical stimuli according to 2 different trial blocking schemes, termed either the successive or the intermixed treatment. In the successive treatment condition, the loudspeakers were evaluated under a given room acoustic condition before moving to the next block of trials, under a different room acoustic condition. In the intermixed treatment condition, the loudspeakers were evaluated under a different room acoustic condition on each trial within a given block of trials. Although loudspeaker preferences did not differ between these two different trial-blocking schemes, there was a significant influence of room acoustics on loudspeaker preferences. The room acoustic variation to which listeners were exposed also significantly influenced the observed patterns of loudspeaker preferences. These results can be summarized simply as follows: The most reflective listening room produced the lowest overall preference ratings, and also enabled listeners to report the largest differences in loudspeaker preference. It was also found that experienced listeners were more discerning of loudspeaker effects, whereas less experienced listeners were more influenced by room effects. The results indicate that both listening experience and room acoustics significantly influence how listeners formulate loudspeaker preferences.

RÉSUMÉ

Les travaux présentés dans cette thèse ont examiné l'influence qu'ont les variations acoustiques des salles d'écoute sur les préférences des auditeurs pour des haut-parleurs utilisés dans des systèmes de reproduction audio multivoie. Ces travaux ont de plus étudié dans quelle mesure ces auditeurs s'adaptent à l'acoustique d'une salle d'écoute. Les préférences des auditeurs parmi cinq systèmes différents à 5 canaux ont été étudiées dans quatre salles d'écoute virtuelles générées à l'aide d'un système de mesure et de reproduction par balayage binaural de salle (BRS). Dans le but d'étudier l'adaptation à l'acoustique d'une salle, deux groupes d'auditeurs ont évalué des stimulus identiques en utilisant deux procédés par blocs d'épreuves différents, soit un traitement dit consécutif et un traitement dit entremêlé. Dans le traitement consécutif, tous les haut-parleurs ont été évalués sous la même condition acoustique avant de présenter le bloc d'épreuves suivant sous une condition acoustique différente. Dans le traitement entremêlé, les haut-parleurs furent évalués sous une condition acoustique différente à chaque épreuve à l'intérieur d'un bloc d'épreuves donné. Bien que les préférences de haut-parleurs n'aient

pas différé entre les deux types de blocs d'épreuves, l'acoustique des salles a eu un effet significatif sur ces préférences. Les variations de l'acoustique auxquelles les auditeurs ont été soumis ont aussi influencé de façon significative les schémas de préférence observés. Les résultats obtenus peuvent se résumer ainsi: les évaluations de préférence d'ensemble les plus basses furent obtenues dans la salle d'écoute la plus réflective. C'est aussi dans cette salle que furent observées les différences les plus grandes dans les préférences de haut-parleurs. De plus, les auditeurs expérimentés ont mieux perçu les différences introduites par les haut-parleurs, tandis que les auditeurs moins expérimentés furent plus influencés par les différences causées par les changements d'acoustique. Ces résultats indiquent que tant l'expérience d'écoute que l'acoustique des salles influencent de façon significative les préférences de haut-parleurs exprimées par les auditeurs.

ACKNOWLEDGEMENTS

I would like to thank my PhD advisors, Dr. William Martens and Dr. Wieslaw Woszczyk for their expertise, guidance, and generous support throughout the course of my PhD studies. Their friendship, humour, and good company have made this a most pleasurable collaboration. Bill Martens deserves special credit for his thoughtful help with the final draft as well as his influential methodological contributions to this research, which forced me to stretch outside the confines of my corporate research culture. The stretch was certainly nice while it lasted. Finally, I wish to thank Dr. René Quesnel for translating the abstract into actual French from the bungled 'Google' version.

I also want to thank my many colleagues at Harman International for their support and important contributions to various aspects of this work. Todd Welti developed the superb binaural room scanning (BRS) calibration used in the research, and together with John Jackson, helped gather the 4,720 BRS measurements needed for this study. John Jackson made a significant contribution to this work, administering the bulk of the listening tests and taking the objective measurements of the loudspeakers. Allan Devantier engineered the adjustable loudspeaker used in this study. Eric Hu re-engineered much of the BRS software in order to meet the demanding requirements of this study. Sean

Hess wrote parts of the listening test software used in these experiments, and along with David Hunt, helped out, both as a listener and in the administration of the tests. Mae Nutley and Kristina Busenitz provided valuable advice in statistical matters, proof-reading, and, along with Phil Muzio, were very supportive throughout. More than 30 listeners, both from within and outside of Harman, participated in these tests and I sincerely appreciate their time and efforts.

Harman International provided 100% of the funding for my PhD studies, and granted me the time, flexibility, and freedom to pursue my goals: For that I am extremely grateful. In this regard, I owe sincere thanks to both my current and former supervisors; respectively, Tim Nind and Dr. Floyd Toole. They have been supportive and patient beyond belief. Floyd's pioneering research into loudspeakers and rooms has been a motivating factor in the author's interest in this field; its influence on the work is undeniable.

My parents, John and Mary Olive deserve credit for fostering and supporting my early interests in music and learning, and for starting me down the correct path in life.

Finally, I want to especially thank my wife Valerie, who, at my advanced age, somehow convinced me to embark on a PhD when most people my age are thinking about golf. She has played a supportive and instrumental role in my success throughout this process, including the editing and proof-reading of this lengthy tome. The dissertation is largely compliant with APA style due to her mastering of its hundreds of mind-boggling rules that only a lawyer like her could begin to interpret correctly. But most of all, I thank her for our two young children, Conor and Mariah, who both “hate the PhD,” and have been patiently waiting for its completion. To Conor and Mariah: I can honestly say, “I am ready to play.”

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ASW	Apparent Source Width
BRIR	Binaural Room Impulse Response
BRS	Binaural Room Scanning
df	Degree of Freedom
ETC	Energy-Time Curve
F	Fisher's F Ratio (also referred to as the F Statistic)
GUI	Graphical User Interface
HRTF	Head-Related Transfer Function
IACC	Interaural Cross-Correlation
KEMAR	Knowles Manikin for Acoustic Research
LEV	Listener Envelopment
r	Pearson's Product-Moment Correlation
p	Probability
Q	Q Value of a Resonance (also referred to as the Q Factor)
RT_{60}	Reverberation Time

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction

We experience most of our music and entertainment today by listening to recordings played through loudspeakers in our homes, and increasingly, in our automobiles. There is general consensus among audio experts that the quality of these recordings and their reproduction is significantly influenced by interactions between the loudspeakers and the room acoustics of these listening spaces.

There are many scientific studies, recently summarized by Toole (2006), that document the physical effects of acoustical interactions between the loudspeaker and listening room. Below a transition frequency of approximately 300 Hz, the room adds its own set of resonances, causing large seat-to-seat variations in the frequency response, which is also dependent on the locations of the loudspeakers. Above the transition region, the quality and proportion of the direct and reflected sounds received at the listener's ears are influenced by the acoustical properties of the loudspeakers and room, as well as the locations of the listeners and loudspeakers. The physical effects of these interactions are easily seen in acoustical measurements made at the listening locations. However, the perception of these loudspeaker-room interactions, and their effect on the quality of recorded and reproduced sounds, are generally not well understood.

Some studies have reported that the position of the loudspeaker can produce preference differences, which, in some cases, are larger than those measured among different models and brands of loudspeakers (Bech, 1993; Olive, Schuck, Sally, & Bonneville, 1994, 1995).[1] However, few studies have shown how listeners' loudspeaker preferences vary when the loudspeakers are compared in different listening rooms (Bech, 1993; Klippel, 1990; Olive et al., 1995; Schuck et al., 1993).

One such study found that switching between four different listening rooms had no significant effect on listeners' loudspeaker preferences (Olive et al., 1995). This was not the expected or desired result, because the researchers were developing a room-adaptive loudspeaker that was aimed at correcting loudspeaker-room interactions (Schuck et al., 1993). In their study, the authors expressed fear that room-adaptive loudspeakers might fall under the same category as exotic audio cables: "If it cannot be demonstrated in a convincing

[1] Few studies have systematically isolated and manipulated the loudspeaker-room interactions that occur below and above the transitional frequency.

fashion that different rooms affect the sound quality of loudspeakers, then it will be difficult to sell a product whose purpose is to correct what cannot essentially be heard" (Schuck et al., 1993, p. 67).

To extricate themselves from this dilemma, the researchers did something quite unexpected to enhance the detection of room effects on loudspeaker preferences. Through the means of a binaural recording-playback system, listeners were able to make contemporaneous comparative ratings of the same loudspeaker among the four rooms without a time gap in between the comparisons. This produced the opposite, but desired, result: The listening room became the dominant variable in listeners' preferences, and the loudspeaker variable had no effect. The researchers were quite happy with the significant room effect, because it meant funding for the room-adaptive loudspeaker project was no longer in jeopardy.

The above experiment had two important findings: (1) listeners were much better at adapting to the room acoustics than the room-adaptive loudspeakers were, and (2) the experimental results could be manipulated by changing the order and context in which the stimuli were compared. Listeners appeared either to adapt 100% to the room acoustics, or not at all, depending on the context under which the rooms and loudspeakers were compared. This result makes clear the need to be careful about how potential room adaptation effects are treated in loudspeaker preference studies.

Room acoustic adaptation is defined as a change in the listener's response to the room acoustics after having spent some period of time in the room. In the context of subjective evaluation of different loudspeakers in different rooms, room acoustic adaptation would be observed as a change in how the room acoustics affect the loudspeaker ratings. Within certain limits, more time spent in the room should produce more adaptation (and consequently less observed room effect), whereas less time spent in the room should produce less adaptation (and consequently greater room effect). However, the time course of adaptation to room acoustics as an influence on the perception of reproduced sound is largely unknown. Therefore, manipulations intended to modulate room acoustic adaptation may or may not produce the desired variation in a listener's state.

The precedence effect (Blauert, 1996; Blauert & Braasch, 2005; Buchholz, Mourjopoulos, & Blauert, 2001; Clifton, 1987; Clifton & Freyman, 1997; Dizon & Colburn, 2006; Djelani & Blauert, 2001; Haas, 1972; Hartmann, 1997; Litovsky, Colburn, Yost, & Gutzman, 1999; Litovsky, Hawley, Fligor, & Zurek, 2000; Litovsky & Shinn-Cunningham, 2001; Lochner & Burger, 1958; Rakerd & Hartmann, 1986, 1992, 1994; Saberi & Antonia, 1990; Saberi & Perrott, 1990;

Shinn-Cunningham, 2000, 2001, 2004; Shinn-Cunningham, Durlach & Held, 1998a, 1998b; Shinn-Cunningham & Ram, 2003; Wallach, Newman, & Rosenweig, 1949; Zurek, 1979) and spectral compensation (Watkins, 1991, 1999, 2005; Watkins & Makin, 1994, 1996a, 1996b, 2007) are two well-known room acoustic adaptation mechanisms that enable humans to perceive the timbre, direction, and intelligibility of a sound source in reflective listening spaces. Both mechanisms involve a central auditory cognitive decision process that uses current and recently received auditory information, and cues from other sensory modalities. Information that is potentially redundant, irrelevant, or implausible is suppressed, indicating that adaptation has occurred (Blauert, 1996).

Research in room acoustic adaptation has important implications for methods used in conducting listening tests, and the design of loudspeakers and rooms used for audio recording and reproduction. The current standards and methods for designing listening rooms (ITU-R BS.775-1, 1994; Producers & Engineers Wing Surround Sound Recommendation Committee, 2004), loudspeakers, and loudspeaker/room-interaction correction products may not adequately account for human perception and adaptation to room acoustics. Indeed, the application of technology to solve loudspeaker-room problems may be superfluous if human perception has already taken care of them.

However, there are many unanswered questions about room acoustic adaptation. The parameters of listening conditions under which room adaptation may or may not be observed are not well established. Room acoustic adaptation has not been studied under multichannel audio listening conditions, where the precedence effect and spectral compensation may also operate differently than expected from the results of studies using fewer loudspeakers. The room acoustics, the number of loudspeakers, the complexity of the listener's task and the attention it requires, and, finally, the test signals and recording techniques, may all influence how well these mechanisms work and how much adaptation occurs.

There are many methodological challenges in room acoustic adaptation research. The independent variables (different loudspeakers, rooms, loudspeaker positions, and room exposure time) must be manipulated in a way that permits real-time, double-blind, comparative evaluations. This is not possible or practical using in situ listening test methods. For this reason, a *binaural room scanning* (BRS) method (Bech, Gulbol, Martin, Ghani, & Ellermeir, 2005; Christensen et al., 2005; Horbach, Karamustafaoglu, Pellegrini, MacKensen, & Thiele, 1999; Olive, Welti, & Martens, 2007; Pellegrini, 2001) was chosen for this study because it allows the independent variables to be captured and stored as a set of *binaural room impulse responses* (BRIR), that may later be used to create virtual room acoustic reproductions through high-quality headphones equipped with a low-latency, head-tracking system.

This dissertation reports on an experimental investigation into the influence of loudspeaker and room acoustical interactions on listener preference for multichannel music imagery, with the potential to reveal the role room acoustic adaptation plays in their perception. The main research questions and experimental methods used to address these questions are defined in the next section.

1.1 Main Research Questions and Experimental Methods

The three main research questions of this dissertation are succinctly defined as follows:

1. To what extent are listeners' preferences for multichannel music imagery influenced by variation in loudspeakers, variation in room acoustics, and interactions between these two variables?
2. Can the effects and interactions of loudspeakers and room acoustics on listeners' preferences be explained by acoustical measurements of the loudspeakers and room acoustics?
3. To what extent has room acoustic adaptation diminished the effect of room acoustics on listeners' preference ratings?

To address these three questions, four surround loudspeaker systems were evaluated in four different listening rooms using three different five-channel music selections. The four loudspeakers differed only in their off-axis frequency response above 300 Hz; meaning that listeners heard loudspeaker variations based only on the differences in the reflected sounds produced by them.

Two groups of listeners were given differential exposure to the listening room acoustics via a BRS measurement and playback system. One group of listeners gave loudspeaker preference ratings while the listening room was held constant throughout the listening session, and only changed between subsequent sessions, a method termed the *successive treatment condition.* A second group of listeners performed the same task under conditions in which the listening room was changed between each trial within the listening session, a method termed the *intermixed treatment condition.* By comparing the differences in preferences between the successive and intermixed treatment conditions, the influence of room acoustic adaptation on sound reproduction can be assessed.

1.2 Original Aspects of This Research

There are at least five original aspects of this research in terms of its subject matter and choice of experimental methodology. To the best of the author's knowledge, it is the first study that has carefully examined:

1. The perception of loudspeaker and room acoustical interactions on multichannel music reproduced through a multichannel (3/2) playback system.
2. The role of room acoustic adaptation on the perception of interactions between loudspeaker and room acoustics.
3. The influence of a loudspeaker's off-axis performance and its interaction with room acoustics on listeners' preference ratings, particularly within the context of multichannel music reproduction. Previous studies have either used monophonic or stereo playback systems, and have not controlled both the direct and the reflected sounds produced by the loudspeaker to allow assessment of their relative importance in different rooms. This is the first study to carefully manipulate the off-axis performance of the loudspeaker above 300 Hz, while keeping the direct sound constant across all room acoustic conditions. It is also one of the first studies to subjectively measure loudspeaker-room interaction effects above 300 Hz, while keeping the effects of room modes below 300 Hz constant.
4. The use of a BRS measurement and playback system to study the perception of loudspeaker and room acoustics. The novel application of BRS technology in this research overcomes most of the experimental and

methodological challenges that have previously thwarted research efforts in this important area.

5. Experimental verification of the performance of a BRS system for this application (see section 3. 9). A key reason for the accurate performance of the BRS system used in this study is the proprietary BRS calibration procedure developed by the author's colleague, Todd Welti (see section 3.8).

1.3 Intended Audience and Applications for This Research

The author believes that scientific research should be judged not only on its scientific merit and originality, but also its real-world application and value to society. The research topic of this dissertation clearly meets both criteria.

A review of the sparse scientific literature on the perception of interactions between loudspeakers and room acoustics (see chapter 2) provides sufficient motivation and justification for this research based on the lack of knowledge about their perception. The potential practical value and application of this research is also quite clear; new knowledge may lead to improvements in the measurement, design, and performance of loudspeakers and listening rooms used in the production and reproduction of audio. Given that a large percentage of society spends a significant amount of time listening to reproduced sound in their cars, homes, and workplaces, there is a substantial audience that could potentially benefit from research of this kind.

Three different groups may potentially benefit from the author's research: (1) audio engineers involved in the production and reproduction of audio; (2) audio consumers; and (3) scientists working in the areas of audio, acoustics, and psychoacoustics. A summary of the various audiences who might be interested in, and benefit from, this research is given below.

1.3.1 The Professional Audio Industry

In 2006, the professional audio industry enjoyed about $4.74 billion (USD) in sales of audio products worldwide.[2] The professional audio industry is well represented by the Audio Engineering Society. The Audio Engineering Society has over 13,000 members worldwide (Furness, 2006). Their members include recording artists, recording producers and engineers, audio scientists, acousticians, and engineers, many of whom are involved in the research and development of technology used for audio recording and reproduction (Pritts, 2007).

The perception and measurement of acoustical interactions between loudspeakers and room acoustics is quickly becoming a significant focus within

[2] These figures were supplied by Paul Chavez, Director of Systems Application Engineering Harman Professional Group (P. Chavez, personal communication, August 20, 2007).

the professional audio industry, making the topic of this dissertation both timely and highly relevant to resolving this prominent issue. The increasingly critical nature of this issue comes into sharp focus when one surveys the recent and significant increase in the number of audio products that are being sold with the primary purpose of correcting the effects of loudspeaker-room interactions (Audyssey, 2007; Bang & Olufsen, 2007; DEQX Pty Limited, 2007; Dirac Research, 2007; Harman/Kardon, 2007; JBL, 2007; Lexicon, 2007; Lyngdorf Audio, 2007; Tact Audio, 2007; Trinnov Audio; 2007).

Similarly, the magnitude of this topic is further reinforced by the ever-increasing number of papers, tutorials and workshops presented on this topic at recent Audio Engineering Society and Acoustical Society of America gatherings (Antsalo, Karjalainen, Mäkivirta, & Välimäki, 2003; Benjamin & Gannon, 2000; Carlstrom et al., 2007; Casimiro et al., 2007; Celestinos & Nielsen, 2005; D'Antonio & Cox, 2001; Fares & Pedersen, 2007; Ferekidis & Kempe, 1996; Goldberg & Mäkivirta, 2003a, 2003b, 2004; Johansen & Rubak, 1999, 2000, 2001; Karjalainen, Antsalo, & Mäkivirta, 2003; Mäkivirta & Anet, 2001; Mäkivirta, Antsalo, Karjalainen, & Välimäki, 2001; Nielsen & Celestinos, 2007; Pedersen, 2003; Pedersen & Kasper, 2007; Santillan, 2001; Toole, 2006; Welti, 2002, 2004; Welti & Devantier, 2006). Although these papers have provided useful physical evidence of loudspeaker-room interactions, there are two equally significant additional issues that have rarely been accounted for in their experimental design and/or conclusions: (1) experimental verification of the perception of loudspeaker-room interactions, and (2) the effect that human adaptation to room acoustics has on these interactions. This dissertation takes a more comprehensive approach to the problem by dealing with all three issues, making it a unique and relevant contribution to the scientific, consumer, and professional audio communities.

1.3.2 The Audio Consumer

The research that forms the foundation of this dissertation is designed to improve our understanding of the perception of loudspeaker-room interactions by utilizing a more comprehensive approach than most of the previous studies in this area. Utilizing an approach that accounts for all three relevant considerations (physical evidence of loudspeaker-room interactions, experimental verification of their perception, and human adaptation to room acoustics) is a necessary first step in reaching a valid and scientifically sound approach to improving the optimization and control of the loudspeakers and rooms used for the monitoring and playback of recordings.

Improvements in the quality and consistency of recordings and their reproduction will directly benefit the audio consumer. Audio consumers are big business in the United States; in 2006, U.S. consumers spent $2.8 billion on their home audio systems (Consumer Electronics Association, 2007b, p. 10), and nearly the same amount ($2.7 billion) on after-market audio systems for their cars (Consumer Electronics Association, 2007b, p. 34).

Since the focus of this dissertation is multichannel music reproduction, it is appropriate to examine the potential utility and impact of its conclusions to the audio consumer. The popularity of multichannel audio products is strong and continues to rise; both in our homes and in our cars. As of 2007, almost 30% of all homes in the US have multichannel audio playback systems, with sales expected to increase significantly in 2007 (Consumer Electronics Association, 2007a, p.10). Additionally, the DVD player, the primary vehicle for multichannel audio playback, has enjoyed unprecedented success in consumer adoption rates, with an 84% penetration rate in U.S. homes as of February, 2007 (Consumer Electronics Association, 2007a, p. 1).

Another significant factor behind the growth of multichannel audio sales is their growing popularity in automobile audio. Sales of multichannel audio systems in cars have increased as consumers spend more time in their cars (an average of 1.63 hours per day) and want both themselves and their rear-seat passengers to be entertained while they drive (Consumer Electronics Association, 2006, p. 22). In fact, when U.S. adults using online services were asked where they listen to most of their audio content, 43% responded “at home,” versus 41%, who responded “in the automobile” (Consumer Electronics Association, 2005, p. 9).

It seems clear that, as multichannel content continues to become more readily available for playback over HDTV, the radio, and the Internet, consumer awareness, interest, and demand for multichannel audio systems will grow. The significant and growing market share of multichannel audio products in both our homes and cars, combined with the growing availability of free multichannel content, are strong indications that the focus of this dissertation is highly relevant to today’s audio market. Furthermore, it is likely that this relevance and applicability will only become stronger and more vital with the passage of time.

1.3.3 Scientists

The study of loudspeaker-room interactions is a relatively new and challenging area of scientific research. One objective of this dissertation is to provide a catalyst for change, with the intention of provoking more students and researchers to embark into this new area of research. The focus of this research should almost certainly receive significantly more attention in the near future. Scientists and engineers have historically tended to focus on the biggest problems faced by society. In the field of audio, this emphasis has translated into a focus on the weakest link in the audio chain, which, for the last 25 years, has been the loudspeaker. As a result, high-quality loudspeakers are now readily available at affordable prices.

Since the problem of loudspeaker quality has largely been solved, the focus has now shifted to the weakest existing link in the sound reproduction chain: The interaction between the loudspeaker and the listening room. As previously discussed, scientific interest in loudspeaker-room interactions is growing rapidly, as evidenced by both the increased frequency of scientific papers on loudspeaker-room interactions, and by the recent proliferation of commercial products aimed at controlling their effects.

Until recently, research in this area was challenged by the lack of affordable, and/or practical, means to perform well-controlled, real-time comparisons of different loudspeakers in different listening rooms. The BRS method configured by the author's colleague, Todd Welti, overcomes this challenge by utilizing a BRS measurement system and a BRS playback system that can be efficiently implemented using equipment that is relatively inexpensive and readily available (see chapter 3).

1.4 The Structure of the Dissertation

The dissertation is divided into six chapters, and the main goals and themes of each are summarized below.

1. **Chapter 1: Introduction**

Chapter 1 introduces the research topic and author's motivation for studying the interactions between loudspeakers and room acoustics. The main research objectives are defined, as well as the potential applications and beneficiaries of the research. The overall structure of the dissertation and content of each chapter is described.

2. **Chapter 2: Loudspeakers, Rooms, and Room Adaptation**

Chapter 2 provides the literature review for the dissertation, summarizing previous and current research into the physical and psychological factors related to loudspeaker-room interactions, as well as their subjective effects on the quality of reproduced sound. A survey of the models used to predict the sound quality of loudspeakers in rooms helps to establish what aspects of their objective measurements most influence the preferences and perceptual attributes of reproduced auditory imagery. Chapter 2 also explains how the modal behaviour of the room dominates the quality of sounds below a certain transitional frequency, above which the room reflections produce many different subjective effects, which are also thoroughly reviewed. Finally, research on room acoustic adaptation is examined, including a summary of the perceptual mechanisms involved.

3. **Chapter 3: The Binaural Room Scanning (BRS) Method**

Chapter 3 describes the BRS method measurement and playback method used for the author's experiments, and explains why it is an important methodological tool for research in loudspeaker-room interactions. A physical description of the measurement and playback systems is given, followed by a discussion of the potential sources of error the BRS system can produce. A calibration method is described that allows removal of many of these errors. Finally, the results of a validation test are presented in which listeners' loudspeaker preference ratings, measured in situ, are compared to those measured using the BRS playback system.

4. **Chapter 4: An Experiment on the Influence of Room Acoustic Variations on Loudspeaker Preference**

Chapter 4 gives an account of the main experiment reported in this dissertation, giving a detailed description of the experimental design, methodology, variables, test setup, and a statistical analysis of the results.

5. **Chapter 5: Discussion of the Experimental Results**

Chapter 5 discusses the results of the experiment in the context of previous studies and the expected outcome as stated in the research hypotheses.

6. Chapter 6: Conclusions

Chapter 6 summarizes the main conclusions of the experimental research, identifies some limitations in its scope, and recommends additional areas of research that warrant further investigation.

Chapter 2 Loudspeakers, Rooms, and Room Acoustic Adaptation

Chapter 2 reviews the relevant scientific literature on loudspeakers, rooms and loudspeaker-room interactions. Additionally, the relevant literature is surveyed concerning the subjective effects of all three, as well as the extent to which listeners adapt to them. This background is necessary for three reasons. First, it will firmly establish the motivation for, and the potential benefits of, conducting research in this important field. Second, it provides the experimental and theoretical framework used to formulate the author's research questions and choice of methodologies. Finally, it provides the necessary background to assess the validity and originality of the experimental results.

Chapter 2 is divided into five main sections. Section 2.1 provides an overview of the acoustic events that occur within the sound field of a domestic-sized listening room. Section 2.2 identifies the important physical parameters of a loudspeaker that affect its perceived sound quality, and the underlying auditory perceptual attributes. A summary of the previous work on loudspeaker models that are designed to predict listeners' loudspeaker preferences in listening rooms is given for the purpose of identifying the most salient physical parameters of a loudspeaker that interact with the room acoustics. Section 2.3 surveys the literature on acoustical interactions between the listening room and loudspeaker, their subjective effects, and perceptual mechanisms that help humans deal with sounds in reflective listening spaces. Section 2.4 focusses on previous research related to room acoustic adaptation. Section 2.5 summarizes the main conclusions of this chapter.

2.1 The Sound Field in a Domestic Listening Room

A loudspeaker radiates sound into a listening room in all directions, creating a complex sound field that builds up and then decays within a few hundred ms. The sound field in a room is often deconstructed into three sound components, categorized according to their time of arrival at the listener's ears: (1) the direct sounds, (2) the early reflected sounds, and (3) the late reflected sounds. The decomposition of a sound field into these three components originates from concert hall studies that have established that each component plays a different role in our perception of the timbre, spatial attributes, loudness, and intelligibility of the sound sources (Ando, 1977, 1998; Ando & Schroeder, 1985; Barron, 1971; Barron & Marshall, 1981; Beranek, 1996, 2003; Bradley &

Soulodre, 1995, 1996; Bronkhorst & Houtgast, 1999; Hess & Blauert, 2005; Hess, Braasch, & Blauert, 2003; Hidaka & Beranek, 2000; Hidaka, Beranek, & Okano,1995; Nielsen, 1993; Okano, Beranek, & Hidaka, 1998; Schroeder, Gottlob, & Siebrasse, 1974). The direct sound provides timbre cues and the dominant localization (direction) cues, whereas the early reflections provide additional cues regarding the distance, spatial extent, and intelligibility of the sound source. The early and late reflections provide auditory cues that contain information about the size of the listening space, as well as contributing to the timbre and loudness of the sound sources (Hameed, Pakarinen, Valde, & Puliki, 2004). This is discussed in more detail in section 2.3.

Many of the same perceptual processes used for auditioning sound in concert halls are applicable to analyzing the sound fields produced by loudspeakers in rooms. By deconstructing and analyzing the qualities of the direct, early, and late reflected sounds radiated by the loudspeakers, insight into the timbral and spatial aspects of the audio reproduction system has been gained. Figure 1 is a graphical illustration of the first reflections in Room R4, one of the rooms used in the experimental investigation. Room R4 is described in further detail in chapter 4.

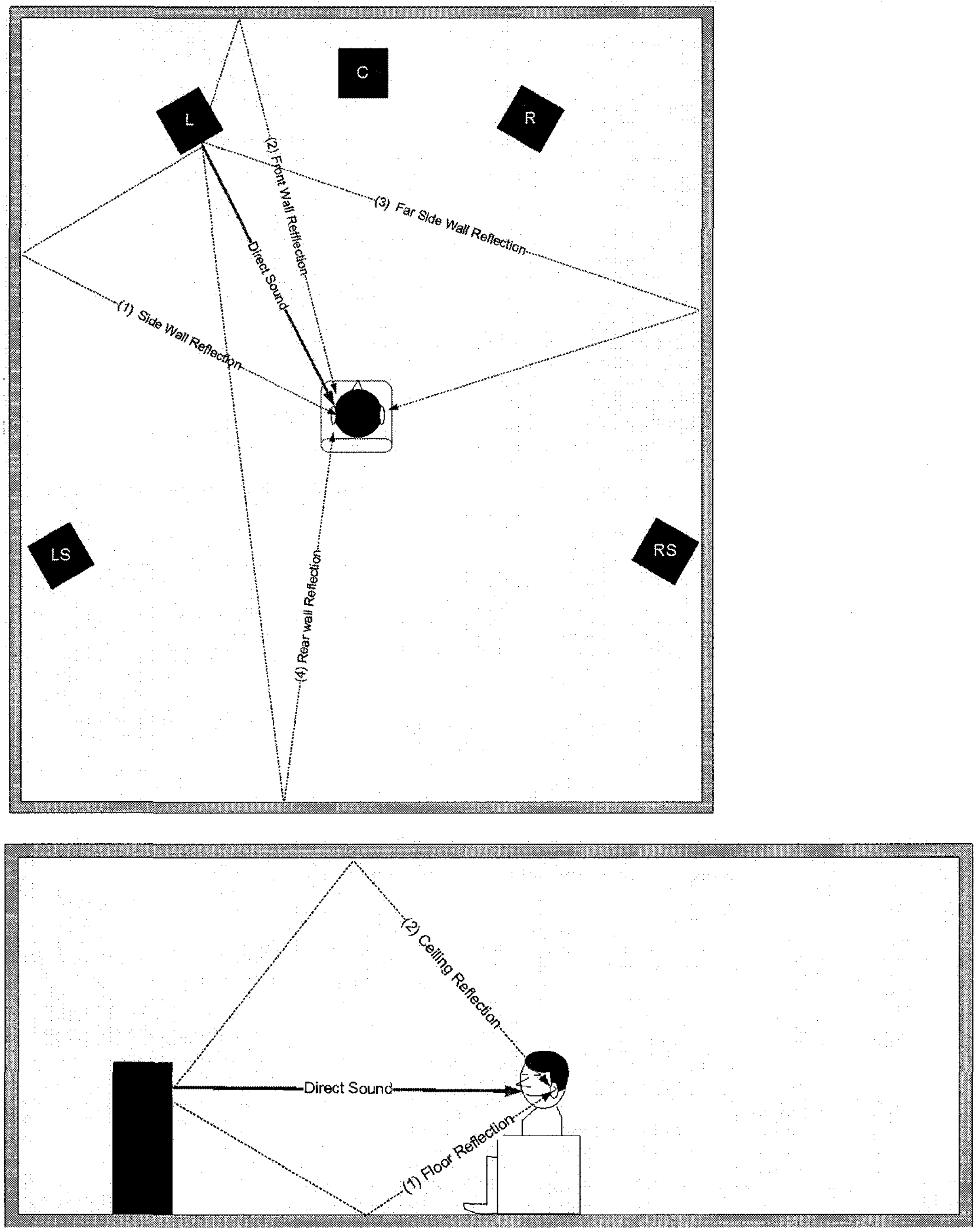


Figure 1. The direct sound and first reflections in a listening room numbered in order of their arrival at the listener's position.

For the sake of clarity, only the direct sounds and first order reflections produced by the left front loudspeaker were depicted in Figure 1. It is important to note that the loudspeakers were symmetrically arranged with respect to the front and side walls of the room. This arrangement ensured that the physical properties of the direct and reflected sounds produced by the left- and right-front channels and the surround channels were the same, provided that the acoustical treatment of the room walls were the same at the reflection points. The center channel had its own unique set of reflection patterns.

As depicted in Figure 1, the first sound to reach the listener was the direct sound from the loudspeaker, which arrived 8.4 ms after the loudspeaker received an input signal.[3] The physical characterization of the direct sound at the listener can be based solely on an anechoic measurement of the loudspeaker taken along its reference listening axis.[4] Additional anechoic measurements of the loudspeaker, taken slightly off axis (around its reference axis, approximately 0° to ±45°), were necessary to predict the direct sound received by listeners sitting slightly off axis to the loudspeaker's reference axis.

The next sounds that arrived at the listener were the first order reflections, which, in Room R4, occurred within 24.4 ms after the arrival of the direct sound. As in most domestic rooms, the first room reflection was the floor bounce, which,

---

[3] This is based on the speed of sound of 344 ms at 21° C sea level.

[4] There will be some attenuation in level at high frequencies due to air absorption.

in this particular room, arrived 2.9 ms after the direct sound. The second reflection came from the ceiling (4.7ms), followed by the front wall (5.7 ms), the left side wall (6.7 ms), the right side wall (16.5 ms), and the rear back wall (24.4 ms).

It is important to note that these reflected sounds originated from the loudspeaker at further off-axis angles than the direct sound. Therefore, to physically characterize and predict the spectral quality of these reflections, comprehensive, fine-resolution anechoic measurements of the loudspeaker had to be taken at significantly off-axis angles in both the horizontal and vertical radiating orbits.

The frequency-dependent absorption coefficients of the surfaces where these reflections occurred must also be considered. Carpets, drapes, and other various absorptive and scattering objects in the room may alter the spectral envelopes, phase, and levels of the reflected sound, which, to some extent, can be predicted if their absorption coefficients are known. A final observation is that the angles at which the different reflections arrive at the listener have important consequences for the perceived timbre, spatial properties, and intelligibility of the reproduced sounds. This is discussed in more detail in a later section.

The second and higher order reflections are not depicted in Figure 1, but they generally arrive after the first reflections. As the radiated sounds from the loudspeaker travel further distances throughout the room and interact with the room's absorptive and scattering objects, they become attenuated in level, particularly at the mid-to-high frequencies, where most sound absorption in rooms typically occurs. Eventually, these late arrival sounds decay some 300-600 ms after the direct sound (Bradley, 1986). Predicting the physical characteristics of these late arrival reflections would require an accurate acoustical model of the listening room as well as fine spatial-frequency resolution of the loudspeaker's frequency response, measured around its horizontal and vertical radiating orbits.

2.2 The Loudspeaker

Section 2.2 reviews previous and current research efforts aimed at modelling and predicting the perceived sound quality of loudspeakers in reflective rooms based on different types of acoustical measurements. An examination of this work will provide insight into which physical parameters of the loudspeaker are most related to its perceived sound quality in rooms, and explain how different variations in room acoustics might alter the perception of those salient loudspeaker parameters.

2.2.1 Different Loudspeaker Measurement Philosophies

Prior to the mid 1980's, little consensus existed among audio researchers about where or how to measure a loudspeaker in order to best characterize its perceived sound quality in a listening room. Toole (1986a) supplied a good historical perspective on the three loudspeaker measurement philosophies that were prevalent during the mid 1980's, with more recent developments documented in a later paper by the author (Olive, 2004a).

During the 1980's, the audio industry produced three divergent philosophies regarding which type of measurements best characterized a loudspeaker's sound quality. The first measurement camp advocated that the total radiated sound power of a loudspeaker best represented its perceived sound quality (Allison, 1974; Bose, 1969; Consumer's Union, 1973; Hirsch, 1982; McShane, 1969; Rosenberg, 1973; Torick, 1977). A second school of thought claimed that the loudspeaker's in-room, steady-state frequency response, measured at the listening location, best represented what listeners perceived in the room (Staffeldt, 1974, 1984). The third loudspeaker measurement contingent, comprised mostly of British loudspeaker researchers (British Broadcasting Corporation, 1979; Colloms, 1985; Harwood, 1976;), argued that the direct sound from the loudspeaker largely characterized what listeners perceived, and that, therefore, loudspeaker measurements taken in an anechoic chamber were the most appropriate.

Each of these three loudspeaker measurement philosophies represented disparate views on: (1) which loudspeaker measurement parameters matter the most to its perceived sound quality; (2) how the listening room interacted with these parameters; and, indirectly, (3) the perceptual importance of the direct, early, and late reflected sounds radiated by the loudspeaker in a room. The sound-power camp largely disregarded the perceptual importance of the direct sound, while placing great emphasis on the role of the reflected sounds in the listening room. As a result, their measurement philosophy created a significant bias towards highly reverberant listening rooms. The direct-sound measurement advocates argued that direct sound dominated our perception of the loudspeaker, a view that resulted in a bias towards anechoic-like listening rooms.

The in-room measurement camp fell somewhere between the other two groups because their resulting measurements included the effects of both the direct and reflected sounds, as well as the acoustical interactions between the loudspeaker and the room. Out of the three measurement approaches, the in-room measurement philosophy has proven to be the least biased towards a particular type of room acoustics because their measurements were taken in the actual listening room.

A common element missing from the loudspeaker measurement debates of the early 1980's was accurate and reliable subjective measurements of loudspeakers. Without proper listening test data on the loudspeakers, no one had the scientific foundation to argue which loudspeaker measurement approach best represented the listener's perception of a loudspeaker and its sound quality.

Since the mid-1980's, accurate and reliable loudspeaker listening test methods have been developed, allowing scientists to correlate subjective and objective loudspeaker measurements (Gabrielsson & Lindström, 1985; Klippel, 1990; Olive, 2004a, 2004b; Toole, 1986a, 1986b). The development of new models aimed at predicting listeners' loudspeaker preferences and their underlying perceptual attributes based on a set of loudspeaker measurements have emerged from these studies. Work in this area is summarized in the next section.

2.2.2 Predictive Models of Loudspeaker Sound Quality

Properly controlled listening tests on loudspeakers are both expensive and time-consuming. As a result, several loudspeaker researchers have developed mathematical models aimed at predicting listeners' loudspeaker sound quality ratings based on a set of objective measurements. These mathematical models are summarized by the author in Olive (2004a, pp. 2-4). The four loudspeaker models discussed below are fundamentally different in five ways:

1. The number of loudspeaker measurements used to represent the quality of the direct, early, and late reflected sounds
2. Where the loudspeaker measurements are taken (e.g. in a room, in an anechoic chamber, or in a reverberation chamber)
3. The frequency resolution of the measurements used in the model (e.g. 1/3-octave or 1/20th octave)
4. Whether or not the loudspeaker model applies a psychoacoustic loudness model to the measurements
5. Most importantly; whether the models have been adequately verified by experimental comparisons of the predicted versus observed sound quality ratings using properly controlled listening tests

2.2.2.1 Loudspeaker models based on sound-power measurements

One of the earliest loudspeaker models was developed by the Consumer's Union (1973) in the early 1970's, and is still used today to predict the sound quality of loudspeakers that are reviewed in *Consumer Reports* magazine. The magazine reports loudspeaker accuracy scores on a 100-point scale, which they calculate from the loudspeaker's sound-power response measured in 1/3-octave bands at fixed center frequencies. The underlying assumption is that an accurate loudspeaker should have flat sound power, and that any deviations from this represent a less than accurate loudspeaker.

Consumer's Union (1973) made several transformations to the raw sound-power response in an effort to account for the low-frequency changes that occurred because of room boundary effects and wall absorption. The raw sound-power response was then adjusted in 1/3-octave bands according to loudness, using Steven's Mark VII scheme (Stevens, 1972). As the loudspeaker deviated from equal loudness over a certain bandwidth, the error was subtracted from its overall 100-point score (Consumer's Union). Although this model has been used for over 30 years to review consumer loudspeakers, Consumer's Union has never published a study that either tested or validated the model's accuracy.

Recently, the author conducted an experiment designed to test the Consumer's Union sound power-based model by selecting 13 loudspeakers tested by the Consumer's Union, and comparing their predicted accuracy scores to their observed preference ratings based on properly controlled, double-blind listening tests (Olive, 2004a, 2004b). The correlation between the observed preference ratings and their predicted accuracy ratings was very low and negative; the *Pearson product-moment correlation* (r) was -.22, and, in this case, the *probability* (p) of falsely concluding the existence of the relation between observed and predicted preference was .44 (i.e., the probability of incorrectly rejecting the null hypothesis here was much greater than is typically allowed in such tests).

In order to determine whether sound power-based loudspeaker models are inherently flawed, the author developed and tested a new sound-power model based on anechoic loudspeaker measurements having fine frequency (2 Hz) and spatial (10°) resolution. The author's new sound-power model differed from the Consumer's Union model in two important ways. First, it used finer frequency resolution measurements than the Consumer's Union model (1/20-octave versus 1/3-octave). Second, the new model considered several different target functions other than the "flat" sound-power target preferred by the Consumer's Union model (Consumer's Union, 1973). The new model considered the slope, the smoothness, and the average narrowband deviations within the sound-power curve as possible criteria for predicting the perceived sound quality of the loudspeaker.

The new sound-power model yielded much more successful predictions of listeners' loudspeaker preference ratings than the Consumer's Union model (r = .87, Olive, versus r = -22, Consumer's Union). When applied to a group of 70 different loudspeakers, the predictive accuracy of the author's model was somewhat lower (r = 0.79).

The contrast in performance between these two sound-power models has provided scientific evidence that demonstrates: (1) that the flat sound-power criterion was not a good indicator of loudspeaker sound quality, and (2) that the commonly used 1/3-octave measurement has insufficient resolution for accurately predicting sound quality. The first point can be explained by the fact that most consumer loudspeakers become increasingly directional at higher frequencies. In order to achieve a flat sound-power, the on-axis, or direct, sound of the loudspeakers must be artificially boosted, making them sound too bright. Experimental evidence has demonstrated that compromising the direct sound to achieve flat sound power results in lower sound quality or preference ratings (Olive, 2004b).

Analysis of these results lends scientific support to the notion that the quality of the direct sound is an important indicator of a loudspeaker's perceived sound quality. The need for higher frequency resolution measurements than 1/3-octave indicates that listeners can perceive audible differences in the loudspeaker's frequency response that are not being accurately represented by a 1/3-octave measurement; the medium and high *Q values (or Q factor) of the resonances* (*Q*)[5] will be particularly misrepresented. Finally, there is some doubt as to whether or not the loudness model used in the Consumer's Union model is

[5] The *Q* value of a resonance (also referred to as both the "*Q*" and the "Q-factor") is defined as the frequency of the resonance divided by its bandwidth.

sufficiently accurate, or even appropriately applied, to predict the perceived sound quality of a loudspeaker.

2.2.2.2 A new loudspeaker model based on comprehensive anechoic measurement data

As discussed in the previous section, the best performance that the author's new sound-power model achieved was a predicted versus observed accuracy of $r = 0.87$ when applied to 13 loudspeakers reviewed by Consumer's Union. Although the performance of the new model was much better than the Consumer's Union model ($r = -0.22$), it was hypothesized that a sound-power model was unable to achieve better performance due to an inherent flaw; a single curve cannot adequately represent the perceived quality of the direct, early, and late reflected sounds in a listening room.

To explore this hypothesis further, the author developed a predictive model based on a set of anechoic measurements that included separate spatial averages used to estimate the direct sound, the early reflected sound, the sound-power response, and a predicted in-room response based on a weighted average of the three. These measurements were based on the original work of Toole (1986a, 1986b). Their application was extended by Devantier (2002), who conducted a survey of stereo and multichannel setups in 15 domestic rooms to determine the average angles at which the direct and early reflected sounds were being radiated from the loudspeakers.

The survey data collected during that study (Devantier, 2002) were used to calculate the frequency response measurements shown in Figure 2. The frequency response curves depicted are, from top to bottom, the on-axis sound, the listening window, the early reflections, the predicted in-room curve, the total radiated sound power, and the directivity indices based on the sound-power and the early reflection curves. The on-axis curves and listening window curves are representative of the direct sound for listeners, sitting on axis or slightly off axis.

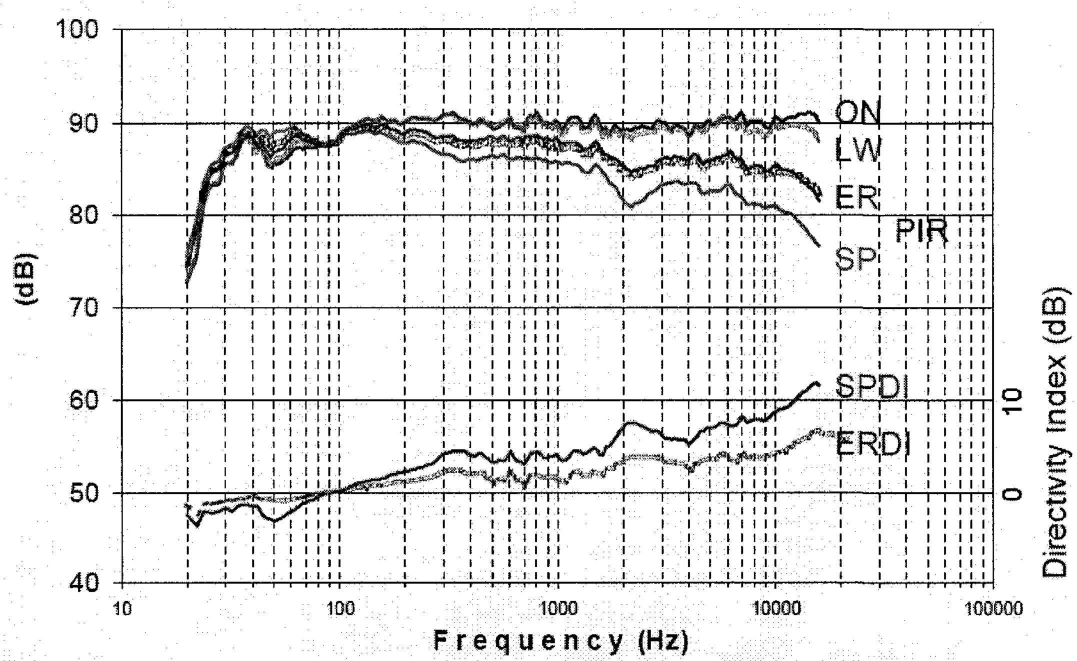


Figure 2. A standard set of anechoic loudspeaker measurements.

The early reflection curve represents the first few reflections a listener would hear in a typical listening room, and the predicted in-room and sound-power curves would best represent the late reflected sounds produced by the loudspeaker.

The author developed several statistical metrics based on amplitude deviations within a set of spatially-averaged anechoic loudspeaker measurements that characterized the predicted qualities of the direct, early, and late reflected sounds in a listening room. Multiple regression and principal component analysis were used to identify which set of weighted metrics and measurements most accurately predicted listeners' loudspeaker preference ratings based on controlled double-blind listening tests.

Based on that analysis, a model was designed that could accurately predict (r = 0.995) preference ratings for the 13 loudspeakers tested by Consumer's Union (Olive, 2004b). When applied to a group of 70 loudspeakers, a slightly modified version of the same model produced somewhat less accurate predictions (r = 0.86) of the observed preference ratings. Both models included: (1) the quality measurements of the direct sound, based on using the on-axis curve; and (2) the early and late reflected sounds, based on the predicted in-room and sound-power curves. The direct and reflected sounds were generally weighted equally (35% each), with the remaining 30% of the model based on measurements of the loudspeaker's low-frequency bass quality and extension.

The fact that bass quality was an important determining factor in listeners' loudspeaker preferences underlines the importance of controlling loudspeaker-room interactions below the transition frequency of the room, where room modes and room boundary effects can wreak havoc on the accurate reproduction of low-frequency sounds.

2.2.2.3 Loudspeaker models based on in-room measurements

Up to this point, this section has focussed on predictive loudspeaker models based on sound-power and anechoic measurements. However, two other types of loudspeaker models have been developed that warrant attention. The first was based on in-room loudspeaker measurements and was proposed by Staffeldt (1974, 1984). The second was proposed by Klippel (1990) and used a combination of in-room and anechoic measurements. Staffeldt (1984) claimed that the timbre of two loudspeakers in two different rooms would be identical, provided that they had identical 1/3-octave spectra measured at the entrance to the ear canal located at the in-room listener location. However, his listening tests were based on only one listener, and the room was relatively large and reverberant. He later proposed a model for predicting the loudspeaker's timbre based on calculating the specific loudness of the 1/3-octave data (Staffeldt, 1984). Again, no data were published to provide the means to test or verify the accuracy of the model.

Klippel (1990) published a perceptually based model for predicting various loudspeaker perceptual dimensions, as well as the overall pleasantness and naturalness of the sounds produced through the loudspeaker. A total of 45 loudspeakers (both real and simulated), three different listening rooms, 13 programs, and 45 listeners were used. The input to the model used a combination of the loudspeaker's direct and reflected sounds, and utilized 1/3-octave in-room measurements or anechoic measurements that were weighted to predict the in-room response. The model examined errors between the measured loudspeaker response and an ideal flat reference; superimposed on this was the spectrum of the music program to help give a better impression of what the listeners heard (Klippel).

Using a modified loudness model (Paulus & Zwicker, 1972), Klippel (1990) calculated the differences in loudness density between the reference and measured curves across each 1/3-octave center frequency using a critical bandwidth filter. The loudness difference was further transformed and weighted for each objective metric that Klippel had developed to predict the various underlying perceptual dimensions that he had identified using multidimensional and factorial analysis.

The correlations between the objective and perceptual metrics were quite high overall (Klippel, 1990). The perceptual attributes most related to overall sound quality were pleasantness and naturalness. Pleasantness could be best predicted with objective measurements of the loudspeaker's colouration and brightness; whereas the prediction of naturalness included a third objective measurement of the loudspeaker's feeling of space (related to its directivity). Klippel noted that the correlations between predicted and observed values of pleasantness and naturalness varied depending on which listening room the loudspeakers were tested in. For pleasantness, the correlations varied from an *r* value of -0.32 to 0.94, and for naturalness, from 0.52 to 0.93. The variation in the loudspeaker model's accuracy due to the independent variable room suggested that this model does not adequately account for the perceived change in loudspeaker sound quality that occurs in different listening rooms (Klippel, 1990).

In their study, Gabrielsson, Lindström and Till (1991) examined the psychophysical relationship between a loudspeaker's frequency response and its perceived sound quality. The frequency responses of 18 different loudspeakers were measured in three different rooms (the listening room, a reverberation chamber, and an anechoic chamber), and were then compared to determine which loudspeaker measurement best correlated with their ratings given by listeners on seven different perceptual scales, and a fidelity scale. The authors concluded that the loudspeaker frequency response measured in the listening room corresponded best with its perceived sound quality (Gabrielsson et al., 1991).

However, unlike Klippel (1990), they failed to: (1) provide statistical evidence to support their claim, or (2) present a loudspeaker model based on the in-room measurements that would predict the sound quality ratings (Gabrielsson et al., 1991). In all likelihood, such a model would not have accurately predicted the loudspeaker sound quality ratings based on the coarse frequency and spatial resolution of their frequency response measurements. The measurements in the listening room were only 1-octave resolution, and the sound-power measurements had only 30 Hz resolution. The anechoic measurements were taken at two angles (0° and 30° in the horizontal plane), meaning that only the direct sound was represented in their analysis (Gabrielsson et al., 1991). The importance of having adequate frequency and spatial resolution in the frequency response measurements used in loudspeaker models will become clear in the following section.

2.2.2.4 Comparison of loudspeaker models based on in-room, sound-power, and anechoic measurements

At first glance, models based on in-room loudspeaker measurements would seem to have an advantage over models based on anechoic measurements. In-room measurements allow the capture of the loudspeaker-room interactions that dominate the loudspeaker's low-frequency behaviour below 300 Hz and they include the room's absorption and interference effects in the captured reflected sounds.

However, as argued by Toole (1986a, 1986b), the in-room measurements of the loudspeaker did not allow easy separation, assessment, and weighting of the direct and reflected sounds, which listeners seemed to naturally perform in their perception of them. The general visual untidiness of in-room frequency response curves seemed to belie a listeners' ability to recognize and compensate for these loudspeaker/room variations, as discussed later in sections 2.3 and 2.4.

To study this question further, the author compared the predictive accuracy of the different loudspeaker models based on comprehensive anechoic measurements, sound-power measurements and in-room measurements (Olive, 2004b). In-room measurements with high-frequency resolution (2 Hz) were taken for 13 different loudspeakers. The measurements were then smoothed into the popular 1/3-octave resolutions used in the loudspeaker models developed by Consumer's Union (1973), Klippel (1990), and Staffeldt (1974, 1984).

The author developed loudspeaker models based on the 1/3-octave and 1/20-octave smoothed measurements that were then compared to the two different sound-power models (Consumer's Union, 1973; Olive, 2004b) described in section 2.2.2.1, and to the model based on comprehensive anechoic measurements (Olive, 2004b) described in section 2.2.2.2. The correlation values from the best to worst models were: the anechoic model ($r = 0.995$), the in-room model with 1/20-octave frequency resolution ($r = 0.91$), the sound-power model, ($r = 0.87$), and, finally, the in-room model with 1/3-octave frequency resolution ($r = 0.75$). These results seem to confirm the results reported by Toole (1986b) below, after he did a similar comparison of in-room, sound-power, and anechoic loudspeaker measurements:

> From the sound-power data alone it was possible to recognize the worst loudspeakers, but there were ambiguities in ranking the better performers. This is essentially the nature of caution issue by Brocine and von Recklinghausen, mentioned earlier. In summary, therefore, it would seem to be imprudent to consider either sound-power or steady-state listening-room measurements as the definitive measure of loudspeaker amplitude response. (p. 346)

2.2.2.5 Summary of the research on loudspeaker models

The most important findings of the research on loudspeaker models aimed at predicting a loudspeaker's sound quality can be summarized as follows:

1. Models based on comprehensive high-resolution anechoic measurements offer the most accurate predictions of listeners' preference ratings. The most likely reason is that this model allows separate quality measurements and weighting of the direct, early, and late reflected sounds. This feature may be important because it better approximates how humans perceive sounds in reflective listening rooms.
2. Models based on in-room loudspeaker measurements give slightly less accurate predictions of listeners' loudspeaker preferences because they do not allow independent quality measurements or weighting of the loudspeaker's direct and reflected sounds. Models based on sound-power measurements give even less accurate predictions of sound quality. They do not allow separate measurements of the direct and reflected sounds, nor do they include any of the loudspeaker-room interaction effects found in the models based on in-room measurements.
3. Accurate models require higher frequency resolution measurements (1/20-octave) than the commonly used 1/3-octave.

4. Listeners prefer loudspeakers that possess the flattest measured on-axis frequency response that also remains similarly smooth off axis. The quality and extension of the bass below 300 Hz is an important predictor of loudspeaker preference, accounting for about 30% of the overall loudspeaker preference rating.

A similar relationship between the smoothness of the measured frequency response of a loudspeaker and its preference rating has been reported with headphones. Moore and Tan (2004) found that the perceived naturalness of headphone reproduction increased as the spectral irregularities in its frequency response were reduced.

None of the loudspeaker models previously described included the nonlinear distortion performance of the loudspeaker in their analysis. In both of the loudspeaker studies reported by Olive (2004a) and Toole (1986b), nonlinear distortion measurements on the loudspeakers generally showed no correlation with listeners' loudspeaker preference ratings. It is well-known that conventional total harmonic distortion measurements have proven to be poor indicators of the detection and perception of loudspeaker distortion (Moore, Tan, Zacharov, & Mattila, 2004; Tan, Moore, & Zacharov, 2003; Voishvillo, 2006) because they do not take into account human perceptual masking. With the exception of smaller, low-powered loudspeakers, most loudspeakers have produced satisfactory sound pressure levels in domestic rooms where the nonlinear distortions were below the threshold of audibility.

Substantial further work is needed in the area of developing and testing models to predict loudspeaker sound quality. The existing models need to be tested and validated in a wider variety of real-world listening rooms, which, in itself, is no small undertaking. A more efficient and reliable subjective evaluation method is needed for comparing multiple loudspeakers in different listening rooms, while controlling their positional effects. The BRS method described in chapter 3 is well suited for this task.

To improve the accuracy of the loudspeaker models and their potential generalizability, they should be extended to include the important perceptual effects related to program material, loudspeaker-room interactions, and possible effects related to room acoustic adaptation. Here again, the BRS method is a valuable tool for conducting this research. There is very little existing experimental work related to the perceptual effects of either loudspeaker-room interactions or room acoustic adaptation. The lack of knowledge in this area is a primary motivation for this dissertation.

2.2.3 Loudspeaker Directivity

A loudspeaker radiates sounds into the room in all directions, with typical loudspeaker designs becoming increasingly directional at higher frequencies (Borwick, 2001). As the directivity of the loudspeaker increases, a greater proportion of its total energy is focussed towards the listening area, with less energy radiated out into the room that would have otherwise been absorbed or re-directed back to the listener as reflected sound. As a result, there is a greater ratio of direct-to-reflected sound received at the listener's ears.

The directivity index of a loudspeaker is the difference in dB between its on-axis frequency response and its total radiated sound power. The directivity index rises monotonically with increasing frequency in most common forward-facing, direct-radiator type loudspeakers (Figure 2). At higher frequencies (above 2 kHz), an increasing proportion of the loudspeaker's total radiated sound is contained within the direct sound and the first reflections.

Considered by itself, the directivity index is not a very useful indicator of loudspeaker sound quality because it indicates nothing about the quality of the direct, reflected, and total radiated sound power. The directivity index only reveals the extent to which the sound-power and the direct sound are similar; both could be similarly good (smooth), and similarly poor (highly irregular), which, in both cases, would be represented by a flat directivity index.

This fact may explain why the loudspeaker's directivity index has never been a particularly good indicator of overall loudspeaker sound quality as found in both the author's study (Olive, 2004b) and in an earlier investigation by Toole (1986b). Toole succinctly makes the following point about the directivity index: "As a description of a specific aspect of loudspeaker performance, the directivity index has some limited use in, for example, public address applications. As a figure of merit it would appear to have little value" (Toole, 1986b, p. 346).

Experimental studies on the subjective effects of loudspeaker directivity have proven difficult; manipulation of the loudspeaker's directivity was difficult to achieve while maintaining a constant frequency response, both on- and off-axis. The few controlled studies on loudspeaker directivity that were done in stereo have indicated that loudspeaker directivity affects the spatial impression of the audio reproduction (Klippel, 1990).

Wider dispersion (i.e. less directional) loudspeakers have been shown to produce stronger lateral reflections that are associated with an increased sense of *listener envelopment* (LEV) (Bradley & Soulodre, 1996; Morimoto & Maekawa, 1989) and a greater *apparent source width* (ASW) (Okano et al., 1998). Both spatial attributes are related to a lower *interaural cross-correlation coefficient* (IACC) value measured at the listener's ears, a feature associated with the preferred sound quality present in concert halls.

Higher values of ASW and LEV have been observed (with an increase in the levels of the early [<80 ms] and late [>80 ms] lateral reflections) in concert halls (Bradley & Soulodre, 1995). More recent surround-sound based studies conducted in rooms have indicated that LEV is related to the overall playback level, as well as to the level and angular distribution of late lateral reflections (Soulodre, Lavoie, & Norcross, 2003). It has been shown that more accurate predictions of LEV must include the frequency-dependent temporal integration rates of the human auditory system as observed in forward masking experiments (Jesteadt, Bacon, & Lehman, 1982; Moore & Glasberg, 1983). The preferred directivity of loudspeakers has also been found to depend on the program signals. Klippel (1990) reported that more directional speakers were preferred for speech compared to music.

The advent of multichannel audio has made loudspeaker directivity a moot point in consumer audio. While stereo playback required wider dispersion loudspeakers to achieve a sense of LEV,[6] this can be easily achieved in multichannel audio by using the side- and rear-channel loudspeakers. The listener's sense of both the LEV and ASW can be manipulated by the recording artist during the recording process itself, or during playback by the consumer through adjustments made in their surround sound processor.

---

[6] The wider dispersion stereo loudspeakers produce a greater sense of LEV by producing stronger lateral reflections.

There are both economical and performance reasons that explain why more esoteric loudspeaker directivities are not commercially successful. More exotic loudspeaker directivities may require additional drivers, electrical parts, acoustic lenses, and more expensive enclosures that significantly increase the cost of the loudspeaker with questionable performance gains. For these reasons, the vast majority of professional and consumer loudspeakers sold today are forward-facing, direct radiator types, with only slight variations in their directivity; defined by the size, number, arrangement, and electrical summation of their drivers. These types of loudspeakers are popular because they offer the best overall technical performance at a reasonable cost. Even the professional recording industry seems to have recognized that forward-facing loudspeakers are the norm, and recommend their use in the production of surround-sound material (ITU-R BS.775-1, 1994; Producers & Engineers Wing Surround Sound Recommendation Committee, 2004).

2.2.4 The Perceptual Attributes of Loudspeakers

Several studies have focussed on identifying the underlying perceptual attributes related to listeners' loudspeaker preferences (Gabrielsson, 1979; Gabrielsson & Lindström, 1985; Gabrielsson et al., 1991; Gabrielsson & Sjögren, 1979; Klippel, 1990; Lorho, 2007). Gabrielsson and Lindström (1985) used factorial analysis to identify eight perceptual attributes that described the perceived quality of loudspeakers: (1) clarity (clearness/distinctness), (2) sharpness/hardness versus softness, (3) brightness versus dullness, (4) fullness versus thinness, (5) spaciousness (feeling of space), (6) nearness, (7) absence of extraneous noises (nonlinear distortions), and (8) loudness. In a later study, Gabrielsson and his colleagues (Gabrielsson et al., 1991) asked experienced listeners to rate 18 different loudspeakers on seven of these perceptual scales (loudness was excluded) and to give an overall fidelity rating. In correlating the ratings of the perceptual scales to different objective loudspeaker measurements, they concluded that: “All perceptual dimensions seemed to be influenced by the frequency response, even spaciousness although the stereo information in the signal was disregarded” (Gabrielsson et al., 1991, p. 718).

Klippel (1990) utilized multidimensional and factorial analysis to identify the perceptual dimensions and their attributes using 45 loudspeakers tested in three rooms. He found seven perceptual dimensions related to listeners’ impressions of loudspeaker sound quality that were correlated with loudspeaker frequency response: (1) clearness (clarity), (2) treble stressing (sharpness), (3) general bass emphasis (loudness), (4) low bass emphasis, (5) feeling of space, (6) clearness in bass, and (7) brightness.

Using a quantitative descriptive analysis method, Lorho (2007) elicited a large number of descriptors for the sound quality attributes of multimedia speakers. However, after principal component analysis was applied, the descriptors were reduced to a set similar to that found by Gabrielsson. The eight loudspeaker sound quality attributes (outlined above) developed by Gabrielsson and Lindström (1985) were used by Toole (1986a, 1986b), who included five additional sound quality methods attributes that were related to spatial qualities, based on less scientific methods. The additional spatial sound quality attributes contributed by Toole were: (1) definition of sound images, (2) continuity of sound stage, (3) abnormal spatial effects, (4) spatial perspective, and (5) width of sound stage and impression of distance/depth.

In a study by the author (Olive, 2004a), the 2,381 comments given by listeners for 13 different loudspeakers were categorized into 80 "unique" adjectives, and 27 of these adjectives account for 80% of all comments. Timbre-related adjectives accounted for 94% of all descriptors, whereas spatial-related and distortion descriptors accounted for only about 3% of all descriptors. Principal component analysis was performed on a subset of 40 comments, which revealed that nine underlying factors explained 95% of the comment variance.

The first factor (31% of the variance) was associated with bass quality, overall spectral balance, and smoothness. The second dimension (19% of the variance) was related to specific timbre colourations in the bass, treble, and midrange. When the loudspeaker preferences were mapped into this two-dimensional space, highly preferred loudspeakers shared common features, such as good bass (not too full or thin) and neutral (low-colouration) timbre, whereas the less preferred speakers had deficiencies in bass (thin) and specific colourations in the bass, treble, and midrange.

There has been very little research into the development of objective metrics that predict the intensities of the loudspeaker perceptual attributes. Gabrielsson tried to predict the intensities of these attributes using in-room measurements, but was not very successful (Gabrielsson, Hageman, Bech-Kristensen, & Lundberg, 1990). Klippel (1990) was able to develop metrics based on the loudspeaker's direct sound and in-room, steady-state response that predicted the intensities of his seven perceptual loudspeaker attributes with reasonable accuracy.

In summary, several investigations into the perceptual attributes of loudspeakers using different methods and loudspeakers have largely reached the same conclusion; that nine different perceptual attributes principally account for the perceived differences among loudspeakers. One attribute is related to loudness, five are related to timbre, and two are related to spatial attributes. The remaining attribute is related to nonlinear distortion.

2.3 The Listening Room

Toole (2000, 2006) gives a good summary of the psychoacoustic problems found in domestic listening rooms intended for stereo and multichannel audio reproduction. Typical domestic listening rooms are relatively small acoustic spaces with absorptive, reflective, and scattering objects distributed unequally throughout the room in close proximity to the loudspeakers and listeners. Therefore, a highly diffusive sound field (such as those found in concert halls) does not exist in domestic rooms, and concepts, such as reverberation time, are not valid when applied to domestic rooms. Rather, it is clear that when significant variations in the spectral, temporal, and directional properties of the sound field exist throughout the domestic listening room, the positioning of both the listeners and the loudspeakers significantly influences these variations.

This dynamic behaviour of sound in domestic rooms lends itself to psychoacoustic analysis of the frequency, time, and spatial domains. There is a significant transitional division between these domains at around 300 Hz. Below 300 Hz, the in-room, steady-state frequency response best represents effects related to room mode and solid angle/boundary effects related to the positioning of the loudspeakers (Toole, 2006).

Mäkivirta and Anet (2001) documented these effects in a survey that measured 372 factory-calibrated loudspeakers installed in 164 professional audio control rooms. On average, they found 12 dB variations (after 1/3-octave smoothing) at 200 Hz, increasing monotonically to 20 dB at 30 Hz. Similarly, Welti (2002) reported up to 40 dB in-room variations below 300 Hz. Even ITU compliant multichannel loudspeaker setups (ITU-R BS.775-1, 1994) have produced large in-room amplitude variations below the transition frequency among the different front and surround loudspeaker positions (Zacharov & Bech, 2000). Although these positional-dependent variations in amplitude response have been minimized with the use of a single bass-managed subwoofer, there have still been 24 dB seat-to-seat variations below 80 Hz (Toole, 2006, p. 471, Figure 19).

In a typically sized domestic room, above about 300 Hz, the density of the room modes becomes sufficiently high such that a steady-state analysis of the room's acoustical behaviour becomes less meaningful. Above this transition frequency,[7] analysis of room reflections (their level, spectrum, direction, and temporal distribution) is useful for assessing their potential subjective effects on the quality of reproduced sounds. Here, a thorough understanding of the perception of reflections and their subjective effects is necessary before

[7] The transition frequency is often referred to as the Schroeder Frequency, defined as the frequency at which there are at least three room resonances that overlap within their half power points or bandwidth (Borwick, 2001).

establishing whether the design of recordings, loudspeakers, and listening rooms should compensate for them.

The following two sections summarize the distinctive acoustical factors that occur below and above the transitional frequency, and their corresponding subjective effects, which influence the sound quality of audio reproduction in a listening room.

2.3.1 Room Modal Effects Below the Transition Frequency

Below the room's transition frequency, the standing waves in a room dominate what can be heard. Every listening room has a set of natural resonant modes that can be calculated from its dimensions. For a simple rectangular room, the standing waves can be calculated using the formula below (Borwick, 2001).

$$f_{(n_l, n_w, n_h)} = \frac{c}{2}\sqrt{\left(\frac{n_l}{l}\right)^2 + \left(\frac{n_w}{w}\right)^2 + \left(\frac{n_h}{h}\right)^2}$$

Where:

c = speed of sound in air, typically (344 m/s)

n_i = integer values 0, 1, 2, 3…

l = length of room measured in meters

n_w = integer values 0,1,2,3...

w = width of room measured in meters

n_h = integer values 0,1,2,3...

h = height of room measured in meters

The axial modes occur along a single dimension of the room (its width, length, and height), the tangential modes occur along two dimensions, and the oblique modes travel across all three room dimensions. As the number of room dimensions increases, the standing wave becomes less energetic. For this reason, the axial modes are generally the most dominant and problematic in achieving consistent bass across a wide seating area. Tangential modes only become an issue if the walls are particularly stiff (mechanically rigid).

The axial modes can be viewed as a series of waves travelling along the length, width, and height of the room, with their pressure maxima occurring at the boundaries of the room. For the odd order modes (1^{st} , 3^{rd}, 5^{th}...), the pressure maxima at each boundary are of opposite polarity to each other at any instant in time, and, at the center of the room, there are pressure minima, or nulls. For the even order modes (2^{nd}, 4^{th}, 6^{th} ...), the pressure maxima at opposite room boundaries have the same polarity, and, in contrast to the odd order modes, there are pressure maxima at the center of the room.

With knowledge of where the low-frequency pressure maxima and minima occur in the room, the locations of the subwoofer and listener can be optimally selected to produce the smoothest low-frequency response at the listening location. For the listener, the optimal location is where the modes are not at either their maximum or minimum pressure.

Figure 3 shows an optimal subwoofer and listener location for one dimension of the room. The listener is located approximately two-thirds down the length of the room, near the location where the peaks and nulls of the first, second, and fourth order modes occur. By placing the subwoofer in the null, or zero pressure point, of the third order mode, this mode is effectively not being excited by the loudspeaker. Therefore, the listener will not hear the peak of this mode. Equalizing the subwoofer at the listener location should yield further improvements to the smoothness of the low-frequency response.

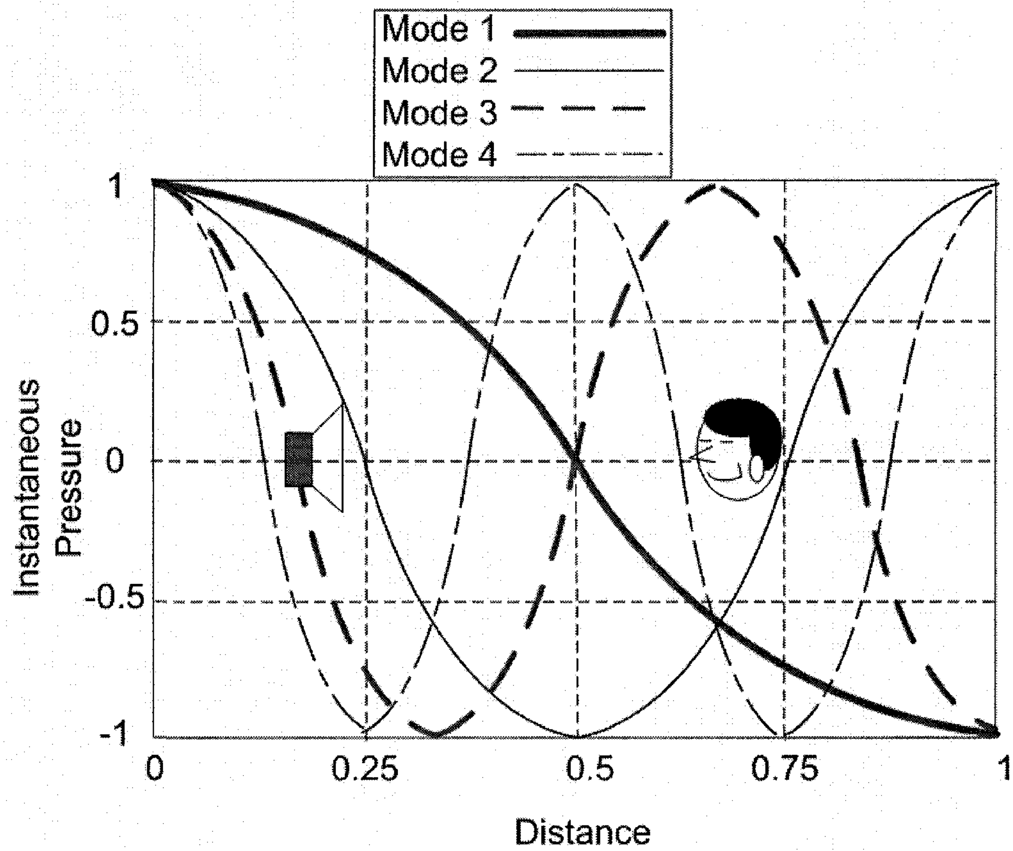


Figure 3. Loudspeaker and listener locations that yield smooth low-frequency response at the listening location.

2.3.2 Controlling Room Modal Effects Using Multiple Subwoofers

For single subwoofer setups, the previous solution has worked well for one listener in a single seat. However, audio entertainment has not usually been a solitary activity, and, in multiple listener situations, the previous solution has obviously fallen far short. To achieve a smoother in-room response across a larger seating area, multiple subwoofers have been employed (Welti, 2002; Welti & Devantier, 2006).

For rectangular rooms, Welti (2002, 2004) has developed some generalized optimal placement of two to four subwoofers in a room for achieving the lowest seat-to-seat variance in the low-frequency response. Some of the better configurations are depicted in Figure 4. The underlying principle was the use of destructive interference to cancel as many room modes as possible. By placing two subwoofers along opposite walls, the odd order modes were effectively cancelled along this dimension because they are opposite polarities. A third subwoofer placed at the null of the odd order modes cancelled the second order even mode, and a fourth subwoofer allowed mode cancellation in two dimensions.

The generalized subwoofer placement rules have not worked well for non-rectangular and asymmetrical rooms. For these types of rooms, Welti and Devantier (2006) developed a subwoofer optimization algorithm called "Sound Field Management," which they based on the in-room measurements of the subwoofers in their potential locations, taken at each listening seat. The algorithm rank ordered all of the possible subwoofer-location combinations based on their seat-to-seat variance in amplitude response, and then calculated the optimal subwoofer gain, time delay, and global equalization needed to achieve the best result.

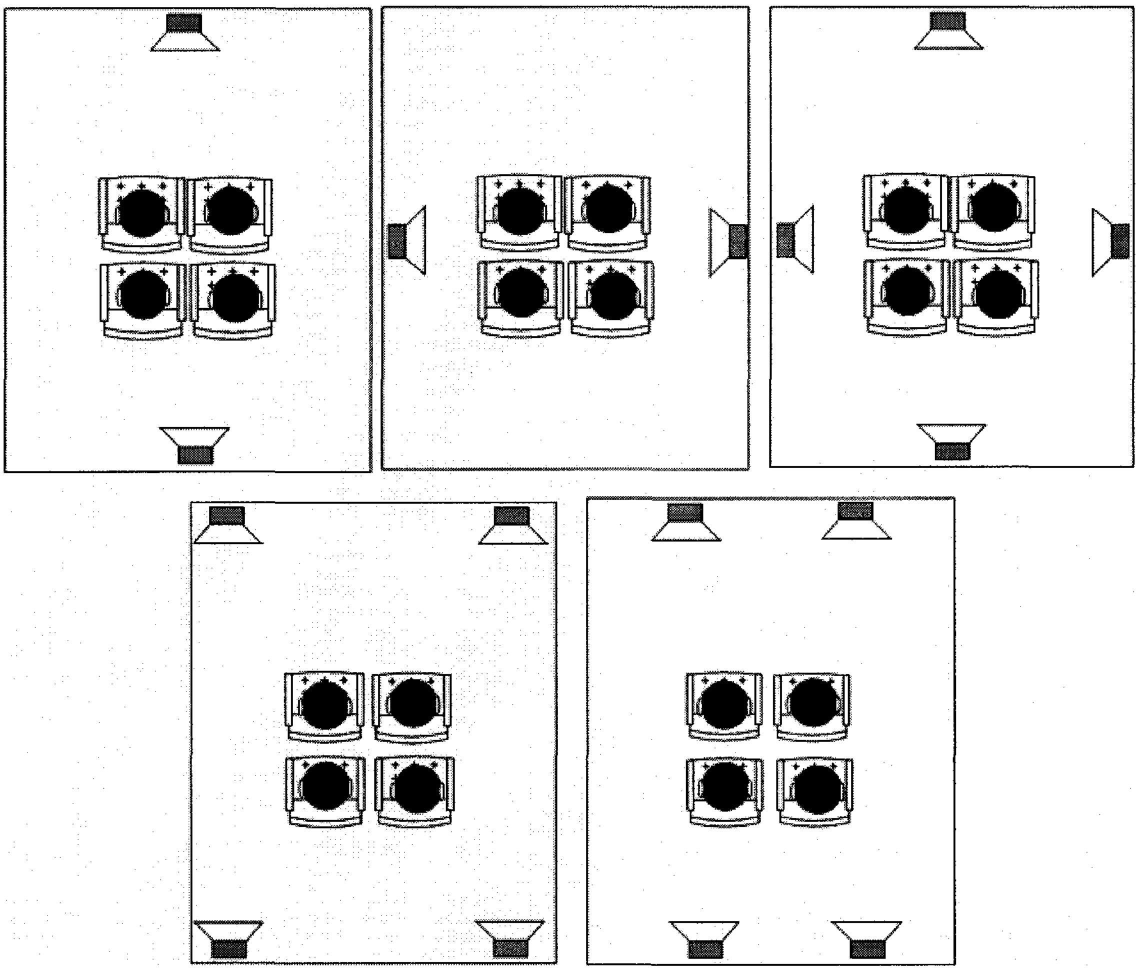


Figure 4. Recommended positions for subwoofers based on Welti (2002).

2.3.3 Subjective Effects Below the Room Transition Frequency

Several studies have reported how different loudspeaker positions within the same room affect listeners' loudspeaker sound quality ratings (Bech, 1993; Klippel, 1990; Olive et al., 1994, 1995). The author demonstrated that the positional effects and their influence on preference ratings were frequently larger than the effects attributed to the loudspeakers themselves (Olive et al., 1994). However, none of the three studies discussed above attempted to isolate the effects of the room below its transition frequency from the effects that occur above it.

The author has found only three studies that specifically evaluated the subjective effects of loudspeaker and room acoustical interactions below the room transition frequency. This was achieved by: (1) using subwoofer-satellite loudspeaker systems and moving the subwoofer(s) to different room locations (Schuck et al., 1993; Welti, 2004); and (2) by using a binaural auralization of a model of a real room, which allowed the experimenters to independently manipulate the effects below the transition frequency (Fazenda, Avis, & Davies, 2005).

Using a satellite-subwoofer loudspeaker system, Schuck et al. (1993, p. 72) found that digital equalization applied below 200 Hz significantly reduced the variance in loudspeaker fidelity ratings related to different positions of the subwoofer within the room. Fazenda et al. (2005) evaluated auralizations of a single loudspeaker in different rooms by splicing the auralized BRIR below 200 Hz to a constant measured BRIR of a real loudspeaker in an actual room. In this way, the loudspeaker-room effects above 200 Hz remained constant, while the effects below 200 Hz varied. While the authors reported that listeners could clearly identify the different rooms below 200 Hz, the subjective effects were highly dependent on the program and its spectral content.

Two informal subjective evaluations of subwoofers located in different positions of the same room (Zacharov, Bech, & Meares, 1995), and/or different rooms (Benjamin & Gannon, 2000), have also been reported in the literature. These investigations were informal in the sense that the listening tests were not performed in a scientifically rigorous manner (e.g. no double-blind controls or statistical analysis). Zacharov et al. (1995) compared a 3/2 surround system with one and two subwoofers (85 Hz cross-over, fourth-order acoustic low-pass filter), placed at different positions in the same room. After viewing different film material, the three listeners, who sat in three different seats, gave informal assessments. The authors stated: “In conclusion, the number, placement, and critical alignment of low-frequency sources (<85 Hz) would appear to be noncritical for the reproduction of 5.1- channel audiovisual material in a domestic environment” (Zacharov et al., 1995, p. 285).

The authors' conclusions were quite misleading because they seemed to have implied that their evaluations could be generalized to all domestic rooms. Given that the subwoofer listening evaluations were informal, sighted, and based on only three listeners (the three authors) in one listening room, the results from this study can be largely discounted. More recent scientific evidence has shown that the number, placement, and alignment of subwoofers in domestic rooms are all critical variables in achieving consistent bass across a wide seating area (Welti, 2002, 2004; Welti & Devantier, 2006), even though a definitive listening test remains to be done.

Benjamin and Gannon (2000) measured the in-room frequency response variation of subwoofers below 200 Hz across a fixed seating area in eight different rooms. They found that the in-room frequency response varied by as much as 30 dB, and that the response was usually dominated by one large peak that could be equalized. The perceptual effects of these variations in frequency response included both uneven loudness among low-frequency notes and audible pitch shifts for high Q value room resonances. Interference between adjacent room modes caused echo-like behaviour with a time scale of 50 to 400 ms. Using a method of adjustment, they found listeners' preferences for subwoofer level settings depended strongly on the program material, and that different listeners prefer different settings, even though the intra-listener reliability was quite good.

Johansen and Rubak (2001) conducted listening tests on three different loudspeaker-room digital correction schemes optimized for a specific seat in the room. The equalizations were performed across the entire audio bandwidth in an attempt to compensate for loudspeaker-room interactions, both below and above the transition frequency. Although no statistical evidence was shown, they reported slight improvements in the overall sound quality, bass quality, and clarity. The exception occurred when the listener sat in a non-optimized seat. In that case, the listener preferred the unequalized loudspeaker system. These results have lent support to the general proposition that the problem with room correction systems is that they cannot be generalized to more than one listening seat in the room.

2.3.3.1 The audibility of low-frequency room resonances

The audibility of low-frequency resonances was investigated by the author and his colleagues in order to gain a better understanding of the perception of room resonances (Olive, Schuck, Ryan, Sally, & Bonneville, 1997). The 70.7% detection thresholds of a single-added resonance (peak) and antiresonance (dip) were measured at several different center frequencies between 63 and 500 Hz.

Thresholds were measured for low *Q* (1), medium *Q* (10), and high *Q* (30) values using two contrasting signals (continuous pink noise and discontinuous 10 ms-wide pulses at 10 Hz) reproduced over headphones.

The mean detection thresholds for the resonant peaks and dips are shown in Figures 5(a)-(b) for pink noise through headphones (Olive et al., 1997, p. 121, Figure 2) and in Figures 5(c)-(d) for impulses through headphones (Olive et al., 1997, p. 122, Figure 3). As previously demonstrated in resonance detection studies (Toole & Olive, 1988), the audibility of a low-frequency resonance was found to be highly signal-dependent (Olive et al., 1997). For pink noise, the low *Q* value resonances and antiresonances produced the lowest overall detection thresholds, which increased approximately 3.2 dB per doubling of the *Q* value. The thresholds were similar for both peaks and dips for the lowest *Q* values. For the discontinuous impulse signal, the effect of the *Q* value on the detection threshold had the opposite effect: the thresholds were lower as the *Q* value of the resonance increased (Olive et al., 1997).

The detection thresholds were also frequency-dependent for medium and higher *Q* values resonances; for the medium *Q* (*Q* = 10) and high *Q* (*Q* = 30) value resonances, the detection threshold increased 0.5-2 dB as the center frequency of the resonance below 500 Hz was halved. For low *Q* value resonances, the threshold remained relative constant regardless of the center

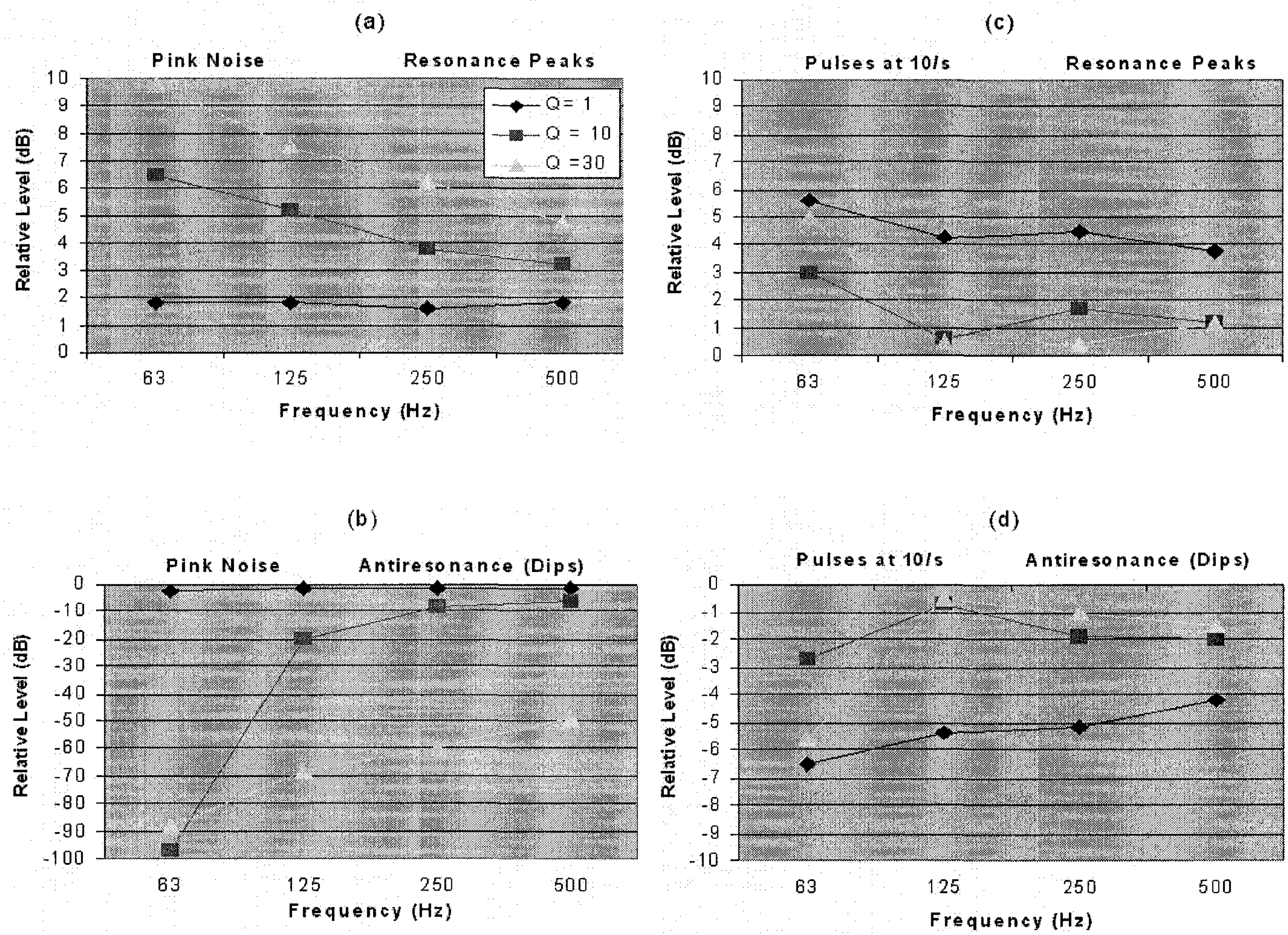


Figure 5. The 70.7% detection thresholds for an added (a)-(b) resonance and (c)-(d) antiresonance for pink noise and pulses.

frequency of the resonance.

More recently, Avis, Fazenda, and Davies (2007) measured the difference limen for the 'Q-factor' of room modes below 200 Hz. The difference limen is defined as the minimum adjustment in the *Q* value of a resonance required to produce an audible change. The difference limens were measured by having subjects listen through headphones to an auralization of a modelled listening room below 200 Hz, while digitally adjusting the *Q* values of the room resonances until there was an audible difference. This was done for two different listening rooms with short (RT_{60} = 0.2 s) and medium (RT_{60} = 0.5 s) reverberation times (Avis et al., 2007).

The measured BRIRs of each room were then spliced together with the auralized model of the room at 200 Hz. For room resonances with low (*Q* = 1), medium (*Q* = 10), and high (*Q* = 30) *Q* values, the difference limens in *Q* values were: 16.6, 11, and 7.4, respectively, and the more reverberant room produced higher difference limens in the *Q* values. The authors concluded that, unless the *Q* values of the room resonances were relatively high, the use of acoustic passive absorbers to treat them would have had little audible benefit (Avis et al., 2007).

Together, the two studies described above indicate that the audibility of room resonances is complicated by several confounding factors. Higher *Q* value resonances tend to be more easily heard when the program signal is discontinuous and the room's reverberation time is very low; under these listening conditions, the ringing of the resonance is most audible. However, the audibility of the ringing decreases, due to temporal masking, as the signal becomes more continuous, and/or the reverberation of the room increases.

Under these listening conditions, lower *Q* value resonances become more audible. Since film sound tracks and music and speech recordings typically contain both continuous and discontinuous signals, the audibility of room resonances will be problematic, regardless of their *Q* value, and must be dealt with. Fortunately, loudspeaker equalization below the transition frequency and/or judicious placement of the subwoofers and listener(s) provide practical solutions to the problems associated with room resonances.

2.3.4 Room Effects Above the Transition Frequency

Above the room's transition frequency, much of the loudspeaker's total energy arrives at the listener as reflected sounds. A scientific understanding of the perception of reflected sounds and their subjective effects is beneficial when considering the acoustical design of the listening room, loudspeakers, and their optimal positioning in the room. Toole (2006) provides a good summary of relevant literature on the effects of reflections in rooms. According to Toole (2006), the effects that reflections produce include:

1. Localization: direction (the precedence effect)
2. Localization: distance

3. Image size and position
4. Sense of space
5. Timbre: comb filtering and repetition pitch
6. Timbre: audibility of resonance
7. Speech intelligibility

2.3.4.1 The Effects on Reflections on Localization: Direction

It is well established that listeners can localize the direction of a sound source in enclosed spaces, even in the presence of numerous reflections that arrive from different directions than the direct sound. The perceptual mechanism responsible for this is known as the "law of the first wave front" (Wallach et al., 1949) or the precedence effect (Blauert, 1996; Blauert & Braasch, 2005; Buchholz et al., 2001; Clifton, 1987; Clifton & Freyman, 1997; Dizon & Colburn, 2006; Djelani & Blauert, 2001; Haas, 1972; Hartmann, 1997; Litovsky et al., 1999, 2000; Litovsky & Shinn-Cunningham, 2001; Lochner & Burger, 1958; Rakerd & Hartmann, 1986, 1992, 1994; Saberi & Antonia, 2003; Saberi & Perrott, 1990; Shinn-Cunningham, 2000, 2001, 2004; Shinn-Cunningham, et al., 1998a, 1998b; Shinn-Cunningham & Ram, 2003; Zurek, 1979).

The initial direction of the sound source is established by both the interaural time and the intensity differences at the two ears, as well as by the monaural spectral cues provided by the direct sound. There is a fusion zone where the reflected sounds from the room are not perceived as separate spatial events, but appear to come from the same direction as the original sound. Although the listener does not hear the reflected sounds as separate events, he does clearly perceive them as contributing to the timbre, loudness, and intelligibility of the sound. Additionally, in some cases, the reflected sounds contribute to the perceived spatial extent of the sound source. Depending on the level, delay time, angle of arrival, the reverberation of the room, and the temporal-spectral properties of the direct and reflected sounds, the reflected sounds may be perceived as separate auditory events or as echoes.

Research over the past 30 years indicates that the precedence effect is the result of evaluation and decision-making processes that humans perform at a higher level of the central nervous system; where, in addition to processing auditory information, cues from other sensory modalities and prior knowledge are taken into consideration (Blauert & Divenyi, 2001; Djelani & Blauert, 2001; Litovsky et al., 1999; Litovsky & Shinn-Cunningham, 2001). Cognitive processes play a key role in our perception of the precedence effect. Auditory cues that are redundant become less relevant to our perception of the source's direction, distance, and spatial impression. When the auditory cues become unstable or contradict more dominant sensory visual cues, listeners rely on the most recent set of learned cues that are "plausible" (Hartmann, 1997; Rakerd & Hartmann, 1992, 1994).

The precedence effect consists of three stages: the initial build up, the suppression, and the breakdown. In unfamiliar acoustical settings, there is a short build-up period, during which listeners apparently listen to both the direct and reflected sounds in order to learn the features of the reflection patterns. After the build-up period, the reflections are suppressed, and are no longer heard as separate spatial events. Spatial suppression of the reflections can persist up to 9 s, even when the sound source is not continuous, after which there is a breakdown of the precedence effect (Blauert & Braasch, 2005). The release of the prior precedence effect and build up of a new precedence effect is often triggered by a sudden physical change in the reflection patterns' level, direction, arrival time, or spectrum (Blauert, 1996; Dizon & Colburn, 2006).

The precedence effect seems to be most effective when the spectra of the direct and reflected sounds are similar (Blauert & Divenyi, 1988; Divenyi, 1992; Shinn-Cunningham, Zurek, Durlach, & Clifton, 1995). Although it remains to be experimentally proven, this suggests that the loudspeaker should have constant directivity, and that the acoustical treatment of the room must be broadband at least down to 200-300 Hz. (Toole, 2006). Finally, the cognitive underpinnings behind the precedence effect are supported by experimental evidence indicating that it is possible to alter the strength of the precedence effect through (1) listener training (Saberi & Antonia, 2003), (2) learning (Litovsky et al., 2000; Shinn-Cunningham, 2000), and/or (3) altering the listeners' expectations (Clifton, Freyman, Litovsky, & McCall, 1994).

2.3.4.2 The Effects of Reflections on Localization: Distance

Early reflections provide important cues for judging the distance of sounds in rooms (Blauert, 1996; Bronkhorst & Houtgast, 1999; Nielsen, 1993; Zahorik, 2002). Through practice, listeners can judge the distance of a sound source by learning the room's reflections patterns and transferring this knowledge to other locations in the room. Additionally, listeners can, to some extent, transfer this knowledge to rooms with similar acoustical features (Schoolmaster, Kopčo, & Shinn-Cunningham, 2003, 2004; Shinn-Cunningham, 2001; Shinn-Cunningham & Ram, 2003). The spectrum and level of the direct sound also provides monaural cues for judging the distance of a sound source.

Many audio recordings strive to realistically simulate a real or artificial space. This is more possible with multichannel technology; the side and rear loudspeakers allow more accurate spatial distribution of the direct and reflected sounds. An important research question raised by Toole (2006) is whether the reflections in the listening room prevent the directional and distance cues in the recording from being properly perceived. Even more problematic are the perceptually more dominant visual cues of the loudspeaker and the room boundary locations; these visible physical objects may prevent listeners from localizing the auditory images beyond these visual boundaries. When listeners are presented with auditory localization cues that are ambiguous, or that conflict with visual information, the most "plausible" cue will generally dictate where the sound is localized (Rakerd & Hartmann, 1986).

2.3.4.3 Reflection Effects on Image Size and Position

Various studies have shown that early lateral room reflections can alter the perceived size and position of auditory images (Bech, 1998; Olive & Toole, 1989; Toole, 2006). The detection thresholds were measured by Olive and Toole for a single lateral reflection in different acoustic spaces, using loudspeakers to simulate the direct and reflected sounds. For reflections under 20 ms, a sense of spaciousness was perceived as the level of the reflection (relative to the direct sound) was increased 10 dB above its absolute detection threshold. At 10 dB above reflection detection threshold, the auditory sensation changed from a sense of spaciousness or LEV, to a broadening of the source, even though the direction remained relatively stable. In concert hall studies, similar effects were reported as an increase in the ASW, which was related to a decrease in the value of the IACC (Okano et al., 1998).

2.3.4.4 Reflection Effects on Spatial Impression

Spatial impression is comprised of two spatial aspects, the ASW and the LEV. These two effects have been associated with strong lateral reflections that produced lower IACC values at the listeners' ears. In concert halls, higher values of LEV and ASW gave listeners a heightened sense of the space, and were considered positive attributes (Ando, 1977, 1998; Ando & Schroeder, 1985; Barron, 1971; Barron & Marshall, 1981; Beranek, 1996, 2003; Bradley & Soulodre, 1995, 1996; Hess & Blauert, 2005; Hidaka & Beranek, 2000; Hidaka et al., 1995; Merimaa & Hess, 2004; Okano et al., 1998; Schroeder et al., 1974).

Within the context of audio reproduction in domestic listening rooms, there is sparse scientific research to indicate whether this is a desirable attribute. Some audio purists would argue that, in order to hear what the recording artist truly intended, the domestic room should acoustically approximate the condition of the professional control room where the recording was mixed. Determining what constitutes the ideal listening room is a complicated research question that most likely depends on a confluence of variables: the genre of music, the number of playback channels, the recording techniques, and the experience and expectations of the listeners.

For multichannel audio systems, a range of spatial impressions can be achieved through judicious adjustment of the salient cues by the recording engineer, or by the listener, using their surround sound processor (Griesinger, 1996; Soulodre, Popplewell, & Bradley, 1989). For stereo setups, the only means available to produce a heightened sense of spatial impression are: (1) through binaural spatial processing of the stereo signals, or (2) by providing the listener strong side wall room reflections.

2.3.4.5 The Effect of Reflections on Timbre and Speech Intelligibility

A popular myth among audiophiles is that most room reflections are undesirable due to the timbral colourations, or comb-filtering, they cause from the summation of two delayed, coherent signals. The "comb effect" refers to the alternating series of peaks and notches in the frequency response that occur from the constructive and destructive interference between the two delayed signals. Visually, the comb effect is most apparent in a frequency response graph when the delayed signals are electrically or acoustically summed at a single microphone. However, a listener equipped with two functioning ears and a brain is capable of perceptually sorting out the direct and reflected sounds in a way that minimizes the timbral colouration, which is counterintuitive to what the in-room frequency response measurement made with the microphone at the listener's ear would seem to indicate.

The colouration effect is greatest when the reflected sound comes from the same direction as the direct sound (Hartmann, 1997; Rakerd & Hartmann, 1986), and additional reflections seem to reduce the colouration (Barron, 1971; Blauert, 1996). Furthermore, there is evidence of a central auditory mechanism that perceptually compensates for distortions in the spectral envelope produced by the room reflections (Watkins, 1991, 1999, 2005; Watkins & Makin, 1994, 1996a, 1996b, 2007). Evidently, listeners are able to learn and adapt to the invariant acoustic features in rooms, such as the reflection patterns, so that they can better hear the true features of the sound source; including its direction, timbre, and intelligibility. The fact that we are able to enjoy, and actually prefer, listening to music in highly reflective spaces, such as concert halls, attests to our ability to adapt to reflections. The relative timbre constancy of our voices as we move from room to room is an everyday example of the adaptation process at work.

Reflections in concert halls provide timbral richness to the sounds of the musical instruments, which they clearly lack when we listen to concerts outdoors. There is experimental evidence that the reflections in domestic rooms also enhance the timbre of reproduced sounds. Adding repetitions to the reproduced sound enhances the audibility of low and medium *Q* value resonances within the recorded instruments (Olive & Toole, 1989).

Finally, there is evidence that a single early reflection (< 30 ms) has little negative effect on the intelligibility of speech in rooms. Intelligibility improves as the time arrival of the reflection decreases from 30 ms, which is the average temporal integration interval of speech (Nakajima & Ando, 1991). As the time arrival of the reflections increase beyond 30 ms, and/or their relative level approaches that of the direct sound, it has a disturbing and negative effect on the intelligibility of speech. The spatial separation of the reflection from the direct sound is also a factor; intelligibility decreases as the separation of the reflection is reduced from 60° (horizontal) to 0°, the direction of the direct sound.

In studies where multiple reflections of different-sized rooms were simulated, reflections that arrived before 50 ms had the same positive effect on intelligibility as increasing the level of the direct sound did (Bradley, Sato, & Picard, 2003; Soulodre et al., 1989). Based on the results of these studies, reflections in most domestic listening rooms would seem to have only positive effects on the intelligibility of reproduced speech.

2.3.4.6 Audible Effects of a Single Reflection

Several studies have investigated the audible effects of a single, lateral reflection in the presence of the direct sound at 0° incidence (Haas, 1972; Lochner & Burger, 1958; Meyer & Schodder, 1952; Olive & Toole, 1989). The thresholds for the different subjective effects of the reflection on speech from these four studies have been summarized in Figure 6. The absolute level detection threshold of a single lateral reflection delayed by 1 ms was about -20 dB relative to the direct sound, and increased to -15 dB over the course of 20 ms, beyond which the threshold rapidly falls (Olive & Toole, 1989). As the level of the reflection increased above the detection threshold, a sense of spaciousness or LEV developed. Approximately 10 dB above the absolute detection, Olive and Toole noted that the auditory image became wider, and shifted in direction towards the reflected sound.

Other studies found that a further 10 dB increase in level of the reflection produced a second auditory image (Lochner & Burger, 1958; Meyer & Schodder, 1952). When the level of the reflection was approximately 10 dB higher than the direct sound, the apparent loudness of the reflection was equal to that of the direct sound, the so-called “Haas effect” (Haas, 1972).

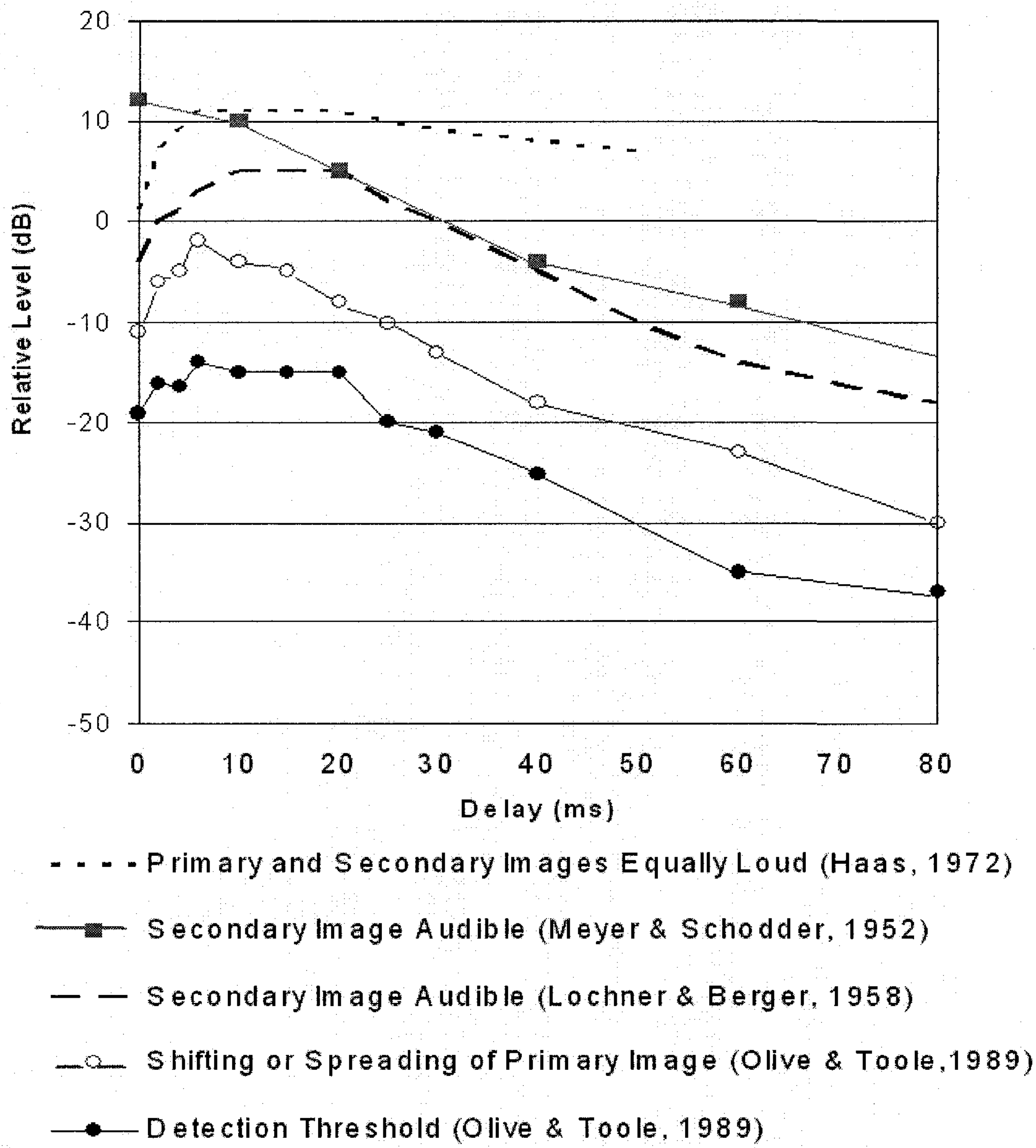


Figure 6. Effects of a single reflection for speech.

In conclusion, a single lateral reflection produces four distinct perceptual effects, separated by approximately 10 dB increments above the absolute detection threshold: (1) spaciousness, (2) image shift and spreading, (3) two distinct sound sources, and (4) equal loudness for the primary and secondary sound sources. It must be emphasized that these results were generated using loudspeakers that simulated the direct sound and the first single reflected sound only; they were not full simulations of real rooms. Reflection thresholds under more realistic acoustic conditions that more closely approximate those typically found in listening rooms are examined in the next section.

2.3.4.7 Reflections in Domestic Rooms

A survey of consumer audio setups in 15 different domestic listening rooms found that the average level of the first six reflections relative to the direct sound fell somewhere between -3 to -8 dB, at delays between 2-10 ms (Devantier, 2002). The earliest and strongest reflections were from the floor and ceiling. The levels of the other four reflections generally fell outside of the threshold range where image spreading and image shifting would have occurred. Based on these findings, it seems safe to state that most reflections in domestic rooms generally occur at levels that produce few, if any, negative effects.

This was, in fact, confirmed by Bech (1995, 1996) who investigated the threshold of audibility for individual reflections produced by a single loudspeaker in a typical domestic listening room. He found that among all of the first order reflections, only the floor reflection (he modelled a reflective, uncarpeted floor) contributed to the timbre of a single loudspeaker when reproducing noise. For speech signals, none of the reflections contributed individually to the timbre of the loudspeaker. His study confirmed that single reflections in typical domestic rooms have little impact on the timbre of a single loudspeaker.

Over the years, some professional recording engineers and acoustical consultants have advocated that, in order to preserve the audible effects of the reflections in a recording, the early reflections must be absorbed, diffused, or re-directed away from the listener. This viewpoint has persisted, and still exists today in various standards. The ITU-R BS.775-1 (1994) document has recommended that all reflections within 15 ms be attenuated -10 dB relative to the direct, to at least above 1 kHz. The Producers and Engineers Wing Surround Sound Recommendation Committee (2004) was less specific, but their intention about rooms used for the production of surround sound was clear:

> To as great a degree as possible, early reflections should be suppressed... in addition, there should be as much diffusion as a budget will allow... To summarize: the more uniform (diffuse) the ambience in the professional mixing environment, the more site-independent the resultant mixes will be. (section 3.1, ¶.2)

Two different studies have examined the detection thresholds of a first reflection accompanied by other reflections in both real rooms (Olive & Toole, 1989) and an accurately simulated room (Bech, 1998). Olive and Toole repeated their single lateral reflection study in three different rooms: (1) an anechoic chamber, (2) a highly reflective large room, and (3) an IEC compliant listening room (IEC 268-13, 1985). Olive and Toole utilized an IEC listening room in both its natural state (RT_{60} = 0.4 s) and in a modified state, where all the first reflections were attenuated. Between these three room extremes, anechoic to highly reflective, the absolute and image shift spreading thresholds that resulted were only modified by 1-5 dB. Only when the lateral reflection delay was 20 ms or greater, did the highly reflective rooms begin to increase the thresholds by 20 dB due to temporal masking.

Bech (1998) examined the detection threshold of a single reflection in the presence of 16 other reflections with simulated reverberation added after 22 ms. He conducted the experiment in an elaborate anechoic setup that was intended to accurately simulate an IEC listening room. He found similar spatial effects at detection levels similar to those found by Olive and Toole (1989), indicating that the same mechanisms and effects that occurred in real rooms also occurred in simulated reflection environments.

In conclusion, the results of these two reflection studies based on both real and accurately simulated listening rooms provide little scientific justification for acoustically treating the early reflections in most domestic listening rooms. In most cases, the reflections will be below the levels at which there are negative effects on the quality of reproduced sound. Furthermore, there is no scientific evidence that the audibility of reflections in recordings during playback is affected by the reflections produced by the loudspeaker in the listening room.

2.3.4.8 The Effect of Program Material on the Detection of Reflections

Studies show that the detection thresholds of reflections and their subjective effects are dependent on the physical characteristics of the program signal (Olive & Toole, 1989). Figure 7 depicts the detection thresholds of a single lateral reflection for several different types of signals, including Mozart symphonic music, speech, castanets recorded in a reverberant room, and electronically generated impulse trains.

The thresholds tend to vary according to whether the program signal is temporally discontinuous (or impulsive) or continuous (such as music). At shorter time delays, the reflection detection thresholds for impulsive sounds are quite high, dropping rapidly as the reflection's time delay is increased. For continuous signals, such as Mozart, the thresholds are lower, and relatively constant, as the delay is increased. Figure 7 shows that adding reverberation to impulsive signals (e.g. castanet signals) has the effect of lowering the threshold to the levels found in music. Like Mozart, the thresholds remain relatively constant until the time delay increases 10-20 ms, beyond which the threshold rapidly drops.

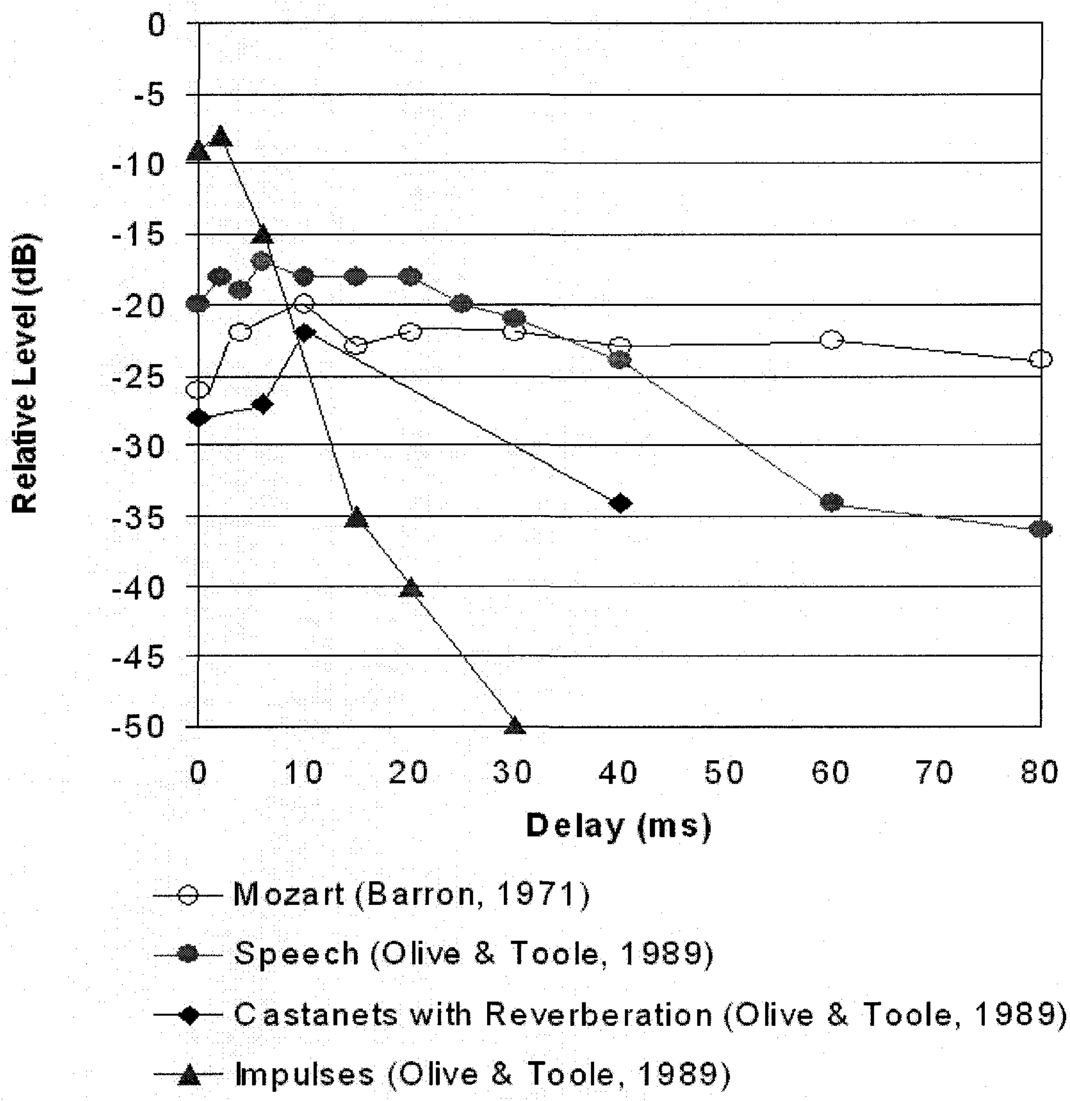


Figure 7. The detection thresholds for a single lateral reflection for different program signals.

In domestic listening rooms, reflections typically have less high-frequency content than the direct sound, due to the off-axis response of the loudspeaker and the acoustical treatment in the room. Low-pass filtering a lateral reflection increases the detection threshold of a lateral reflection (Olive & Toole, 1989). However, even with the high-frequency energy removed above 1 kHz, the reflection was still audible.

2.3.4.9 Preferences for Room Reflections in Sound Reproduction

Ando (1977) has shown that, for live music performance spaces, listeners and musicians preferred lateral reflections that are 10 dB higher in level than those typically found in a domestic listening room (Devantier, 2002). Based on this evidence, Toole (2006) argued that there may be justification for leaving lateral reflections in domestic listening rooms well enough alone because, in domestic listening rooms, they have typically occurred at levels where they produce only neutral or positive effects. To illustrate, he asserted that added lateral energy will enhance the positive spatial attributes (LEV and ASW) associated with highly preferred concert halls.

To the best of the author's knowledge, no definitive study has ever shown what the preferred reflections in a listening room are. Listener reflection preferences are likely complicated by several factors, including: (1) the number of channels used for recording and reproduction, (2) the recording techniques employed (Rumsey, 2001), (3) the genre of music (Voelker, 1985), (4) the directivity of the loudspeakers (Klippel, 1990), and, evidently, (5) the audio training and expectations of the listener (Rumsey, 1999; Rumsey, Zielinski, Kassier, & Bech, 2005b).

Evidence exists that lateral reflections improve the sense of LEV and ASW in stereo setups (Griesinger, 1996). Whether the same applies to multichannel setups has not yet been studied. Many early multichannel recordings are criticized for being too "discrete," implying that the loudspeakers are too easily localized, with little continuity of sound stage between the different loudspeaker channels.

The interaction effects among loudspeaker directivity, listening room acoustics, program material, and listener preference were noted in at least two studies. For stereo reproduction, Klippel (1990) found that listeners preferred less directional speakers for the reproduction of speech and pop music programs, and more directional speakers for the reproduction of classical music. Voelker (1985) reported similar findings with respect to preferred control room acoustics used for stereophonic reproduction of different music genres. In his study, 90 listeners gave their preference choices from amongst four different professional control rooms that each had different degrees of absorption added to them. Binaural recordings of different two-channel loudspeakers (ostensibly equalized to be the same at the listening location), reproducing different program selections, were made in each of the four rooms, and listeners evaluated the recordings through headphones.[8]

Voelker (1985) found that the preferred acoustical treatment was strongly program dependent. For chamber and organ music, a more reflective room (RT_{60} = 0.7s) was preferred, and for pop and disco music, a less reflective live-end-dead-end room (RT_{60} = 0.4 s) was preferred, and the second choice was a nearly reflection-free control room (RT_{60} = 0.2 s), which was preferred for solo drums. The common physical factor seemed to have been the average duration, or auto-correlation interval, of the signal. The perceptual explanation was that, as the signals become more sustained over time, the precedence effect's fusion zone (during which the room reflections were suppressed) was extended (Voelker).

It is important that both Klippel (1990) and Voelker (1985) used stereophonic, not multichannel, loudspeaker setups in their studies. Voelker also set up a 24-channel multichannel loudspeaker system in a reflection-free room to see if he could simulate the sound field heard in the most reflective control room. By adjusting the levels and delays of the signals sent to each of the loudspeakers, he essentially replicated the stereophonic experience heard in the

---

[8] There were at least two major methodological problems with this study. The first was the use of different models of loudspeakers in each of the four rooms. Second, the 1/3-octave equalization of the different loudspeakers at the listening location had insufficient frequency resolution, and it would not have normalized effects related to differences in loudspeaker directivity (Voelker, 1985).

reflective room. By turning the levels and delays up and down, Voelker tailored the multichannel reproduction system according to the preferred taste for the particular genre of music.

Whether he did so unconsciously or presciently, Voelker effectively demonstrated one of the major advantages of multichannel audio over stereo reproduction: The ability of the recording artist or the consumer to adjust the ratio of direct-to-reflected sound to suit the musical genre requirements. The inherent flexibility of multichannel audio has largely eliminated the need to optimize the acoustics of a listening room for a particular genre of music.

2.4 Room Acoustic Adaptation

In chapter 1, room acoustic adaptation is defined as a change in the listener's response to the room acoustics after having spent some period of time in the room. In the context of the subjective evaluation of different loudspeakers in different rooms, room acoustic adaptation is observed as a change in how the room acoustics affect the loudspeaker ratings. Within certain limits, more time spent in the room should produce more adaptation (less room effect), whereas less time spent in the room should produce less adaptation (greater room effect).

The time course of adaptation to room acoustics as an influence on the perception of reproduced sound is largely unknown. Therefore, a manipulation that is intended to modulate room acoustic adaptation may or may not work. Despite the lack of research in room acoustic adaptation, a general consensus exists among many prominent acousticians (Blauert, 1996, Gilford, 1979; Toole, 2006) that listeners appear to adapt to, and ignore many of, the acoustical features of rooms that, in objective terms, would seem to indicate that the sound quality should be significantly compromised.

Gilford (1979) observed this cognitive process at work while reviewing the acoustics of 120 different recording studios and 160 listening rooms used by the British Broadcasting Corporation. After evaluating the monaural recordings made in different studios in both domestic and the 160 listening rooms, Gilford found that the acoustics of the different studios were generally apparent and well-preserved, regardless of which listening rooms they were reproduced in. He concluded:

> The fact that the listening room does not have a predominant effect of quality is very largely due to the binaural mechanism. This is particularly true of rooms used for listening to music programs, since objectively speaking, the standing-wave effects in the room can accentuate particular notes by as much as 8 or 10 dB relative to their neighbours. (Gilford, 1979, p. 28)

Gilford (1979) noted that there were some interaction effects between the studios where the recordings were made and the acoustics of the listening rooms used for their reproduction. Recordings of speech made in six different studios were played back in four different listening rooms, where a panel of recording engineers judged their preference for the sound of the different studios. Gilford found that, while the over-reverberant listening room did not change the rank order of preference for the recordings, it increased the inconsistency in the engineer's ratings. Listening rooms with longer reverberation times tended to favour studios with heavy bass cuts, whereas acoustically dead listening rooms favoured studios with longer reverberation times (Gilford).

2.4.1 Neurological Evidence of Room Acoustic Adaptation

Current scientific thinking is that there are two known perceptual mechanisms related to room acoustic adaptation: the precedence effect and spectral compensation (as previously discussed in section 2.3.4.1). In both cases, it appears that a form of adaptation to the effects of the reflections happens based on learning the reflection patterns in the room that are time invariant. Once the patterns are learned, perceptual compensation or suppression of the reflections occurs.

Recent neurological studies have shown that certain areas of the brain are used to extract reverberation cues, which, in turn, are employed to discriminate amongst different sizes of rooms (Schoenwiesner, Braasch, & Zatorre, 2006). By observing changes in the functional magnetic resonance imaging responses of listeners when certain features of sounds were varied over time, while others were held constant, the researchers learned that certain regions of the brain were associated with specific auditory processes. By manipulating the sound source and the size of the room, they found that the superioral temporal sulcus and the right posterior superior temporal sulcus responded to changes in the sound source, while the posterior superior temporal sulcus was sensitive to changes in room size (Schoenwiesner et al., 2006).

The same researchers also found that another part of the brain was specific to the localization of the sound source (planum temporale). When the room reflection patterns were held constant, they observed reduced activity in the region of the brain that was responsible for processing these features (Schoenwiesner et al., 2006). In conclusion, their study provided neurological evidence that specific parts of the brain were responsible for analyzing patterns in room reflections, and that these areas became less active as the patterns become more constant. While the study did not provide direct evidence of room acoustic adaptation, it did produce evidence that the brain is well equipped to adapt to room acoustics.

The author is only aware of one prior study wherein evidence of adaptation to listening room acoustics was found in the context of evaluating loudspeaker preferences. This study is discussed in some detail in the next section.

2.4.2 Loudspeaker and Room Preference Study by Olive et al. (1995)

In a previous study, the author and his colleagues (Olive et al., 1995), conducted a series of listening experiments, where 20 listeners gave preference ratings for three high-quality loudspeakers, located in three positions, in four different domestic-sized listening rooms. Separate ratings were given for four different music programs, reproduced through a single (monophonic) loudspeaker.

The loudspeaker/room evaluations were done using two different playback methods (in situ listening versus binaural reproductions of the loudspeakers over headphones), and two different stimulus ordering schemes (within-room and between-room judgments) (Olive et al., 1995). The binaural recording and playback device utilized in the experiment allowed for manipulation of factors that were likely related to room adaptation effects. For example, the context and time interval over which the loudspeakers and listening rooms being evaluated could be held constant within a listening session.

The within-room method required listeners to make multiple comparisons amongst the three loudspeakers in each trial, with the independent variable listening room held constant throughout a block or session of trials (Olive et al., 1995). In four separate sessions, the same loudspeakers were evaluated in four different rooms.[9] In contrast, the between-room method allowed listeners to compare the same loudspeaker among the four different rooms, within the same trial. In the following trial, a different loudspeaker-room combination would be evaluated. The method and order in which the listeners evaluated the loudspeakers, positions, rooms, and programs was randomized. The within-room method provided maximum opportunity for room adaptation to occur, and the between-room method afforded minimal opportunity for adaptation to occur. All of the evaluations were performed double-blind. The main conclusions of this study (Olive et al., 1995) were:

---

[9] The *within-room* method described in the study by Olive et al. (1995) is a *successive* treatment condition, the same method that is used in the current study, as described in chapter 4. In both experiments, the room remained constant throughout a block of trials in which different loudspeakers were compared. The room was only changed at the onset of a new block, or session, of trials.

1. For within-room comparisons, loudspeaker preferences were relatively independent of the room (i.e. the room was not a significant effect), which provided evidence that room adaptation was at work during these trials.
2. For between-room comparisons, the listening room was the dominant effect on listener preference ratings (i.e. the loudspeaker was not a significant effect), providing evidence that room adaptation was not effective during these trials.
3. The results of the test were essentially the same regardless of which playback method was used; in situ or binaural.

It is important to point out some of methodological limitations of this study (Olive et al., 1995) that have been largely addressed in the current dissertation study:

1. Accuracy of the binaural device: The binaural reproduction utilized static head position, with dynamic head-tracking headphones, but without BRS.
2. Loudspeaker positional effects: There was no real-time positional substitution of the loudspeakers, although these positional biases were balanced across all of the loudspeakers in each of the four rooms.
3. Loudspeaker variables were not constant: While all three loudspeakers were high-quality, forward-facing, direct radiators, there was no attempt to

keep one factor constant (e.g. direct sound), while manipulating the others.

4. Absence of independent listening groups: All 20 listeners participated in both the within- and between-room tests. Ideally, independent groups of listeners should have been used for the within and between-room experiments to avoid a confluence of biases related to the effects caused by both order and learning the acoustics of the rooms.

2.5 Conclusions from the Literature Review

This section briefly summarizes the main conclusions reached in the literature review.

2.5.1 Summary of the Sound Fields in a Domestic Room

A loudspeaker radiates sound in all directions in a room, some portion of which arrives at the listener's ears as direct, early, and late reflected sounds. Each component plays a different role in our perception of the sound source and its surrounding space. The direct sound establishes the direction of the sound source and its timbre. The early reflections provide perceptual cues related to the source's spatial extent and distance; and improve the intelligibility of the source. The late reflections provide information about the size of the room, and contribute to the loudness and timbre of the source. The quality and proportion of the direct and reflected sounds depend on the acoustic properties of the listening room and loudspeakers, as well as on the positioning of the loudspeakers and listeners in the room. All of these factors physically interact with each other in ways that affect the perceptual attributes of the reproduced sound.

2.5.2 Summary of the Loudspeaker's Effect on the Quality of Sound Reproduction

To a substantial degree, it is possible to predict the overall sound quality of a loudspeaker above 200-300 Hz in a typical reflective listening room using a model based on a set of comprehensive, high-resolution, anechoic frequency response measurements of the loudspeaker. These measurements provide important information about the quality of the direct sound, as well as the early and late reflected sounds; whereas models based on in-room, or sound-power, measurements do not. Consequently, models based on in-room and sound-power measurements less accurately predict listener loudspeaker preferences, compared to models based on comprehensive anechoic data.

However, anechoic measurements lack information about: (1) the room's modal behaviour below 200 Hz; and (2) the reflective and absorption characteristics of the listening room, which affects the properties of the reflected sounds received at the listener's ears. The loudspeaker's low-frequency performance accounts for approximately 30% of its predicted preference rating, indicating that loudspeaker-room interactions below 200 Hz may affect the quality of sound reproduction.

The vast majority of loudspeakers sold today have similar directivities, for both economic and performance reasons. Loudspeakers with more exotic directivities require additional drivers, electrical parts, acoustic lenses, and more expensive enclosures that significantly increase the cost of the loudspeaker with questionable performance gains. With the advent of multichannel audio, the extra side and rear channels provide a more flexible means to achieve many of the same spatial benefits of loudspeakers with special directivities, and do so at a significantly lower cost to the consumer.

Several investigations into the perceptual attributes of loudspeakers using different methods and loudspeakers have largely reached the same conclusion. The generally accepted view is that nine different perceptual attributes account for the principal differences perceived by listeners in sound quality among different loudspeakers. One perceptual attribute relates to loudness, five to timbre, and two to spatial attributes. The remaining attribute relates to nonlinear distortion.

2.5.3 Summary of the Perceived Effects in Domestic Listening Rooms

Typical domestic listening rooms are 60-151 m^3 and have reverberation times of around 0.4 s. The concept of a highly diffuse sound field, like those found in concert halls, does not exist in domestic rooms, due primarily to their relatively low ceilings and higher proportion of sound absorbing and scattering objects distributed non-uniformly about the room. Due to the absence of an isotropic reverberant field, the sound field in a domestic room can consequentially be varied significantly by simply moving the listeners or loudspeakers only a few inches.

2.5.3.1 Summary of perceived effects below the room's transition frequency

Below about 300 Hz, the room's axial resonant modes dominate the quality of reproduced sounds and cause ±15 dB seat-to-seat variations, which change according to the placement of the loudspeakers in the room. The use of a single, equalized subwoofer can reduce these variations for a single seat, and multiple subwoofers can reduce the variations among multiple seats.

This author is aware of only three studies (Fazenda et al., 2005; Schuck et al., 1993; Welti, 2004) that have specifically measured the subjective effects of low-frequency variations on sound quality, and the effectiveness of various loudspeaker-room correction schemes, indicating a significant opportunity for more research in this important area. Studies on the detection of low-frequency resonance peaks and notches indicate that their audibility depends on the frequency, *Q* value, level, and type of program signal. The higher *Q* peaks and dips are more audible on impulsive (discontinuous) signals, but less audible than lower *Q* values for continuous signals, such as music and noise. The reverberation time of the room above 200 Hz decreases the audibility of higher *Q* value resonances.

2.5.3.2 Summary of perceived effects above the room's transition frequency

Above 300 Hz, the off-axis performance of the loudspeaker becomes important because these sounds are received at the listener's ears as room reflections that are modified by the acoustic properties of the room's boundaries. The detection thresholds and perceptual effects of room reflections are generally well-known. Surveys of typical loudspeaker setups in domestic rooms indicate that the levels of the strongest early reflections fall below the thresholds at which negative subjective effects occur (e.g. echoes or secondary images, image shifting or spreading). The remaining effects are all positive ones, and correlate positively to improvements in the intelligibility, loudness, timbre, and LEV (particularly the early lateral reflections).

Two well-known perceptual mechanisms, the precedence effect and spectral compensation, help humans perceive the true direction and timbre of sound sources in highly reflective rooms. Both mechanisms appear to operate at the central stage of auditory perception, involving cognitive aspects associated with learning and adaptation. Adaptation to room reflections occurs when features in the reflection patterns are perceived as invariant and "plausible," and is based on previous auditory cues and information from other sensory modalities (such as vision).

2.5.4 Summary of Room Acoustic Adaptation

Room acoustic adaptation refers to a change in the effect of the room on listener's perception over some period of time. The precedence effect and spectral compensation are two perceptual mechanisms involved in room acoustic adaptation. There is recent neurological evidence that specific areas of the brain are affected by changes in the reflection patterns of rooms, and these areas become less active as the patterns become invariant over time.

Several acousticians have observed that audio reproduction through loudspeakers seems to be relatively immune to the acoustics of the listening room. However, there has been very little research on how we adapt to room acoustics in the context of multichannel audio reproduction. Only one previous study (Olive et al., 1995) specifically addressed how room acoustic adaptation influences listener loudspeaker preferences, and it used monophonic reproductions of music programs reproduced through a single loudspeaker.

When listeners evaluated the loudspeakers in successive treatment conditions of four different listening rooms, the room acoustics had no significant effect on the loudspeaker preference ratings, indicating that room acoustic adaptation may have been the underlying reason. When listeners made comparisons of the loudspeakers among the four rooms using a binaural recording/playback system, the room acoustic effects became the dominant influence on listener preference rating, and the loudspeaker effects were no longer a significant factor (Olive et al., 1995).

In conclusion, this review of the scientific literature on loudspeakers, listening rooms and room acoustic adaptation has uncovered several gaps in our scientific knowledge about this subject. While there is a wealth of information about the perceptual mechanisms (e.g. the precedence effect) that are most likely related to room acoustic adaptation, no one has studied how these mechanisms operate in the context of multichannel reproduction of music in listening rooms. Similarly, while there is an abundance of physical evidence of acoustical interactions between loudspeakers and rooms, no single study has yet measured their subjective effects in a carefully controlled manner using today's multichannel audio setups. There are clearly many challenging and exciting research opportunities; the justification and motivation for this dissertation.

Chapter 3 The Binaural Room Scanning Method

Researchers studying the perception of acoustical interactions between loudspeakers and rooms face several methodological challenges. The first challenge is how to present and control the variables being tested: the different listening rooms, loudspeakers, and the time line over which the listener is exposed to the different room acoustics. The next challenge is whether or not to evaluate different listening rooms and loudspeakers by comparing each sequentially (using a single-stimulus method). Due to a considerable time gap between comparisons of the stimuli, using this single-stimulus approach might reduce the observed sensitivity and reliability of listeners' judgments as compared to real-time, multiple comparison methods (Soulodre, 2003). The final challenge is whether to perform in situ listening tests of different rooms and loudspeakers using a real-time, double-blind, multiple comparison method, which in the author's opinion, is not logistically feasible.

The dimensions, geometry, and absorption characteristics of a room cannot be easily changed in real-time, and requires the use of some means to keep listeners from becoming aware of what is going on. Each of the perceptually important loudspeaker variables (its directivity and frequency responses, both on- and off-axis) cannot be easily manipulated without affecting change in the other loudspeaker variables. Real-time, in situ comparisons of different loudspeakers cannot be done without the use of an elaborate and expensive speaker-mover[10] to remove known strong positional biases (Olive, Castro, & Toole, 1998). For all of these reasons, in situ tests are simply not an option.

BRS is a method of capturing and reproducing the BRIRs of one or more loudspeakers located in an acoustic space. Using the BRS method solves the methodological problems related to research in the perception of acoustical interactions between loudspeakers and rooms by allowing the manipulation of loudspeaker and room before the responses are captured and stored as a set of BRIRs. The captured BRIR sets can be reproduced later and compared in real-time through calibrated headphones equipped with a head-tracker. This allows real-time, double-blind, multiple comparisons of the independent variables that

---

[10] A speaker-mover, or "speaker-shuffler," is an automated, mechanical device that positions each loudspeaker under test in the exact same location of the listening room during comparative subjective evaluations of different loudspeakers. In this way, the acoustical interactions between the room and the loudspeakers are held constant, thus controlling any loudspeaker-position-related biases. For an example of such a device see Olive, Castro, and Toole (1998).

would, otherwise, not be possible using in situ methods. For these reasons, the author chose the BRS method to study the acoustical interactions between loudspeakers and rooms. To the best knowledge of the author, this is the first study to use a BRS method to look at the perception of the acoustical interactions between multichannel loudspeaker systems and room acoustics. It is also one of the first studies to complete a validation of the performance of a BRS measurement and playback system based on how well it can replicate the in situ measurements of listeners' preferences of different loudspeakers evaluated in a reflective listening room.

The following sections describe the BRS method, its potential applications, and the equipment and software required to assemble a BRS system. All of the known errors and limitations of the BRS method are identified, many of which are handled in this experiment through the use of a calibration procedure developed by Todd Welti (Olive et al., 2007). Finally, a validation test is performed to determine the accuracy of the calibration and the performance of the BRS system.

3.1 Description of the Binaural Room Scanning (BRS) Method

BRS is a method of capturing (typically for the sake of reproducing) the BRIRs of one or more loudspeakers located in an acoustic space. Using microphones placed at the ears of a manikin designed to match a human listener, the BRIRs are measured for different head orientations at the location of the listener in the reproduction space (Horbach et al., 1999). Upon playback, the music (or other audio signal) is dynamically passed through the head-orientation appropriate BRIR filter set using a real-time, convolution engine. The BRIR-processed music is reproduced through high-quality headphones equipped with a low-latency, head-tracking system. This approach maximizes the similarity between the listener's experience of the original loudspeaker reproduction system and the headphone-based reconstruction of the loudspeaker signals by ensuring that the loudspeaker signals are presented from spatially-stabilized locations as listeners turn their heads.

3.2 Possible Applications for the Binaural Room Scanning (BRS) Method

As a research tool, BRS allows psychoacoustic investigations, and listening tests using arbitrary sound sources, in different acoustical spaces that would otherwise not be practical, or even possible, using conventional in situ tests. The use of BRS makes it possible to implement psychoacoustic investigations in virtual acoustic reproductions for a number of important applications. Some of these applications include comparative, real-time, subjective evaluations of:

1. Different concert halls, other performance venues, recording studios, and listening rooms
2. Different automotive audio systems (Bech et al., 2005)
3. Different loudspeakers in the same or different acoustical spaces—the application for this dissertation study (see also Olive & Martens, 2007)
4. Different acoustical treatments applied to a listening room

The BRS method makes these evaluations both practical and possible with some additional benefits. Once the acoustical spaces are captured and stored, the listening tests can be quickly and inexpensively repeated at any time, and in any place that has a BRS playback system. The use of BRIR convolution means that any arbitrary program signal can be used for listening tests. This degree of flexibility means that listening tests can essentially be out-sourced to anywhere in world, without the need to build additional (potentially expensive) listening rooms.

3.3 Description of the Binaural Room Scanning (BRS) System

Figure 8 shows a block diagram of the BRS measurement and playback system, which are each discussed separately in the following sections.

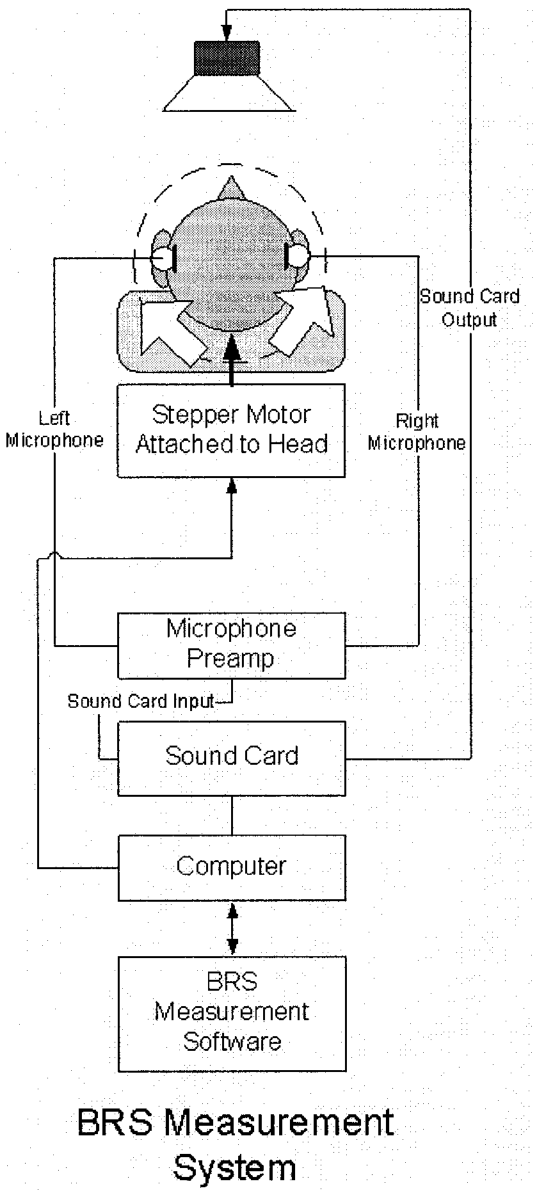


Figure 8. Block diagram of the binaural room scanning (BRS) measurement and playback system.

3.3.1 The Binaural Room Scanning (BRS) Measurement System

The BRS measurement system utilized in this dissertation experiment consisted of a Knowles Electronic Manikin for Acoustic Research, type 45BA (Knowles Electronics, 2007), equipped with average male human pinnae (Knowles Electronics model DB-065/DB-066). A pair of high-quality, miniature 3 mm microphones (DPA model 4060-BM connected to the Sound Devices, model USBPre 1.5 preamplifier) were placed at the blocked entrance of the ear canal of the binaural manikin. This location was chosen because it produces the least inter-listener variability in *head-related transfer function* (HRTF) measurements (Blauert, 2005, p. 230).

The neck of the manikin was mechanically modified so that its head could be rotated within the horizontal plane. This was done in 1° increments, by a stepper motor that was controlled by the computer using the BRIR measurement software. An external, eight-channel, digital sound card (RME HDSP9652 with ADI-8 DS interface) provided the audio interface between the computer, loudspeaker, and binaural microphones. The computer (Hewlett-Packard xw4400) ran a custom Matlab measurement program that generated the test signals and captured the measured sets of BRIRs. The test signal was a log-based chirp that gave an optimal signal-to-noise ratio at low frequencies (Farina, 2000). The signal-to-noise was further improved by averaging a number of repeated measurements.

For each loudspeaker in the room, a set of BRIRs was captured at the listening location over a horizontal angular range of ±60°. The number of BRIRs required depended on the desired angular resolution within this range, which could have been as small as 1° increments. At this study's resolution, a total of 59 BRIR measurements were made for each loudspeaker at 2° increments over a horizontal angular range of ±60°.

3.3.2 The Binaural Room Scanning (BRS) Playback System

The BRS playback system consisted of a pair of high-quality headphones (Sennheiser model HD 600). These open-air, free-air equivalent headphones had low acoustic source impedance, making them appropriate for reproduction of binaural signals captured at the blocked meatus (Møller, Hammershøi, Jensen, & Sørensen, 1995). Attached to the headphones was a low-latency (<35 ms), ultra-sonic, head-tracking system (Logitech Ultrasonic Tracker). The head-tracker monitored the angular position of the listeners' head and sent these coordinates to the BRS playback software, which then switched to the corresponding set of BRIR filters for that angle.

The same computer (Hewlett-Packard xw4400) was used for playback. A custom Java-based software program provided a *graphical user interface* (GUI) (see Figure 9) that was used to: (1) load the appropriate library of BRIR filter sets required for a listening experiment, (2) configure the sound card, (3) monitor the listeners' head position, and (4) switch between the different sets of BRIR filters.

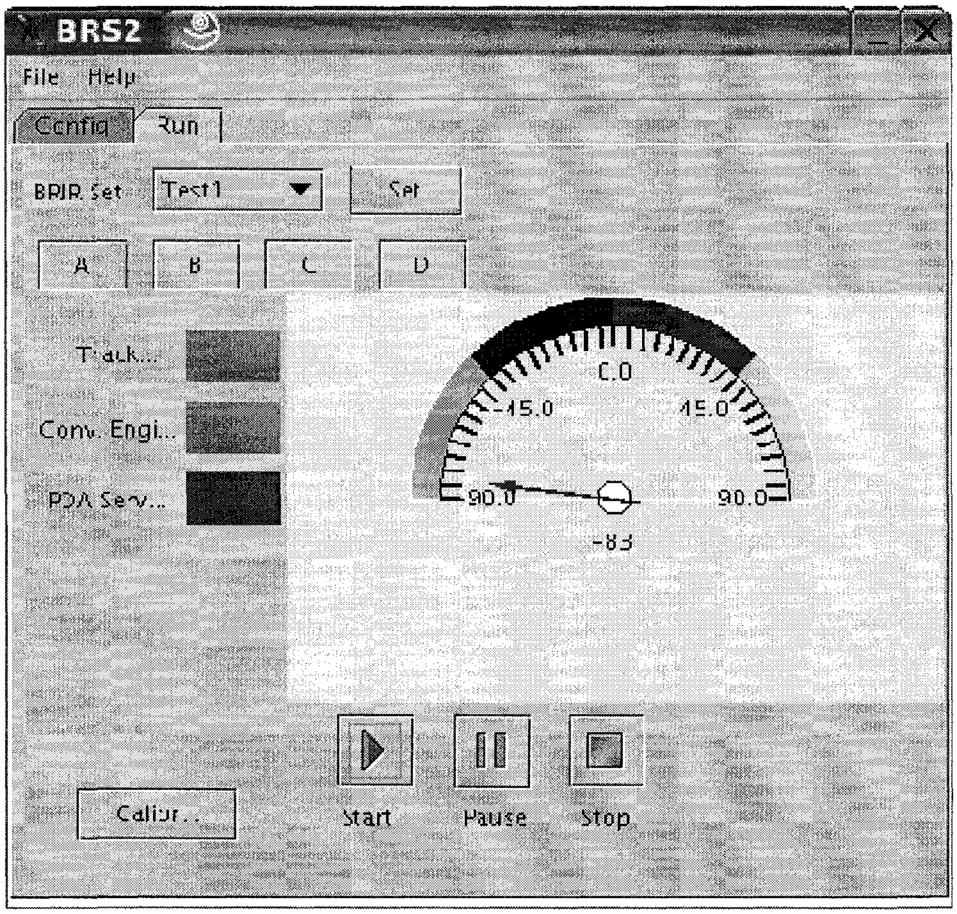


Figure 9. The graphical user interface (GUI) for binaural room scanning (BRS) playback software.

Switching the filter sets was remotely controlled from another computer via the transmission of the communication protocol/internet protocol (TCP/IP) commands sent over a network. This feature has proven useful when the BRS playback equipment cannot be put in the same room as the listener due to its fan noise, space limitations (e.g. automobile audio listening tests), or potential to bias the listener's responses.

3.4 The Importance of Head Tracking

Over the years, psychoacoustic studies of loudspeakers (Olive et al., 1994; Toole, 1991), listening rooms (Olive et al., 1995), concert halls (Schroeder et al., 1974), and automobile sound systems (Azzali, Farina, Rovai, Boreanaz, & Irato, 2004; Farina & Ugolotti, 1999; Mikat, 1996; Otto, 1997; Shively, 1998) have been conducted using binaural recording and playback systems. However, these studies did not use BRS measurements of the acoustic spaces coupled with head-tracking playback.

Given the significant additional costs associated with the combined use of BRS and head tracking (additional equipment and time spent taking measurements), there must be a compelling reason to use them. Many studies have shown that the main advantage of using head tracking in binaural reproduction systems is that it reduces errors in the perceived direction of the sound source, particularly when the sound source is located within the cone-of-confusion[11] or directly in front of or behind the listener (Begault, Wenzel, & Anderson, 2001; Minnaar, 2001; Minnaar, Olesen, Christensen, & Møller, 2001a, 2001b; Wightman & Kistler, 1999).

In natural hearing, humans move their heads (often unconsciously) towards the sound source for two reasons; to improve its intelligibility and to better enable them to localize and identify it. Some of the improvements that result from moving our heads are related to visual cues, but some important improvements in the perception of auditory cues also occur. Head movements produce changes in the interaural and monaural localization cues that are evaluated by the listener in order to better ascertain the direction of the source (Blauert, 1996, pp. 178-191; Perrett & Noble, 1997). Head movements are particularly helpful for disambiguating the location of a sound source in front of the listener from a sound source behind the listener, where binaural cues are particularly weak for stationary listener-source orientations.

In binaural reproductions that do not employ head tracking, where these dynamic, head-movement-related, localization cues are unavailable, increased errors in the perceived direction and distance of the source take place. Sound

[11] The cone-of-confusion refers to the directional ambiguity of sound sources that are located within a cone-shaped area to the side of the listener's head where the interaural time differences (ITDs) produced by a source are the same (Moore, 1997, p. 227).

sources may become difficult to externalize, and are often perceived by the listener as being very close to, or even inside of, their head. This happens most often in the presence of narrow band signals, and/or in the absence of room reflections. Sources located either in front or behind the listener are often observed to be reversed.

The use of head tracking with the binaural reproduction system significantly diminishes the problems described above by providing dynamic localization cues. However, the advantages of head-tracking binaural display appear to diminish when room reflections are included in the reproduction, and/or when individualized HRTFs are used (Begault et al., 2001).

3.5 Binaural Room Scanning (BRS) Measurement and Playback Errors

There are several sources of error in any BRS measurement and playback system including:

1. **Measurement Errors:** Errors relating to the repeatability and accuracy of BRS measurements and to the playback of binaural signals.

2. **Anatomical Errors:** Errors relating to the differences in the shape and size of the manikin's head/torso/pinna versus the listener's same measurements.

3. **Positional Errors:** Errors relating to positional differences between the manikin's head, at the time of the BRS measurement, and the listener's head, observed in situ.

4. **Cognitive-Related Errors:** Errors occurring from inaccurate BRS reproduction of important non-auditory (e.g. tactile and visual) cues, and other cognitive-related factors (e.g. room acoustic adaptation, stimuli order, and context) that influence listeners' judgments.

For the purpose of calibration, it is useful to categorize these errors as either (1) directionally dependent or directionally independent errors, or (2) individualized or non-individualized errors. The reasons for classifying the errors in this manner will become clear in the following sections.

3.5.1 Categories and Sources of Binaural Room Scanning (BRS) System Errors

Figure 10 graphically illustrates the 11 different sources of errors related to BRS measurement, playback, and calibration. Figure 10 broadly categorizes the 11 different sources of errors as either: (1) directionally dependent or directionally independent, and further, as either (2) individualized or non-individualized. Directionally dependent errors are measurement errors that change as a function of the angle of the sound incidence at the binaural manikin's ears. All of the directionally dependent errors fall into the individual classification (errors 1-3 in Figure 10). Directionally independent errors are not impacted by the location of either the listener's ears or the sound source, and are further classified as either individualized (errors 4-7 in Figure 10) or non-individualized (errors 8-11 in Figure 10). Each of the individual errors is briefly discussed below.

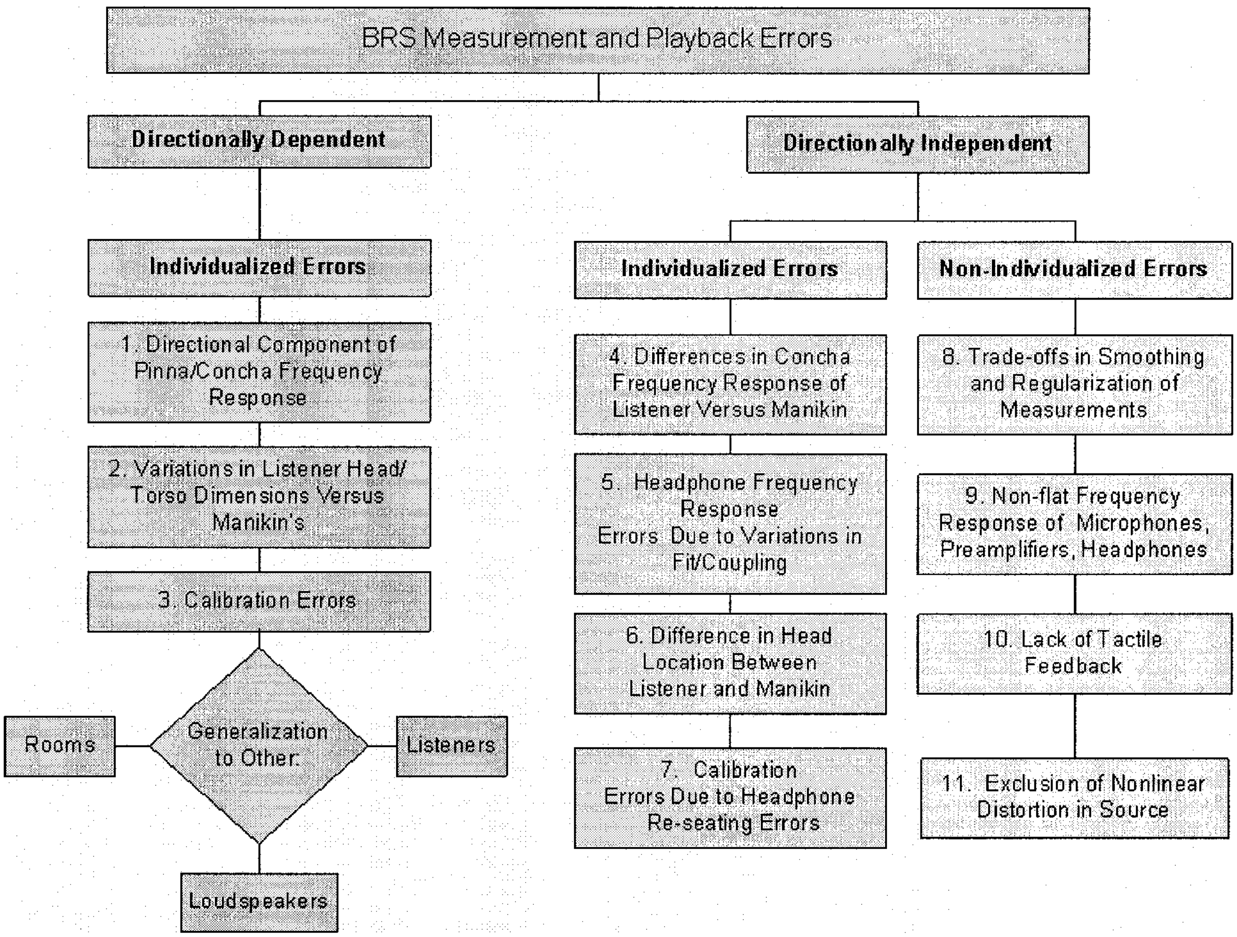


Figure 10. Categories and sources of binaural room scanning (BRS) system errors.

Two of the three sources of directionally dependent, individualized BRS system errors (see the left side of Figure 10) are related to anatomical differences between the manikin and the listener. The first source of error stems from differences between the size and shape of the manikin's pinnae/concha and those of the individual listener's (error 1 in Figure 10). Similarly, the second source of error originates from differences between the size and shape of the manikin's head/torso and those of the individual listener's (error 2 in Figure 10). Since the size and shape of listeners' heads, torsos and ears naturally vary from one person to another, the resulting errors are classified as individualized errors, and they cannot be corrected during calibration using a global filter.

The third and final source of directionally dependent, individualized errors is the BRS in situ calibration process itself (error 3 in Figure 10), which is described in greater detail in section 3.6. The calibration process may not work as expected when applied to other loudspeakers and rooms that are not a part of the original calibration process. Here, differences in the directivity of the loudspeakers, and/or the reflective properties of the rooms may be the source of these directionally dependent, individualized errors. This error is classified as individualized because it varies from loudspeaker to loudspeaker, and from room to room, depending on their respective acoustical properties.

Directionally independent BRS errors include both individualized and non-individualized error types (see the right side of Figure 10). The four sources of directionally independent errors that are also further categorized as individualized are: differences in the transfer function between the manikin and individual listeners (error 4 in Figure 10), headphone frequency response errors due to variations in fit and coupling (error 5 in Figure 10), differences between the location of the manikin's ears during measurement versus the location of individual listener's ears during the in situ listening tests (error 6 in Figure 10), and, finally, calibration errors due to headphone re-seating errors (error 7 in Figure 10).

The most significant source of these directionally dependent, individualized errors is headphone playback response differences that are caused by headphone fit and coupling problems at the listener's ears (inter-listener variance), that is further complicated by re-seating issues (an intra-listener variance). These issues combine to produce both intra- and inter-listener variances that cause significant inconsistencies in the measured in-ear frequency response, particularly at lower and higher frequencies (Møller, Hammershøi, Jensen, & Sørenson, 1995; Toole, 1984).

The remaining sources of directionally independent errors are all classified as non-individualized (see the third column of Figure 10) and include: errors due to trade-offs in the smoothing and regularization of the measurements taken during calibration (error 8 in Figure 10), errors caused by the non-flat frequency response of the microphones, preamplifiers, and headphones used for the BRS measurements (error 9 in Figure 10), errors related to the absence of tactile feedback (error 10 in Figure 10), and errors in the BRS measurements related to nonlinear distortions in both the loudspeakers and the rooms (error 11 in Figure 10).

From the calibration perspective, certain categories of BRS errors are relatively simple, and much less costly, to correct than others. One of the most significant sources of directionally independent, non-individualized errors can be corrected using a single calibration filter. The sources of this type of error that can be removed through calibration generally include the non-flat frequency responses of the microphones, preamplifiers, and of the headphones (error 9 in Figure 10). The non-flat headphone frequency response in headphones can occur from limitations in their design, or from their intended diffuse-field calibration (Thiele, 1986). These errors are considered non-individualized because they are an inherent part of the headphone design, and are not related to how they couple to the ears of the individual listener.

On the other hand, all directionally dependent errors and all individualized errors (errors 1-7 in Figure 10),[12] present a greater challenge, and cannot be corrected using a single global filter. The two most significant and challenging sources of errors in these categories stem from (1) physical differences between the binaural manikin's concha/pinnae and those of the individual listeners (error 1 in Figure 10), and (2) headphone frequency response errors due to variations in fit and coupling (error 5 in Figure 10). Both of these errors must be addressed with a custom-designed filter, individualized for each listener. Naturally, a comprehensive application of this type of correction is both costly and time-consuming. Therefore, the costs of performing this process must be carefully weighed against its performance benefits before considering this cumbersome application. The benefits of individualized calibrations for binaural playback systems are considered in the next section via a review of the relevant perceptual literature.

3.6 Calibration of the Binaural Manikin

In the previous section, the binaural manikin was identified as a major source of both directionally dependent and individualized errors. For this PhD study, it was not considered feasible to use individually molded pinnae for each

[12] *All* directionally-dependent errors fall within the classification of individualized. In contrast, directionally-independent errors can be *either* individualized or non-individualized.

test subject. In view of this, an examination of the relevant perceptual literature is required to explore what effects, if any, this might have had on the results of this study.

There have been several detailed studies with regard to the need for individualized pinnae (Begault et al., 2001; Minnaar, 2001; Minnaar et al., 2001a; Møller, Hammershøi, Jensen, & Sørenson, 1999; Møller, Sørensen, Jensen, & Hammershøi, 1996; Spikofski & Fruhman, 2001; Wenzel & Begault, 1999; Wightman & Kistler, 1999). Most of these studies did not employ head-tracking systems, which made them less relevant to this application. Generally speaking, the studies demonstrated that individualized pinnae yielded some performance improvements, particularly with respect to both front/back localization errors and median-plane localization errors. The few studies that did utilize head-tracking playback to directly explore this issue generally found no significant improvements in performance derived from the use of individualized pinnae (versus non-individualized pinnae). Horbach et al. (1999) found no significant differences in horizontal localization when comparing sound sources in situ to BRS playback of the sources captured with non-individualized pinnae. Begault et al. (2001) compared localizations errors using generic and individualized HRTFs coupled with head tracking, and found no significant difference for speech signals.

A recent study of perceived quality of binaural reproduction (Usher & Martens, 2007) compared the naturalness of speech sounds using individualized versus non-individualized HRTFs, and the results suggested that listeners will choose, as most natural, those sources that are processed using HRTFs similar to their own, but not necessarily identical to their own. In fact, Usher and Martens (2007) found that listeners could regard imagery produced using non-individualized HRTFs as more natural than that produced using their own HRTFs, as long as the non-individualized HRTFs were not too different from their own.

The BRS system employed in this dissertation study did not capture and playback different head rotations in the vertical plane, which raised a question as to whether this had adversely affected the perception of loudspeakers in rooms during the experiment. Not many studies have addressed this issue, but informal observations have indicated that, for some listeners, the perceived location of the image may be slightly elevated and slightly diffused. Furthermore, it has been shown that these errors can be offset by adding visual cues, which have assisted listeners in disambiguating the location of the source, even under normal listening conditions (Begault, et al., 2001). Placing the loudspeakers in the listening room during BRS playback has helped to "anchor" the perceived locations of the loudspeakers reproduced by the BRS playback system.

A listener's familiarity with the acoustics in which the binaural recording was made has also improved the performance and realism of the reproduction (Plenge, 1972). It has been reported that listeners adapt to non-individualized binaural recordings, and that increasing the length or number of repetitions of the test has helped the listener resolve many of the initially perceived errors (Minnaar et al., 2001a; Zahorik, 2002). Finally, the horizontal head tracking employed by the BRS setup may aid the listener in vertical location (Perrett & Noble, 1997).

3.7 Calibration of BRS Playback

The errors that result from using a non-individualized manikin head can be corrected, to some extent, by making individualized calibrations on each subject. Individualized calibrations can also be utilized to correct for errors resulting from different measured responses of the headphones on each subject. The drawback to this corrective procedure is the increased time required for calibration. Informal listening by the author and his colleagues has indicated that this is not mandatory, although further research is needed.

3.8 General Calibration Procedure

A calibration method developed by Todd Welti (Olive et al., 2007) corrects for the entire BRS measurement system and playback chain by employing a set of calibration correction filters for this purpose. Figure A1 in Appendix A outlines how the calibration filters were derived.

The general principle behind this calibration method is that the BRS measurement and playback system should capture and reproduce the same signals that would be measured at the listener's ears sitting in the same sound field, facing the same direction. This is accomplished by comparing what is measured in situ at the listeners' blocked meatus, to what is measured in their blocked meatus during playback, through the (uncalibrated) BRS playback system. Any observed differences are corrected.

Since average (non-individualized) pinnae were used on the binaural manikin in this study, it was not possible to simultaneously correct the free- and diffuse-field responses for individual listeners. The calibration method applied in this study corrected the diffuse-field (or more correctly quasi-diffuse field), an approach that has been generally supported by other investigations (Larcher, Jot, & Vandernoot, 1998). In addition to correcting headphone errors, this method was able to partially correct for differences between the individual listener's and manikin's pinna/concha response.

Figure 11 is an illustration of the in-room frequency response of a loudspeaker, measured at the blocked ear of a listener (dark curve) and reproduced through headphones after the BRS system has been calibrated (light curve). Ideally, the two curves should be the same, and any differences indicate errors. The differences between the two curves are relatively small, indicating that the calibration is effective at removing most of the errors in the BRS measurement and playback system.

The errors that occur below 100 Hz are mainly due to directionally independent, individualized headphone fit and re-seating errors (errors 5 & 7 in Figure 10). The errors that occur above 8 kHz are mostly directionally dependent, individualized errors, related to in-ear microphone placement during calibration (error 3 in Figure 10) and from anatomical differences between the mankin's pinnae and those of the individual listener's (error 1 in Figure 10). There are also some smaller, directionally independent, individualized errors between 3-6 kHz that result from frequency response differences between the concha of the manikin and the listener (error 4 in Figure 10).

The magnitude of these minor calibration errors is significantly smaller than measured differences among the loudspeakers, among the rooms and among the loudspeaker-room interactions that were evaluated in this PhD study. This factor, in combination with the successful validation test performed (discussed in the next section), indicates that the performance of the BRS system is not a significant variable in this study.

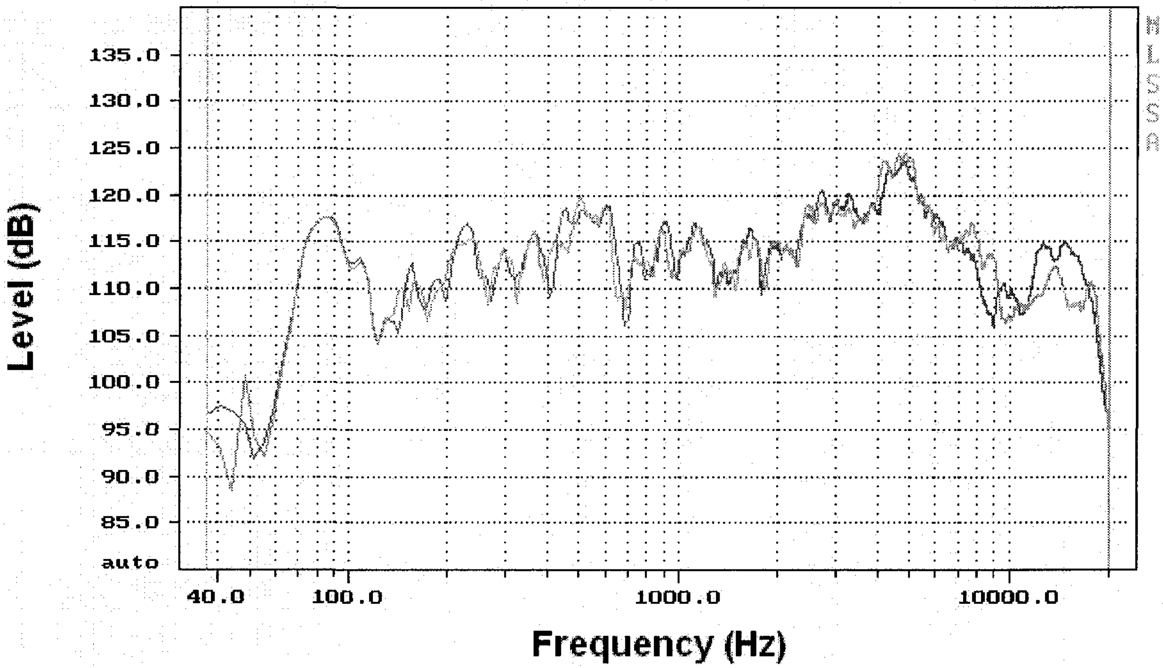


Figure 11. The in-room frequency response of loudspeaker measured at the ear of the listener in situ (dark curve) versus what is measured at the headphones during binaural room scanning (BRS) playback (light curve).

3.9 Validation of Binaural Room Scanning (BRS) Calibration

Currently, no international standards or recommendations exist for testing and validating the performance of BRS systems. There are many possible approaches. In the author's opinion, one of the best possible approaches would be to conduct a listening test where listeners are required to report on a specific perceptual attribute (spatial or timbral) for imagery created using the BRS system, and then the same test is repeated in situ. Ideally, the results for both would be the same.

The author is only aware of two published BRS system validation tests (Bech et al., 2005; Olive et al., 2007) as opposed to "virtual audio" playback systems.[13] One BRS system validation test compared the impairment ratings of low bit-rate audio codecs, evaluated in situ, using a BRS playback method (Bech et al.). No significant differences in ratings were found between the in situ and BRS playback methods.

The second validation test compared the listeners' preference ratings of four different loudspeakers in a domestic-sized reflective listening room, using both in situ and BRS playback methods (Olive et al., 2007). The results of this test have been summarized below. It should be noted that these results were not reported in Olive et al. (2007), but were based on a repeat of the same tests, using a slightly improved calibration procedure, based on the headphone responses measured for a different subject.

A total of 8 trained listeners participated in two separate sessions where the four loudspeakers were evaluated in situ or through headphone-based BRS

---

[13] A distinction is made here between "virtual audio" and BRS playback systems. BRS playback systems reproduce BRIRs based on actual measurements of the source in the room. Virtual audio systems reproduce simulated BRIRs based on a model of the room; not actual measurements. Several BRS validation tests have been done in studies that have utilized the "virtual audio" approach as reviewed by Pedersen and Minnaar (2006).

reproductions of the loudspeakers. A session consisted of eight trials, wherein, during each trial, the subject would give each of the four loudspeakers a preference rating on an 11-point preference scale. This was repeated using four different programs, with one repeat of each program. All presentations of the stimuli were performed using a double-blind, multiple comparison method (four loudspeakers at a time) that utilized a randomized presentation order for both loudspeakers and programs. The order in which the BRS and in situ tests were completed was equally distributed among the subjects to minimize possible order biases.

The test results were statistically analyzed using a repeated measures *analysis of variance* where the between-subjects factor was playback method (in situ or BRS playback method). The independent variables included loudspeakers (four levels), programs (four), repeat (one), and playback method (two); and preference rating was the dependent variable. The alpha level for all tests was 0.05. A summary of the analysis of variance table, shown in Table 1, indicated that the only significant effect was due to the loudspeaker; with a *Fisher's F ratio or F-statistic* (F) (3, 21) = 101.3, $p < 0.001$. More importantly, there were no significant effects or interactions related to the playback method. Loudspeaker preference ratings were essentially the same for the in situ and BRS playback methods. This is graphically shown in Figure 12, where the mean loudspeaker preference ratings are shown for each method. The graph demonstrates that there were no significant interactions between the loudspeaker ratings and the method.

Table 1

The Analysis of Variance Table for the

Binaural Room Scanning (BRS) Validation Test

Source	*df*	*MS*	*F*	*p*	Lambda	Power
Subject group (S)	7	45.802				
Method (M)	1	1.466	.528	.4911	.528	.095
M x S	7	2.778				
Loudspeaker(L)	3	809.654	101.291	<.0001*	303.872	1.00
L x S	21	7.993				
Program (P)	3	.112	.732	.5433	2.197	.176
P X S	21	.153				
M X L	3	4.251	1.029	.3997	3.066	.234
M X L X S	9	4.13				
M X P	63	.035	.287	.8344	.860	.095
M X P X S	3	.121				
L X P	9	.313	.508	.8631	4.575	.225
L X P X S	63	.615				

Note. $p^* \leq .05$

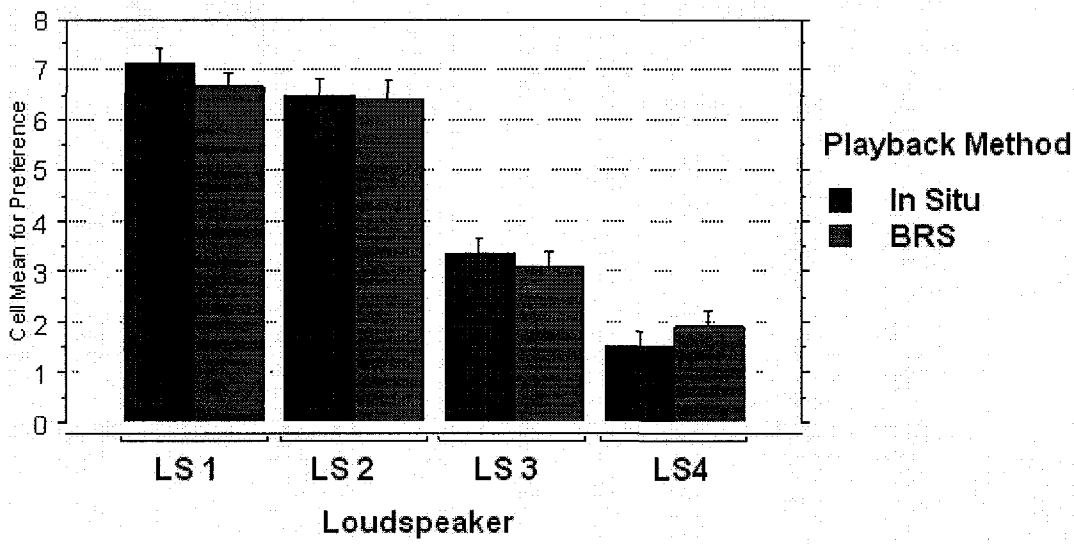


Figure 12. The mean loudspeaker preference ratings and 95% confidence intervals for both the in situ and the binaural room scanning (BRS) methods.

The question of whether individualized calibrations are required for BRS-based evaluations of loudspeakers can be answered, in part, by comparing the individual listener loudspeaker ratings for each of the two playback methods to see if they are in agreement. Here, the individual listener loudspeaker preference results are essentially the same for both in situ and BRS playback methods.

In conclusion, the BRS calibration method utilized in this study is sufficiently accurate to allow valid and reliable preference ratings of different loudspeakers in domestic-sized, reflective listening rooms. Furthermore, the results support the conclusion that individualized calibrations are not necessary to reliably measure listeners' loudspeaker preferences in rooms. This finding is both important and relevant to the validity of the experiment described in the next chapter of this dissertation, where the BRS method is used to measure preference ratings of different loudspeakers in several different domestic-sized, reflective rooms.

3.10 Limitations of the Binaural Room Scanning (BRS) System

All of the limitations of the BRS system have been described in section 3.5, and are also summarized in Figure 10. For applications related to testing loudspeakers, the main limitations of the BRS system are that it does not take into account either tactile information or non-linear distortion in the loudspeaker. The results of the validation study described in section 3.9 indicate that the absence of tactile information did not play a significant role in that experiment's results. However, in listening situations where significantly more tactile information is present (e.g. automobiles, motion-platforms, and setups using near-field subwoofers), failure to include such tactile information in the playback experience may produce different results. This is an area that requires more research.

The absence of non-linear distortion in BRS measurements is also not an issue in loudspeaker listening tests, provided that the loudspeakers have no significant audible distortion at the playback levels that they are tested at. It is possible to confirm this a priori with objective measurements and informal listening. Fortunately, most of the major loudspeaker studies have reported that non-linear distortion is not a significant factor in listeners' loudspeaker preferences, with the exception of the smallest loudspeakers (see section 2.2.2.5).

The use of systems based on BRS for psychoacoustic investigations is a relatively new development, and many opportunities for further validation of this method exist. Although the validation test performed in this study establishes that BRS provides an accurate and reliable approach to measuring loudspeaker preference, more validation tests of the BRS system are required to determine its accuracy and reliability for measuring specific timbral and spatial attributes of auditory imagery. These tests should use both stereo and multichannel playback systems in different listening rooms and automobiles. Until then, the performance of the BRS system cannot be generalized to perceptual testing applications that deviate far from the one described above.

Chapter 4 An Experiment on the Influence of Room Acoustic Variations on Loudspeaker Preference

This chapter describes a listening experiment designed to answer several research questions relating to the effects of loudspeaker-room interactions on listener preference ratings made in response to auditory imagery associated with multichannel music reproduction, and the extent to which listeners adapt to these effects. This chapter focusses on the experiment that forms the basis of this dissertation study and is divided into three main sections. Section 4.1 defines the research questions and hypotheses, followed by a description of the experimental design, and the details of the method and test setup are given in section 4.2. Finally, section 4.3 explains the statistical analysis of the results.

4.1 Research Questions and Hypotheses

The experiment was designed to address the following three research questions:

1. To what extent are listeners' preferences for multichannel music imagery influenced by different loudspeakers, room acoustics, and the interaction between loudspeakers and room acoustics?
2. Can the effects and interactions of loudspeakers and room acoustics on listeners' preferences be explained by acoustical measurements of the loudspeakers and room acoustics?
3. To what extent has room acoustic adaptation diminished the effect of room acoustics on listeners' preference ratings?

The three research questions outlined above lead to the formulation of four hypotheses, stated as follows:

1. Loudspeakers with more irregular off-axis frequency responses will receive lower preference ratings.
2. The preference ratings associated with the tested loudspeakers will decrease as the reflectivity of the room increases.
3. Longer exposure times to a particular room acoustic condition across successive rating trials will produce room acoustic adaptation, yielding a relative decrease in the effect of the room on the loudspeaker preference ratings.

4. Varying the room acoustic conditions between rating trials will produce less room acoustic adaptation than the case in which room acoustic conditions are held constant throughout a block of rating trials, producing a relative increase in the effect of the room acoustics on the loudspeaker preference ratings (in the intermixed treatment condition compared to the successive treatment condition).

4.2 Experimental Design

To answer these research questions, the experiment was designed to measure listener preferences for multichannel music imagery reproduced through four different loudspeakers, located in four different listening rooms. The program selections consisted of three different five-channel music selections. The stimuli were evaluated by two separate groups of listeners, using two different trial ordering schemes that provided a differential exposure to the room acoustics, and, hence, a potential change in room acoustic adaptation. The independent variables and their levels are summarized in Table 2.

The experimental design utilized a 2 x 4 x 4 x 3 x 3 repeated measures analysis of variance, where the between-subjects variable was the trial ordering method. More details on each independent variable are given in the following sections.

Table 2

The Experiment's Independent Variables and Their Levels

Independent variable	Levels of variable
Trial ordering method	Successive, intermixed
Loudspeaker	L1, L2, L3, L4
Room	R1, R2, R3, R4
Program	JV, LL, SD
Observation	O1,O2,O3

4.2.1 Independent Variables

The following sections describe the selection and physical properties of the independent variables in this experiment.

4.2.1.1 Listening rooms

Four rectangular listening rooms were used in these experiments, all located at Harman International in Northridge, California. The rooms have been previously used for both formal and informal subjective evaluations of audio technology, and were designed to acoustically mimic typical domestic listening spaces.

All four rooms had sufficiently low background noise (< NC-25), and fell within typical domestic room sizes, ranging in volume from 60 to 151 m^3 (Bradley, 1986). Dimensional diagrams for the four rooms and their loudspeaker setups are shown in Appendix B. The room dimensions and volumes are summarized in Table 3.

Apart from their dimensional differences, three of four rooms (R1, R2, and R3) had similar acoustic treatments. Specifically, the distribution, absorptive qualities, and placement of objects about the room were similar. The acoustic treatments were mostly provided by common objects that are found in most households, such as carpeting, curtains, chairs, and bookshelves.

Table 3

Listening Room Dimensions and Volumes

Room	Dimensions (m) (L x W x H)	Volume (m^3)
R1	5.84 x 4.29 x 2.4	60.19
R2	6.2 x 5.08 x 2.74	105.55
R3	9.04 x 6.58 x 2.54	151.2
R4	7.26 x 6.32 x 2.74	126.03

In contrast to the other three rooms, Room R4 was the most reflective. This was achieved by removing most of its furniture (6 leather tub chairs) and several fibreglass panels (measuring 1.22 m [height] x 1.22 m [width] X 7.6 cm [depth]) from its two side walls. After doing so, the two side walls were essentially 100% reflective. The hardwood floor was also reflective, except for a 2.7 cm thick woollen rug with a 2.7 cm thick foam underlay, which only covered 30% of the floor, and was placed under the seating area. Other than this area rug, the only significant absorption that remained in Room R4 was provided by some fibreglass panels that covered 42% of the total area of the front and rear walls. In total, only 17% of the total area of the room's surfaces was treated to absorb sound above 300 Hz. Some limited diffusion above 1 kHz was provided by RPG Skyline Diffusors (RPG Incorporated, 2007) mounted near the ceiling along the four walls.

A feature common to all four rooms was a reflective surface where the first lateral side wall and ceiling reflections occurred at the listening location. The properties of these reflections varied among the different rooms, based on the positions of the listener and loudspeakers in relation to the side walls. In the two smallest rooms (R1 and R2), the first lateral reflections were approximately the same.

4.2.1.2 Loudspeakers

Several criteria were employed for selecting the loudspeakers in these experiments. First, constant on-axis and low-frequency performance among all four speakers was desired, with the only difference being in their measured off-axis performance. By controlling the direct sound, loudspeaker preferences could be directly related to the physical properties of the loudspeaker's off-axis radiated sound, as well as to the reflective properties of the listening rooms. Furthermore, the acoustic similarities among the loudspeakers would minimize stimulus recognition biases that might have otherwise swamped any room effects, and possibly inhibited room acoustic adaptation.

To improve the degree to which this study's results can be generalized, those loudspeaker designs most commonly found in consumer and professional audio setups were selected. A custom 'adjustable' loudspeaker was designed and fabricated, using forward-facing, direct-radiator, electro-dynamic components, which make up the majority of loudspeakers sold. On each of the four sides of the loudspeaker, an array of drivers were configured to simulate a common three-way bookshelf or tower loudspeaker, with and without a waveguide on the tweeter. The final two loudspeakers were configured as typical three-way, center-channel loudspeakers that contained two midrange drivers arranged either vertically or horizontally, on either side of the tweeter. More details on the four loudspeakers used in this study are given in Table 4.

Table 4

Description of the Four Loudspeakers Used in the Experiments

Loudspeaker	Configuration	Cross-over frequencies
L1	Three-way, with 25 mm tweeter without waveguide, 165 mm midrange, 203 mm woofer	2 kHz, 150 Hz, 80 Hz
L2	Three-way, with 25 mm tweeter with waveguide, 165 mm midrange, 203 mm woofer	Same as above
L3	Three-way vertical array with 25 mm tweeter, 2 x 102 mm mid-range	3 kHz, 150 Hz, 80 Hz
L4	Three-way horizontal array with 25 mm tweeter, 2 x 102 mm mid-range	Same as above

The loudspeakers had an active, programmable cross-over processor (DBX DriveRack 260) that provided driver equalization. A subwoofer (JBL HTPS400) was common to all four loudspeakers, and the arrays were driven by a three-channel amplifier (Proceed Amp 3).

The anechoic frequency response measurements (see Figure C1 in Appendix C) show that each loudspeaker had virtually identical flat (± 1 dB, 30 Hz-20 kHz) on-axis frequency response. The only differences occurred in their off-axis responses above 300 Hz, as indicated in their spatially averaged response curves (early reflection, sound-power, and directivity index responses). The off-axis curves of the four loudspeakers became progressively worse in the following order: L2 (best), L1, L3, and L4 (worst). It is arguable whether L4 was worse than L3, but it is safe to say that L3 and L4 did not perform as well off axis as either L2 or L1. A photograph of the loudspeaker and its associated equipment is shown in Figure 13.

Figure 13. A photograph of the adjustable loudspeaker showing side L1.

4.2.1.3 Music programs

The three musical selections used in these experiments all originated from commercially available multichannel DVD-A and DVD discs (see Table 5). The music tracks were transferred from a DVD to a digital editing software program (Sonar 6, Producer) and edited into short 20-30 s loops. Each loop was saved as five-channel (16 bit, 48 kHz) Microsoft extensible (*.wav) file. All three tracks were well recorded with full-range, spectrally-dense auditory imagery distributed across all five channels. This ensured that all of the programs would be as equally revealing and sensitive to possible spectral and spatial differences related to the different loudspeaker-room combinations.

Table 5

Program Selections Used in the Experiments

Program	Source
JV (Jazz Vocal)	Gordon Goodwin's Phat Pack, "The Phat Pack," CD-DVD (2006). Track 2: Too Close for Comfort (featuring Dianna Reeves)
LL (Lyle Lovett)	Lyle Lovett, "Joshua Judges Ruth," DVD Audio/DTS (2002). Track 2: Church
SD (Steely Dan)	Steely Dan, "Two Against Nature," DVD Audio (2001). Track 1: Gaslighting Abbie

4.2.1.4 Successive and intermixed treatment conditions of room

The experimental context in which auditory stimuli are presented has been shown to influence preference choices or ratings (Rumsey, 2006). The context under which different loudspeakers (Olive, 2004a; Olive, 2004b) and listening rooms (Olive et al., 1995) are compared has also been shown to potentially affect how listeners scale their preferences. In a similar manner, it has been demonstrated that the trial ordering of music programs has influenced preference choices for different multichannel microphone techniques (Kim, de Francisco, Walker, Marui, & Martens, 2006; Martens, Marui, & Kim, 2006). Some researchers have argued that context influences listeners' preferences to such a degree that measures of preference should be avoided altogether in audio research (Zielinski, 2006).

The conclusion that preference should be avoided due to the influence of context is a hasty supposition because it ignores two highly significant facts: (1) that there are effective methods for controlling contextual effects in preference experiments, and (2) that humans make preference choices everyday where context is an integral factor in their decisions. To remove context and preference from sound-quality listening tests severely limits the external validity and generalizability of their results to how humans respond in the real-world. To fully understand the complex cognitive decision processes behind listeners' loudspeaker preference choices, scientists should not exclude the influence of context in listening tests. For these reasons, context should be viewed, not as a nuisance variable, but as integral factor worthy of investigation (Baird, 1997; Martens, 2006).

To study how preference and room acoustic adaptation might be influenced by context, two different trial ordering schemes were used to evaluate the stimuli. The two trial ordering schemes, referred to as the successive and intermixed treatment conditions (Keppel & Wickens, 2004), are graphically illustrated in Figure 14. The trial and session orders were randomized in a balanced way for both the successive and intermixed treatment conditions to control for any order-related biases.

A successive treatment condition, where the room was held constant throughout a single session, and only changed between subsequent listening sessions, was utilized for the first group of listeners as they evaluated the stimuli. The second group of listeners evaluated the stimuli using an intermixed treatment condition, where the room was always changed in subsequent trials within the listening session.

The differential exposure to the room acoustics was expected to produce differences in room acoustic adaptation. Based on a similar study where successive treatments of room were used (Olive et al., 1995), it had been anticipated that the successive treatment condition would produce more room acoustic adaptation (i.e. less room effect on the preference ratings) due to the longer exposure time (nine trials) to the same room acoustics. Correspondingly, it had been expected that the intermixed treatment condition would produce less room acoustic adaptation (i.e. more room effect on the preference ratings) due to the shorter exposure time (one trial) to the same room acoustics.

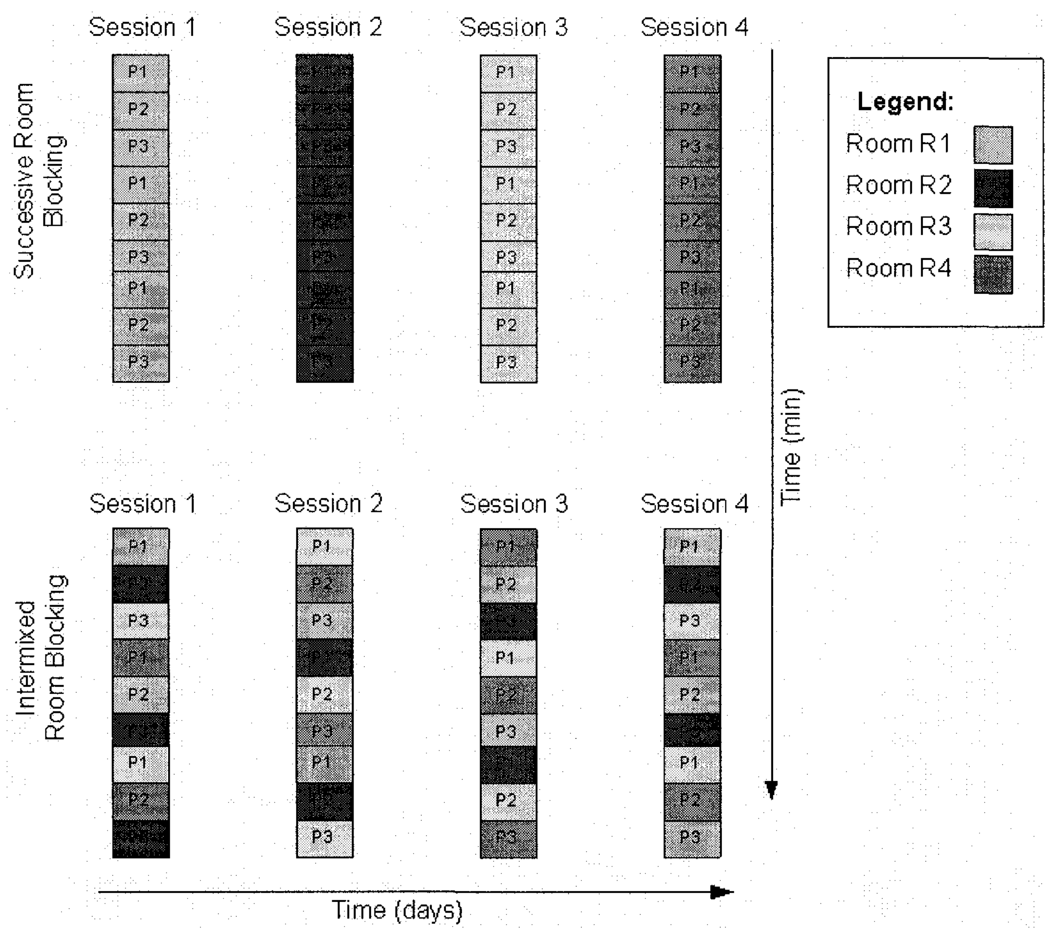


Figure 14. The successive and intermixed room blocking schemes with the three different programs used in this experiment designated as P1, P2, and P3.

4.2.2 Dependent Variables (Preference Rating and Preference Choice)

All stimuli in these experiments were evaluated on an 11-point preference scale that included semantic descriptors on every second interval to encourage the listener to specify his/her intensity level of like or dislike for the scaled stimulus (see lower diagram within Appendix D). Each listening session began with a block of six trials, where the listener made preference choices between paired comparisons of the four loudspeakers (see upper diagram within Appendix D).

The purpose of these pre-trials was to familiarize the listener to the stimuli during an easier task before moving to the more challenging multiple-preference rating task (Kousgaard, 1987). The paired-comparison task provided the means to study how the complexity of the listener's task affected their performance and preference ratings. The simpler pre-test took much less time to complete, and possibly produced less room acoustic adaptation, than the multiple-comparison preference rating task. The results of the preference choice trials are not reported in this paper because they are outside the scope of this dissertation.

4.2.3 Selection of Listeners

A total of 28 paid volunteers participated in the listening experiments. The listeners were current or former employees of Harman International, who were screened for normal audiometric hearing (ISO 389-1, 1998; ISO 8253-1, 1989). Each subject passed a training task that required them to identify spectral distortions added to music (Olive, 2001).

Of the 28 total listeners, 12 (43%) were considered experienced, having had 2 or more years of experience in controlled loudspeaker listening tests. The remaining 16 listeners (57%) were relatively inexperienced, having had less than 6 months of experience in formal loudspeaker listening tests.

The ages of the listeners ranged from 20 to 48 years (median age = 32.5, SD = 9.7 years), and 93% were male. Previous studies (Toole, 1986b; Toole & Olive, 1994) have demonstrated that gender is not a significant variable in loudspeaker preference ratings. Therefore, the high percentage of male participants in this study should not have unduly impacted the extent to which the results can be generalized to females.

All of the listeners were given written listener instructions (Appendix D) before beginning the experiments. Prior to participating in the experiment, the potential subjects were required to read the prospective participant instruction form (Appendix E), and to sign the participant content form (Appendix F). Prior to commencement of the experiment, the author qualified for and was granted a certificate of ethical acceptability of research involving humans from McGill University (Appendix G).

4.2.4 Loudspeaker Setup and Binaural Room Scanning (BRS) Measurements

Loudspeaker and listener positioning in the room (the listener's physical position within both the room and in relation to the loudspeaker) were treated as nuisance variables in this study. In order to control loudspeaker and listener positional biases, all of the loudspeakers were scanned in the same positions in the room from the same listening position. In all four rooms, the loudspeakers were set up symmetrically according to ITU-R BS.775-1 (1994).

The symmetrical arrangement ensured that no loudness, timbre, or spatial imbalances were present between the left and right channels in any of the rooms. Controlling these significant nuisance variables made it possible to trace any perceptual effects to the physical differences in the loudspeakers and/or their interaction with the room acoustics. The exclusive use of symmetrical setups and rectangular rooms in this study may have somewhat limited the extent to which these results can be generalized to loudspeaker setups more commonly found in consumers' homes.

In each of the four different listening rooms, BRIR measurements were taken at the listening location, for each loudspeaker, at each of the five positions at 2° angular resolution over a horizontal azimuth of ±60°. In total, over 4,720 BRIR measurements (4 [loudspeakers] x 5 [positions] x 4 [rooms] x 59 [head positions] = 4,720) were taken to capture all of the experimental test conditions.

The distances and angular positioning of the loudspeakers relative to the listener were compliant with ITU-R BS.775-1 (1994):

1. ± 30° for the front left and right loudspeakers
2. 0° for the center loudspeaker
3. ± 115° for left and right surround channels

The loudspeakers were symmetrically arranged in a circle with each loudspeaker positioned equidistant to the listening location in the middle of the circle, the so-called listening 'sweet spot.' Precise calibration of the loudspeakers' positions in each room was performed through the use of a laser pointer, attached to each loudspeaker, and a small white reflective dot on the binaural manikin's nose. Further confirmation of loudspeaker positional accuracy was achieved by inspecting the in-room frequency responses and the BRIRs of each loudspeaker to ensure that they matched the previous measurement. Precise loudspeaker positioning was necessary to ensure that the direct sound was constant across all loudspeakers, positions, and rooms in terms of its: (1) frequency response, (2) level, (3) relative time arrival among the five channels, and 4) angle of sound incidence at the listener location. The BRS measurement parameters are summarized in Table 6.

Table 6

Binaural Room Scanning (BRS) Measurement Parameters for the Experiments

Parameter	Value
Sampling rate	48 kHz
Bit rate	16 bits
Average in-room S/N	75 dB
Impulse length (ms)	500 ms
Spatial resolution	2^{o} resolution ± 60^{o} horizontal azimuth

4.2.5 Binaural Room Scanning (BRS) Playback of Loudspeakers and Rooms

The BRS playback of the different loudspeakers and rooms was done using the calibrated playback system that is detailed in Figure 15. The server computer (HP xw8200) controlled the entire listening test through its custom listening test software. The output from its digital sound card (M-audio Firewire 410) was sent to a master volume control (Lexicon MC-12 Balanced), and then to the eight-channel sound card (RME HDSP 965 with ADI-8 DS interface). The eight-channel sound card was connected to the BRS playback convolution engine. The output from the BRS engine was sent to a headphone amplifier (AKG phone amp V6HP) that powered the listener's headphones (Sennheiser HD 600). The azimuth of the listener's head position was monitored with a Logitech Ultrasonic Head-tracker that transmitted the current angle to the BRS playback engine. Upon receiving the current head position angle, the BRS engine switched to the corresponding set of BRIR filters.

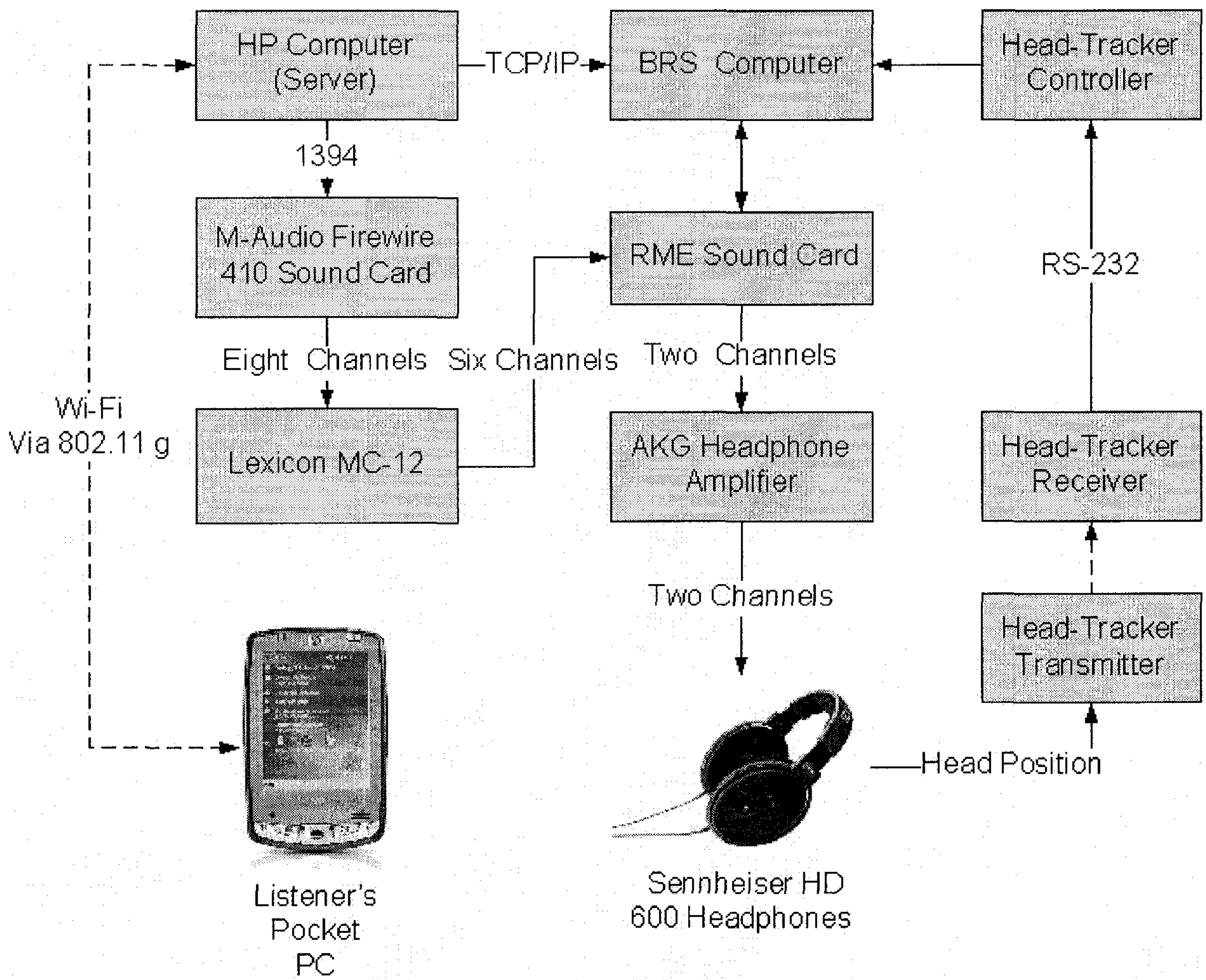


Figure 15. The playback system used for the experiments.

The listener was able to switch at will between the different surround sound loudspeaker systems using a hand-held, wireless Pocket PC (HP model hx2490) that was equipped with a custom listener software application known as the "Pocket Car Evaluator." The Pocket PC sent control commands wirelessly to the server computer, which instantly switched to the set of BRIRs for the specified loudspeaker surround system. BRS playback of the stimuli was conducted in the Harman International Reference Room, one of the four rooms binaurally scanned for this experiment.

To enhance the realism and presence of the BRIR playback, a set of high-quality loudspeakers were set up in the identical ITU-R configuration used for the BRIR measurements (see Figure 16)[14]. The listener sat in the same 'sweet-spot' where the binaural manikin was placed for the BRS measurements. The average playback level of the program selections at the listeners blocked meatus was a comfortable 76 dBA (slow). The playback level remained fixed for the duration of the experiment.

[14] While Figure 16 is an accurate depiction of the loudspeaker configuration, the acoustical treatment is not representative of the room during the BRS measurement.

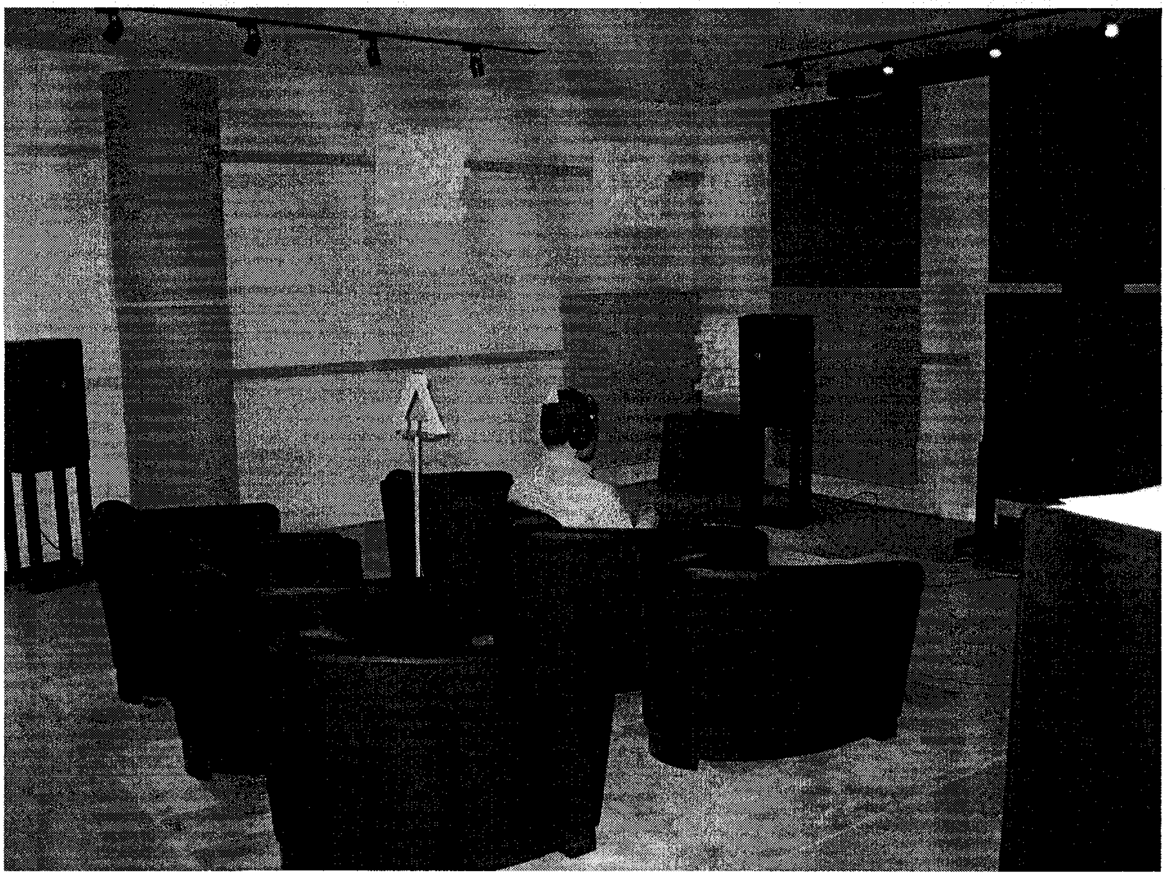


Figure 16. A listener in the Harman Reference Room using the BRS playback system.

4.2.6 Listener's Task

Each listening session began with a series of six pre-trials where the listener made a simple preference choice between different pairs of surround sound loudspeakers for one program selection (LL). Using a Pocket PC equipped with the Pocket Car Evaluator software, listeners could switch at will between the stimuli and enter their responses (see upper diagram within Appendix D). Once the final ratings were entered and stored in the database, the next trial automatically began.

Listeners generally completed all six preference choice pre-trials within 5-10 min and then began their multiple-comparison preference rating task. Listeners were required to complete all nine trials and enter their ratings using the GUI (see lower diagram within Appendix D). The software automatically checked for tied ratings, forcing the listeners to make a preference choice between the loudspeakers. Typically, the entire listening session took about 25-30 min to complete. Listeners were only allowed to participate in a maximum of two sessions per day, with a break between the morning and afternoon sessions.

4.3 Experimental Results: Preference Rating Test

The analysis of data from the experiment described in the previous section is reported in this section.

4.3.1 Statistical Analysis

A repeated measures analysis of variance for a 2 x 4 x 4 x 3 x 3 full factorial model was used for analysis of the independent variables: trial order method (two levels), loudspeakers (four levels), rooms (four levels), programs (three levels), and observation (three levels). The dependent variable was preference rating. An alpha level of 0.05 was used for all statistical tests.

4.3.2 Main Effects: Room

The analysis of variance presented in Table 7 indicated that room was the main effect; F (3, 78) = 3.89, p = 0.01. There were significant interactions between the independent variables of room and loudspeaker; F(9,234) = 3.22, p = 0.001, and also between the independent variables of loudspeaker and program; F(6,156) = 2.15, p = 0.05. The mean preference ratings and the 95% confidence intervals are shown in Figure 17 for the variable room. In order of decreasing preference, the mean ratings for each listening room were: R1 = 5.22, R2 = 5.07, R3 = 5.05, and R4 = 4.76.

Table 7

The Analysis of Variance Table for the Preference Rating Experiment

Source	*df*	*MS*	*F*	*p*	Lambda	Power
Method (M)	1	12.146	.109	.744	.109	.061
Subject group (S)	26	111.851				
Room (R)	3	37.841	3.890	.0120*	11.671	.813
R x M	3	20.039	2.060	.1123	6.181	.500
R x S	78	9.727				
Loudspeaker (L)	3	3.632	.333	.8018	.998	.111
L x M	3	7.449	.682	.5656	2.047	.184
Lx S	78	10.919				
Program (P)	2	7.115	1.041	.3602	2.083	.214
P x M	3	14.458	2.116	.1308	4.232	.403
P x S	52	6.832				
R x L	9	16.237	3.219	.0011*	28.968	.983
R x L x S	234	5.045				
R X P	6	2.517	.795	.5754	4.768	.303
R x P x S	156	3.167				
L x P	6	5.624	2.149	.050*	12.892	.757
L x P x S	156	2.617				

Note. $p^* \leq .05$

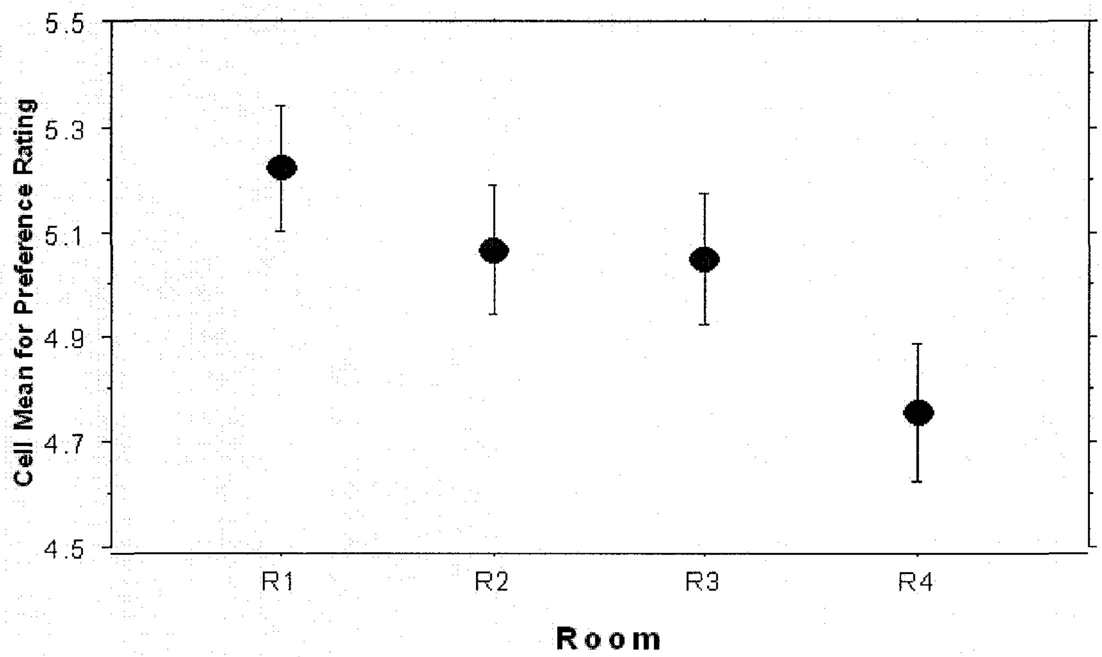


Figure 17. The mean preference ratings and the 95% confidence intervals for room.

A Scheffe post-hoc test was performed to determine if the differences in ratings between each pair of rooms were statistically significant (see Table 8). The results of this test indicate that Rooms R1, R2, and R3 were all preferred over Room R4. However, there were no significant preferences among the three rooms. Lastly, there were no statistically significant effects related to the variables loudspeaker or the trial ordering method (successive versus intermixed). This will be discussed further in chapter 5.

Table 8

The Scheffe Post-Hoc Test Results for the Variable Room.

Room comparison	Mean difference	Critical difference	*p*
R1,R2	.156	.251	.3872
R1,R2	.171	.251	.3018
R1,R4	..464	.251	< .0001 *
R2,R3	.015	.251	.9987
R2,R4	.308	.251	.0081 *
R3,R4	.293	.251	.0138 *

Note. $p* \leq .05$

4.3.3 Interaction Effects

The following sections describe the interaction effects that were present in this experiment.

4.3.3.1 Loudspeaker and room interaction

The analysis of variance in Table 7 demonstrated that there was a significant interaction between the independent variables loudspeaker and room. The mean preference ratings and 95% confidence intervals for this interaction are plotted in Figure 18 below. The graph in Figure 18 clearly establishes that loudspeaker preference ratings were dependent on the room in which the ratings were given. In Rooms R4 and R2, there were significant differences in the loudspeaker preference ratings, whereas, in Rooms R1 and R3, there were none. Overall, the loudspeakers were rated the lowest in Room R4, and the highest in Room R1. The loudspeakers with the largest variance in preference ratings because of room acoustic variations were L2 and L4. The preference ratings for Loudspeaker L1, and, to a lesser extent, L3, were the least influenced by room acoustic variations.

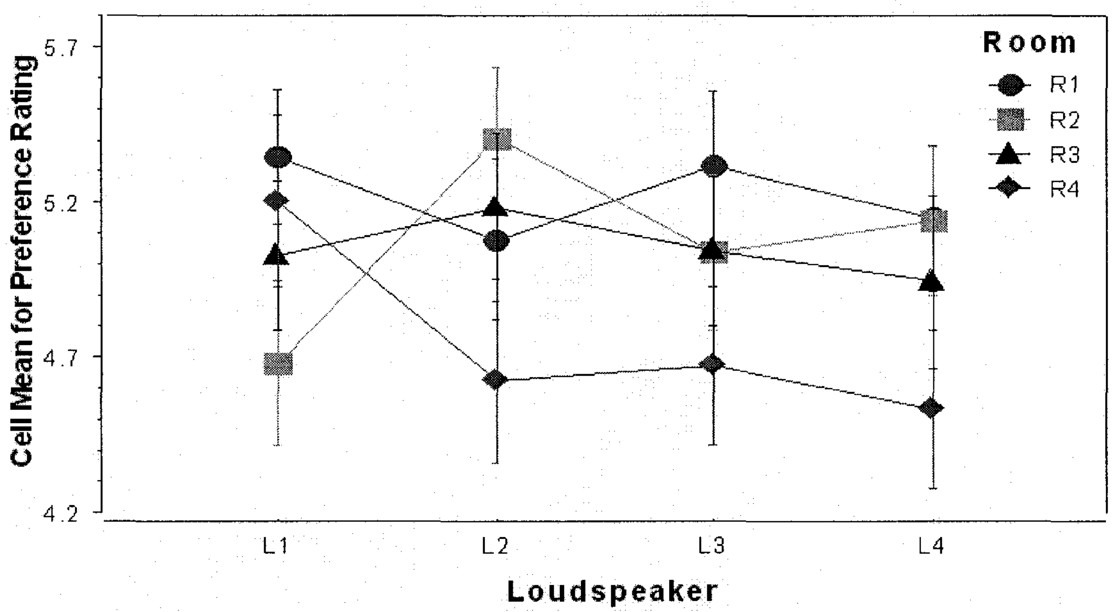


Figure 18. The mean preference ratings and 95% confidence intervals for loudspeaker and room.

4.3.3.2 Program and loudspeaker interaction

The analysis of variance table indicates that a second significant interaction between the independent variables program and loudspeaker took place. Figure 19, which plots the interaction between loudspeaker and program, shows that Program JV (jazz band with female vocalist) and Loudspeaker L4 are largely responsible for this interaction effect. When Loudspeaker L4 was auditioned using Program JV, it received significantly lower ratings.

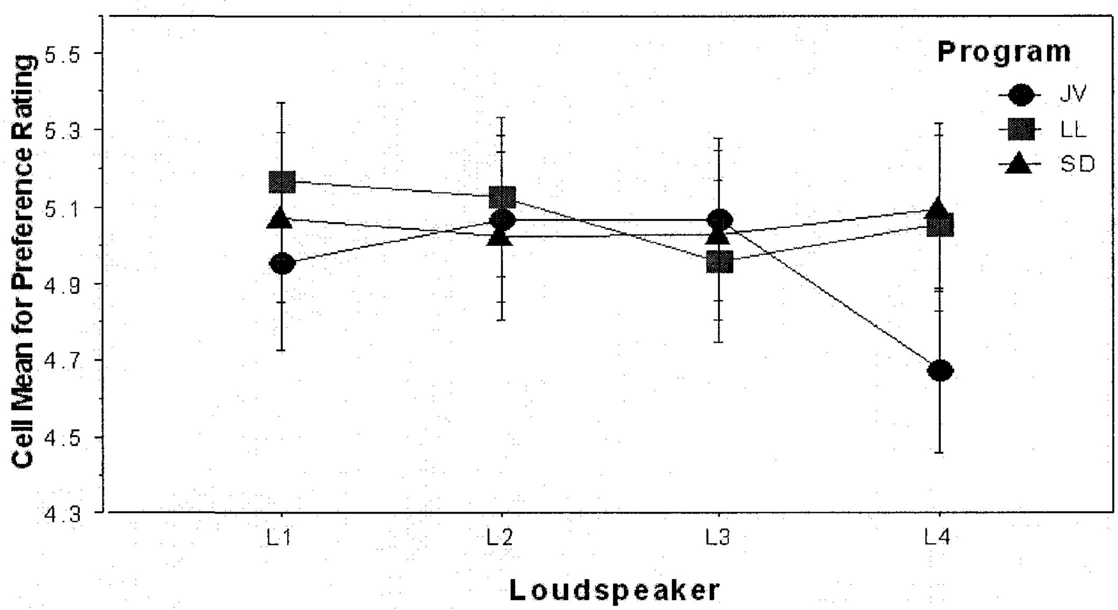


Figure 19. The mean preference ratings and 95% confidence intervals for the interaction between loudspeaker and program.

4.3.4 Effects Related to Prior Experience in Loudspeaker Listening Tests

Previous studies on the subjective evaluation of loudspeakers have shown that there are significant differences in the performances of trained and untrained listeners (Bech, 1992; Gabrielsson, 1979; Gabrielsson et al., 1990; Olive, 2003). Although the resulting overall rank ordering of loudspeaker preferences among trained and untrained listeners have essentially remained the same (Olive, 2003), the general consensus is that trained listeners give more reliable and discriminating loudspeaker preference rating.

The influence of a listener's training and experience was examined post hoc in this study to determine what role, if any, it had on the reliability and discrimination of a listener's loudspeaker preference ratings. A second goal was to determine whether a prior history of participation in controlled listening tests had any effect on loudspeaker preference, room preference, and/or room acoustic adaptation.

Although all of the subjects in this study had been screened to ensure that they possessed normal audiometric hearing, there were significant differences among the listeners regarding their level of experience, in terms of time, participating in formalized loudspeaker tests. Of the total number of listeners who participated in this study, 57% were recent recruits, having had less than 6 months of experience participating in these types of tests. The remaining subjects (43%) possessed 2-15 years of experience in formalized loudspeaker listening tests. Using this criterion, the subjects were divided into two groups based on whether they had two or more years of experience. Although the two groups were not equal in overall size, an equal number of experienced subjects were placed in both successive and intermixed treatment conditions for the purposes of the statistical analysis.

A repeated measures analysis of variance of the experimental data was performed, with experience as the between-subjects factor. Experience was shown to be a main effect; $F\,(1, 26) = 9.27$, $p = 0.005$. The interaction effect between experience and loudspeaker was almost statistically significant; $F\,(3, 78) = 2.37$, $p = 0.08$). This was also true for the interaction between listening experience and room; $F\,(3, 78) = 2.28$, $p = 0.09$). Figure 20 plots the mean preference ratings and upper 95% confidence intervals for experienced listeners (2 or more years of experience) and inexperienced listeners (less than 2 years of experience).

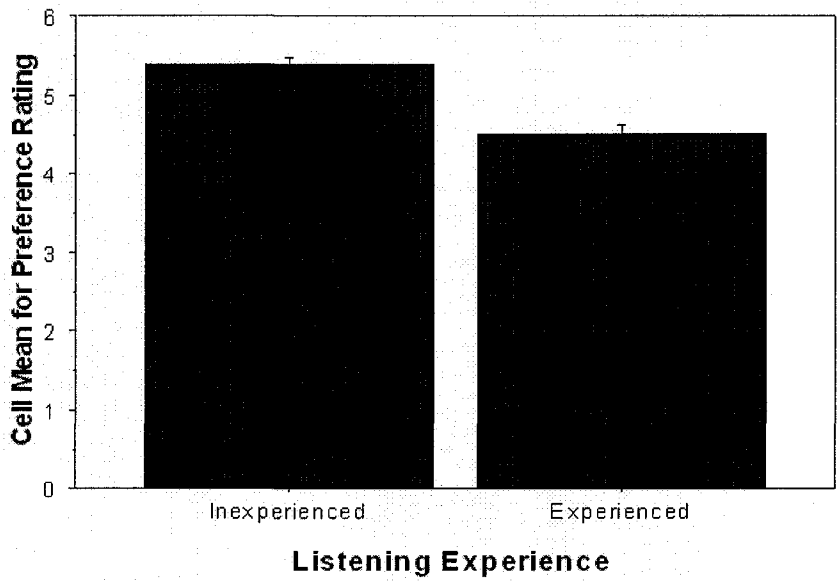


Figure 20. The mean preference ratings and 95% confidence intervals for experienced and inexperienced listeners.

Figure 20 reveals a familiar scaling effect related to listener experience and training as previously observed in Olive (2003). Experienced listeners tended to use the lower part of the preference scale, and spread their ratings further apart on the scale, ostensibly, to register their more critical and discriminating sentiments.

The individual listener's *F* has been utilized by researchers to assess and compare the discrimination and reliability of a listener's loudspeaker preference ratings (Bech, 1992; Olive, 2003). A two-way analysis of variance was performed on each individual listener to calculate their *F* statistic for the independent variables loudspeaker and room. The mean individual listener's *F* statistic for these independent variables is shown in Figure 21 for both experienced and inexperienced listeners.

The large confidence intervals, particularly for the experienced subjects, reflect the small sample size (12 subjects) and range of differences in their performances. However, the graph illustrates an interesting trend. The higher loudspeaker F-statistic for the experienced listeners (versus inexperienced listeners) indicates that they are more discerning and/or reliable in rating loudspeakers than the inexperienced listeners.

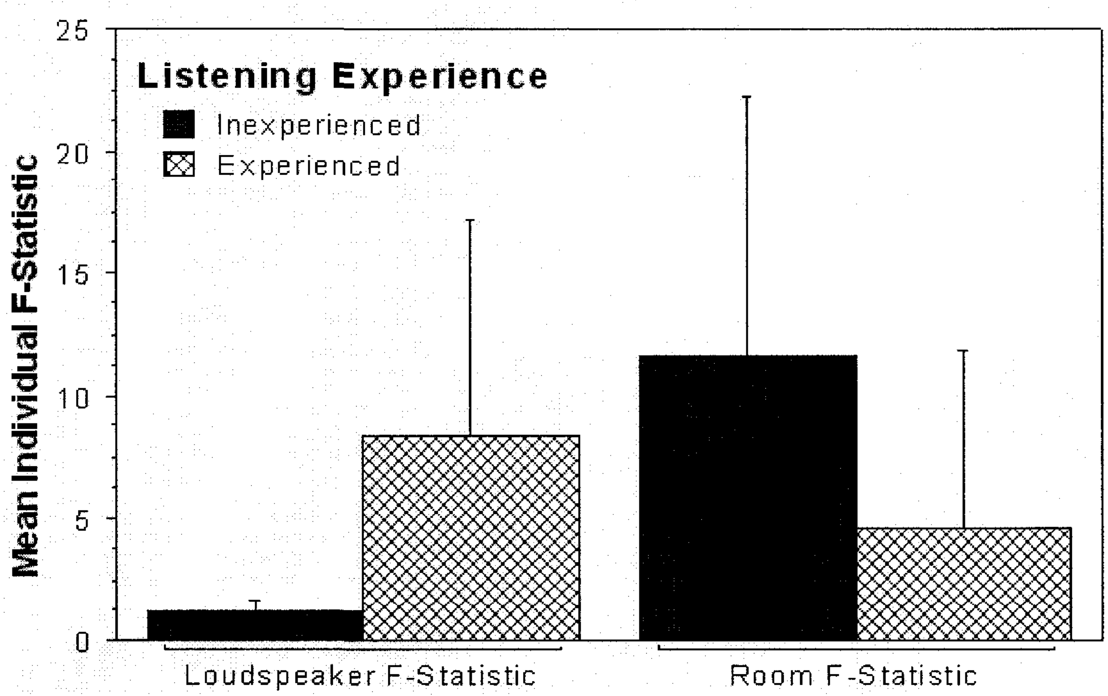


Figure 21. The mean individual listener's F-statistic.

For the variable room, the opposite trend is found; inexperienced listeners are more discerning and/or reliable than the experienced listeners when rating room effects. This contrast seems to suggest that experienced listeners may focus more on loudspeaker differences, whereas inexperienced listeners focus more on room effects. To explore this idea further, the mean loudspeaker and room preference ratings are plotted for both experienced and inexperienced listeners in Figures 22 and 23, respectively.

As noted earlier, the experienced listeners in this study gave lower preference ratings on an absolute scale. However, Figure 22 shows that they were more discriminating amongst the loudspeakers than the inexperienced listeners; particularly as regarded Loudspeaker L4, which they rated lower than the other three loudspeakers. On the other hand, the inexperienced listeners had no significant loudspeaker preferences.

Figure 23 illustrates that the inexperienced listeners were more discriminating among the different rooms than the experienced listeners. They were much more disapproving of the highly reflective room, Room R4, than the experienced listeners. In contrast, the experienced listeners had no strong room preferences, although they tended to give lower ratings than the inexperienced listeners in Room R3, the room with the highest ratio of direct-to-reflected sound.

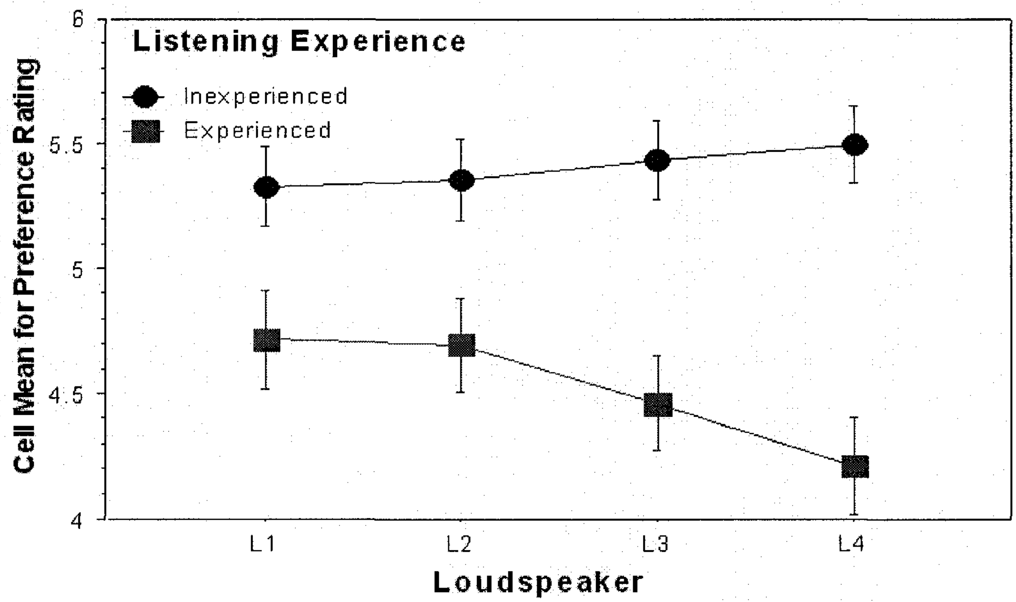


Figure 22. The mean preference ratings and 95% confidence intervals for experienced and inexperienced listeners.

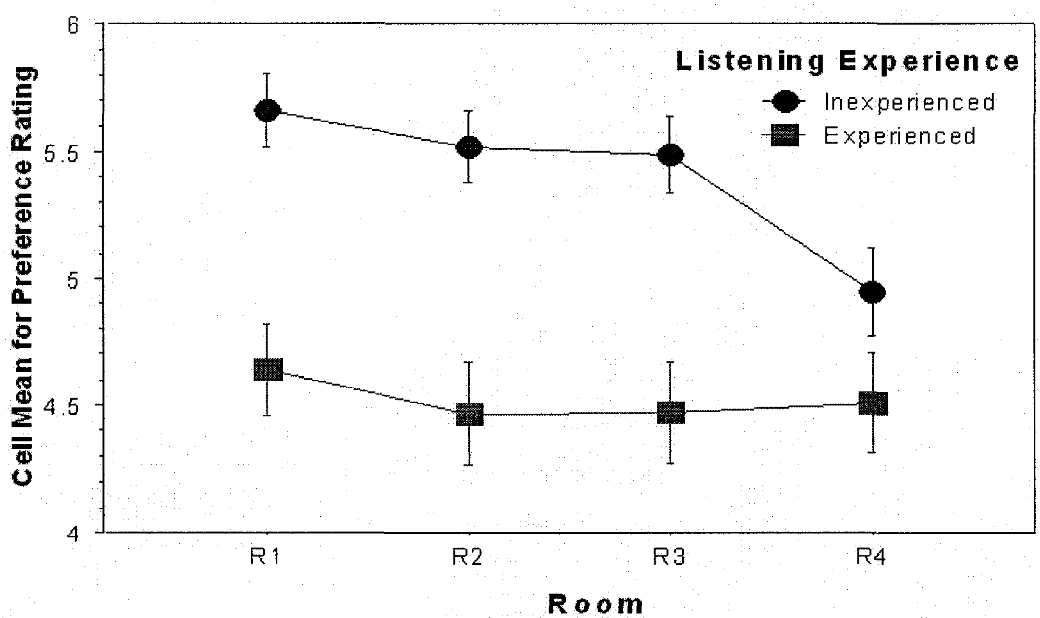


Figure 23. The mean preference ratings and 95% confidence intervals for room; experienced and inexperienced listeners.

Other studies have reported parallel differences in how trained versus untrained listeners weigh the relative importance of timbre and spatial attributes. Ando (1977) found that trained musicians preferred lower levels of reflected sounds than non-musicians. In the same vein, Rumsey and his colleagues (Rumsey, Zielinski, Kaissier, & Bech, 2005a, Rumsey et al., 2005b) found that, in surround sound reproduction listening tests, expert listeners assigned more weight to the timbre and frontal stage spatial fidelity than the untrained listeners, who cared more about the envelopment of the surround sound channels.

Rumsey (1999) also found similar responses from audio experts, who preferred the original stereo material to a five-channel, up-mixed version of the same program. Evidently, untrained listeners were generally more willing to sacrifice some timbral accuracy and frontal stage spatial fidelity in exchange for the added sense of envelopment that the lateral energy from the surround speakers and room reflections provided.

The same underlying principles may well have been at work in this study. If we assumed that the respective loudspeaker and room effects are timbrally, rather than spatially, related this could explain why experienced listeners were more influenced by loudspeaker (i.e. timbre) effects, and inexperienced listeners were more influenced by room (i.e. spatial) effects. The experience gained in years of formal loudspeaker listening tests clearly helped the experienced listeners to focus on the more subtle loudspeaker effects, rather than the more obvious room effects, which the inexperienced listeners gravitated towards in their preference ratings. The dichotomy between experienced and inexperienced listeners in terms of how they weigh the relative importance of timbral and spatial attributes in sound reproduction is a fascinating topic that requires further research.

Chapter 5 Discussion Of Experimental Results

Chapter 5 discusses the experimental results reported in the previous chapter within the context of the three main research questions posed at the onset of this dissertation study (see section 1.1). In order to further validate the conclusions drawn from these results, the findings of the current study are compared to those of related prior studies, as well as to current theories on the perception of loudspeaker-room interactions.

For the sake of clarity, the three main research questions of this dissertation are restated as follows:

1. To what extent are listeners' preferences for multichannel music imagery influenced by variation in loudspeakers, variation in room acoustics, and interactions between these two variables?
2. Can these effects and their interactions on listener preference be explained by acoustical measurements of the loudspeakers and room acoustics?

3. To what extent does room acoustic adaptation diminish the effect of room acoustics on listeners' preference ratings?

The experimental results reported in this dissertation provide the clearest answer to the first research question presented above. The obtained results support the following conclusions:

1. Room acoustics have a significant effect on preferences for multichannel auditory imagery.
2. The specific room acoustic conditions in which listeners are exposed affect the pattern of their loudspeaker preferences.
3. Program selection influences listener loudspeaker preference.
4. Having a prior history of participating in loudspeaker preference tests influences how listeners weigh the relative importance of loudspeaker and room effects when formulating their preference ratings.

Section 5.1 discusses these four findings and shows that these conclusions generally agree with the results of previous related studies as well as with current theories on the perception of loudspeaker-room interactions. Section

5.2 addresses the second research question posed in this dissertation: Can these effects and their interactions on listener preference be explained by acoustical measurements of the loudspeakers and room acoustics? Acoustical measurements of the same loudspeaker taken in each of the four rooms show that there were significant differences among the respective physical patterns of reflections arriving at the listening location. Based on the experimental results, it seems clear that the reflectivity of the different rooms played an important factor in the room and loudspeaker/room interaction effects on preferences observed in this experiment.

Section 5.3 addresses the third research question of this dissertation: To what extent does room acoustic adaptation diminish the effect of room acoustics on listeners' preference ratings? It was hypothesized that the longer exposure times to a particular room acoustic condition across successive treatment conditions would produce room acoustic adaptation, yielding a relative decrease in the effect of the room on the loudspeaker ratings. On the other hand, the intermixed treatment condition was expected to produce less room acoustic adaptation, yielding an increase in the effect of the room on the loudspeaker ratings. The analysis of variance of the experimental results, as discussed in chapter 4, failed to produce conclusive statistical evidence that room acoustic adaptation had occurred. However, further analysis of the successive and intermixed treatment conditions provides evidence that room acoustic adaptation may have occurred in this study.

A previous study by the author provides stronger evidence that room acoustic adaptation does occur under some experimental conditions (Olive et al., 1995). The two studies are compared to determine what differences in their experimental design and methodologies, with particular emphasis on the differences in their selection of the independent variables (loudspeaker, room, program, and use of different trial ordering schemes), accounted for their differing conclusions. Furthermore, a post-hoc analysis of the experimental results presented in this section suggests that, although it was not conclusively established by the overall experimental results, relatively rapid room acoustic adaptation may well have occurred during the experiment reported in this dissertation.

Finally, section 5.4 discusses the importance of the dissertation findings in view of current knowledge in the perception and measurement of acoustical interactions between loudspeaker and rooms. Additionally, this section discusses the broad implications of this study's experimental results for the future design, measurement, evaluation, and specifications of both loudspeakers and listening rooms.

5.1 The Influence of Variations in Room Acoustics, Loudspeakers, Programs, and Listening Experience on Loudspeaker Preference Ratings

The four key findings of this experimental investigation are that: (1) room acoustics have a significant effect on preferences for multichannel auditory imagery; (2) the specific room acoustics to which listeners are exposed affect the pattern of their loudspeaker preferences; (3) program selection influences listener loudspeaker preference; and (4) having a prior history of participating in loudspeaker preference tests influences how listeners weigh the relative importance of loudspeaker and room effects when formulating their preference ratings. Each of these findings is discussed separately in the four sections below.

5.1.1 Room Acoustics Influence Preferences for Multichannel Auditory Imagery

The first key finding of this dissertation study is that room acoustics influence listener preference ratings for multichannel auditory imagery. The room effect itself was largely due to presentations made in one particular room, R4, where listeners, on average, gave lower loudspeaker preference ratings than in the other three rooms.

No known studies have investigated the relationship between the reflectivity of a listening room and its influence on sound quality ratings for multichannel audio reproduction. Two previous related studies (Bech, 1993; Olive et al., 1995) compared loudspeakers in rooms using single loudspeakers reproducing monophonic signals, rather than a multichannel setup as used in this experiment. This raises several concerns regarding whether the results of the previous studies that used mono setups in different rooms can be compared with the multichannel setups used in the current investigation.

There are at least three concerns bearing on the validity of comparisons between the results of the current study to the studies by Bech (1993) and Olive et al. (1995). The first concern is that both mono and stereo loudspeaker tests may be biased towards wider dispersion loudspeakers that produce strong lateral room reflections, a prerequisite for LEV. For multichannel evaluations of loudspeakers, the use of loudspeakers exhibiting wider dispersion may not be necessary because the additional surround loudspeakers can easily replicate the increased LEV produced by wider dispersion loudspeakers in mono. This issue is discussed in more detail in a later section.

The second concern to be considered when comparing the current study to the previous two studies is that the pattern and ratio of the direct-to-reflected sounds in a multichannel loudspeaker setup is likely to be fundamentally different from what would be measured in the same room using a mono or stereo setup. As Toole (2006) has pointed out, domestic rooms do not have isotropic diffuse fields due to their low ceilings and unequal distribution of absorptive, reflective, and scattering objects. However, increasing the number and distribution of loudspeakers around the room creates a greater likelihood that the listener will hear a more "diffuse" sound in the room than what might be heard in the same room with a single loudspeaker. This could change how the room, the loudspeakers, and their interactions affect listeners' sound quality ratings.

The third concern that should be addressed when comparing the current study to the previous two studies is that comparisons across multichannel and mono/stereo loudspeaker evaluations may be invalid because of the perceptual mechanisms (the precedence effect and spectral compensation) that suppress and compensate for room reflections have not yet been investigated within the context of multichannel audio reproduction in reflective rooms. The initial build up, fusion, and breakdown of the precedence effect depends on cognitive factors that are sensitive to the expectations of the listener as well as to abrupt changes in the direction, level, and spectrum of the direct and reflected sounds.

Multichannel recordings are potentially more capable than mono/stereo recordings of producing such disruptions, which may or may not result in a change in the audibility of the reflections from the listening room. Whether or not the effects of the listening room and the off-axis response of the loudspeakers are more audible under mono versus multichannel listening conditions is an important question that needs further research.

The related study by Olive et al. (1995) reported no significant room effects on the preference ratings for the successive treatment condition of the room variable. However, none of the rooms in that study were as large or as reflective as Room R4 used in the current study. In fact, the authors reported that all four of the rooms in the prior study fell within the reverberation times found in domestic homes: RT_{60} of 0.4 s with a standard deviation of 0.1 s from 80 Hz to 4 kHz (Bradley, 1986). The authors concluded: "It is sufficient to say that all of the four rooms meet or closely approach some, or all of the IEC specifications, and that none are atypical of average, real-world, domestic rooms" (Olive et al., 1995, p. 4).

Bech (1993) evaluated four loudspeakers in three different rooms, and found that the loudspeaker fidelity ratings tended to be higher in one of the three rooms. The in-room measurements of the four loudspeakers were not published to provide a possible explanation of the nature of the effect. However, reverberation measurements of the rooms were provided. The preferred room had a relatively constant RT_{60} of 0.4 s between 125 Hz and 8 kHz, which is typical for most domestic listening rooms (Bradley, 1986). The second room had significantly less reverberation (RT_{60} = 0.25 s), whereas the third room had significantly more reverberation, between 160 Hz-500 Hz.

The results of the Bech (1993) study support the findings of the current study; that overly reflective listening rooms produce lower loudspeaker preference ratings. Furthermore, his study also warns that having too few reflections in the room can lead to lower loudspeaker sound quality ratings as well. The benefits of room reflections in listening spaces (established in section 2.3.4) include an increase in loudness, intelligibility, spatial extent (LEV and ASW), and timbral richness.

A common finding in the related previous studies (Bech, 1993; Olive et al., 1995), partially confirmed by this dissertation study, is that listeners prefer loudspeakers in listening rooms that are neither too reflective nor too absorptive. Whether the loudspeakers are auditioned in mono (as in the previous studies), or in multichannel configurations (as in this dissertation study), listeners prefer rooms that approximate the acoustic conditions of an average domestic listening room, as surveyed by Bradley (1986). This makes the design of listening rooms rather straightforward, and potentially very cost-effective. Other than furniture, carpet, and a well-stocked bookshelf or wine rack, no additional esoteric room acoustic treatment or acoustic expertise are necessarily required to set up a listening room that acceptably approximates those typically found in domestic settings.

5.1.2 Room Acoustics Influence Loudspeaker Preferences

The second key finding of the current study is that the specific room acoustic conditions in which listeners are exposed affect the pattern of their loudspeaker preferences. The interaction between loudspeaker and room acoustics is illustrated in Figure 24 (which also appears in chapter 4 as Figure 18).

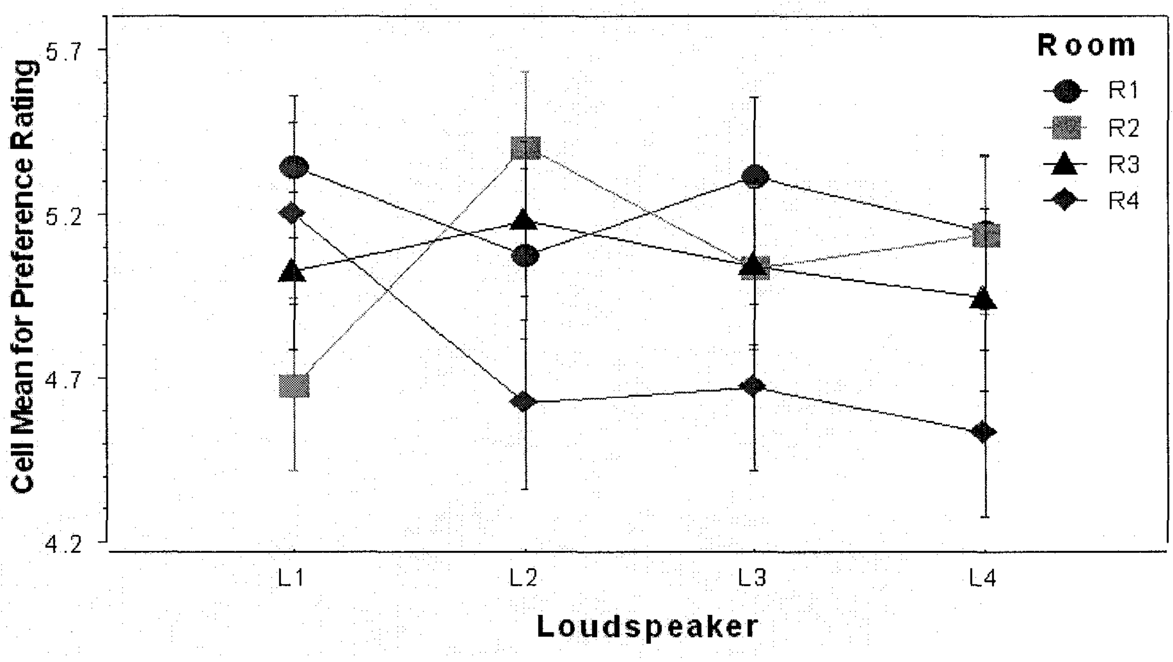


Figure 24. The mean preference ratings and 95% confidence intervals for the loudspeaker-room interaction.

Examination of the mean loudspeaker ratings across the four rooms establishes that the reflectivity of the room is an influential factor that enables listeners to better hear differences in, and to formulate preferences among, the loudspeakers. This conclusion is based on the following observations:

1. The more reflective rooms in this study (R4 and, to a lesser extent, R2) generally produced larger differences in mean loudspeaker ratings than in the less reflective Room R3.
2. The loudspeakers with the poorest measured off-axis responses (L3 and L4; see Figure C1) generally received significantly lower ratings in the most reflective room (R4).

It had been anticipated that the loudspeakers with the poorest off-axis frequency responses (L3 and L4) would be less preferred in the more reflective room (R4) due to the relatively higher proportion of off-axis sounds that arrived at the listener compared to the less reflective room (R3). This hypothesis was generally confirmed because Loudspeakers L3 and L4 received comparatively lower ratings in Room R4 than in the less reflective Room R3. In fact, in the room (R3) that produced the highest ratio of direct-to-reflected sounds at the listening location, there were no significant loudspeaker differences whatsoever: Indicating that listeners were less able to discriminate among the loudspeakers in rooms that were less reflective. The absence of significant loudspeaker differences was particularly pronounced in Room R3, where the listeners received the lowest levels of early lateral reflections than in any of the other rooms, due to the significantly greater distance of its side walls from the loudspeakers and listener.

There were two notable exceptions in this study's results to the general trend that the off-axis loudspeaker colourations were more audible in reflective rooms, as depicted in Figure 24. The first exception occurred in Room R4, where one of the loudspeakers with a smooth off-axis response (L2) was rated lower than expected. The second exception occurred for Loudspeaker L1, which had a similarly good off-axis response, but was rated lower than expected in Room R2. Clearly, the room and loudspeakers (and perhaps the listeners) were interacting in ways that cannot be explained by the anechoic loudspeaker measurements alone. Further analysis of the in-room loudspeaker measurements will hopefully unravel this mystery. The analysis will not be shown here because they fall outside scope of this dissertation, and will be presented in a future paper.

The effect of room reflectivity on loudspeaker preferences as reported in this PhD study confirms findings from a previous study (Olive et al., 1995, p. 8). In that study, the smallest and most acoustically dead room (Room A) produced the smallest differences in loudspeaker preferences. The authors attributed this effect to the lack of room reflections at mid and high frequencies, where the dominant loudspeaker colourations were present. The largest and most reflective room in their study (Room L) was found to be much better for revealing loudspeaker preferences. As stated by the authors (Olive et al., 1995):

> In the case of Room L, speaker-position interactions were not a factor. Therefore, there was likely some acoustical characteristic about the room that made audible differences between the speakers subjectively more apparent. Perhaps, it was a combination of the room's liveness and asymmetry that produced a greater variety of reflections arriving at different angles of incidence and times. This would have certainly placed more extreme demands on the off-axis performances of the loudspeakers, which the measurements indicate were as different as their on-axis ones. (p. 13)

While these comments are in agreement with the findings of the present study, they are admittedly speculative in nature; the authors were not able to confirm the relationship between the off-axis frequency response of the loudspeakers and the rooms' reflectivity because the loudspeakers used in the study were variable in both their direct and reflected sounds.

In conclusion, the results of this dissertation study confirm that the reflectivity of the room is an important influence on listener loudspeaker preferences. Furthermore, this study provides direct evidence that the off-axis frequency response of the loudspeaker becomes an increasingly important factor in listeners' preference ratings as the reflectivity of the listening room increases.

No previous study has been able to confirm this for either mono, stereo, or multichannel setups because they did not experimentally control the direct and reflected sounds radiated by the loudspeaker. Therefore, this dissertation study makes both an original and important contribution to the scientific field of audio.

5.1.3 Program Influences Loudspeaker Preference

Previous studies have reported that program or music selections can significantly influence listeners' sound quality ratings of: (1) different loudspeakers (Gabrielsson & Lindström, 1985; Klippel, 1990; Olive, 2003), (2) different listening rooms (Bech, 1993; Olive et al., 1995), and, even (3) recordings made with different multichannel microphone techniques (Kim et al., 2006).

In the current study, a small, but statistically significant, interaction between program and loudspeaker was shown (see Table 7). The interaction was largely isolated to Program JV (jazz band with female vocalist) and Loudspeaker L4, which had the poorest off-axis frequency response. The finding suggests that the loudspeaker colourations were more audible for this particular program. To confirm this supposition, a separate experiment would be required to measure the intensities of the perceptual attributes of each loudspeaker for each program.

Program interactions can also occur when the program itself contains a distortion that is complementary to a similar, but opposite, one in the loudspeaker. For example, a 'bright' program can make a 'dark' loudspeaker sound neutral.

Interactions between the directivity of the loudspeaker and the "genre" of music have also been reported by Klippel (1990) and Voelker (1985). Klippel found that listeners preferred more directional loudspeakers for speech, and less directional loudspeakers for music. Similar program interactions with room acoustics have been found as well. In a study on the preferred acoustics of professional control rooms, Voelker (1985) found that the more reflective control rooms (RT_{60} = 0.7s) were preferred for chamber and organ music, whereas the less reflective control rooms (RT_{60} = 0.4 s) were preferred for pop and disco music. Finally, a nearly reflection free room (RT_{60} = 0.2 s) was preferred for solo drum music. The notion that certain acoustic spaces are preferred for different genres of music is well established among concert hall designers, and composers have even written music with specific performance spaces in mind.

A more recent study by Weisser and Rindel (2006) also found that the reverberation characteristics preferred by listeners in small rooms are largely dictated by program. Listeners made comparative ratings of a single loudspeaker reproducing monophonic programs in seven different small rooms, via binaural recordings, reproduced over headphones. Subjective ratings included overall sound quality, boominess, and boxiness (also referred to as colouration). The preferred reverberation time for speech was as low as possible, and was 0.3 – 0.5 s for music. Objective metrics based on the rooms' 1/3-octave reverberation times were able to predict the subjective ratings with reasonably good accuracy.

5.1.4 A Listener's Experience Influences How They Weigh the Relative Importance of Loudspeaker and Room Effects in Formulating Preference

Having a prior history of participating in loudspeaker preference tests influences how listeners weigh the relative importance of loudspeaker and room effects when formulating their preference ratings. Listeners with 2 or more years of prior experience in formal loudspeaker listening tests gave lower ratings than the inexperienced listeners with less than 6 months of prior experience.[15] Furthermore, experienced listeners were more discerning of loudspeaker effects, whereas inexperienced listeners were more discerning of room effects in their respective formulations of preference ratings. This effect of experience was

---

[15] As previously discussed in chapter 4, a similar finding was reported in a study by Olive (2003); that experienced listeners use lower preference ratings than inexperienced listeners when evaluating loudspeakers. While the ratings of the experienced listeners were found to be more discriminating and reliable than the inexperienced listeners, the overall rank ordering of loudspeaker preferences was the same.

previously discussed in some detail in section 4.3.4, where several studies were cited that generally confirm these results.

5.1.5 Differences in Loudspeaker Preference Due to Number of Channels

This dissertation study is the first known study to examine listener loudspeaker preferences based on evaluations of multichannel music programs. All of the previous loudspeaker preference studies were based on comparisons of either one (mono) or two (stereo) loudspeakers. An important issue, as discussed earlier in this chapter, is whether the results of the previous studies, done exclusively in either mono or stereo, can be validly generalized to today's multichannel setups? At the heart of this inquiry lies the issue of whether listeners become less discerning and discriminating of loudspeaker sound quality as the number of loudspeakers is increased. If this hypothesis is true, it would imply that, as the number of loudspeakers increase, a lower quality loudspeaker would become more accepted.

Toole (1983, 1986a) has reported that listeners were less discriminating when evaluating loudspeakers in stereo than they were when evaluating loudspeakers in mono, even though the rank ordering of loudspeaker fidelity ratings tended to be the same. Over the years, different explanations for this phenomenon have been proposed that remain largely untested.

A possible explanation posited by the author is that, as the number of loudspeakers increases, so do the spatial and timbral complexities, and variability, of the sound field. The variability comes from the wider dynamic range of directional cues in the recording itself (Choisel & Wickelmaier, 2006; Gustavino & Katz, 2004; Rumsey, 2001), and from more complex acoustical interactions between the room and the multiple sound sources, which are themselves more sensitive to variations in the program and listening position than monophonic evaluations (Toole, 2006; Welti, 2002; Welti & Devantier, 2006). These increasing, added complexities and variations in the sound field require the listener to divide his/her attention between the timbral and spatial attributes of the auditory imagery that they are evaluating. The increased complexity of the listeners' task naturally leads to less discriminating and reliable ratings than would occur during mono evaluation.

Evidence from the current study and previous studies (see section 4.3.4) establishes that experienced listeners tend to weigh the importance of certain timbral and spatial attributes differently than inexperienced listeners. This relative difference may exert the strongest effect on loudspeaker preference ratings when the program is stereo, and an even a greater effect for multichannel programs because they contain a wider variation and range of spatial effects than mono programs. This consequentially creates greater intra-listener variance in the ratings due to potentially larger loudspeaker-program interactions, and more inter-listener variance in the ratings related to experience, when compared to tests that utilize monophonic recordings.

Toole (1983, 1986a, 1986b) has indicated that the amount of stereophonic versus monophonic information in the recording used in stereo loudspeaker evaluations is an important nuisance variable. According to him, loudspeaker discrimination in stereo was vastly improved if the stereo recordings had a strong monophonic component, as commonly found in multi-track recordings made with single microphones (Toole, 1983).

More recently, Toole (F. Toole, personal communication, July 14, 2007) has postulated that the differences in mono-versus-stereo loudspeaker preference ratings may be related to listeners making trade-offs between the loudspeaker's timbral accuracy and its perceived spaciousness. When different loudspeakers are evaluated in mono, listeners may overlook loudspeaker off-axis colourations, within certain limits, if the loudspeaker has wider dispersion (i.e. lower directivity) that produces the stronger lateral reflections necessary for the perception of LEV.

Toole further observed that this spatial advantage becomes less apparent when the same group of loudspeakers are evaluated with stereo recordings that have a strong inter-channel phase difference as is commonly found in recordings made with widely-spaced pairs of microphones. The spatial benefit accrued from wider dispersion that is highly appreciated in mono becomes less apparent in stereo because the strong inter-channel phase differences in the stereo recording produces a similar sense of LEV through all of the loudspeakers under test, regardless of their directivities. At this point, the loudspeaker's timbral accuracy may become a more important determining factor in listeners' loudspeaker preferences (F. Toole, personal communication, July 14, 2007). How this interaction effect operates with multichannel loudspeakers and recordings is largely unknown, and requires more research and experimental verification.

5.2 The Relationship Between the Acoustical Measurements of the Room and Listener Preference Ratings

The acoustical differences among the four listening rooms in terms of their reflective energy are plotted in Figures 25 (a)-(d). The *energy-time curves* (ETC) of each room are displayed. Normally, the ETC is calculated from the impulse response (Heyser, 1971a, 1971b, 1971c; Vanderkooy & Lipshitz, 1990), but here the BRIR (left ear only) is used. Consequently, the ETCs depicted in Figure 25 contain the frequency response and directional properties of the binaural manikin.

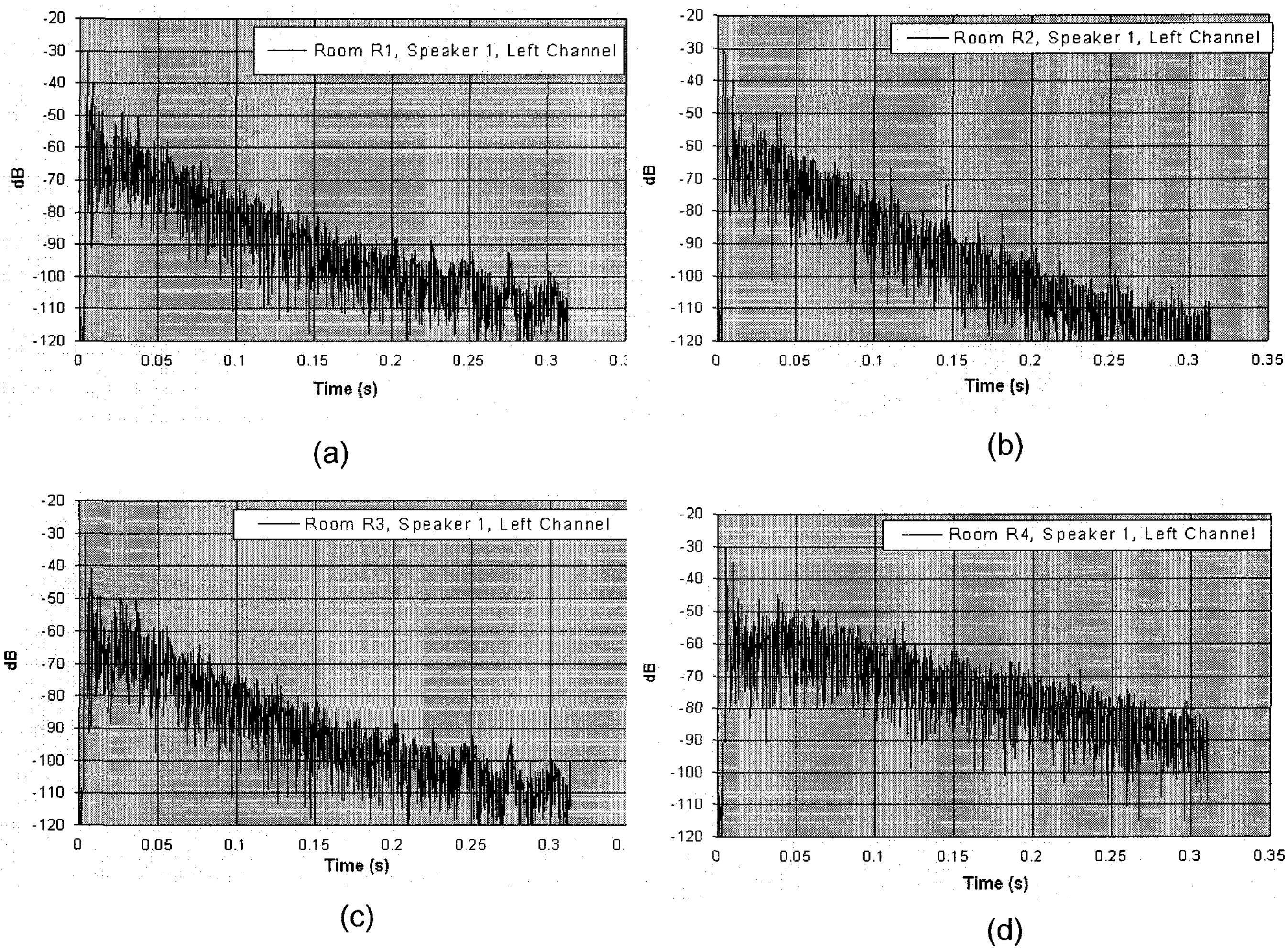


Figure 25. The energy-time curves (ETC) for rooms: (a) R1, (b) R2, (c) R3, and (d) R4. The ETC is calculated from the BRIR (left ear only) measured at the listening location using the left front loudspeaker.

The BRIR was measured at the listening location using Loudspeaker L1, located in the left front position. Only the first 300 ms of the ETCs are shown because, at that point, most of the energy in the room has decayed. Looking at

Figure 25, the energy decay patterns were very similar in Rooms R1, R2, and R3, with some small differences in their finer details.

The general similarity of the ETCs for Rooms R1-R3 may partially explain why the loudspeaker preference ratings did not vary significantly among these three rooms, even though the differences in their reverberation times were outside the mid-frequency RT_{60} difference limens (0.042 s ± 0.015 s) for small rooms (Niaounakis & Davies, 2002). This confirms the results of a previous study (Olive et al., 1995) that found that the reverberation time of a listening room is not necessarily a reliable predictor of how the room will affect the sound quality of loudspeakers.

On the other hand, Room R4 has significantly higher levels of late reflected energy beginning at 60 ms that are 15-20 dB higher than the levels of later arriving reflections that were found in the other three rooms. These late-arriving, higher-level reflections likely contributed to the lower preference ratings observed in Room R4.

A formal scientific investigation into the underlying perceptual attributes of the four different listening rooms utilized in this dissertation study has not yet been initiated. However, informal comments solicited from listeners revealed that Room R4 sounded "excessively reverberant, very reflective, overly live, and too bright, with emphasis on the midrange frequencies." These comments were related to the clarity and spectral balance of the auditory imagery, and may have provided the listeners' justifications for the lower preference ratings that they gave in this room. Interestingly, there was a general absence of listener criticisms related to the early reflections. For example, adjectives associated with a description of auditory image shift and/or a wider ASW were not given by the subjects. The lack of criticisms regarding image shift and/or a wider ASW confirm that the levels of the lateral reflections in these rooms were below the threshold at which these reported effects occur (Okano et al., 1998; Toole, 2006).

Later arrival lateral reflections (>80 ms) produce higher LEV values (Soulodre et al., 2003), a desirable attribute for concert halls (Bradley & Soulodre, 1995, 1996). However, in listening rooms, these late arriving, temporally dense reflections may produce a loss of clarity for discontinuous signals, such as speech, due to increased temporal masking. It is important not to attenuate early reflections arriving before 50 ms because they have been shown to actually improve speech intelligibility, particularly when listeners are outside the main axis of radiation of the loudspeakers, or at listening distances where the direct sound is weak (Bradley et al., 2003; Nakajima & Ando, 1991; Soulodre et al., 1989).

To verify the exact physical cause of the lower ratings in Room R4, additional experiments would be required that systematically manipulated the reflective properties within the room. In addition to the ETC, the author is considering other acoustical measurements of the room and loudspeaker that might further explain the influence of the loudspeaker-room interactions on listener preference.

5.3 Did Room Acoustic Adaptation Occur in This Study?

It had been hypothesized that successive treatment conditions would produce more room acoustic adaptation than the intermixed treatment conditions due to prolonged exposure to the same room acoustics. In this study, the differential exposure to the varying room acoustics administered through the successive and intermixed treatment conditions failed to produce a statistically significant difference in listeners' loudspeaker preference ratings due to the effect of room. Stated differently, the interaction between method (successive and intermixed treatment conditions) and room was not statistically significant; $F\,(3, 78) = 2.06$, $p = 0.11$. Therefore, no direct statistical evidence was produced that proved that room acoustic adaptation occurred.

The fact that this study did not produce statistical evidence to support this hypothesis does not necessarily mean that room acoustic adaptation did *not* occur. One possibility is that adaptation may have occurred for the *both* the successive and intermixed treatment conditions if the time course over which the intermixed treatments were evaluated was simply too long for listeners to avoid adaptation to the room acoustics; this potentially explains why the results for both treatment conditions are the same.

In the previous study (Olive et al., 1995), evidence for room acoustic adaptation was based on a different test paradigm than the one utilized in the current study. In the previous study, the breakdown in room acoustic adaptation occurred when listeners made contemporaneous comparative ratings among the four rooms via a binaural playback device, which produced significant room effects that had not occurred in the successive treatment conditions of the listening rooms. The contemporaneous comparisons of the rooms meant that listeners literally had just a few seconds to adapt to the room acoustics before they changed. This test paradigm would have enhanced the audibility of room acoustic differences in the current study by making room acoustics the changing feature in the listening trial, while the loudspeaker remained constant.[16] The time course employed by this paradigm was much shorter than the 1-2 min time course used for evaluating the intermixed treatment conditions in this dissertation study. These observations give further credence to the notion that room acoustic

---

[16] By experimentally manipulating one component of an audio signal over a short time course, the features in that component become more audibly prominent as compared to when they are time-invariant (Moore, 1996, p. 200).

exposure times must be significantly reduced in order to produce a breakdown in room acoustic adaptation.

Additional evidence exists that suggests that room acoustic adaptation may have been at work in the current study. The mean preference ratings and 95% confidence intervals for the four listening rooms for both the successive and intermixed treatment conditions are plotted in Figure 26. Although the interaction

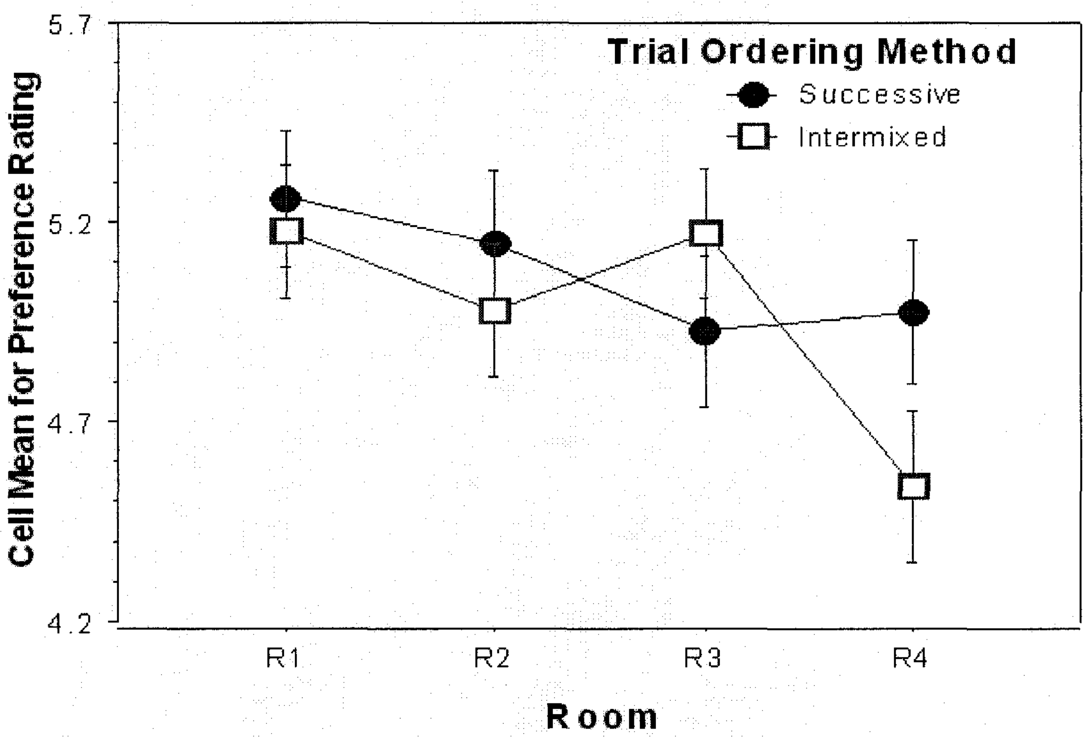


Figure 26. The mean preference ratings and 95% confidence intervals for the four listening rooms for the successive and intermixed treatment conditions.

is not statistically significant; $F(3, 78) = 2.06$, $p = 0.11$, it is considered for the purpose of determining whether further investigation is warranted. There are two notable differences between the successive and intermixed treatment conditions in the pattern of preferences related to the room acoustics. First, the preference ratings for the intermixed treatment conditions were lower than the successive treatment conditions for three of the four rooms. Secondly, the intermixed treatment condition was more sensitive to the effects of the room acoustic variations than the successive treatment condition. The effect was most pronounced in Room R4, where the intermixed treatment condition produced significantly lower preference ratings than were measured in the other three rooms, and significantly lower ratings than measured in the same room using a successive treatment condition. This provides evidence that the intermixed treatment conditions may have produced less room acoustic adaptation than the successive treatment conditions.

The efficiency and speed with which listeners adapt to different acoustic environments may seem surprising at first glance, until one considers that humans frequently adapt to room acoustics, often without being aware of it. Every day we walk through different acoustical spaces, at home or at work, where the acoustical properties often vary significantly. Yet, we hardly notice the effect of the different rooms or environments on the timbre, spatial properties, or intelligibility of the people we converse with. The timbral and localization constancy of sound sources within these different reflective spaces indicates that powerful, room-compensating, perceptual processes are at work. It is only when we move into in a large reflective space, such as a concert hall, that we become aware of how the room changes the timbre, intelligibility, and spatial aspects of the sounds around us. The fact that we can adapt to and actually prefer these effects in concert halls suggests that our learning, experience, and expectations all play an important role in how we adapt to and prefer certain room acoustics.

5.4 The Importance of the Findings of This Dissertation

The findings reported in this dissertation have broad implications on the design, measurement, and evaluation of both the loudspeakers and listening rooms used for multichannel audio reproduction. These findings also have important implications for the design of listening tests and the interpretation of their results. They are summarized as follows:

1. Listening room designs that produce high levels of late reflected sounds should be avoided because they produce lower preference ratings for multichannel music reproductions.
2. The off-axis frequency response of the loudspeaker has a significant influence on loudspeaker preference, and this becomes increasingly

important as the reflectivity of the listening room increases. The design, measurement, evaluation, and specification of loudspeakers should carefully account for the off-axis performance of the loudspeaker.

3. The program can significantly influence the preference rating of the loudspeakers; therefore, this variable needs to be carefully controlled when evaluating loudspeakers.

4. Listeners with a prior history of participation in formal listening tests are more discerning of loudspeaker differences, whereas less experienced listeners are more influenced by room effects. This finding has important implications in the selection and training of listeners used in both scientific experiments and consumer audio product marketing research.

The majority of consumer and professional listening rooms will generally have lower levels of late arrival reflections than the levels found in Room R4, where listeners gave the lowest preference ratings. R4 had only 17% of its total surface area covered with sound absorptive material (see section 4.2.1.1), and very few sound scattering objects in the room (only a chair and a low cabinet). The current study, interpreted together with the prior relevant studies, indicates that, provided that the listening room is not overly reflective or too absorptive, listeners will generally be satisfied with its effect on the quality of audio reproductions.

Three of the four rectangular rooms used in this study had acoustics that approximated a range of typical domestic listening rooms, and produced uniform, and higher, preference ratings than Room R4. This finding suggests that approximating the acoustics of an average domestic room is a good design target and bestows an additional benefit because it indicates that no special acoustical treatment or expertise is required to produce a room that is suitable for high-quality multichannel audio reproduction.

The results of this study do not support the notion that the early reflections (arriving within 15 ms after the direct sound) should be avoided, as advocated in several current professional audio recommendations (ITU-R BS.775-1, 1994; Producers & Engineers Wing Surround Sound Recommendation Committee, 2004). In fact, the smallest and most preferred room, Room R1, had the strongest and earliest first arrival side wall and ceiling reflections. Therefore, it would seem that the results of the current study have the potential to contradict this popular view, although further investigation into the respective rooms' acoustical measurements would be required before this could be confirmed.

The importance of the off-axis frequency response of a loudspeaker has been generally well accepted among loudspeaker researchers. However, this is the first known study that has proven, beyond doubt, that it influences listeners' loudspeaker preferences. Unlike previous studies, this study carefully controlled the direct sound while manipulating only the off-axis response above 300 Hz. Loudspeaker positional biases were not a factor since each loudspeaker was evaluated in the same position using a BRS measurement and playback system.

Given the importance of the off-axis performance of the loudspeaker demonstrated in this study, loudspeaker manufacturers, audio magazine reviewers, consumer and professional audiophiles, and the loudspeaker standards organizations need to reconsider its importance in the design, measurement, evaluation, and specifications of loudspeakers. The off-axis performance of a loudspeaker can only be assessed by analysis of a set of comprehensive anechoic measurements with fine spatial and frequency resolution, such as those shown in Appendix C. Loudspeaker measurements based on a single curve done in the listening room, or even in a reverberation chamber, will not suffice.

For loudspeaker manufacturers, this finding underlines the importance of subjectively evaluating loudspeakers in rooms that are sufficiently reflective, so that the audible nature of their timbre colourations can be accurately assessed under the most sensitive listening conditions. In this way, the loudspeaker manufacturer gains some degree of assurance that the loudspeaker will be well received by consumers, regardless of the degree of reflectivity in their listening rooms.

The influence of program on loudspeaker preferences is well documented in the scientific audio literature. While more research is needed to explain the exact nature of the program-loudspeaker interaction in this study, the results emphasize the value of carefully selecting programs, as well as the importance of using several different programs to balance any potential bias effects on listener loudspeaker preferences. The selection of programs for evaluation of multichannel audio systems is particularly challenging because there are considerable variations among the recordings in their use of the front, center, and surround channels to create the auditory imagery (Rumsey, 2001).

Finally, the effect of listening experience on the results of this study confirms the need to carefully consider the selection and training of listeners when designing listening experiments. Although experienced listeners are highly desirable due to their more reliable and discerning sound quality ratings, the exclusion of less experienced listeners in a study may reduce the extent to which the results can be generalized to populations of inexperienced listeners. Therefore, the effect of listening experience has important implications on the listening test design and interpretation of the results.

Chapter 6 Conclusions

This final chapter summarizes the conclusions of this experimental investigation into the influence of room acoustic variations, and room acoustic adaptation, on listeners' loudspeaker preferences. This chapter is divided into five sections. Section 6.1 surveys the aspects of the literature review that motivated this research. Section 6.2 is a review of the BRS measurement and reproduction system, without which this research would have been possible. The conclusions of the experimental results are summarized in section 6.3. Section 6.4 discusses the limitations of the experimental results in terms of their generalizability to real-world, multichannel audio setups and listening rooms. Section 6.5 identities the areas for future research that can potentially address some of the most important questions raised by this investigation.

6.1 Summary of the Literature Review

The central purpose of the literature review is to identify the psychophysical variables that influence audio reproduction in rooms; specifically those relating to loudspeakers, listening rooms, and acoustical interactions between them. A summary of the literature has been previously provided in section 2.5, so that only the central themes that motivated this research are highlighted.

A review of several previous listening test studies where the nuisance variables were well controlled established that loudspeakers affect the quality of sound reproduction (see section 2.2). Other studies have further confirmed that the overall sound quality of the loudspeaker is related to nine perceptual attributes that fall under the categories of timbre, spatial, loudness, and nonlinear distortion (see section 2.2.4). Researchers are in agreement that all of the perceptual attributes (except nonlinear distortion) relate to the frequency response and directivity of the loudspeaker. In sharp contrast, little consensus exists about how to best measure the frequency response of the loudspeaker in a way that most accurately characterizes its perceived sound quality in a listening room (see section 2.2.1).

To directly answer this question, the author conducted a previous study (see section 2.2.2.4) that compared the accuracy of different models designed to predict loudspeaker preference ratings based on set of frequency response measurements. The models were compared using the three most common types of loudspeaker frequency response measurements believed to most accurately represent its sound quality: (1) in-room frequency responses; (2) sound-power frequency responses; and (3) a set of comprehensive, spatially-averaged, anechoic measurements that allow separate measures of the direct and reflected sounds radiated by the loudspeaker.

In his previous study, the author performed a series of well-controlled loudspeaker listening tests to determine the loudspeaker preference ratings. These observed ratings were then compared to the predicted ratings of the three models, each using a different type of measurement to arrive at their respective predictions. The model based on anechoic measurements produced the most accurate predictions of loudspeaker preference ($r = 0.86$). Models based on the in-room measurements provided significantly less accurate predictions than the anechoic measurements, and the sound-power measurements provided the least accurate predictions of the observed preference ratings. Both the in-room and sound-power measurements integrated the loudspeaker's direct and reflected sounds into a single curve, which could not adequately represent how the listener perceives the sounds from the loudspeaker.

In conclusion, the above study provides evidence that accurate predictions of listener loudspeaker preferences requires separate measures of the direct, early, and late reflected sounds radiated by the loudspeaker Still, the perceptual importance of the loudspeaker's off-axis frequency response can only be inferred because its radiated direct and reflected sounds were not isolated, or independently evaluated, in a controlled way. Furthermore, the listening tests were completed in one listening room only, clearly limiting the extent to which these results can be generalized to other rooms.

To adequately address these issues, a more direct way of validating the importance of the loudspeaker's off-axis performance in different rooms is required. The practical and methodological challenges of testing loudspeaker-room interactions using conventional in situ listening test methods are, practically-speaking, insurmountable (see section 3.0), and this provides the most probable explanation for why such an experiment has not yet been done. Yet, such an experiment would greatly improve our scientific understanding of the perception of acoustical interactions between loudspeakers and rooms. The fact that such a study has not been done is a key motivating factor behind the experimental investigation of this dissertation.

Parallel studies in concert halls have demonstrated the perceptual importance of the direct, early, and late reflected sounds. The physical properties of these three components affect our perception of a sound source's timbre, its localization (direction, distance, spatial extent), and its intelligibility. This suggests that the physical properties of the off-axis sounds of the loudspeaker, which primarily arrive at the listener as reflected sounds, are perceptually important to its perceived sound quality.

Much research has focussed on why listeners are able to accurately identify the timbre, direction, distance, and intelligibility of a sound source, even in the presence of interfering reflections (see section 2.3.4). Listeners are aided by two auditory perceptual mechanisms (the precedence effect and spectral compensation) that both suppress, and compensate for, the interfering effects of room reflections. Cognitive processes occurring in the central auditory stage of the brain use auditory information and cues from other sensory modalities to decide which reflections should be suppressed, and which ones should be compensated for. If the reflection patterns are redundant, time invariant, and plausible, the reflections will generally be fused with the direct sound.

Sudden physical changes to the pattern of direct and reflected sounds will cause a breakdown in the fusion of the direct and reflected sounds. Whether or not this breakdown could be triggered by the manner in which the signals are mixed in a multichannel recording is not known. Studies on the precedence effect indicate that the suppression of a reflection decreases as the spectra of the direct and reflected sounds become more dissimilar. The off-axis frequency response of the loudspeaker and the frequency-dependent absorption in the room can both significantly alter the spectra of the reflected sounds, potentially reducing the effectiveness of the precedence effect, and the extent to which listeners adapt to the room acoustics.

Both the precedence effect and spectral compensation have not been studied within the context of multichannel audio reproduction in listening rooms, which raises some questions regarding their operation under these conditions. This also motivated the author to embark on the current investigation.

The literature review in chapter 2 established that the listening room is a major source of variability in the quality of reproduced sound. Many studies have presented physical evidence of the acoustical interactions between the loudspeaker and the room, but few have experimentally validated their perception (see section 2.3).

At low frequencies (below approximately 300 Hz), the modal behaviour of the room dominates what the listeners hear (see section 2.3.1). Additionally, changes in low-frequency gain due to solid angle effects and room boundary effects occur that are related to the position of the loudspeaker in the room. Single or multiple subwoofers can be equalized and judiciously placed to reduce the typical ± 15 dB low-frequency variations that have been reported (see sections 2.3.2 and 2.3.3).

Above 300 Hz, the physical properties of the room reflections (their level, arrival time, direction, and spectrum) relative to the direct sound significantly influence the listener's perception of the auditory imagery (see section 2.3.4). A thorough review of the room reflection literature relating to domestic listening rooms can be summarized in a few salient points:

1. Most of the reflections in domestic rooms are generally below the detection threshold at which negative effects occur (e.g. image shift, secondary images or echoes, comb-filtering, and/or poor intelligibility). In other words, the precedence effect and spectral compensation are able to operate effectively in these rooms, at least for mono and stereo loudspeaker setups.
2. Certain reflections, particularly lateral ones, produce subjective effects that are considered positive and lead to improved intelligibility, loudness, timbre richness, and enhanced spatial impression (ASW and LEV). Spatially-deprived stereophonic setups may benefit from these additional lateral room reflections. However, the preferred amount of reflections seems to depend on the genre of the music.
3. Multichannel audio provides more flexible control of the spatial and timbral qualities of the reproduced auditory imagery due to the greater spatial distribution of loudspeakers around the listener; this arguably relaxes the

various requirements for specific room acoustics and loudspeaker directivities for different genres of music.

There have been very few systematic studies of the perceptual effects of loudspeaker-room interactions. Several studies have shown that the loudspeaker position in the room can significantly affect its perceived sound quality due to effects above and below the room's transition frequency. Only three studies have evaluated how the loudspeaker preference ratings of loudspeakers varied across different rooms; but two of those studies were limited to monophonic comparisons of single loudspeakers, and the third was done in stereo.

In conclusion, many studies have provided ample physical evidence of acoustic interactions between the loudspeaker and the room, but only a few studies, limited to mono or stereo reproductions, have experimentally validated how these interactions affect the quality of reproduced sound. The lack of experimental data on the perception of loudspeaker-room interactions related to multichannel audio reproduction makes this study an original, and long overdue, contribution to the scientific field of audio research.

An important research objective of this dissertation is to examine the role of room acoustic adaptation in the perception of room acoustic variations, and their subsequent influence on loudspeaker preference. Only one previous study (by the author and several colleagues) on room acoustic adaptation is related to the current one (see section 2.4.2). In that study both in situ and binaural methods were used to evaluate three different loudspeakers (in mono), located in three different positions within four different rooms. When listeners evaluated the different loudspeakers in successive blocks of rooms, the room variable had no effect on the loudspeaker preferences ratings, while the loudspeaker effects were highly significant. The lack of room effect was thought to have been caused by room acoustic adaptation.

To test this further, the authors used a binaural record-playback device that allowed listeners to make contemporaneous, comparative ratings of the same loudspeaker among the four rooms, without a time gap in between the comparisons. By manipulating the context in which the different loudspeakers or rooms were compared and evaluated, the loudspeaker and room effects on listener preferences were reversed in that study. While the study provided the first experimental evidence that room acoustic adaptation occurs, the questions left unanswered by this study provided further motivation for the current dissertation study.

6.2 Summary of the Binaural Room Scanning (BRS) Method

Research in the perception of loudspeaker-room interactions is challenged by several practical and methodological constraints (see section 3.0). Controlled, double-blind, real-time, multiple comparisons among the salient loudspeaker and room parameters are not practical, or even possible, using conventional in situ tests. For this reason, the author chose an alternative method, known as BRS, for this dissertation's experimental investigation. BRS allows the different loudspeaker setups and rooms to be adjusted in a very controlled way, which are captured and stored as a set of BRIRs. The multichannel music is subsequently convolved in real-time with the stored BRIR filters, and then reproduced through high-quality headphones equipped with a low–latency head-tracker. This maximizes the similarity between the reproduced headphone signals, measured at the listeners' ears during BRS playback, and the signals measured during an in situ listening test, where the listener was allowed to move their head (see sections 3.1 to 3.4).

The BRS measurement and playback system contains several sources of errors that have been categorized in this dissertation as either individualized or non-individualized, and further, as either directionally dependent or directionally independent (see section 3.5). To a significant degree, many of these errors can be minimized through proper calibration. A calibration procedure was described that aims to reproduce the same signals at the listeners' ears during playback through headphones as would be measured in situ at the listeners' ears (see sections 3.6 to 3.8).

An important question is whether an individualized calibration is required for loudspeaker preference tests conducted in virtual rooms. Prior to beginning the main experiment that forms the basis of the study reported in this dissertation, the literature was surveyed (see sections 3.4 and 3.5) and a validation test of the BRS system was conducted in order to answer this key question (see section 3.9). A total of 8 trained listeners gave loudspeaker preferences in situ, as well as through headphones reproducing the captured BRIRs of the loudspeakers, both situated in the same reflective room. The loudspeaker ratings were virtually identical for the in situ and BRS system playback. On the basis of those results, the author concluded that the BRS system was a sufficiently accurate and reliable vehicle for the subsequent investigation into room acoustic adaptation that forms the basis for this PhD study.

6.3 Main Conclusions of the Experiment on the Interactions Between Loudspeakers and Listening Rooms

A listening experiment was designed to answer several research questions relating to the effects of loudspeaker-room interactions on listener preference ratings made in response to auditory imagery associated with multichannel music reproduction, and the extent to which listeners adapt to these effects. The three main research questions were:

1. To what extent are listeners' preferences for multichannel music imagery influenced by different loudspeakers, room acoustics, and the interaction between loudspeakers and room acoustics?
2. Can the effects and interactions of loudspeakers and room acoustics on listeners' preferences be explained by acoustical measurements of the loudspeakers and room acoustics?
3. To what extent has room acoustic adaptation diminished the effect of room acoustics on listeners' preference ratings?

Listener preference ratings were measured for auditory imagery associated with three different five-channel music selections reproduced through four different 3/2 surround loudspeaker systems set up in four different listening rooms. The loudspeakers differed only in their off-axis frequency responses above 300 Hz, meaning that the loudspeaker variations were only present in the reflected sounds received at the listeners' ears; the direct sound remained constant across all loudspeaker and room conditions.

Listener preferences among the four different five-channel loudspeaker systems were observed in four different virtual listening rooms via a binaural room scanning (BRS) measurement and playback system. To study room acoustic adaptation, two groups of listeners evaluated identical stimuli according to two different trial blocking schemes, termed either the successive or the intermixed treatment. In the successive treatment condition, all of the loudspeakers were evaluated under a given room acoustic condition before moving to the next block of trials under a different room acoustic condition. In the intermixed treatment condition, loudspeakers were evaluated under a different room acoustic condition on each trial within a given block of trials.

The results were analyzed as a 2 x 4 x 4 x 3 repeated measures analysis of the variance model where the dependent variable was preference rating, and the independent variables were: (1) trial ordering method (two levels; successive and intermixed treatment conditions), (2) loudspeakers (four), (3) room (four), and (4) program (three). The trial order method was treated as a between-subjects factor. All statistical tests were performed at an alpha level of 0.05. The four main conclusions that resulted from this analysis are summarized in the following sections.

6.3.1 Room Acoustics Influence Listener Preferences for Multichannel Auditory Imagery

The variation in room acoustics had a significant effect on listener preference ratings for the multichannel auditory imagery. Lower preference ratings were observed in Room R4, which had higher levels of late arrival reflections than the other three rooms. This was confirmed by examining the ETC of Room R4, which had 15-20 dB more energy than the other three rooms in its reflections that arrived 60-300 ms after the direct sound. Informal comments from listeners further confirmed what the ETC had already revealed; that Room R4 was "too reverberant," "overly live," "exceedingly bright," and "lacked clarity." The other three rooms had similar looking ETCs, and there were no significant differences among the mean preference ratings given in those three rooms.

6.3.2 Room Acoustics Influence Loudspeaker Preference

The specific room acoustic conditions in which listeners are exposed affect the pattern of their loudspeaker preferences. This study provides direct evidence that the source of this variation is related to both the reflectivity of the room and the off-axis frequency response of the loudspeaker. As the reflectivity of the listening room increases, the off-axis colourations in the loudspeaker become easier to detect, and this leads to lower preference ratings.

The largest range of loudspeaker preference ratings were observed in the most reflective room (R4). In contrast, there were no significant loudspeaker preference ratings in the least reflective room (R3). In the most reflective room (R4), listeners generally preferred those loudspeakers that had the smoothest off-axis frequency responses based on an analysis of their anechoic measurements (see Appendix C).

6.3.3 Program Selection Influences Loudspeaker Preference

A significant interaction was found between program and loudspeaker that was largely isolated to Program JV (jazz band with female vocalist). This program produced lower ratings for Loudspeaker L4 than it received when coupled with the other two programs. Loudspeaker L4 had an irregular off-axis frequency response and generally received the lowest preference ratings across the board as well. Program interactions with loudspeakers have been confirmed in several previous studies (see section 5.1.3). While the exact cause of this interaction in the present study is still under investigation, it is likely that the unique spectral and temporal characteristics of this program revealed colourations in the loudspeaker that were not apparent with the other two program signals.

6.3.4 Prior Participation in Loudspeaker Tests Affects the Influence of Loudspeaker and Room Effects on Listener Preference

Having a prior history of participating in loudspeaker preference tests influences how listeners weigh the relative importance of loudspeaker and room effects when formulating their preference ratings. The listener's level of experience was also a factor in how sensitive their responses were to the loudspeaker and room effects. Listeners with over 2 years of experience had, on average, higher loudspeaker *F* ratios than the inexperienced listeners, indicating that experienced listeners were more discerning of loudspeaker differences. On the other hand, the inexperienced listeners had higher room *F* ratios, indicating they were more discerning of room effects than the experienced listeners. Previous studies, where spatial versus timbral trade-offs were made in multichannel reproductions, have shown that experienced listeners tend to weigh timbre effects more heavily than inexperienced listeners. Inexperienced listeners, on the other hand, weigh certain spatial effects more heavily than experienced listeners (see section 4.3.4).

When viewed within the context of this study, the author posits that experienced listeners, through their conditioning, are more responsive to the loudspeaker effects, which are more timbre-related. Likewise, the inexperienced listeners are more responsive to the room effects, which are more spatial-related. Experienced listeners use lower preference ratings overall, which confirms the findings of a previous study by the author that compared the preferences of trained to those of untrained listeners.

6.3.5 Relationship Between Preference and Acoustical Measurements of Room

This section summarizes results related to the effects above and below the transition frequency of the listening room, which differ significantly.

6.3.5.1 Above the room's transition region

Room preferences related to acoustic effects above the transition frequency (300 Hz) can be partially explained by the ETC of the BRIR. The room (R4) that received the lowest preference ratings had significantly higher level, later arrival reflections than found in the other three rooms. In comparison to the other three rooms, the later arrival reflections in Room R4 were 15-20 dB higher over a time period of 60-300 ms after the arrival of the direct sound.

6.3.5.2 Below the room's transition region

Acoustical measurements of loudspeaker-room interactions below the room transition region (about 300 Hz) are not reported in this paper because they are beyond its scope. Analysis of these effects will be the topic of a future paper.

Preliminary analysis of the frequency response of the loudspeakers measured at the listening position in each room reveal significant differences among the rooms below 200 Hz. The highest preference ratings were observed in Room R1, which produced more bass at the listening position between 70-100 Hz than the other three rooms. While listeners informally commented that there were differences in bass qualities among the rooms, the dominant room effects appear to be related to differences in their reflectivity above the transition frequency. The most and least reflective rooms (R4 and R3, respectively) were the primary source of variance in the preference ratings. An alternative, and perhaps controversial, explanation is that listeners may have adapted to the room effects below 200 Hz. Very few perceptual studies exist on this important topic. More research is clearly warranted to explain why listeners' preference ratings are unaffected by the significant variations between the rooms in their low-frequency responses, measured in room.

6.3.6 The Effect of Room Acoustic Adaptation on Loudspeaker Preference

Differential exposure to the different room acoustics failed to produce a significant effect on listeners' preference ratings. Therefore, no direct statistical evidence was produced that proves either proposition: (1) that room acoustic adaptation occurred, or (2) that it didn't occur. Most likely, the length of time (1-2 min) that was required to complete a trial in the intermixed treatment condition was simply too long to prevent room acoustic adaptation from occurring. If this is true, listeners appear to adapt very quickly to the listening room's acoustics.

Evidence that room adaptation may have occurred can be derived by comparing the preference ratings given to the individual rooms for both the successive and the intermixed treatment conditions (see Figure 26). The intermixed treatment conditions were more influenced by variations in the room acoustics than the successive treatment conditions were. Overall, the intermixed treatment conditions produced lower preference ratings than the successive treatment conditions. The highly reflective room (R4) produced particularly lower preference ratings from the intermixed treatment condition when compared to the ratings from the successive treatment condition. While this evidence is not conclusive, it clearly mandates more research, and also prevents the author from concluding that room acoustic adaptation did *not* occur.

6.4 Generalizability of This Study

The experimental findings of this study may, or may not, apply to listening conditions outside the ones tested here. The generalizability of a study, or its external validity, describes the degree to which the findings will be true over a range of conditions outside of those tested in the laboratory. Below are some possible factors in this study that may limit its generalizability and/or external validity.

1. **Rooms:** The listening rooms in this study are all simple, windowless, rectangular-shaped rooms, with dimensions that fall within a range of real-world values. While rectangular rooms are very common, many domestic listening rooms have an irregular shape, others have several doors and windows, many have an open floor plan, and some even feature a cathedral ceiling. Differences between the rooms utilized in this experiment and the vast range of features in domestic rooms naturally limit the generalizability of this study's experimental findings. However, the author is hopeful that some of these types of room will be included in his future studies.
2. **Acoustic Treatment of Rooms**: Three of the rooms in this study have reverberation and acoustical treatments that fall within the range of typical

domestic rooms (Bradley, 1986). Room R4 is unusually reverberant, and is likely outside of the range of most domestic rooms its size.

3. **Loudspeakers:** A conscious decision was made not to include poor quality loudspeakers in this study in order to avoid possible stimulus recognition biases that might have swamped any room effects and/or room acoustic adaptation. All four loudspeakers in this study were among the best examples of their kind sold today. Speakers with unusual directivities were purposefully not represented in this study as justified in section 2.2.3.
4. **Loudspeaker Setups:** A limiting factor in the generalizability of this study was the use of only ITU-R BS.775-1(1994) setups. Perfect, symmetrically arranged circles of equidistant loudspeakers are not common in consumer's homes, since practical, aesthetic, or spousal-related considerations usually take priority. ITU loudspeaker setups are most commonly used by audio academics, scientists, and serious audiophiles. More data are needed for loudspeaker setups that are arranged asymmetrically, diagonally in a corner, and/or with different angular (vertical and horizontal) positioning of the surround speakers.
5. **Programs:** The temporal and spectral envelope of the program affects the detection of reflections and resonances, as well as other spectral irregularities in loudspeakers and rooms (see section 2.3 for a review).

This study used three five-channel selections of pop and jazz vocal music that all contained both temporally continuous and discontinuous signals. Not included in this study were highly impulsive signals, such as clicks or percussive transients, that are known to produce lower detection thresholds for late arrival reflections. In highly reflective rooms, the precedence effect and room acoustic adaptation may not be as effective in suppressing reflections using these types of signals.

6. **Room Acoustic Adaptation**: There may be other factors that influence room acoustic adaptation that are not tested in this study. These factors may include: (1) the nature and complexity of the listening task, (2) the physical properties of test signals, (3) the number of loudspeakers used, (4) the listeners' training and experience, (5) the exposure time to the room acoustics, and (6) the acoustic properties of the loudspeakers and rooms. Shorter room acoustic exposure times (< 30 s) than used in this study are recommended for future studies in order to produce less room acoustic adaptation.

6.5 Future Research

As is often the case in scientific research, this study raises many new questions regarding the perception of acoustical interactions between loudspeakers and rooms, and the role played by room acoustic adaptation. This study provides evidence that room acoustic adaptation occurs very quickly, however the exact time course over which it most effectively operates is still unknown, and requires more study.

This study did not attempt to identify the underlying perceptual attributes affected by the acoustical interaction between the loudspeakers and rooms, and how they relate to the observed change in listeners' preference ratings. Informal feedback from listeners in this study indicates that both timbral and spatial attributes are affected by the acoustical interactions between loudspeakers and rooms. Further research is needed to identify the exact number and nature of these perceptual attributes.

Additional research is needed to test and validate whether the evaluations of loudspeakers made using multichannel recordings can be generalized to evaluations made using mono or stereo recordings. The experience of the listener should be included as a variable in this research, since this study, along with previous ones, confirms that listeners weigh the importance of certain timbral and spatial attributes differently in the formulation of their preferences, depending on their prior listening experience.

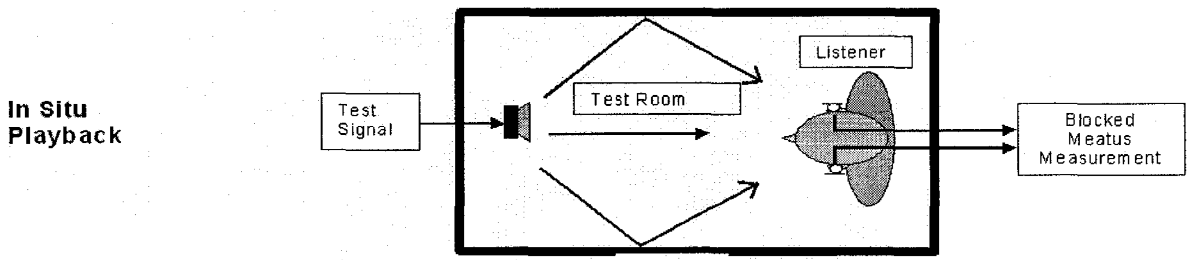
Finally, this study brings to light the need to identify the relative perceptual importance of acoustical effects that occur above the room's transition frequency compared to the effects that occur below the transition frequency. Additional listening tests and analysis of the BRIRs captured at the listening location may provide answers to this question. If listener preference ratings are primarily affected by loudspeaker-room interactions that occur at frequencies below 300 Hz, then an effective loudspeaker-room correction scheme may be able to effectively correct the problem.

This dissertation has only begun to scratch the surface of the myriad of challenging and ground-breaking research opportunities in this new area of scientific audio research. The author is hopeful that this dissertation will generate significant interest and fruitful results in the near future.

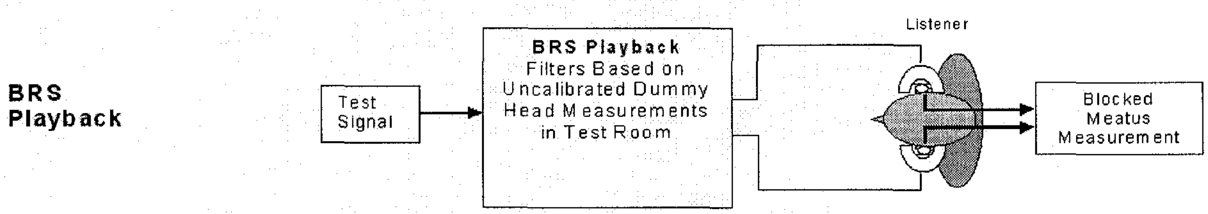
Appendices

Appendix A

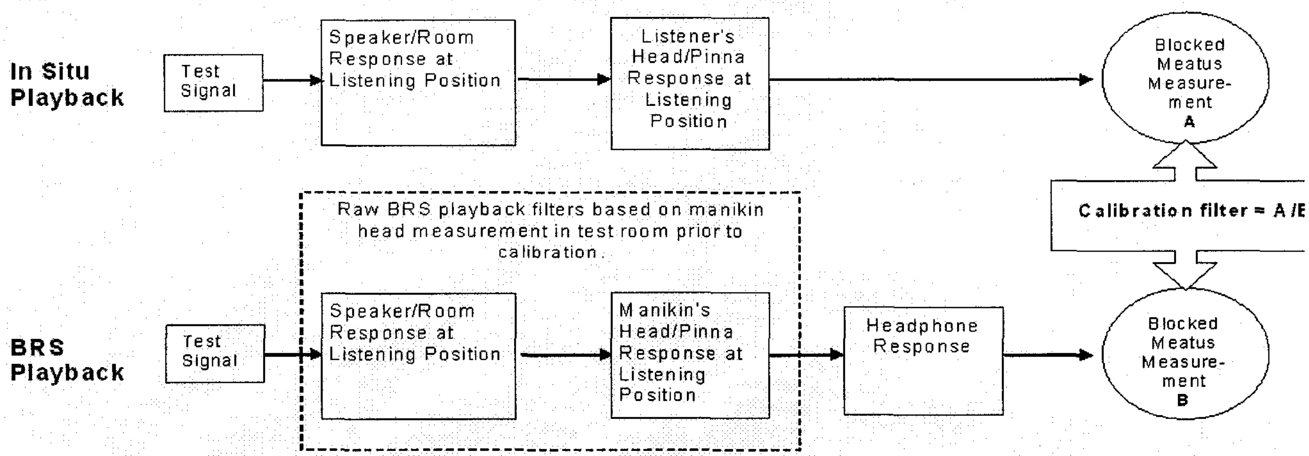
Binaural Room Scanning (BRS) Calibration Procedure



(a)



(b)



(c)

Figure A1. An overview of the calibration method used in this experiment, depicted as follows; (a) in situ playback and measurement for calibration, (b) BRS playback (uncalibrated) and measurement for calibration, and (c) schematic representation of calibration process. A is the signal that *should* appear at the listener's ears in the test room, B is the signal that *does* appear using the uncalibrated BRS playback.

Appendix B

Room Dimensions and Loudspeaker Setup

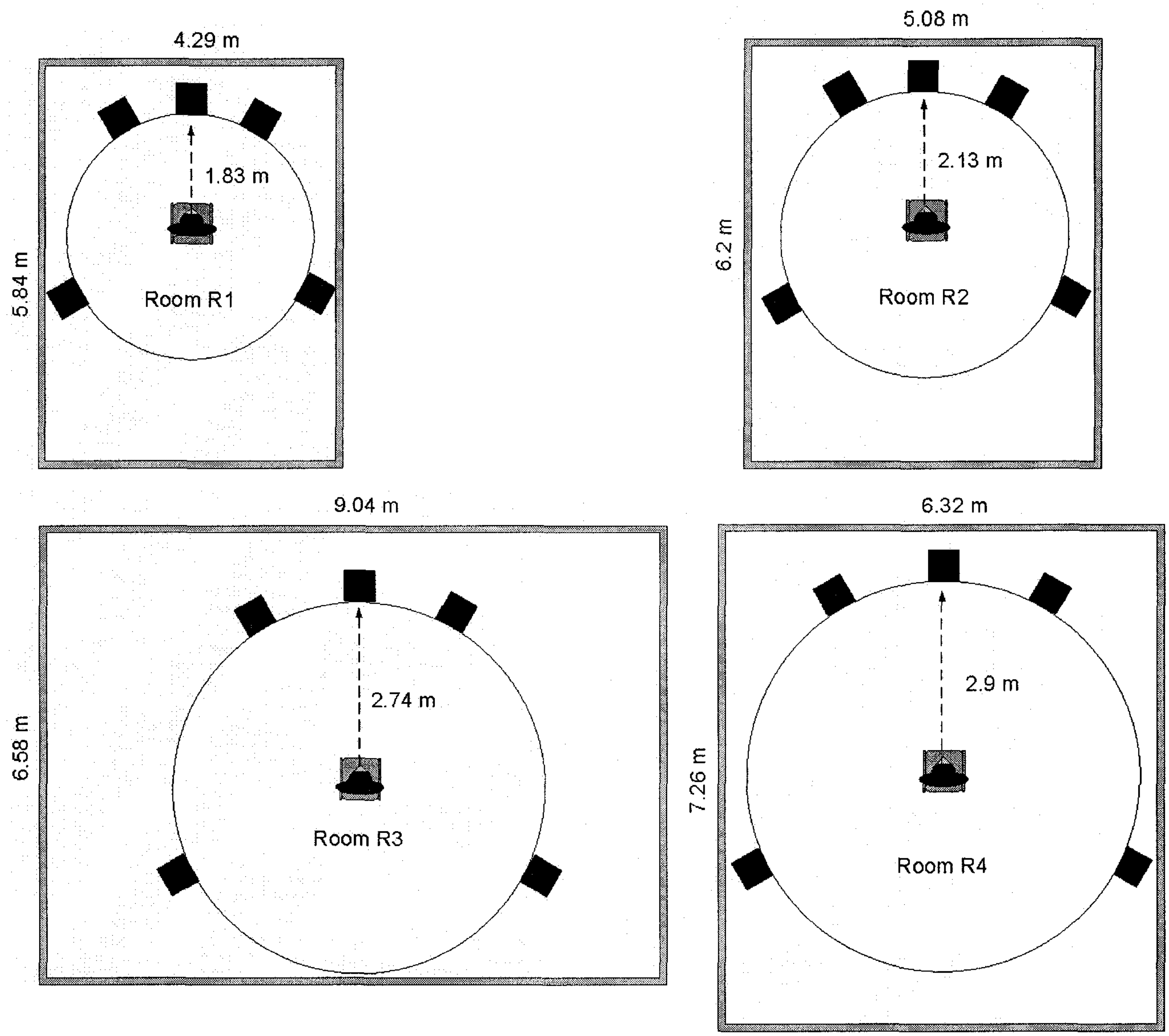
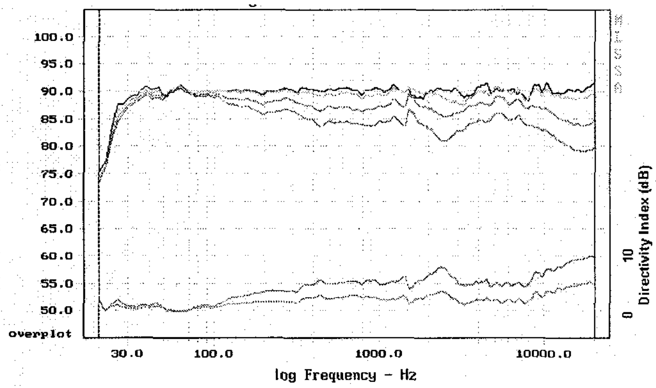


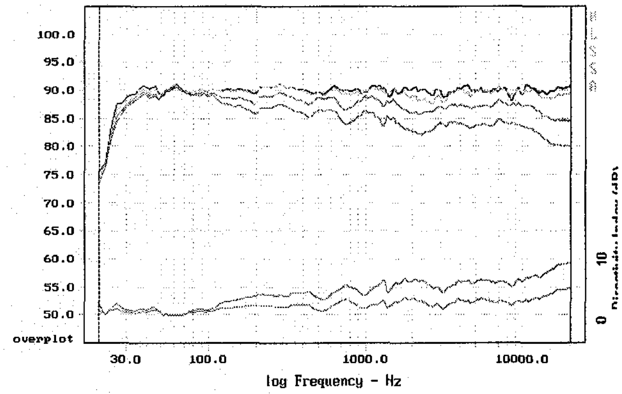
Figure B1. Diagrams of the loudspeaker setups in each of the four listening rooms used in this study (R1, R2, R3, and R4). All setups are compliant with ITU-R BS.775-1 (1994); with the left and right front channels at ±30°, the center loudspeaker at 0°, and the left and right surrounds at ±115°. See Table 3 for more information on the rooms.

Appendix C

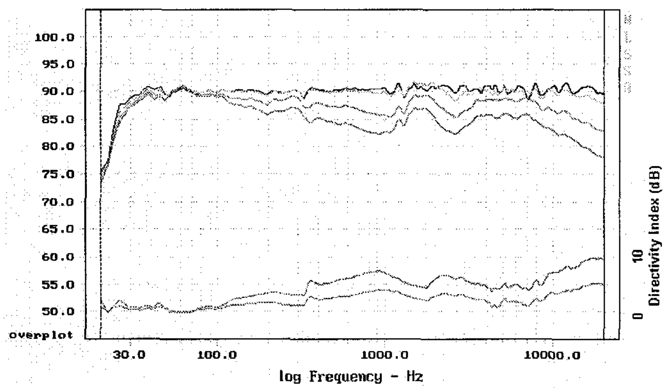
Measurements of Loudspeakers Used in This Study



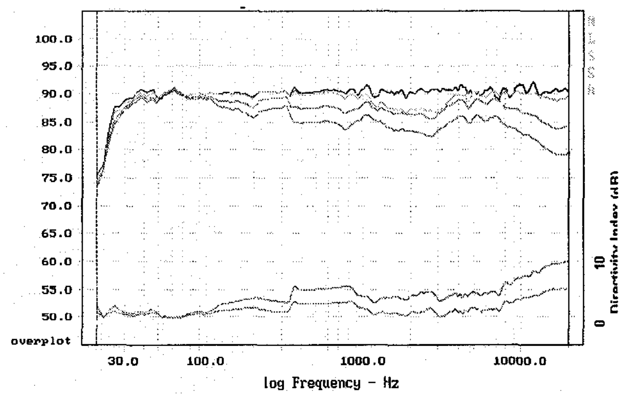
(a)



(b)



(c)



(d)

Figure C1. The anechoic frequency response measurements of the four loudspeakers used in this study: (a) L1, (b) L2, (c) L3, and (d) L4. The curves in each graph represent from top to bottom: the on-axis response (black curve), the average direct sound (green curve), the early-reflections (red curve), and the sound-power response (blue curve). The two bottom curves represent the directivity indices based on the sound power (blue curve) and the early reflections (red curve). See Table 4 for more details on the loudspeakers, and Devantier (2002), for more information on the derivation of the curves.

Appendix D

Listener Instructions

In these tests you will be evaluating binaural headphone reproductions of different 5.1 channel surround sound systems. Your preferences among the different surround sound system should reflect your overall personal tastes and preferences in sound reproduction quality. In formulating your preference choices and ratings, consider both the spectral and spatial qualities of each surround sound system.

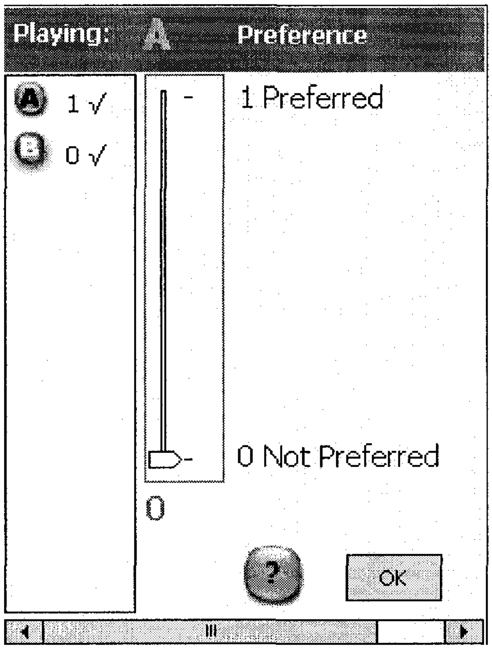
Paired Preference Choice Task

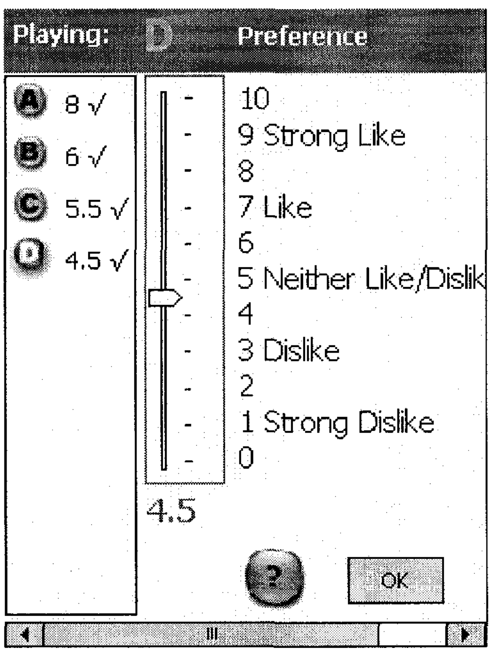
You are asked to first complete 6 trials, where you are presented pairs of different surround systems. You switch between the two surround sound systems by clicking on Buttons A and B using your stylus (see Figure D1) and listen as long as you like. Enter a rating of 1 for the system you prefer, and a rating of 0 for the system that is not preferred. The rating can be changed by moving the slider with your stylus or using the Up-Down buttons on the PDA. Once your final preference choices are made click on the Buttons OK, and then DONE (note: If you give tied ratings, you will be prompted to re-enter your ratings).

After you have hit DONE, your results will be automatically saved and the next trial will be loaded. The same program selection (Jazz Vocal) is used for all 6 trials. When you have completed all 6 trials, the experimenter will instruct you to begin the preference rating test.

Preference Rating Task

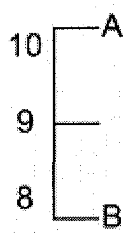
In this test, you will complete 9 trials where four different surround sound systems are presented in each trial. Three different program selections will be used in this test. Like the previous test, you can listen and switch among the four different surround systems as long as is necessary by clicking on Buttons A through D (see Figure D2). Using the slider and/or Up-Down PDA buttons, enter a rating for each surround system using the 11-point preference shown below.



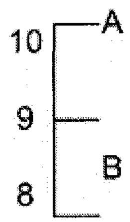


We are more interested in the relative differences in preferences among the different surround sound systems, rather than how you feel about them on an absolute basis of overall sound quality. Therefore, we encourage you to use as much of the scale as necessary in order to discriminate and express your relative preferences among each surround sound system presented. We give the following guidelines in how you express the relative differences in preference among two different surround sound systems.

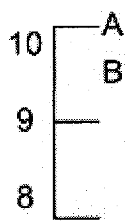
A strong preference between two surround systems (e.g. A and B) is expressed as separation of 2 or more ratings.



A moderate preference is a separation of 1 to 2 ratings.



A slight preference is a separation of 0.1 to 0.5 rating.



Appendix E

Information Supplied to Prospective Participants

Dear Participant,

You are invited to participate in a study entitled *"The Effects of Room Acoustic Adaptation on Loudspeaker Preferences in Multichannel Music Reproduction"* which is being conducted by Sean Olive as part of his PhD research at McGill University.

Purpose of the research: These listening tests are carried out in order to gain information on the human perception of multichannel music recordings reproduced via different loudspeakers setups in different listening rooms; and in particular, how these factors affect your sound quality preference ratings.

What is involved in participating: In this study, I would like you to listen to and compare short multichannel music samples that have been reproduced through different five-channel loudspeaker systems setup in different listening rooms. All samples represent typical music recordings that are reproduced at normal listening levels, so there is no chance for hearing damage. You can switch between each sample by clicking on the buttons on the graphical user interface (GUI) of the Pocket PC. Please make your responses using the buttons and sliders provided on the screen. After you have confirmed your input response by clicking on the provided button, the next sound sample will be presented. No feedback will be given to the listener at any point during the study, but questions about the study will be answered at the end of the study.

Each session will be completed in approximately 20 minutes, and after the session you can take a break. Your participation will require less than two hours total per day. Though the whole experiment consists of 4 sessions over a time period of 2 weeks, you may discontinue participation in this study at any point during the process. Your name will not be disclosed at any time. In oral and written presentations of the results of this study, your responses will be assigned a number and will be grouped with those of the other participants for later analysis. The aim of this test is not to evaluate your individual performance, but to gather findings about human auditory responses to reproduced sound. For participation in this study, you will receive a $15 gift certificate (reimbursable at Amazon.com or MyHarman.com) at the end of each listening session. At the completion of the fourth session, you will also receive a set of Harman Kardon EP-730 headphones (retail value of $199.)

If you would like to learn about the results after the end of the study, please contact Sean Olive at (818) 895-5776 or via email at solive@harman.com. I would like to thank you for your interest in this investigation.

Should you have any further questions or concerns regarding this study you may also contact my PhD supervisor, Dr. William Martens at (514) 398-4535 x089795 or by email at wlm@music.mcgill.

Sincerely,

Sean Olive
Manager Subjective Evaluation, R&D Group
Harman International

Appendix F

Participant Consent Form

Participant Consent Form

My participation in the study *"The Effects of Room Acoustic Adaptation on Loudspeaker Preferences in Multichannel Music Reproduction,"* conducted by Sean Olive is voluntary. I understand that I may discontinue participation at any point during the experiment and that my name will not be disclosed at any time during the analysis or the dissemination of the findings. I have read the above information and the Information for Prospective Participants, and I agree to participate in this study.

Participant's Signature: ____________________ Researcher's signature: ______________

Participant's Name: ______________________Date: __________________________

(Please print)

Appendix G

Certificate of Ethical Acceptability of Research Involving Humans

McGill

Research Ethics Board Office
McGill University
845 Sherbrooke Street West
James Administration Bldg., rm 419
Montreal, QC H3A 2T5

Tel: (514) 398-6831
Fax: (514) 398-4644
Ethics website: www.mcgill.ca/researchoffice/compliance/human/

Research Ethics Board II
Certificate of Ethical Acceptability of Research Involving Humans

REB File #: 140-0107

Project Title: The effects of room acoustic adaptation on loudspeaker preferences in mulit-channel music reproduction

Principal Investigator: Sean Olive

Department: Schulich School of Music

Status: Ph.D. student

Supervisor: Prof. W. Martens

Funding Agency and Title (if applicable): N/A

This project was reviewed on January 16, 2007 by

Expedited Review ✓
Full Review ____

[signature]

Blaine Ditto, Ph.D.
Chair, REB II

Approval Period: January 16, 2007 **to** January 15, 2008

This project was reviewed and approved in accordance with the requirements of the McGill University Policy on the Ethical Conduct of Research Involving Human Subjects and with the Tri-Council Policy Statement: Ethical Conduct For Research Involving Humans.

---

* All research involving human subjects requires review on an annual basis. A Request for Renewal form should be submitted at least one month before the above expiry date.
* When a project has been completed or terminated a Final Report form must be submitted.
* Should any modification or other unanticipated development occur before the next required review, the REB must be informed and any modification can't be initiated until approval is received.

References

ALLISON, R. F. (1974). The influence of room boundaries on loudspeaker power output. *Journal of the Audio Engineering Society,* 22, 314-320.

ANDO, Y. (1977). Subjective preference in relation to objective parameters of music sound fields with a single echo. *Journal of the Acoustical Society of America,* 62, 1436–1441.

ANDO, Y. (1998). *Architectural Acoustics: Blending sound sources, sound fields, and listeners.* AIP series in modern acoustics and signal processing. New York: AIP Press.

ANDO, Y., & SCHROEDER, M. R. (1985). *Concert hall acoustics.* Springer series in electrophysics, vol. 17. Berlin, Germany: Springer-Verlag.

ANTSALO, P., KARJALAINEN, M., MÄKIVIRTA, A., & VÄLIMÄKI, V. (2003, March). Comparison of modal equalizer design methods. [Preprint 5844]. *Proceedings of the 114th International Convention of the Audio Engineering Society,* Amsterdam.

AUDYSSEY (2007). Audyssey Sound Equalizer [loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.audyssey.com/soundequalizer.html

AVIS, M. R., FAZENDA, B. M., & DAVIES, W. J. (2007). Thresholds of detection for changes to the Q factor of low-frequency modes in listening environments. *Journal of the Audio Engineering Society,* 55(7/8), 611-622.

AZZALI, A., FARINA, A., ROVAI, G., BOREANAZ, G., & IRATO, G. (2004, October). Construction of a car stereo audio quality index [Preprint 6306]. *Proceedings of the 117th International Convention of the Audio Engineering Society,* San Francisco.

BAIRD, J. C. (1997). *Sensation and judgment: Complementarity theory of psychophysics.* Scientific psychology series. Mahwah, NJ: Lawrence Erlbaum Associates.

BANG & OLUFSEN (2007). Adaptive Bass Control [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.bang-olufsen.com/page.asp?id=125&callFlashSoundPicture=Room%20adaption

BARRON, M. (1971). The subjective effects of first reflections in concert halls—The need for lateral reflections. *Journal of Sound and Vibrations,* 15, 475–494.

BARRON, M., & MARSHALL, A. H. (1981). Spatial Impression due to early reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibration.* 77, 211-232.

BECH, S. (1992). Selection and training of subjects for listening tests on sound-reproducing equipment. *Journal of the Audio Engineering Society,* 40(7/8), 590-610.

BECH, S. (1993, June). The influence of the room and of loudspeaker position on the timbre of reproduced sound in domestic rooms [Paper 12-007]. *Proceedings of the 12th International Conference of the Audio Engineering Society: The Perception of Reproduced Sound,* Copenhagen, Denmark.

BECH, S. (1995). Timbral aspects of reproduced sound in small rooms, I. *Journal of the Acoustical Society of America,* **97**, 1717-1726.

BECH, S. (1996). Timbral aspects of reproduced sound in small rooms, II. *Journal of the Acoustical Society of America,* **99**, 3539-3549.

BECH, S. (1998). Spatial aspects of reproduced sound in small rooms. *Journal of the Acoustical Society of America,* **103**, 434-445.

BECH, S., GULBOL, M-A., MARTIN, G, GHANI, J., & ELLERMEIER, W. (2005, May). A listening test system for automotive audio - part 2: Initial verification [Preprint 6359]. *Proceedings of the 118th International Convention of the Audio Engineering Society,* Barcelona, Spain.

BEGAULT, D. R., WENZEL, E.M., & ANDERSON, M. R. (2001).Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society,* **49**(10), 904-916.

BENJAMIN, E., & GANNON, B. (2000, September). Effect of room acoustics on subwoofer performance and level setting [Preprint 5232]. *Proceedings of the 109th International Convention of the Audio Engineering Society,* Los Angeles.

BERANEK, L. L. (1996) *Concert and opera halls, How they sound.* Woodbury, NY: Published for the Acoustical Society of America through the American Institute of Physics.

BERANEK, L. L. (2003). Subjective rank-orderings and acoustical measurements for fifty-eight concert halls. *Acta Acustica united with Acustica,* **89**, 494-508.

BLAUERT, J. (1996). *Spatial hearing: The psychophysics of human sound localization* (Rev. ed.). Cambridge, MA: MIT Press.

BLAUERT, J. (2005). *Communication Acoustics.* Berlin, Germany: Springer-Verlag.

BLAUERT, J, & BRAASCH, J. (2005). Acoustic communication: The precedence effect. *Proceedings of Forum Acusticum 2005,* Budapest, Hungary.

BLAUERT, J., & DIVENYI, P. L. (1988). Spectral selectivity in binaural contralateral inhibition. *Acustica,* 66, 267–274.

BORWICK, J. (2001). *Loudspeaker and headphone handbook* (3rd ed.). Oxford, England: Focal Press.

BOSE, A. G. (1968, October). On the design, measurement and evaluation of loudspeakers. [Preprint 622]. *Proceedings of the 35th International Convention of the Audio Engineering Society,* New York.

BRADLEY, J. S. (1986). Acoustical measurements in some Canadian homes. *Canadian Acoustics,* 14, 14-20.

BRADLEY, J. S., SATO, H., & PICARD, M. (2003). On the importance of early reflections for speech in rooms. *Journal of the Acoustical Society of America,* 113, 3233–3244.

BRADLEY, J. S., & SOULODRE, G. A. (1995). The influence of late arriving energy on spatial impression. *Journal of the Acoustical Society of America,* 97, 2263-2271.

BRADLEY, J. S., & SOULODRE, G. A. (1996). Acoustics 1995: Listener envelopment: An essential part of good concert hall acoustics. *Journal of the Acoustical Society of America,* 99, 22.

BRITISH BROADCASTING CORPORATION (1979, November). *Design of high-level studio monitoring loudspeakers type LS 5/8* (Rpt. BBC-RD-1979/22).

London: Mathers, C. D.

BRONKHORST, A. W., & HOUTGAST, T. (1999). Auditory distance perception in rooms. *Nature,* **397**, 517–520.

BUCHHOLZ, J. M., MOURJOPOULOS, J., & BLAUERT, J. (2001, May). Room masking: Understanding and modelling the masking of room reflections [Preprint 5312]. *Proceedings of the 110th International Convention of the Audio Engineering Society,* Amsterdam.

CARLSTROM, D., CLARK, D., DEETZ, K., FIELD, D., KLACZA, R., TREVENA, A., ZIEMBA, M., & NOUSAINE, T. (Chair). (2007, October). *Uneven Bass Reproduction in Automobiles [Session W18].* Workshop to be conducted at the 123rd International Convention of the Audio Engineering Society, New York.

CASIMIRO, N., GEDDES, E., HOLMAN, T., LIPSHITZ, S., MIHELICH, R., OLIVE, S., PEDERSON, J., & NIND, T. (Chair). (2007, October). *FIR or IIR? That is the Question! [Session W20].* Workshop to be conducted at the 123rd International Convention of the Audio Engineering Society, New York.

CELESTINOS, A., & NIELSEN, S. B. (2005, October). Optimizing placement and equalization of multiple low frequency loudspeakers in rooms [Preprint 6545]. *Proceedings of the 119th International Convention of the Audio Engineering Society,* New York.

CHOISEL, S., & WICKELMAIER, F. (2006, May). Relating auditory attributes of multichannel sound to preference and to physical parameters [Preprint 6684]. *Proceedings of the 120th International Convention of the Audio Engineering Society,* Paris.

CHRISTENSEN, F., MARTIN, G., MINNAAR, P., SONG, W.-K., PEDERSEN, B., & LYDOLF, M. (2005). A listening test system for automotive audio – part 1:

System description. *Proceedings of the 118th International Convention of the Audio Engineering Society,* Barcelona, Spain.

CLIFTON, R. (1987). Breakdown of echo suppression in the precedence effect. *Journal of the Acoustical Society of America,* 82, 1834-1835.

CLIFTON R. K., & FREYMAN, R. L. (1997). The precedence effect: Beyond echo suppression. In Gillkey, R. H., & Andersen, T. R. (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 233-256). Mahweh, NJ: Lawrence Erlbaum and Associates.

CLIFTON, R. K., FREYMAN, R. L., LITOVSKY, R. Y., & MCCALL, D. (1994).Listeners' expectations about echoes can raise or lower echo threshold. *Journal of the Acoustical Society of America,* 95, 1525-1533.

COLLOMS, M. (1985). *High-performance loudspeakers* (3rd ed.). New York: J. Wiley.

CONSUMER ELECTRONICS ASSOCIATION (2005, March). *Audio consumption study.* Arlington, VA: Author.

CONSUMER ELECTRONICS ASSOCIATION (2006, April).*Consumer technologies in the automotive environment.* Arlington, VA. Author.

CONSUMER ELECTRONICS ASSOCIATION (2007a, April). *CEA market research report: 9th annual household and teen CE ownership and market potential study.* Arlington, VA: Author.

CONSUMER ELECTRONICS ASSOCIATION (2007b, July). *US consumer electronics sales and forecasts: 2003-2008.* Arlington, VA: Author.

CONSUMER'S UNION (1973, August). Loudspeaker Accuracy: CU's Tests.

Consumer Reports, 38, 456-457.

D'ANTONIO, P., & COX, T. J. (2001, May). Determining optimum room dimensions for critical listening environments: A new methodology [Preprint 5353]. *Proceedings of the 110th International Convention of the Audio Engineering Society,* Amsterdam.

DEQX PTY LIMITED (2007). DEQX PDC-2.6P [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.deqx.com/

DEVANTIER, A. (2002, October). Characterizing the amplitude response of loudspeaker systems [Preprint 5638]. *Proceedings of the 113th International Convention of the Audio Engineering Society,* Los Angeles.

DIRAC RESEARCH (2007). Dirac Live [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.dirac.se/products.asp

DIVENYI, P. L. (1992). Binaural suppression of nonechoes. *Journal of the Acoustical Society of America,* 91, 1070-1084.

DIZON, R. M., & COLBURN, H. S. (2006). The influence of spectral, temporal, and interaural stimulus variations on the precedence effect. *Journal of the Acoustical Society of America,* 119, 2947-2964.

DJELANI, T., & BLAUERT, J. (2001). Investigations into the build-up and breakdown of the precedence effect. *Acta Acustica united with Acustica,* 87, 253–261.

FARES, E., & PEDERSEN, J. A. (2007, September). Natural timbre in room correction systems (part II). *Proceedings of the 32nd International Conference* *of the Audio Engineering Society: DSP for loudspeakers,* Hillerød, Denmark.

FARINA, A. (2000, February). Simultaneous measurement of impulse response and distortion with a swept-sine technique [Preprint 5093]. *Proceedings of the 108th International Convention of the Audio Engineering Society,* Paris.

FARINA, A., & UGOLOTTI, E. (1999, April). Subjective comparison between stereo dipole and 3D ambisonic surround systems for automotive applications [Paper 16-048]. *Proceedings of the 16th International Conference of the Audio Engineering Society: Spatial Sound Reproduction,* Artikum, Finland.

FAZENDA, B. M., AVIS, M. R., & DAVIES, W. J. (2005). Perception of Modal distribution metrics in critical listening spaces-dependence on room aspect ratios. *Journal of the Audio Engineering Society,* 53(12), 1128-1141.

FEREKIDIS, C., & KEMPE, U. (1996, May). Room mode excitation of dipolar and monopolar low-frequency sources [Preprint 4193]. *Proceedings of the 100th International Convention of the Audio Engineering Society,* Copenhagen, Denmark.

FURNESS, R. (2006). *AES more active than ever.* Retrieved August 23, 2007, from http://www.aes-press.org/pms/en_interview_rkf.rtf.

GABRIELSSON, A. (1979). Dimension analysis of perceived quality of sound reproduction systems. *Scandinavian Journal of Psychology,* 20, 159–169.

GABRIELSSON, A., HAGEMAN, B., BECH-KRISTENSEN, T., & LUNDBERG, G. (1990). Perceived sound quality of reproductions with different frequency responses and sound levels. *Journal of the Acoustical Society of America,* 88, 1359-1366.

GABRIELSSON, A., & LINDSTRÖM, B. (1985) Perceieved sound quality of high-fidelity loudspeakers. *Journal of the Audio Engineering Society,* 33(1/2), 33-53.

GABRIELSSON, A., LINDSTRÖM, B., & TILL, O. (1991) Loudspeaker frequency response and perceived sound quality. *Journal of the Acoustical Society of America,* 90, 707-719.

GABRIELSSON, A., & SJÖGREN, H. (1979). Perceived sound quality of sound-reproducing systems. Journal of the Acoustical Society of America, 65, 1019-1033.

GILFORD, C. L. S. (1979). The acoustic design of talks studios and listening rooms. *Journal of the Audio Engineering Society,* 27(1/2), 17-31.

GOLDBERG, A., & MÄKIVIRTA, A (2003a, March). Automated In-Situ Frequency Response Optimisation of Active Loudspeakers [Preprint 5730]. *Proceedings of the 114th International Convention of the Audio Engineering Society,* Amsterdam.

GOLDBERG, A., & MÄKIVIRTA, A. (2003b, May). Statistical analysis of an automated in-situ frequency response optimisation algorithm for active loudspeakers [Paper 18]. *Proceedings of the 23rd International Conference of the Audio Engineering Society: Signal Processing in Audio Recording and Reproduction,* Helsingør, Denmark.

GOLDBERG, A., & MÄKIVIRTA, A. (2004, May). Performance comparison of graphic equalization and active loudspeaker room response controls. [Preprint 4193]. *Proceedings of the 116th International Convention of the Audio Engineering Society,* Berlin, Germany.

GRIESINGER, D. (1996, November). Multichannel matrix surround sound decoders for two-eared listeners [Preprint 4402]. *Proceedings of the 101st International Convention of the Audio Engineering Society,* Los Angeles.

GUASTAVINO, C., & KATZ, B. F. G. (2004). Perceptual evaluation of multi-dimensional spatial audio reproduction. *Journal of the Acoustical Society of America,* **116**, 1105–1115.

HAAS, H. (1972). The influence of a single echo on the audibility of speech (Doctoral dissertation, University of Gottingen, Germany, 1949). Translation printed in *Journal of the Audio Engineering Society,* 20(2), 146–159.

HAMEED, S., PAKARINEN, J., VALDE, K., & PULKKI, V. (2004, May). Psychoacoustic cues in room size perception [Preprint 6084]. *Proceedings of the 116th International Convention of the Audio Engineering Society,* Berlin, Germany.

HARMAN/KARDON (2007). AVR 745 [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.harmankardon.com/product_detail.aspx?Region=USA&Country=US&Language=ENG&cat=REC&prod=AVR%20745&sType=C

HARTMANN, W. M. (1997) Listening in a room and the precedence effect. In Gilkey, R.H., & Andersen, T.R. (Eds.), *Binaural and spatial hearing in real and virtual environments.* Mahweh, NJ: Lawrence Erlbaum and Associates.

HARWOOD, H. D. (1976, May). Some factors in loudspeaker quality. *Wireless World,* **82**, 45-54.

HESS, W., & BLAUERT, J. (2005). Evaluation of auditory spatial impression in performance spaces. *Proceedings of Forum Acusticum 2005,* Budapest, Hungary.

HESS, W., BRAASCH, J., & BLAUERT, J. (2003, October). Acoustical evaluation of virtual rooms by means of binaural activity patterns [Preprint 5864]. *Proceedings of the 115th* International Convention *of the Audio Engineering*

Society, New York.

HEYSER, R. C. (1971a). Determination of loudspeaker signal arrival times: Part 1. *Journal of the Audio Engineering Society,* **19**(9), 734-743.

HEYSER, R. C. (1971b). Determination of loudspeaker signal arrival times: Part 2. *Journal of the Audio Engineering Society,* **19**(10), 829-834.

HEYSER, R. C. (1971c). Determination of loudspeaker signal arrival times: Part 3. *Journal of the Audio Engineering Society,* **19**(11), 902-905.

HIDAKA, T. & BERANEK, L. L. (2000). Objective and subjective evaluations of twenty-three opera houses in Europe, Japan, and the Americas. *Journal of the Acoustical Society of America,* **107**, 368-383.

HIDAKA, T., BERANEK, L. L., & OKANO, T. (1995). Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls. *Journal of the Acoustical Society of America,* **98**, 988-1007.

HIRSCH, J. (1982, August). Testing speakers. *Stereo Review,* **47**, 24-25.

HORBACH, U., KARAMUSTAFAOGLU, A., PELLEGRINI, R., MACKENSEN, P., & THIELE, G. (1999, May). Design and applications of a data-based auralization system for surround sound [Preprint 4976]. *Proceedings of the 106th International Convention of the Audio Engineering Society,* Munich, Germany.

IEC 268-13 (1985). *Listening tests on Loudspeakers (listening tests and objective methods for rating the quality of the transmission systems).* International Electrotechnical Commission. Geneva, Switzerland.

ISO 389-1 (1998). *Acoustics -- Reference zero for the calibration of audiometric* *equipment -- Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones,* International Organization for Standardization, Geneva, Switzerland.

ISO 8253-1 (1989). *Acoustics -- Audometry methods -- Part 1: Basic pure tone air and bone conduction threshold audiometry,* International Organization for Standardization, Geneva, Switzerland.

ITU-R BS.775-1 (1994). *Multichannel stereophonic sound system with and without accompanying picture.* International Telecommunication Union/ ITU Radiocommunication Sector, Geneva, Switzerland.

JBL (2007). Synthesis [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.jblsynthesis.com/technology/tech3.aspx?Language=ENG

JESTEADT, W., BACON, S.P., & LEHMAN, J. R. (1982). Forward masking as a function of frequency, masker level, and signal delay. *Journal of the Acoustical Society of America,* 71, 950-962.

JOHANSEN, L. G., & RUBAK, P. (1999, May). Target functions and preprocessing techniques in digital equalization design [Preprint 4896]. *Proceedings of the 106th International Convention of the Audio Engineering Society,* Munich. Germany.

JOHANSEN, L. G., & RUBAK, P. (2000, February). A new framework for unified analysis and evaluation of loudspeakers and listening rooms [Preprint 5096]. *Proceedings of the 108th International Convention of the Audio Engineering Society,* Paris.

JOHANSEN, L. G., & RUBAK, P. (2001, May). Listening test results from a new loudspeaker/room correction system [Preprint 5323]. *Proceedings of the 110th*

International Convention of the Audio Engineering Society, Amsterdam.

KARJALAINEN, M., ANTSALO, P., & MÄKIVIRTA, M. (2003, May). Modal Equalization by Temporal Shaping of Room Response [Paper 10]. *Proceedings of the 23rd International Conference of the Audio Engineering Society: Signal Processing in Audio Recording and Reproduction,* Helsingør, Denmark.

KEPPEL, G., & WICKENS, T. D. (2004). *Design and analysis: A researcher's handbook (4th ed.).* Upper Saddle River, NJ: Pearson Prentice Hall.

KIM, S., DE FRANCISCO, M., WALKER, K., MARUI, A., & MARTENS, W. L. (2006, June/July). An examination of the influence of musical selection on listener preferences for multichannel microphone technique [Paper 9-2]. *Proceedings of the 28th International Conference of the Audio Engineering Society: The Future of Audio Technology – Surround and Beyond,* Pitea, Sweden.

KLIPPEL, W. (1990) Multidimensional relationship between subjective listening impression and objective loudspeaker parameters. *Acustica,* 70, 45-54.

KNOWLES ELECTRONICS (2007). KEMAR Manikin Type 45BA [Knowles Electronics Manikin for Acoustic Research]. G.R.A.S. Sound and Vibration. Specifications retrieved August 20, 2007, from http://www.gras.dk/00012/00330/

KOUSGAARD, N. (1987, August/September). The application of binary paired comparisons to listening tests. *Proceedings of a Symposium on Perception of Reproduced Sound,* Gammel Avernaes, Denmark.

LARCHER, V., JOT, J.-M., & VANDERNOOT, G. (1998, September). Equalization methods in binaural technology [Preprint 4858]. *Proceedings of the 105th*

International Convention of the Audio Engineering Society, San Francisco.

LEXICON (2007). MC- 12 HD [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.lexicon.com/products/specs.asp?ID=19

LITOVSKY, R. Y., COLBURN, H. S., YOST, W. A., & GUTZMAN, S. J. (1999). The precedence effect. *Journal of the Acoustical Society of America,* **106**, 1633-1654.

LITOVSKY, R. Y., HAWLEY, M. L., FLIGOR, B. J., & ZUREK, P. M. (2000). Failure to unlearn the precedence effect. *Journal of the Acoustical Society of America,* **108**, 2345-2352.

LITOVSKY, R. Y., & SHINN-CUNNINGHAM, B. G. (2001). Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression. *Journal of the Acoustical Society of America,* **109**, 346-358.

LOCHNER, J. P. A., & BURGER, J. F. (1958). The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech. *Acustica*, **8**, 1–10.

LORHO, G. (2007, May). Perceptual evaluation of mobile multimedia loudspeakers [Preprint 7050]. *Proceedings of the 122nd International Convention of the Audio Engineering Society,* Vienna, Austria.

LYNGDORF AUDIO (2007). RP-1 [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.lyngdorf.com/index.php?option=com_content&task=view&id=21&Itemid=45

MÄKIVIRTA, A., & ANET, C. (2001, September). A survey study of in-situ stereo and multi-channel monitoring conditions [Preprint 5496]. *Proceedings of the 111th International Convention of the Audio Engineering Society,* New York.

MÄKIVIRTA, A., ANTSALO, P., KARJALAINEN, M., & VÄLIMÄKI, V. (2001, September). Low-Frequency Modal Equalization of Loudspeaker-Room Responses [Preprint 5480]. *Proceedings of the 111th International Convention of the Audio Engineering Society,* Amsterdam.

MARTENS, W. L. (2003). Perceptual evaluation of filters controlling source direction: Customized and generalized HRTFs for binaural synthesis. *Acoustical Science and Technology,* 24, 220-232.

MARTENS, W. L. (2006, April). Contextual effects in sensory evaluation of spatial audio: Integral factor or nuisance? *Proceedings of the 1st International Workshop on Spatial Audio & Sensory Evaluation Techniques,* Guildford, England.

MARTENS, W. L., MARUI, A., & KIM, S. (2006, June/July). Investigating Contextual Dependency in a Pairwise Preference Choice Task [Paper 9-1]. *Proceedings of the 28th International Conference of the Audio Engineering Society: The Future of Audio Technology – Surround and Beyond,* Pitea, Sweden.

MCSHANE, C. L. (1969). The Meaning of Quantitative Loudspeaker Measurements [Preprint 651]. *Proceedings of the 36th International Convention of the Audio Engineering Society,* Hollywood, CA.

MERIMAA, J., & HESS, W. (2004, October). Training of listeners for evaluation of spatial attributes of sound [Preprint 6237]. *Proceedings of the 117th International Convention of the Audio Engineering Society,* San Francisco.

MEYER, E., & SCHODDER, G. R. (1952). Uber den Einfluss von Schallr¨uckw¨urfen auf Richtungslohalisation und Lautast¨arke bei Sprache [On the influence of reflected sound on directional localization and loudness of speech]. *Nachr. Akad. Wiss. Gottingen, Math. Phys. Klasse IIa,* 6, 31–42.

MIKAT, D. C. (1996, November). Subjective evaluations of automotive audio systems [Preprint 4340]. *Proceedings of the 101st International Convention of the Audio Engineering Society,* Los Angeles.

MINNAAR, P. (2001). *Simulating an acoustical environment with binaural technology – investigations of binaural recording and synthesis.* Doctoral dissertation, Aalborg University, Copenhagen, Denmark.

MINNAAR, P., OLESEN, S. K., CHRISTENSEN, F., & MØLLER, H. (2001a). Localization with binaural recordings from artificial and human heads. *Journal of the Audio Engineering Society,* 49(5), 323-336.

MINNAAR, P., OLESEN, S. K., CHRISTENSEN, F., & MØLLER, H. (2001b, July/August). The importance of head movements for binaural room synthesis. *Proceedings of the International Conference on Auditory Display,* Espoo, Finland.

MØLLER, H., HAMMERSHØI, D., JENSEN, C. B., & SØRENSEN, M. F. (1995). Transfer characteristics of headphones measured on human ears. *Journal of the Audio Engineering Society,* 43(4), 203-217.

MØLLER, H., HAMMERSHØI, D., JENSEN C. B., & SØRENSEN, M. F. (1999). Evaluation of artificial heads in listening tests. *Journal of the Audio Engineering Society,* 47(3), 83-100.

MØLLER, H., JENSEN, C. B., HAMMERSHØI, D., & SØRENSEN, M. F. (1995). Design criteria for headphones. *Journal of the Audio Engineering Society,* 43(4), 218-232.

MØLLER, H., SØRENSEN, M. F., HAMMERSHØI, D., & JENSEN, C. B. (1995) Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society,* 43(5), 300-321.

MØLLER, H., SØRENSEN, M. F., JENSEN, C. B., & HAMMERSHØI, D. (1996). Binaural technique: Do we need individual recordings? *Journal of the Audio Engineering Society,* 44(6), 451-469.

MOORE, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). San Diego, CA: Academic Press.

MOORE, B. C. J., & GLASBERG, B. R. (1983). Growth of forward masking for sinusoidal and noise maskers as a function of signals delay; implications for suppression in noise. *Journal of the Acoustical Society of America,* 73, 1249-1259.

MOORE, B. C. J., & TAN, C.-T. (2004) Development and validation of a method for predicting the perceived naturalness of sounds subjected to spectral distortion. *Journal of the Audio Engineering Society,* 52(9), 900-914.

MOORE, B. C. J., TAN, C.-T., ZACHAROV, N., & MATTILA, V.-V. (2004) Measuring and predicting the perceived quality of music and speech subjected to combined linear and nonlinear distortion. *Journal of the Audio Engineering Society,* 52(12), 1228-1244.

MORIMOTO, M., & MAEKAWA, Z. (1989). Auditory spaciousness and envelopment. *Proceedings of the 13th International Congress on Acoustics,* Belgrade, Yugoslavia.

NAKAJIMA, T., & ANDO Y. (1991). Effects of a single reflection with varied horizontal angle and time delay on speech intelligibility. *Journal of the Acoustical Society of America,* 90, 3173–3179.

NIAOUNAKIS, T. I., & DAVIES, W. J. (2002). Perception of reverberation time in small listening rooms. *Journal of the Audio Engineering Society,* 50(5), 343-350.

NIELSEN, S. H. (1993). Auditory distance perception in different rooms. *Journal of the Audio Engineering Society,* 41(10), 755-770.

NIELSEN, S. B., & CELESTINOS, A. (2007, September). Time based room correction systems for low frequencies using multiple loudspeakers. *Proceedings of the 32nd International Conference of the Audio Engineering Society: DSP for Loudspeakers,* Hillerød, Denmark.

OKANO, T., BERANEK, L. L., & HIDAKA, T. (1998). Relations among interaural cross-correlation coefficients ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls. *Journal of the Acoustical Society of America,* 104, 255-265.

OLIVE, S. (2001, May). A new listener training software application [Preprint 5384]. *Proceedings of the 110th International Convention of the Audio Engineering Society,* Amsterdam.

OLIVE, S. E. (2003). Differences in performance and preference of trained versus untrained listeners in loudspeaker tests: A case study. *Journal of the Audio Engineering Society,* 51(9), 806-825.

OLIVE, S. E. (2004a, May). A multiple regression model for predicting loudspeaker preference using objective measurements: Part I—listening test results [Preprint 6113]. *Proceedings of the116th International Convention of the Audio Engineering Society,* Berlin, Germany.

OLIVE, S. E. (2004b, October). A multiple regression model for predicting loudspeaker preference using objective measurements: Part II—development of the model [Preprint 6190]. *Proceedings of the 117th International Convention of the Audio Engineering Society,* San Francisco.

OLIVE, S. E., CASTRO, B., & TOOLE, F. E. (1998, September). A new laboratory for evaluating multichannel audio components and systems [Preprint 4842]. *Proceedings of the 105th International Convention of the Audio Engineering Society,* San Francisco.

OLIVE, S. E., & MARTENS, W. L. (2007, October). Interaction between loudspeakers and room acoustics influences loudspeaker preferences in multichannel audio reproduction [Preprint 7196]. *Proceedings of the 123rd International Convention of the Audio Engineering Society,* New York.

OLIVE, S. E., SCHUCK, P. L., RYAN, J. G., SALLY, S. L., & BONNEVILLE, M. E. (1994). The detection thresholds of resonances at low frequencies. *Journal of the Audio Engineering Society,* 45(3), 116-128.

OLIVE, S. E., SCHUCK, P. L., SALLY, S. L., & BONNEVILLE, M. E. (1994). The effects of loudspeaker placement on listener preference ratings. *Journal of the Audio Engineering Society,* 42(9), 651-669.

OLIVE, S. E., SCHUCK, P., SALLY, S. L., & BONNEVILLE, M. (1995, October). The variability of loudspeaker sound quality among four domestic sized rooms [Preprint 4092]. *Proceedings of the 99th International Convention of the Audio Engineering Society,* New York.

OLIVE, S. E., & TOOLE, F. E. (1989, July/August). The detection of reflections in typical rooms. *Journal of the Audio Engineering Society,* 37(7/8), 539-553.

OLIVE, S. E., WELTI, T. S., & MARTENS, W. L. (2007, May). Listener loudspeaker preference ratings obtained in situ match those obtained via a binaural room scanning measurement and playback system [Preprint 7034]. *Proceedings of the122nd International Convention of the Audio Engineering Society,* Vienna, Austria.

OTTO, N. C. (1997, September). Listening test methods for automotive sound quality [Preprint 4586]. *Proceedings of the103rd International Convention of the Audio Engineering Society,* New York.

PAULUS, E., & ZWICKER, E. (1972) Programme zur automaticschen bestimmung der lauheit aus terzpegeln oder frequenczgruppen-pegeln [Computer programs for calculating loudness from third-octave band levels or from critical band levels]. *Acustica,* 27, 253-266.

PEDERSEN, J. A. (2003, October). Adjusting a loudspeaker to Its acoustic environment—the ABC system [Preprint. 5880]. *Proceedings of the 115th International Convention of the Audio Engineering Society,* New York.

PEDERSEN, J. A., & KASPER, T. (2007, September). Fully automatic loudspeaker-room adaptation - The room perfect system. *Proceedings of the 32nd International Conference of the Audio Engineering Society: DSP for Loudspeakers,* Hillerød, Denmark.

PEDERSEN, J. A., & MINNAAR, P. (2006, May). Evaluation of a 3D-audio system with head tracking [Preprint 6654]. *Proceedings of the 120th International Convention of the Audio Engineering Society,* Paris.

PELLEGRINI, R. S. (2001). *A virtual reference listening room as an application of auditory virtual environments.* Doctoral dissertation, Ruhr-Universitat Bochum, Bochum, Germany.

PERRETT, S., & NOBLE, W. (1997). The effect of head rotations on vertical plane sound localization.*Journal of the Acoustical Society of America,* **102**, 2325-2332.

PLENGE, G. (1972). Uber das problem der im-kopf-lokalisa-tion [On the problem of in-head-localization] *Acustica,* **26**, 213-221.

PRITTS, R. (2007). *A global look at audio education.* Retrieved August 23, 2007, from http://www.aes.org/education/pritts4.cfm

PRODUCERS & ENGINEERS WING SURROUND SOUND RECOMMENDATIONS COMMITTEE (2004). Recommendations for surround sound production. National Academy of Arts and Sciences. Retrieved August 23, 2007, from http://www.grammy.com/PDFs/Recording_Academy/Producers_And_Engineers/5_1_Rec.pdf

RAKERD, B., & HARTMANN, W. M. (1986). Localization of sound in rooms, II: The effects of a single reflecting surface. *Journal of the Acoustical Society of America,* **78**, 524-533.

RAKERD, B., & HARTMANN, W. M. (1992). Precedence effect with and without interaural differences - sound localization in three planes [Abstract]. *Journal of the Acoustical Society of America,* **92**, 2269.

RAKERD, B., & HARTMANN, W. M. (1994). More on the precedence effect in three planes [Abstract]. *Journal of the Acoustical Society of America,* **95**, 2917.

ROSENBERG, U. (1973). Loudspeaker measurement and consumer information [Preprint F-4]. *Proceedings of the 44th International Convention of the Audio Engineering Society,* Rotterdam, The Netherlands.

RPG INCORPORATED (2007). Skyline Diffusors. Specifications retrieved August 20, 2007, from http://www.rpginc.com/products/skyline/index.htm

RUMSEY, F. (1999). Controlled subjective assessments of two-to-five-channel surround sound processing algorithms. *Journal of the Audio Engineering Society,* 47(7/8), 563-582.

RUMSEY, F. (2001). *Spatial Audio.* Oxford, England: Focal Press.

RUMSEY, F. (2006, April). Spatial audio and sensory evaluation techniques – context, history and aims. *Proceedings of the 1st International Workshop on Spatial Audio & Sensory Evaluation Techniques,* Guildford, England.

RUMSEY, F., ZIELINSKI, S, KASSIER, S., & BECH, S. (2005a). On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality. *Journal of the Acoustical Society of America,* 118, 968-976.

RUMSEY, F., ZIELINSKI, S, KASSIER, S., & BECH, S. (2005b). Relationships between experienced listener ratings of multichannel audio and naïve listener preferences. *Journal of the Acoustical Society of America,* 117, 3832-3840.

SABERI, K., & ANTONIO, J. V. (2003). Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval. *Journal of the Acoustical Society of America,* 114, 420-429.

SABERI, K., & PERROTT, D.R. (1990). Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. *Journal of the Acoustical Society of America,* 87, 1732-1737.

SANTILLAN, A. O. (2001). Spatially extended sound equalization in rectangular rooms. *Journal of the Acoustical Society of America,* 110, 1989-1997.

SATO, H., BRADLEY, J. S., & MASAYUKI, M. (2005). Using listening difficulty ratings of conditions for speech communication in rooms. *Journal of the Acoustical Society of America,* 117, 1157-1167.

SCHOENWIESNER, M., BRAASCH, J., & ZATORRE, R. (2006, September). *Representation of the acoustic environment in the human cerebral cortex* [27]. Poster session presented at the International Conference on the Auditory Cortex, Grantham, England.

SCHOOLMASTER, M., KOPČO, N., & SHINN-CUNNINGHAM, B. G. (2003). Effects of reverberation and experience on distance perception in simulated environments [Abstract]. *Journal of the Acoustical Society of America,* 113, 2285.

SCHOOLMASTER, M., KOPČO, N., & SHINN-CUNNINGHAM, B. G. (2004). Auditory distance perception in fixed and varying simulated acoustic environments [Abstract]. *Journal of the Acoustical Society of America,* 115, 2459.

SCHROEDER, M. R., GOTTLOB, D., & SIEBRASSE, K. F. (1974). Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustic parameters. *Journal of the Acoustical Society of America,* 56, 1195-1201.

SCHUCK, P. L., OLIVE, S. E., RYAN, J. G., TOOLE, F. E., SALLY, S. L., BONNEVILLE, M. E., MOMTAHAN, K. L., & VERREAULT, E. S. (1993, June). Perception of reproduced sound in rooms: Some results from the Athena Project [Paper 12-006]. *Proceedings of the 12th International Conference of the Audio Engineering Society: The Perception of Reproduced Sound,* Copenhagen, Denmark.

SHINN-CUNNINGHAM, B. (2000, April). Learning reverberation: Considerations for spatial auditory displays. *Proceedings of the International Conference on Auditory Display,* Atlanta, GA. Retrieved April 21, 2007, from Boston University, Department of Cognitive and Neural Studies Web site: http://www.cns.bu.edu/~shinn/pages/pdf/ICAD_Learn_Reverb.pdf

SHINN-CUNNINGHAM, B. (2001, May). Localizing sound in rooms. *Proceedings of the ACM SIGRAPH and EUROGRAPHICS Campfire: Acoustic Rendering for Virtual Environments,* Snowbird, UT. Retrieved April 21, 2007, from Boston University, Department of Cognitive and Neural Studies Web site: http://www.cns.bu.edu/~shinn/pages/pdf/SIGGRAPH.pdf

SHINN-CUNNINGHAM, B. G. (2004, April). The perceptual consequences of creating a realistic, reverberant 3-D audio display. *Proceedings of the International Congress on Acoustics,* Kyoto, Japan.

SHINN-CUNNINGHAM, B. G., DURLACH, N. I., & HELD, R. M. (1998a). Adapting to supernormal auditory localization cues I: Bias and resolution. *Journal of the Acoustical Society of America,* 103, 3656-3666.

SHINN-CUNNINGHAM, B. G., DURLACH, N. I., & HELD, R. M. (1998b). Adapting to supernormal auditory localization cues II: Constraints on adaptation of mean response. *Journal of the Acoustical Society of America,* 103, 3667-3676.

SHINN-CUNNINGHAM, B., & RAM, S. (2003, July). Identifying where you are in room: Sensitivity to room acoustics. *Proceedings of the International Conference on Auditory Display,* Boston, MA.

SHINN-CUNNINGHAM, B. G., ZUREK, P. M., DURLACH, N. I., & CLIFTON, R. K. (1995). Cross-frequency interactions in the precedence effect. *Journal of the Acoustical Society of America,* 98, 164-171.

SHIVELY, R. (1998, October/November). Subjective evaluation of reproduced sound in automotive spaces [Paper 15-011]. *Proceedings of the 15th International Conference of the Audio Engineering Society: Audio, Acoustics and Small Spaces,* Copenhagen, Denmark.

SOULODRE, G. (2003, October). Overview of ITU-R BS. 1534 (The MUSHRA Method). [Powerpoint presentation]. In N. Zacharov (Chair), *Listening tests in practice [TS-17].* Tutorial session conducted at the115th International Convention of the Audio Engineering Society, New York. Retrieved August 23, 2007, from http://www.aes.org/technical/pseas/

SOULODRE, G. A., LAVOIE, M. C., & NORCROSS S. G. (2003). Objective measures of listener envelopment in multichannel surround systems. *Journal of the Audio Engineering Society,* 51(9), 826-840.

SOULODRE, G. A., POPPLEWELL, N., & BRADLEY, J. S. (1989). Combined effects of early reflections and background noise on speech intelligibility. *Journal of Sound and Vibration,* 135, 123–133.

SPIKOFSKI, G., & FRUHMAN, M. (2001, June). Optimization of binaural room scanning (BRS): Considering inter-individual HRTF –characteristics [Paper 1907]. *Proceedings of the 19th International Conference of the Audio Engineering Society: Surround Sound: Techniques, Technology and Perception,* Schloss Elmau, Germany.

STAFFELDT, H. (1974). Correlation between subjective and objective data for quality loudspeakers. *Journal of the Audio Engineering Society,* 22(6), 402-415.

STAFFELDT, H. (1984). Measurement and prediction of the timbre of sound reproduction. *Journal of the Audio Engineering Society,* 32(6), 410-414.

STEVENS, S. S. (1972). Perceived level of noise by Mark VII and decibels (*E*). *Journal of the Acoustical Society of America,* 51, 575-601.

TACT AUDIO (2007). RCS 2.2 XP [loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.tactlab.com/Products/RCS22XP/index.html

TAN, C.-T., MOORE, B. C. J., & ZACHAROV, N. (2003). The effect of nonlinear distortion on the perceived quality of music and speech signals. *Journal of the Audio Engineering Society,* 51(11), 1012-1031.

THEILE, G. (1986). On the standardization of the frequency response of high-quality studio headphones. *Journal of the Audio Engineering Society,* 34(12), 956-969.

TOOLE, F. E. (1983, October). Subjective measurements of loudspeakers: A comparison of stereo and mono listening [Preprint 2023]. *Proceedings of the 74th International Convention of the Audio Engineering Society,* New York.

TOOLE, F. E. (1984, May). The acoustics and psychoacoustics of headphones [Preprint C1006]. *Proceedings of the 2nd International Conference of the Audio Engineering Society: The Art and Technology of Recording,* Anaheim, CA.

TOOLE, F. E. (1986a). Loudspeaker measurements and their relationship to listener preferences: Part 1. *Journal of the Audio Engineering Society,* 34(4), 227-235.

TOOLE, F. E. (1986b). Loudspeaker measurements and their relationship to listener preferences: Part 2. *Journal of the Audio Engineering Society,* 34(5), 323-348.

TOOLE, F. E. (1991, October). Binaural record/reproduction systems and their use in psychoacoustic investigations [Preprint 3179]. *Proceedings of the 91st International Convention of the Audio Engineering Society,* New York.

TOOLE, F. E. (2000, September). The acoustics and psychoacoustics of loudspeakers and rooms--The stereo past and the multichannel future [Preprint 5201]. *Proceedings of the 109th International Convention of the Audio Engineering Society,* Los Angeles.

TOOLE, F. E. (2006). Loudspeakers and rooms for sound reproduction – A scientific review. *Journal of the Audio Engineering Society,* **54**(6), 451-476.

TOOLE, F. E., & OLIVE, S. E. (1988). The modification of timbre by resonances: Perception and measurement. *Journal of the Audio Engineering Society,* **36**(3), 122-142.

TOOLE, F. E., & OLIVE, S. E. (1994, November). Hearing is believing vs. believing is hearing: Blind vs. sighted listening tests, and other interesting things [Preprint 3894]. *Proceedings of the 97th International Convention of the Audio Engineering Society,* San Francisco.

TORICK, E. (1977, October), In the loudspeaker lab. *High Fidelity,* **27**, 69-73.

TRINNOV AUDIO (2007). Optimizer [Loudspeaker room correction processor]. Specifications retrieved August 20, 2007, from http://www.trinnov.com/product_Optimizer.php

USHER, J., & MARTENS, W. J. (2007, June). Perceived naturalness of speech sounds using personalized versus non-personalized HRTFs. *Proceedings of the International Conference on Auditory Display,* Montreal, Quebec, Canada.

VANDERKOOY, J., & LIPSHITZ, S. P. (1990). Uses and abuses of the energy-time curve. *Journal of the Audio Engineering Society,* **38**(11), 819-836.

VOELKER, E.-J. (1985). Control room for music monitoring. *Journal of the Audio Engineering Society,* 33(6), 452-462.

VOISHVILLO, A. (2006, October). Assessment of nonlinearity in transducers and sound systems – from THD to perceptual models [Preprint 6910]. *Proceedings of the 121st International Convention of the Audio Engineering Society,* San Francisco.

WALLACH, H., NEWMAN, E., B., & ROSENWEIG, M. R. (1949). The precedence effect in sound localization. *Journal of Experimental Psychology,* 27, 339-368.

WATKINS, A. J. (1991). Central, auditory mechanism of perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America,* 90, 2942–2955.

WATKINS, A. J. (1999). The influence of early reflections on the identification and lateralization of vowels. *Journal of the Acoustical Society of America,* 106, 2933–2944.

WATKINS, A. J. (2005). Perceptual compensation for effects of reverberation in speech identification. *Journal of the Acoustical Society of America,* 118, 249-262.

WATKINS, A. J., & MAKIN, S. J. (1994). Perceptual compensation for speaker differences and for spectral-envelope distortion. *Journal of the Acoustical Society of America,* 96, 1263-1282.

WATKINS, A. J., & MAKIN, S. J. (1996a). Effects of spectral contrast on perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America,* 99, 3749-3757.

WATKINS, A. J., & MAKIN, S. J. (1996b). Some effects of filtered contexts on the perception of vowels and fricatives. *Journal of the Acoustical Society of America,* **99**, 588–594.

WATKINS, A. J., & MAKIN, S. J. (2007). Steady-spectrum contexts and perceptual compensation for reverberation in speech identification. *Journal of the Acoustical Society of America,* **121**, 257-266.

WEISSER, A., & RINDEL, J. H. (2006). Evaluation of sound quality, boominess, and boxiness in small rooms. *Journal of the Audio Engineering Society,* **54**(6), 495-511.

WELTI, T. (2002, May). How many subwoofers are enough [Preprint 5602]. *Proceedings of the 112th International Convention of the Audio Engineering Society,* Munich, Germany.

WELTI, T. S. (2004, October). Subjective comparison of single channel versus two channel subwoofer reproduction [Preprint 6322]. *Proceedings of the 117th International Convention of the Audio Engineering Society,* San Francisco.

WELTI, T., & DEVANTIER, A. (2006). Low-frequency optimization using multiple subwoofers. *Journal of the Audio Engineering Society,* **54**(5) 347–364.

WENZEL, E. M., & BEGAULT, D. R. (1999). Are individualized head related transfer functions required for auditory information displays? [Abstract]. *Journal of the Acoustical Society of America,* **105**, 1035.

WIGHTMAN, F. L., & KISTLER, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *Journal of the Acoustical Society of America,* **105**, 2841-2853.

ZACHAROV, N., & BECH, S. (2000, September). Multichannel level alignment, part IV: The correlation between physical measures and subjective level calibration [Preprint 5241]. *Proceedings of the109th International Convention of the Audio Engineering Society,* Los Angeles.

ZACHAROV, N., BECH, S., & MEARES, D. (1995). The use of subwoofers in the context of surround sound program reproduction. *Journal of the Acoustical Society of America,* 46, 276-287.

ZAHORIK, P. (2002). Assessing auditory distance perception using virtual acoustics. *Journal of the Acoustical Society of America,* 111, 1832-1846.

ZIELINSKI, S. (2006, April). On some biases encountered in modern listening tests. *Proceedings of the 1st International Workshop on Spatial Audio & Sensory Evaluation Techniques,* Guildford, England.

ZUREK, P. M. (1979). Measurements of binaural echo suppression. *Journal of the Acoustical Society of America,* 66, 1750–1757.