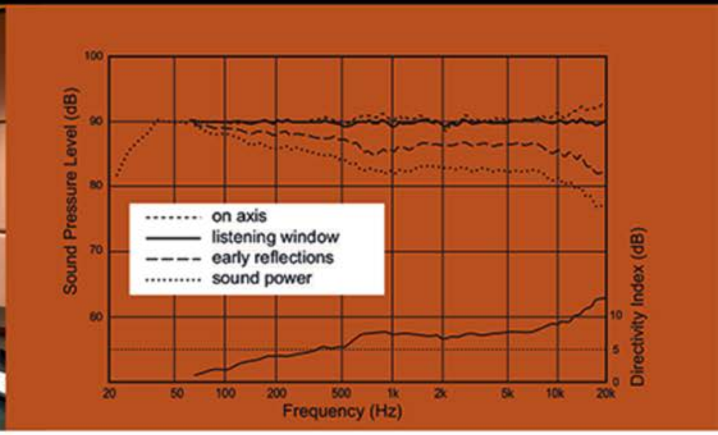
AUDIO ENGINEERING SOCIETY PRESENTS …

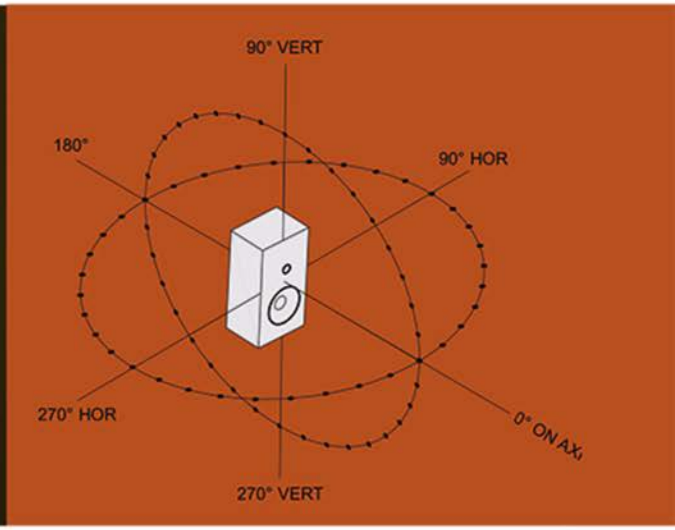
Audio Engineering Society Presents



SOUND REPRODUCTION

THIRD EDITION

The Acoustics and Psychoacoustics of LOUDSPEAKERS AND ROOMS



FLOYD E. TOOLE

ROUTLEDGE

Sound Reproduction

Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms, *Third Edition* explains the physical and perceptual processes that are involved in sound reproduction and demonstrates how to use the processes to create high-quality listening experiences in stereo and multichannel formats.

Understanding the principles of sound production is necessary to achieve the goals of sound reproduction in spaces ranging from recording control rooms and home listening rooms to large cinemas. This revision brings new science-based perspectives on the performance of loudspeakers, room acoustics, measurements and equalization, all of which need to be appropriately used to ensure the accurate delivery of music and movie sound tracks from creators to listeners.

The robust website (www.routledge.com/cw/toole) is the perfect companion to this necessary resource.

Floyd E. Toole was a research scientist at the National Research Council of Canada and, more recently, the corporate vice president of acoustical engineering at Harman International. Now retired, he is a consultant to Harman. He has received the Audio Engineering Society Silver and Gold Medal awards, and lifetime achievement awards from CEDIA and ALMA International. He is a Fellow of the AES, the Acoustical Society of America and CEDIA, and is in the Consumer Technology Hall of Fame.

Audio Engineering Society Presents (AES)
www.aes.org/

Officers of the Technical Council include:

- Francis Rumsey—Chair
- Juergen Herre—Vice Chair
- Michael Kelly—Vice Chair
- Bob Schulein—Vice Chair

Publications:
Handbook for Sound Engineers, Fifth Edition
(9780415842938)
Authored by Glen Ballou
Focal Press, April 1, 2015

Audio Production and Critical Listening, Second Edition
(9781138845947)
Authored by Jason Corey
Routledge, September 1, 2016

Recording Orchestra and Other Classical Music Ensembles
(9781138854536)
Authored by Richard King
Routledge, December 8, 2016

Sound Reproduction

The Acoustics and Psychoacoustics of Loudspeakers and Rooms

Third Edition

Floyd E. Toole

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Third edition published 2018
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

First edition published by Focal Press 2008
Second edition published by Routledge 2013

Library of Congress Cataloging-in-Publication Data
Names: Toole, Floyd E., author.
Title: Sound reproduction : the acoustics and psychoacoustics of loudspeakers and rooms / Floyd Toole.
Description: Third edition. | New York ; London : Routledge, 2017.
Identifiers: LCCN 2017007635| ISBN 9781138921375 (hardback) | ISBN 9781138921368 (pbk.)
Subjects: LCSH: Loudspeakers. | Acoustical engineering. | Sound—Recording and reproducing—Equipment and supplies. | Psychoacoustics. | Recreation rooms.
Classification: LCC TK5983 .T66 2017 | DDC 621.382/84—dc23
LC record available at https://lccn.loc.gov/2017007635

ISBN: 978-1-138-92137-5 (hbk)
ISBN: 978-1-138-92136-8 (pbk)
ISBN: 978-1-315-68642-4 (ebk)

Typeset in Times New Roman
by Apex CoVantage, LLC

Visit the companion website: www.routledge.com/cw/toole

Contents

Acknowledgments

Special thanks are due to the National Research Council of Canada and to Harman International Industries, Inc., who for 27 and 25 years respectively provided funding and facilities for my colleagues and me to investigate the physical and psychoacoustic factors involved in sound reproduction. That we were able to do public presentations of the research findings and to publish them is commendable. Harman engineering staff generously provided many of the measurements seen here. The numerous researchers around the world whose efforts have contributed to the science of audio and to the contents of this book have both my respect and my gratitude.

My thanks to Todd Welti and the late Brad Wood for their constructive criticism of the manuscript, and to my extremely supportive wife Noreen, who shared in the creative process on a daily basis.

I dedicate this book to my father, Harold Osman Toole, who set a high moral standard, taught me the skills to follow in his footsteps as a competent woodworker and handyman, and who insisted that I start life with a good education. He has been a wonderful father and friend—and as this is written, he still is, at 104 years old.

Introduction to the Third Edition

This book is about science in audio, the acoustics and psychoacoustics of loudspeakers, rooms and their combined effects on what listeners hear. It contains many references to work done by researchers all over the world, but among them are references to work done by my research colleagues and me over the years. Consequently, in this introduction to the new edition, I will also introduce myself, my motivations, and my approach to examining aspects of audio.

The first edition of the book was clearly oriented to explaining the science underlying the acoustics and psychoacoustics of loudspeakers, rooms and the listeners who derive pleasure from the combinations. What was called the second edition was a labeling error associated with a change in publishers. The book was unchanged. This edition is substantially new. I have tried to adopt a more linear approach to explaining how the art, technology and science combine to create listening experiences and how we perceive them. Nevertheless, it is a long book with more information than most people need, so I anticipate that it will be read in pieces, dipping into it as appropriate for individual readers. For that reason, certain explanatory details are repeated for clarity.

Readers will find some thoughts that run contrary to conventional audio tradition because new scientific knowledge has superseded incomplete reasoning that provided the original foundations. Some audio folklore needs to be retired. This will not happen overnight, especially when the art is so intertwined with technology and science.

Audio is entertainment, but doing it well may require some homework. Understanding how it works may make the result even more pleasurable.

I was born in 1938 in Moncton, New Brunswick, in eastern Canada. I grew up as a hi-fi enthusiast through the eras of 78s, LPs, open-reel tape, cassettes, tube/valve electronics, and so on. At that time audio was a "participatory" hobby. My father was a consummate do-it-yourselfer, and I followed in his footsteps, building preamps and power amplifiers—from scratch in the beginning, using a lot of war surplus parts, and later from Heath and Eico kits.

My father and I built loudspeaker enclosures in our woodworking shop from designs published in hobbyist magazines. Thick catalogs were full of electronic and audio-related components—so many choices and not a shred of useful data to tell us how they might sound, even if we could understand the data. It was a time of unrestrained opinion, and trial and error. There were aftermarket add-on tweeters, loudspeaker drivers for home-built systems constructed using "universal" enclosures and "universal" crossover networks. The sound quality by today's standards was seriously lacking. The famous acoustician Dr. Leo Beranek once said something like "the sound quality of a home

FIGURE 0.1 *The audio industry of today might wish for this kind of enthusiasm for its mainstream products. Cartoon by Simon Ellinas, www.caricatures.org.uk. Originally published in* Hi-Fi News and Record Review, *1981. Reproduced with permission.*

built loudspeaker increases with the effort put into the hand-rubbed finish." It's true. There were tweaks that claimed to improve the performance of playback hardware and electronics, some of which might even have worked. And always there were maintenance chores, to keep records clean, needles replaced, tubes tested. It was called "hi-fi," high fidelity, a term so abused that it has lost its meaning.

When the term "high fidelity" was coined in the 1930s, it was more a wishful objective than a description of things accomplished—many years would pass before anything resembling it could be achieved. While recreating a live performance was an early goal, and remains one of the several options today, the bulk of recordings quickly drifted into areas of more artistic interpretation.

The essence of high fidelity, the notion of "realism" and the uncolored reproduction of music, dominated almost every discussion of home audio equipment. However, commercial recordings

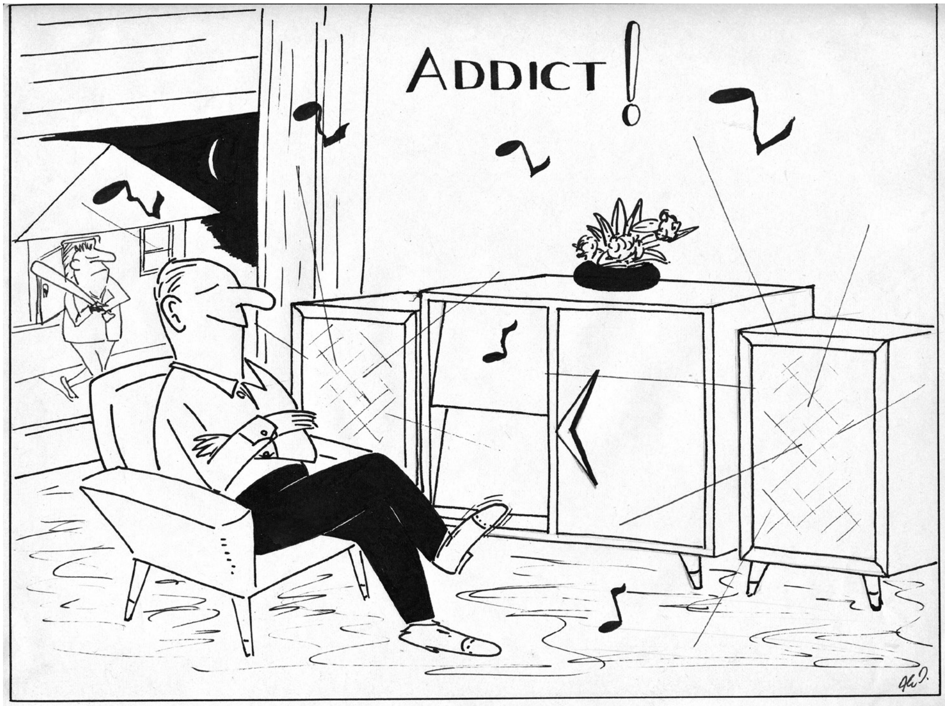


FIGURE 0.2 *As the author saw himself ca. 1958. The stereo cabinetry was self-styled and built in the home workshop mostly by my craftsman father. The enclosures were small Karlsons (a terrible acoustical design), with Goodmans Axiom 12-inch "whizzer-coned" drivers. A Garrard RC-88 changer with a GE VRII mono cartridge, later replaced with an Acos stereo cartridge, drove a homemade preamp (a variation of a Fisher design, as I recall) and homemade Williamson 6L6 power amplifiers, later modified to use Acrosound Ultralinear output transformers. It was loud, and I was proud.*

themselves betrayed the growing divide between the ideals of high fidelity and the reality of what happened in the recording studio.

(Morton, 2000, p. 39)

I went on to study electrical engineering, first at the University of New Brunswick, Canada, then at the Imperial College of Science and Technology, University of London, from which I graduated with a PhD in 1965. My research topic was multidisciplinary, evolving from discussions with Prof. Colin Cherry, known for his expertise in human communication and creator of the term "cocktail party effect," referring to binaural discrimination in complex listening situations. Stereophonic sound and the directional and spatial effects it yielded intrigued him. He was also an electronics engineer, which explains the origins of my own thesis project: an investigation of binaural hearing—sound localization—employing signal generating, processing and data gathering electronics of my own design. In those days, most psychoacoustics research was done without the benefits of modern electronics and acoustical knowledge, so it was interesting to see where these new experimental capabilities took us. The engineering methodology was greatly advantageous, with new experiments being created in days, not months. A thesis and papers resulted (Sayers and Toole, 1964; Toole and Sayers, 1965a, 1965b). It is gratifying to see the results discussed in what is arguably the reference text on spatial hearing (Blauert, 1996). I became captivated by science, finding answers to questions that interested me, and at the same time that seemed to fill a need for a larger audience. It suited my temperament.

Contemporary science is based on a method of inquiry developed in seventeenth-century Europe. The *scientific method* involves observing the natural world, questioning what is seen (or heard, in this context) and then conducting experiments to gather measurable evidence to provide insights or answers. It is a tedious process, requiring care and repetition to ensure that the data are reliable, and a disciplined approach to designing the experiments to ensure that the data bear on the question being asked without being influenced by extraneous factors, including the person asking the question. The answers need to be free from bias. It is not simple, but as will be seen, the results are worth the effort, yielding insights that allow us to now design and predict many aspects of "good sound."

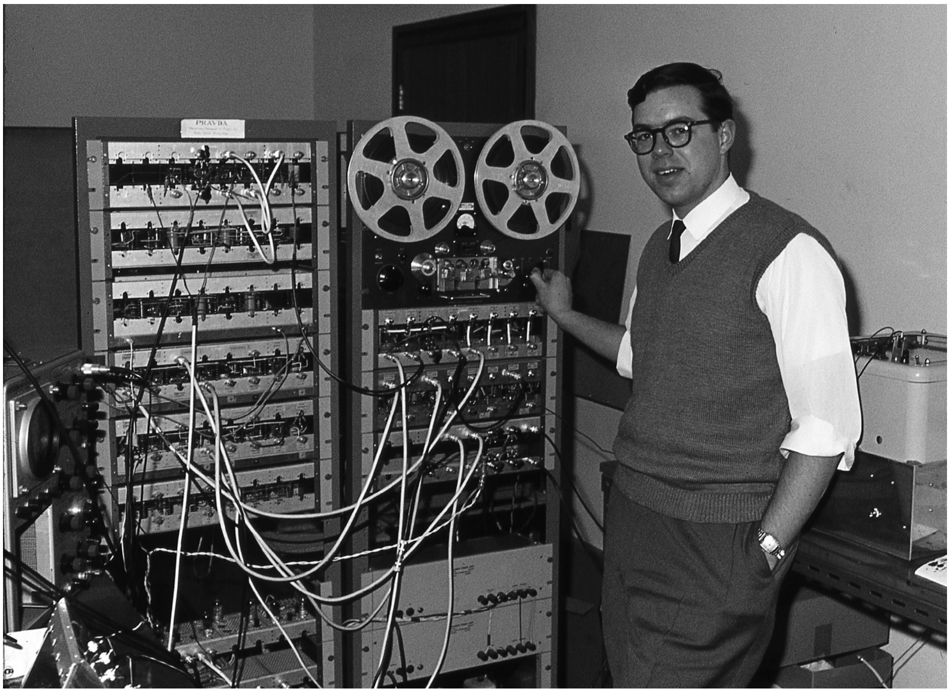


FIGURE 0.3 *Psychoacoustics research, engineering style. Racks of Toole designed and fabricated germanium transistor and tube electronics, including a four-track pulse-width-modulated (DC—200 Hz) analog tape system for control of randomized experimental parameters, and storing listener responses. Results were printed out on an automated X/Y plotter.*

Imperial College, London, ca. 1963.

It is so much easier just to offer an opinion. The problem is that there are so many of them, and they keep changing. It was not surprising to find that the variations were caused by more than the sound. We humans are very susceptible to non-auditory influences that bias our perceptions. Fortunately non-auditory factors are easily controlled.

Upon graduation I was employed as a research scientist at the National Research Council of Canada (NRCC) in Ottawa. My job was to ask questions and find answers by applying the scientific method. I was in the Applied Physics Division, so the emphasis was on real-world issues. A major performance metric was peer-reviewed publications, but also, for my chosen line of investigation, evidence that industry benefitted from the work—the NRCC was taxpayer funded. There could not have been a better place to engage in this kind of research. First, I worked among, and was tutored by, some of the best acoustical scientists in the world. We had access to excellent anechoic and reverberation chambers, as well as all of the latest measuring equipment necessary to quantify sound. There was also the budget and space to create listening rooms for subjective evaluations. The research was successful, and publications resulted.

The embryonic Canadian loudspeaker industry rented the NRCC measurement and listening facilities to design products, and, importantly, Canadian audio magazines paid for the facilities to perform product reviews—anechoic measurements and double-blind listening tests. The products that were designed and reviewed became part of the database for the research, and everyone benefitted from the knowledge as it emerged. A small staff was hired, and I traveled widely, telling the science story to Audio Engineering Society (AES) audiences and to interested manufacturers. The relatively unknown Canadian loudspeaker manufacturers used the credibility of the NRCC and the research to help gain recognition (some are now well known and respected international suppliers). All was as it should be.

One day in early 1991, the phone rang. It was a headhunter offering the possibility of an interesting job with a major audio corporation, Harman International Industries. After 26 years of research, I was intrigued by this opportunity to get directly involved with applying the science to product development—moving closer to the “real” world. Soon I was hired as the corporate vice president of Acoustical Engineering, but very quickly it became more, because I was able to convince the company leaders that we could afford, and indeed needed, a corporate research group that was not attached to the brands and that did not develop audio products. Knowledge was the product, and obviously some of it would migrate into products if it proved to be of value. Harman generously permitted us to publish freely, following the scientific tradition of free exchange of knowledge at AES conventions, conferences and in the journal (some corporations do not allow this).

Harman spent large sums on improved engineering facilities and innovative listening rooms for product evaluation. The benefits were soon seen in improved consistency and quality of sound from the products. Nevertheless, there were arguments from some sales and marketing people who may not have had the same faith in science as we did. Good sound does not guarantee good sales. There are many factors involved in that aspect: appearance, price, size, marketing and retail distribution. These all fall outside the domain of engineering. Still, there was a resolute effort to ensure that products at all price points were competitive in sound quality.

A program of measurements and double-blind evaluations of competing products was set up and continues. However, it has been difficult to maintain for all products because of tight schedules, the numbers of new products being developed, and the decentralization of design and manufacturing as Harman grew into an enormous worldwide, diversified corporation. When I joined Harman in 1991, sales were about $500M, we had a few thousand employees, and we were primarily an audio company headquartered in California. Things changed. Now sales are about $7B, there are about 26,000 employees worldwide, and audio is just part of what Harman does. It is a different company.

I retired in 2007, but I have remained in a consulting role since then. I continue to learn, publish and teach. I had the great pleasure of turning my audio hobby into my profession, and I can truly say that I feel that I have never had a "job."

SCIENCE IN AUDIO

Early in my career I came to grips with some fundamental truths about the role of science in audio. Scientific explanations of the physical world and new technologies have allowed us to enjoy the emotions and aesthetics of music, whenever or wherever the mood strikes us. Music is art, pure and simple. Composers, performers and creators of musical instruments are artists and craftsmen. Through their skills, we are the grateful recipients of sounds that can create and change moods, that can animate us to dance and sing, and that form an important component of our memories. Music is part of all of us and of our lives.

However, in spite of its many capabilities, science cannot describe music. There is nothing documentary beyond the crude notes and symbols on a sheet of music. Science has no dimensions to measure the evocative elements of a good tune. It cannot technically describe why a famous tenor's voice is so revered, or why the sound of ancient Cremona violins has been held up as an example of how it should be done. Nor can science differentiate, by measurement, the mellifluous qualities of trumpet intonations by a master, and those of a music student who simply hits the notes. Those are distinctions that must be made subjectively, by listening. A lot of scientific effort has gone into understanding musical instruments, and as a result, we are getting better at imitating the desirable aspects of superb instruments in less expensive ones. In fact, recent blind evaluations are indications of success (see the box). We are also getting better at electronically synthesizing the sounds of acoustical instruments. However, the determination of what is aesthetically pleasing remains firmly based in subjectivity.

It is at this point that it is essential to differentiate between the production of a musical event and the subsequent reproduction of that musical event. Subjectivity—pure opinion—is the only measure of whether music is appealing. That will necessarily vary among individuals. Analysis of music involves issues of melody, harmony, lyrics, rhythm, tonal quality of instruments, musicianship and so on. In a recording studio, the

SCIENCE, PSYCHOACOUSTICS, MUSICAL INSTRUMENTS AND MUSICIANSHIP

This book discusses the science of sound reproduction. Others have been applying scientific methods to musical instruments and concert venues. Concert hall acoustical investigations have been numerous and well publicized, but those pertaining to the instruments themselves, much less so. Recent papers have stirred interest and controversy by challenging some widely held beliefs. A paper by Fritz et al. (2014) reports results of elaborate blind evaluations of six new and six Old Italian violins (including five by Stradivari). The players were "significant" professional soloists and the evaluations were done in a rehearsal room and in a small concert hall. The result was that when asked to choose a violin to replace their own for a hypothetical tour, 6 of the 10 soloists chose new violins. In the individual ratings, a single new violin was chosen four times, a single Stradivari three times, and two new violins and a Stradivari once each. Tracking those instruments that soloists rejected as unsuitable, the new violins prevailed by a 6:1 ratio. So, to 10 performers, seven of whom routinely play Old Italian violins, the new—much less expensive—alternatives were very attractive.

Bissinger (2008) delved into the details of violin acoustics, and summarized common remarks about the best violins: "they are more 'even' across the measured range, and strong in the lowest range." As an audio person, I interpret this as "flat frequency response and good bass," which seems reasonable. Campbell (2014) provides additional perspective on scientific contributions to several musical instruments—interesting reading for lovers of music with a technical inclination.

On a very different, but related, topic, Tsay (2013) tested the popular notion that "sound is the most important source of information in evaluating performance in music." He found that both novices and professional musicians were able to identify the winners of prerecorded live music competitions better when viewing a video of the event in silence than when listening only, or viewing and listening together. The evaluation of musical performances was found to be dominated by the visual impact of gesticulations, not audible output. Remarkable. It is no wonder that the visual aesthetics of loudspeakers precondition our reactions to the sounds they produce.

recording engineer is an additional major contributor to the art. All of the many electronic manipulations used to create the final stereo mix are judged subjectively, on the basis of whether it reflects the artists' intent and, of course, how it might appeal to consumers.

The evaluation of reproduced sound should be a matter of evaluating the extent to which any and all of these elements are accurately replicated or attractively reproduced. It is a matter of trying to describe the respects in which audio devices add to or subtract from the desired objective. A non-trivial problem is that we, the listeners, were not in the control room at the time the final mix was approved. We don't know what the creators heard, but we still have opinions about what we like and dislike. We don't know who or what to praise or criticize. Often the playback apparatus carries more than its share of responsibility.

When making audio product evaluations, the terminology appropriate to describing the music itself is either insufficient or inappropriate. A different vocabulary is needed. Most music lovers and audiophiles lack this special capability in critical listening, and as a consequence art is routinely mingled with technology. In subjective equipment reviews, technical audio devices are often imbued with musical capabilities. Some are described as being able to euphonically enhance recordings—others to do the reverse. It is true that characteristics of technical performance of playback devices must be reflected in the musical performance, but the technical performance attributes are fixed, and music is infinitely variable. Consequently the interactions are unpredictable. This does not help our efforts to investigate and improve sound reproduction.

Add to this the popular notion that we all "hear differently," that one person's meat might be another person's poison, and we have a situation where a universally satisfying solution might not be possible. Fortunately reality is not so complex, and although tastes in music are demonstrably highly personal, enormously variable, we discover that recognizing the most common deficiencies in reproduced sounds is a surprisingly universal skill when listeners are given a chance to reveal their unbiased opinions. More good news is that most people can do it, even those who think they have "tin ears." Inexperienced listeners take more time, make more mistakes along the way, but in the end, their opinions generally agree with those of the experts. Only those with hearing loss routinely depart from the norm. To a remarkable extent we seem to be able to separate the evaluation of a reproduction technology from that of the program. It is not necessary to be familiar with or to enjoy the music to be able to recognize that it is or is not well reproduced.

How do listeners approach the problem of judging sound quality? Most likely the dimensions and criteria of subjective evaluation are traceable to a lifetime accumulation of experiences with live sound, even simple conversation. If we hear things in reproduced sound that do not occur in nature, or that defy some kind of perceptual logic, we seem to be able to identify it. By that standard, the best sounding audio product is the one that exhibits the fewest audible flaws. Perhaps this is how we are able to make such insightful comments about sound quality based on recordings that either had no existence as a "live" performance, or that we have no personal experience with.

Figure 0.4 shows that in live performances things are relatively uncomplicated. The musicians radiate sound into a performance space and two ears and a brain interpret the combination. These are "reference" listening experiences—there is literally nothing between you, the listener, and the performance. The very complex sound field of a live performance can be sampled by some number of microphones, taken to a recording studio, and manipulated to sound good through some number of channels and loudspeakers. Naturally the spatial complexity of the live performance cannot be replicated through two channels; multichannel schemes can be more persuasive. It is the skill of the recording engineer that allows these compromised reproductions to sound as good and as realistic as they do.

The center of Figure 0.4 illustrates the origin of most of the popular music and jazz that we enjoy. Performers sing and/or play in a studio, together or separately, and their contributions are stored in "tracks," perhaps many of them. Then the recording engineer(s) and musicians "mix" the final product, adjusting the contributions of individual performers, perhaps altering the timbre of voices and instruments with equalization, and

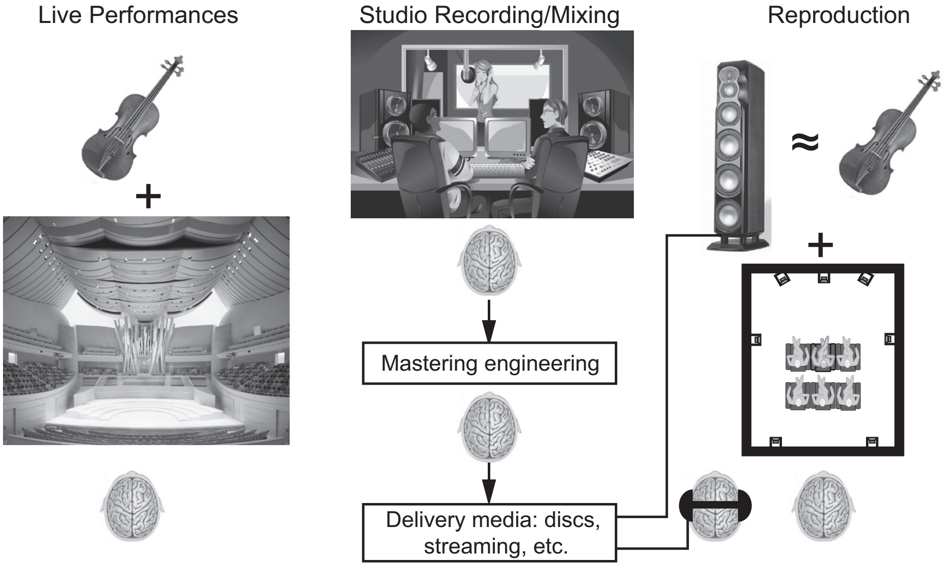


FIGURE 0.4 *A conceptual view of the parallel universes of live performance "originals," and the sound recording and reproduction "replicas," all interpreted by and influenced by two ears and a brain at several points along the way.*

Disney Hall photograph by Federico Zignani.

adding spatial effects: reflections, reverberation and so forth. These days even the pitch of off-tune singing can be corrected. All of this is evaluated as it happens by human listeners—ear/brain combinations—probably several of them, while listening to loudspeakers in a control room. The resulting mix may then be passed to a mastering engineer who applies personal subjective judgment to the musical product through different loudspeakers in a different room, changing it as necessary to fit the chosen delivery format(s). In the days of LPs the manipulations were substantial. The engineer tried to anticipate how the music may sound to the majority of customers in the majority of listening situations while embracing the properties of the delivery media. It would be ideal if there were no limitations to bandwidth and dynamic range, but not all media permit it, and few customers are equipped, or motivated, to enjoy such recordings. Often the audio product is tailored for modest playback equipment, possibly anticipating listening with a certain amount of background noise, and the sound is appropriately preconditioned. Consumers can know none of this, so what we hear from recordings, related to sound and spatial qualities, involves a measure of chance.

On the right of Figure 0.4 is the final step for those fortunate enough to have a listening room. Here is yet another opportunity for changing the audio product: different loudspeakers in different places in different rooms, using different electronics that may impose different signal processing and equalizations to the sounds radiated into the rooms. Again it is the human ear/brain system that generates the perceptions, and the opinions. Of course only a small percentage of people have such playback facilities.

Most make do with modest audio gear, or listen through headphones, the latter yielding a totally different sound experience from stereo recordings that are created for loudspeaker playback.

So, reviewing the process, there are several opportunities for significant personal input into the content of music recordings and how they sound on playback. Serious music lovers should attempt to experience live—unamplified—music performances if they wish to hear "reality." *No* recording, through any number of channels, can perfectly duplicate that reality, however good the playback apparatus. The vast majority of music recordings originate in studios, and before they reach us they have been subject to numerous subjectively guided manipulations, evaluated using loudspeakers of varying pedigree. Even supposedly "pure" classical recordings are massaged in the mixing and mastering processes to make them more pleasant when auditioned through two loudspeakers. The signal from a microphone focused on an acoustically weak instrument or voice may be spatially enhanced so that it fits into the acoustical setting of the entire orchestra. When well done, the trickery is not detected.

This scene may seem hopelessly disorganized, but it works most of the time. Music is very durable, and very likely all of us have spent hours enjoying music through seriously compromised audio systems. However, when a truly good recording is auditioned through a truly good playback system, it can be a spine-chilling experience. It isn't "reality," but it can be absolutely superb music and high-quality entertainment. A general principle for the audio industry should be: if in doubt, at least make it sound good.

Whatever happens in the continuing evolution of audio, it is helpful to understand the basics of the technology, the principles of sound propagation and the psychoacoustics of perception, because it is unlikely that any of those will change. The goal of this book is to provide knowledge and guidance for high-quality reproduction of existing audio formats, and to prepare the way for future developments.

CHAPTER 1

Sound Production vs. Sound Reproduction

Before getting into sound reproduction, the title of this book, it is good to have a look at where it all begins: in live performances—sound production. We may like to think that our audio systems are capable of reconstructing such experiences, but that is simply not possible. Even today, with nearly unlimited bandwidth available, two-channel stereo is the default format. There is no doubt that stereo can be greatly entertaining, and at times can make us feel close to the real experience. But it is sad to say that many recordings can be well described as: left loudspeaker, right loudspeaker and phantom center. This is mono, mono and double mono. Listened to through headphones it becomes left ear, right ear and center of the head. The essence of the music can be conveyed, but any semblance of acoustical space and ambiance is missing. Skillful microphone setup and signal processing by recording engineers can improve things, but at best stereo remains a directionally and spatially deprived format, and an antisocial one, requiring a sweet spot.

We need more channels to capture, store and reproduce even the essential perceptions of three-dimensional sound fields. This is what the movie world has known for decades, and now cinemas have as many as 62 channels in the immersive sound formats. That is excessive for musical needs, but more than two would be nice. Fortunately there are examples of excellent multichannel music, and indications that a binaural version of it will be a part of virtual reality systems. Stay tuned.

On the scientific side, the origins of modern acoustics lie largely in the domain of halls for the performance of classical music. Whether this music appeals to a person or not, the basic perceptions generated by these live performances are generously shared within all recorded music, whatever the genre. Reverberation, spaciousness, envelopment and so on are all simply pleasant perceptual experiences, and recording engineers have been provided with elaborate electronic processors allowing them to be incorporated into any kind of music, adding to the artistic palette. The future is sounding good.

1.1 LIVE CLASSICAL MUSIC PERFORMANCES—SOUND PRODUCTION

I have an approximation of a state-of-the-art sound reproducing system at home, and have always had such systems, gradually improving over the years. As good as they have been, the real thing is a very different and more satisfying auditory event. Dressing up, driving, perhaps eating out, crowds of similar-minded people, the overall visual atmosphere all add to the experience, but the aural components of the experience are the real treat.

I attend about a dozen live concert hall performances a year. Sitting, watching people take their seats and the orchestra assembling on stage, I am aware of a pleasing sense of a large space—I hear the space that my eyes see. This does not happen at home. As the musicians tune up and practice difficult passages, the timbres are enriched by countless reflections—repetitions that give our hearing system more opportunities to hear subtleties. I can localize individual musicians, and even though the sound begins at those locations, the timbres linger in the decaying reverberation all around me. When the music starts, it all magnificently comes together. The sound of the hall is an inseparable part of the performance: rich timbres combined with enveloping space. I am "in" the performance. This is a complex listening experience, allowing one to begin to discover what contributes to an engaging final product. However, it is interesting to consider that with different musicians, instruments, conductors and halls, the "reference" is really not a constant entity.

> In fact, one can correctly assert that a live concert hall performance is what it is at the time, and may never be repeated again. It is sound production.

It is interesting to note that, even in different halls, the essential timbres of voices and musical instruments remain remarkably constant. We have considerable ability to separate the sound of the source from the sound of the hall. In other words, we appear to adapt to the room we are in and "listen through" it to hear the sound sources. A variation on this interpretation is that we engage in what Bregman (1999) calls "auditory scene analysis," and we "stream" the sound of the voices and instruments as significantly separate from the sound of the room. We do this to such an extent that one can focus on the sound from one section of the orchestra, suppressing others. Two ears and a brain are remarkable. If a concert hall performance seems to lack bass, as some do, the inclination is to blame the hall, not the musicians or their instruments. We instinctively know where the blame lies.

Later in the book we will discuss the elements of these auditory events from both perceptual and measurement perspectives. For now, it is sufficient to note that reproducing a concert hall experience means delivering both the timbral and the spatial components. This is not easy.

Achieving satisfactory reproductions of these performances is partially determined by hardware: the electronics, the loudspeakers and the rooms. These are things

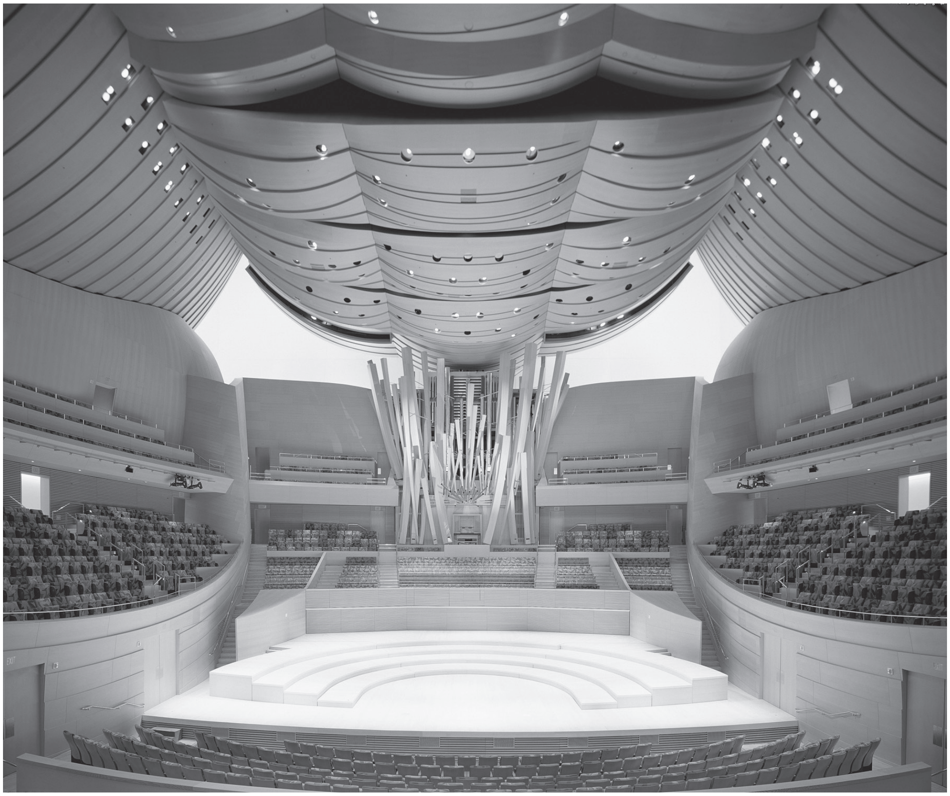


FIGURE 1.1 *Los Angeles County Music Center's Walt Disney Concert Hall Auditorium. Photo by Federico Zignani.*

consumers can choose and manipulate to some extent. But the most important contribution comes from the "system":

- The number of channels and the placement of the loudspeakers and listeners for playback.
- The microphone choice and placement, mixing/sound design, and mastering performed in the creation of the recording.

If the "system" does not allow for certain things to be heard, disappointment is inevitable. The existing system evolved within the audio industry itself and is being implemented by professionals operating within it. These are key factors in our impressions of direction and space, and there is persuasive evidence that spatial perceptions are comparable in importance to timbral quality in our overall subjective assessments of reproduced sound.

Time-domain information, the reverberation, is an important clue to the nature of the performance space. It can be conveyed by a single channel—mono—but it is a spatially "small" experience, with all sound localized to the single loudspeaker. Adding more channels allows for a soundstage, conveyed by the front channels: the lateral spread of the orchestra, with depth. There is also a more subtle component: apparent source width (ASW), or image broadening, wherein sounds acquire dimension, and an acoustical setting; some call it "air" around the instruments. Two channels suffice for a single listener in the symmetrically located "stereo seat," but multiple listeners or a single listener not in the stereo seat require the addition of a center channel to prevent the phantom image soundstage from distorting and ultimately collapsing to the nearer loudspeaker. The perception of being "in" the space with the performers—envelopment—is hinted at in good stereo recordings, absent from many, but is more persuasively delivered by multichannel recordings, with side and other channels providing long-delayed lateral reflections that contribute to both image broadening and envelopment. It is the length of the delays that creates the impressions of large space, something that limits the spatial augmentation possible with multidirectional loudspeakers in small listening rooms. In fact, envelopment is, according to some authorities, the single most important aspect of concert hall performance.

As good as stereo can be, more channels are better for a single listener, and most definitely for multiple listeners. Unfortunately the 5- or 7-channel options almost universally used in movies have been not been commercially successful in the music domain, even though they are capable of more engaging reproductions. The new "immersive" formats, employing even more channels and loudspeakers, were created for movies, where they provide exciting spatial dynamics, but some demonstrations of music programs provide compelling impressions of real concert halls or cathedrals, even as one moves around the listening room. The two-ears, two-channels relationship works for headphones, but is spatially deprived for loudspeaker reproduction. This topic will be discussed in more detail later in Chapter 15.

1.2 LIVE POPULAR MUSIC PERFORMANCES—SOUND PRODUCTION

As pleasurable as the classics are, the majority of our entertainment falls into the numerous subdivisions of what is collectively called "popular" music. The recording methods used are also shared with most jazz we hear. Most of what we listen to is captured by microphones located unnaturally close to voices and musical instruments, or electrically captured without any acoustical connection, and then mixed and manipulated in the control room of a recording studio. The "performance," the art, is what is heard through the monitor loudspeakers during the mixing and mastering operations. The era of large, expensive recording facilities is fading, as more recordings are done in converted bedrooms or garages in homes. The wide availability of sound mixing and processing programs, and the low cost of powerful computers, has given almost anybody access to capabilities that once were the exclusive domain of elaborate recording studio facilities. This paradigm shift is an important factor, expanding the repertoire of recorded music, changing the business model of the music industry itself, and liberating creative instincts that previously had been "damped by dollar signs."

Big-name artists engage in elaborate tours, spanning the globe in some cases. They became popular through recordings created in studios, and reconstructing the essence of that recorded "sound" in a concert situation is sometimes a goal. Occasionally, excerpts of studio recordings creep into the live performances. None of this is a problem because it was always an artificial creation, owing little or nothing to any live unamplified performance. It might seem like cheating, but there are some effects that cannot be duplicated in live performances. In the end, the delivered "art" is what matters.

Figure 1.2 shows that in live, amplified/sound reinforced, popular music performances, the front-of-house (FOH) engineer is in control of the performance heard by the audience. It is an artistic creation in real time, and can vary enormously with different engineers, their tastes and, interestingly enough, how well they hear. I left one show at intermission because the sound was far too loud, and not very good. I learned later that the FOH engineer was known by insiders to have serious hearing loss, but had a long relationship with the performer. Pity. This is a situation in which, for a variety of acoustical, technical and personal reasons, live popular music performances are variable.

One can correctly assert that a live popular music performance is what it is at the time, and may never be repeated again. It is sound production.

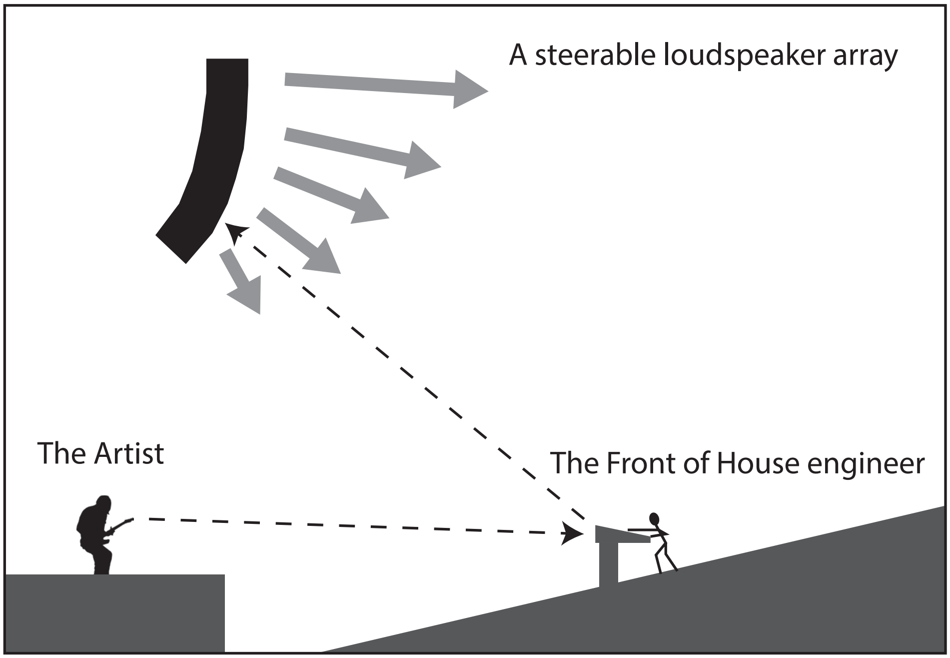


FIGURE 1.2 *A functional diagram of a tour sound system. The microphone and direct-wired inputs from the artists on stage are mixed and manipulated by the front-of-house (FOH) engineer, based on what is heard from the loudspeaker array, the sound of which is shared with the audience.*

1.3 REPRODUCED SOUND—THE AUDIO INDUSTRY

Sound *re*production is different. At some time and place, an original performance occurs, and the objective is to reproduce that performance with as much accuracy as possible whenever and wherever someone chooses to press a "play" button. Most of our listening experiences involve recordings, broadcasts or streaming audio that is reproduced through loudspeakers in a room, loudspeakers in a car or through headphones. This is the audio industry, as shown in Figure 1.3.

Clearly the audio industry is a complex operation, requiring extensive standardization if all of the devices within these different operations and business units are to be compatible with the signals moving through them. More importantly, there is the matter of what listeners in these varied situations hear. Is the result of the mixing and mastering engineering accurately delivered to the customers' ears?

> A fallacy: That reference to a "live" sound is the only way to judge sound quality. Reason: microphones capture only a sample of the live sound field that we would hear in a live performance. Parts are missing. Recording engineers can manipulate the mix to sound something like the live sound, to create a totally artificial experience or anything in between.

Figure 1.4 illustrates the enormous contrasts in scale involved in reproduced sound, and it is reasonable to think that the differences might be insurmountable. Can movies created for large cinemas be credible in small home theaters or television sets in tiny apartments? The delivery systems, the film and music formats also differ, and new options continue to evolve. Making it all work involves complex engineering of many

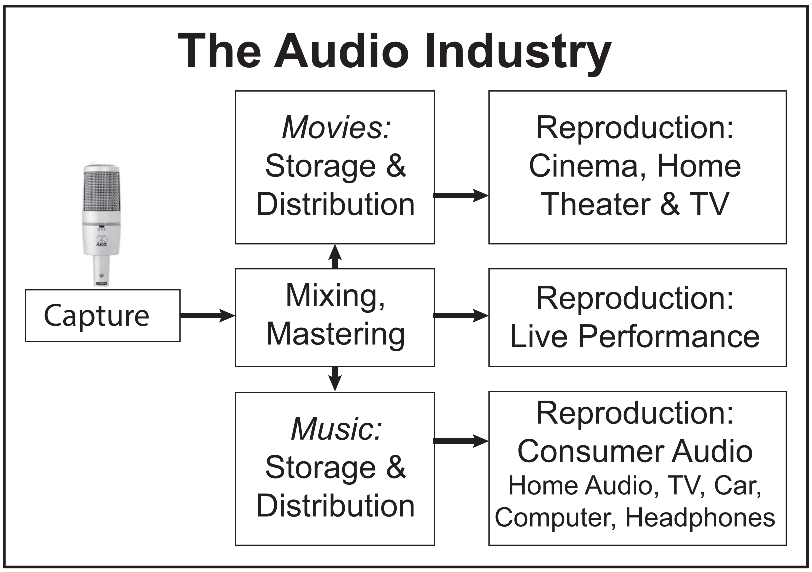


FIGURE 1.3 *The audio industry as we know it.*

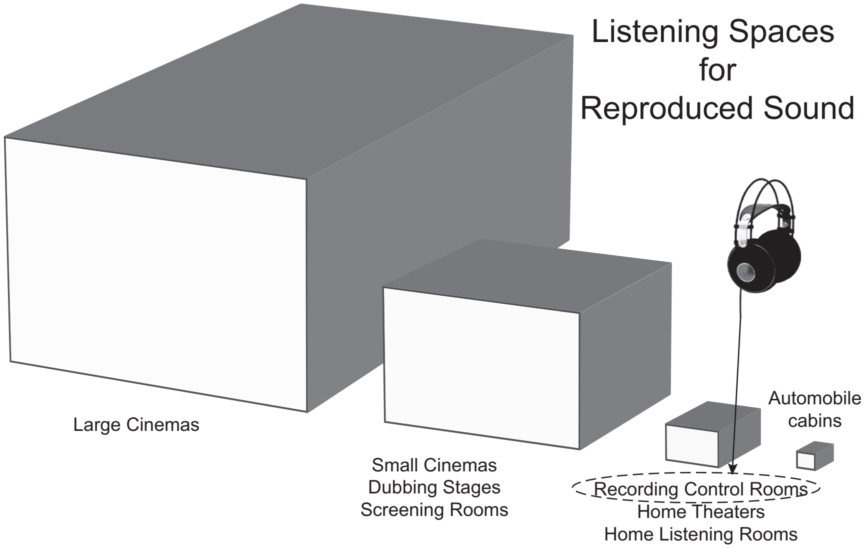


FIGURE 1.4 *The range of listening spaces within which accurate sound reproduction is desired.*

kinds, and some of it requires research to optimize the processes that deliver sounds to listeners' ears, making them appropriate for the circumstances.

Fortunately, cinemas are designed to have reverberation times that are not very different from domestic rooms. Combined with the large directional loudspeakers normally used, listeners end up in fundamentally similar sound fields and the experiences are acoustically more similar than might have been expected. This is discussed in detail later on.

Headphones normally replay stereo sound tracks created using loudspeakers in recording control rooms. Although it is possible to achieve a good timbral match, the spatial presentation is fundamentally different from that delivered by loudspeakers, sounds being localized primarily within or close to the head. For optimum headphone experiences, binaural (dummy head) recordings are needed, but we seem to have adapted well to the gross spatial distortions commonly experienced. The music survives.

The portrayal in Figure 1.5 could very well be wrong in detail, but there is little doubt that the trend is correct. The circumstances within which we are primarily entertained by audio programs have not had the proportional benefit of scientific research. As a result opinions, traditions and outright folklore have had and still have significant influence in these domains.

Audiophiles seeking excellence in sound reproduction are frequently advised by journalists and reviewers who have little or no awareness of the research that has been done. Opinions substitute for fact-based guidance. Most product reviews are done without the benefit of measurements, a situation that creates great uncertainty. Some publications perform basic measurements, but they are necessarily compromised in accuracy, resolution and comprehensiveness. Nevertheless, I respect their efforts, because they

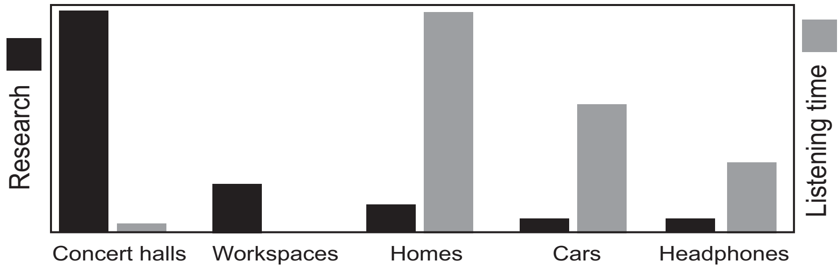


FIGURE 1.5 *A crude estimate of the quantity of scientific research devoted to various acoustical spaces and listening situations compared to the amount of time spent listening in each of them.*

are legitimate attempts to gain some technical insights. Facilities capable of anechoic or useful quasi-anechoic measurements are rare.

Some audio journalists are hostile to the very notion that audio is amenable to scientific investigation—asserting that only subjective opinion, preferably theirs, matters. There is nothing wrong with subjective opinions. As will be seen, it is these that have allowed for deliberate and productive research into what matters in sound reproduction. However, the opinions must be collected under circumstances in which the ever-observant human brain is deprived of information that could bias the opinion about the sound. The tests must be blind—the listeners must not know prices, brands, sizes and so forth. Ideally they should be double-blind, so the experimenter cannot bias the results. There should also be comparison sounds from competing products readily available, because humans are notoriously forgetful when it comes to recalling the details of sound quality. When this is done the subjective data begins to resemble technical measurements, in that they become impressively repeatable and, more importantly, generalizable to a large population. It is this discipline that has been lacking in audio in both the consumer and professional domains.

A motto for the audio industry:

Science in the Service of **Art**
is our **Business**

Good sound is our product.

Inspiration, invention and trial-and-error experimentation have taken us a long way in audio. But there is a growing body of scientific understanding that can take the industry even further. The great benefit of scientific knowledge is that it allows for predictable, repeatable results. Knowing what matters and what does not matter permits us to optimize the design of audio components that can deliver predictably good sound at affordable prices, and excellent sound at higher prices. If we acknowledge the message in the preceding motto, it is the *art* that matters, and because it is audio art, sound quality is an essential deliverable. Nobody gets chills down the spine from the metal box containing an amplifier, or the hand-rubbed finish on an expensive loudspeaker, but the perception of truly excellent sound can be a greatly pleasurable experience. That is our product. Providing the means for successful creation and delivery is the challenge.

1.4 PRESERVING THE ART—THE CIRCLE OF CONFUSION

If "good sound" is our product, how do we know what is "good"? Trying to answer that question plunges us immediately into what I have called "the circle of confusion." A little thought tells us that if consumers are to hear what the artists, musicians and recording engineers created, they should have similar loudspeakers and rooms. If not, they will be hearing something different. With no standards for monitor loudspeakers in recording control rooms, and no standards for consumer loudspeakers, sound quality is not predictable—it is a gamble.

When we listen to recorded music, we are listening to the cumulative influences of every artistic decision and every technical device in the audio chain. Many years ago I created the cartoon shown in Figure 1.6 for my tutorial presentations, illustrating how the never-ending cycle of subjectivity can be broken.

The presumption implicit in this illustration is that it is possible to create measurements that can describe or predict how listeners might react to sounds produced by the device being tested. There was a time when this presumption seemed improbable, and even now some people claim that we cannot measure what we hear. The reality is that with research and the development of newer and better measurement tools, it has been possible to move the hands of the "doomsday clock" to the point where detonation is imminent. In fact, it would be correct to say that the explosion has begun. Some aspects of audible sound are now more reliably revealed by technical data than by the *normal kinds* of subjective evaluations.

The consequences of this circle of confusion are apparent in both loudspeakers and microphones. I recently read of a microphone that had been "voiced" by a well-known industry person, but there was no mention of what loudspeaker he was listening to when voicing, so what does this mean? Many years ago, a loudspeaker manufacturer had a product that was favored as a monitor by a well-known classical music label. The loudspeaker was not neutral, and those recordings revealed a distinctive coloration when reproduced through more neutrally balanced loudspeakers. For a period of time, most of the loudspeakers made by this

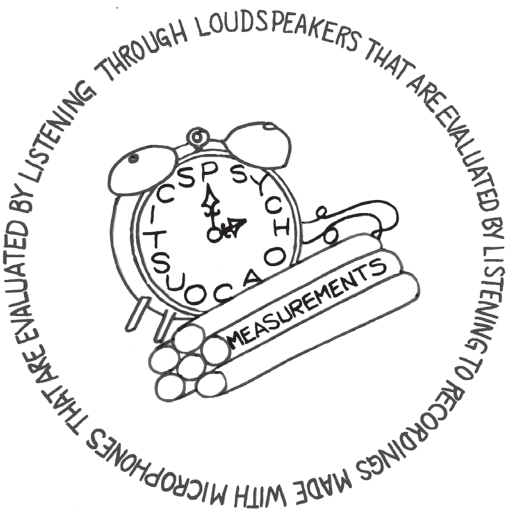


FIGURE 1.6 *The first version of the "circle of confusion" illustrating the key role of loudspeakers in determining how recordings sound, and of recordings in determining our impressions of how loudspeakers sound. The central cartoon suggests that the circle can be broken by using the knowledge of psychoacoustics to advance the clock to the detonation time when the "explosive" power of measurements will be released to break the circle.*