Tampere University of Applied Sciences

Headphone Development Research

Teemu Kinnunen

MASTER’S THESIS
May 2022

Master´s Degree Programme in Emerging Media

ABSTRACT

Tampereen ammattikorkeakoulu
Tampere University of Applied Sciences
Master´s Degree Programme in Emerging Media

KINNUNEN, TEEMU:
Headphone Development Research

Master's thesis 85 pages
May 2022

Modern headphones have been manufactured since the first half of the 20th century, but there are still uncertainties in the industry on how neutral-sounding headphones should sound and measure. Measuring devices that meet current industry standards are expensive, and devices from different manufacturers give varying results. The challenge of measuring headphones is one of the reasons why it is difficult to find a consensus.

This master's thesis is part of headphone product development for a company called MojoLab. The objective of the thesis was to find out whether existing headphones could be modified to make them more neutral in timbre. This would allow them to be used more widely in professional audio work. The thesis first goes through the human auditory system and how the understanding of it has developed over time. The history of the headphones is also reviewed, as well as how different measurement techniques and devices have evolved. The thesis will explore studies on different target curves and the listener preferences on them.

As a part of the research, different types of commercial headphones will be measured and there will also be a comparison of headphones measured by different methods. Listening tests will be used to see how listening perceptions work in conjunction with measurements. In-room responses are also used as a reference for headphones. In the main part of the thesis, everything that has been learned is applied and one headphone model is modified based on the measurement results and feedback from the users.

The modified headphones proved to be a success, and 50 units were produced during the project mainly for pro audio users. According to the feedback, the headphones are serving their purpose. They reproduce the entire hearing range, have a balanced sound, are comfortable to wear and are affordable. The goal for the future is to take advantage of the lessons learned from this project and design headphones from the scratch.

Key words: headphones, target curves, pro audio, measuring devices

CONTENTS

1 INTRODUCTION 5
2 HUMAN HEARING AND HEADPHONES 7
  2.1. Human hearing 7
    2.1.1 Equal-loudness contours 8
    2.1.2 Anatomy of the ear 9
    2.1.3 Ear resonances 11
    2.1.4 Spatial hearing 13
  2.2. Headphones 18
    2.2.1 History of headphones 18
    2.2.2 Types of headphone design 24
    2.2.3 General design criteria for headphones 25
3 MEASURING HEADPHONES 28
  3.1. International standards-making bodies 28
  3.2. Headphone measurement device standards 29
    3.2.1 Standards on using the MIRE technique 29
    3.2.2 Standards on ear simulators and HATS 30
  3.3. Hearing simulators and other measuring devices 31
    3.3.1 Evolution of hearing simulators 31
    3.3.2 Common simulators for testing headphones 34
  3.4. Target curves for headphones 36
    3.4.1 Free field response 37
    3.4.2 Diffuse field response 39
    3.4.3 Harman target curves 40
4 PRODUCT DEVELOPMENT PROJECT 46
  4.1. Equipment 46
    4.1.1 Measurement device 47
    4.1.2 Software 48
  4.2. Background study of headphones and in-room responses 50
    4.2.1 Comparison of measuring devices and methods 53
    4.2.2 Preference in-room responses 56
  4.3. Headphone modification project 63
    4.3.1 Revisions of the modified headphones 64
    4.3.2 Results 67
5 CONCLUSIONS AND DISCUSSION 71
REFERENCES 77

ABBREVIATIONS AND TERMS

AES	Audio Engineering Society
AGC	Automatic gain control
ANC	Active noise cancelling
ANSI	American National Standards Institute
BTE	Behind the ear
CENELEC	European Committee for Electrotechnical Standardization
CTF	Common transfer function
DB	Decibel
DF	Diffuse field
DIN	German Institute for Standardization
DTF	Directional transfer function
DRP	Eardrum reference point
DSP	Digital signal processing
EEP	Ear canal entrance point
ERP	Ear reference point
ERB	Equivalent rectangular bandwidth
FF	Free field
FIR	Finite impulse response
HATS	Head and torso simulator
HRTF	Head-related transfer function
HZ	Hertz
IEC	International Electrotechnical Commission
IHL	In-head localization
ILD	Interaural level difference
ISO	International Organization for Standardization
ITD	Interaural time difference
ITU	International Telecommunication Union
MIRE	Microphone in the ear
SRF	Semi reflective field

1 INTRODUCTION

This master´s thesis is a part of the headphone product development that was done at the request of MojoLab (2022). MojoLab is a company in the field of audio mastering and acoustic consulting services. MojoLab was founded in 2008, when the author graduated from Pirkanmaa University of Applied Sciences with a bachelor's degree in culture and art. The bachelor’s thesis was about starting up a company to this field of business. This master’s thesis brings together the acoustic expertise and critical hearing skills developed over the years and further refines them in interesting new directions.

Many hi-fi and pro audio headphones have been used in MojoLab's mastering studio, but none of them have really been used for anything other than to spot edit cracks or other low-level noises. Most of the headphones have been overly colored in timbre or have been deficient in frequency bandwidth. Coloration may be desirable in hi-fi listening, where the goal is more of just an enjoyable listening experience, but the problem for professionals is that if the headphones make the sound too pleasant and hide the truth, they may not make the right decisions. Reliable headphone monitoring would also be needed to allow for minor editing or fine-tuning outside the studio. It would save time if some changes could be made without always having to go to the studio. These issues formed the objective of the thesis, whether the existing headphones could be modified in some way to make them more suitable for MojoLab.

The purpose and practical measures to achieve the objective of the thesis was to study human hearing, the acoustics of headphones and what measurement methods exist and whether there is a common target curve used by manufacturers. The standardized measuring devices used in research in this field are expensive, but can these hearing simulators accurately simulate the complex human auditory system? To understand how neutral and truthful headphones should sound, one should first understand how neutral loudspeakers should sound in a well acoustically treated listening room. Since we already have some background knowledge on this, this aspect will only be covered shortly that it does not overstep the scope of this thesis.

By studying the acoustics of headphones, and through trial and error, different solutions were found to modify the sound of headphones in many different areas. Although the initial idea was to customize the headphones only for the use of MojoLab, many other audio professionals were also interested in the headphones because they had similar needs. So, prototypes were also made for other professionals in the same field. This also provided a broader view and some research data on if the people’s preferences differ greatly when looking for neutral headphones for mixing and mastering in addition to studio loudspeakers.

2 HUMAN HEARING AND HEADPHONES

The human auditory system is a complex system. Hearing and its associated organs are also individual and therefore affect the way a person hears things. Headphones can be used to listen to a signal source more privately than loudspeakers. The acoustics of headphones and how the sound is perceived are also slightly different from that of loudspeakers. This chapter discusses the hearing system and its characteristics, as well as headphones and how headphones have evolved over the years.

2.1. Human hearing

The human auditory system can hear frequencies from about 20 Hz to 20 000 Hz. The human ear is most sensitive between 2 000 – 5 000 Hz, largely due to resonances in the ear canal. Sensitivity of hearing varies with frequency, but also loudness affects sensitivity. Sound pressure is measured on a decibel scale (dB). Normal hearing is defined as a hearing threshold of -10 to 15 dB at all frequencies. The ear can detect sound pressure levels of up to 130 dB, at which point it becomes painful, such a high sound pressure may be created by, e.g., a fighter jet’s afterburner.

Frequencies below 20 Hz are called infrasonic frequencies. Humans can hear those if the level is sufficiently high. According to Henrik Møller and Christian Pedersen from Aalborg University: “The ear is the primary organ for sensing infrasound, but at levels somewhat above the hearing threshold it is possible to feel vibrations in various parts of the body.” (Møller & Pedersen 2004, 37.)

The upper hearing limit usually decreases with age. This age-related hearing loss is called presbycusis. Age-related hearing loss occurs equally in both ears, and the loss is gradual, so you might not realize that hearing is impaired. Presbycusis commonly arises from changes in the inner ear as we age, but the causes can also be the changes in middle ear or along the nerve pathways from ear to the brain. (NIDCD 2016, 1.)

2.1.1 Equal-loudness contours

An equal-loudness contour means that a constant sound pressure level is perceived when the listener is presented pure steady tones over the frequency spectrum (Suzuki & Takeshima 2004). In 1927 at Bell Telephone Laboratories, B.A. Kingsbury made first research on equal loudness contours (Suzuki et al. 2003). But it was PhD Harvey Fletcher and Wilden Munson, also from Bell, who conducted research on equal loudness contours in 1933 that became standard for many years. In their research, tests were done with only 11 subjects, and the test were done with headphones with test tones ranging from 62 to 16 000 Hz. Although the tests were done with headphones, the data was converted to correspond to what would have been measured in a free field (Fletcher & Munson 1933, 87–89). Fletcher and Munson pioneering work in 40 phons contours are used as the basis of the A-weighting function and the term Fletcher-Munson curves is still used when people are talking about equal loudness contours.

Robinson and Dadson (1956) made their own research about the subject in 1956, and the results were standardized as ISO 226. In their study they had a larger number of subjects, 90 persons ranging from 16 to 63 years old. The test range was 25 – 15000 Hz, and the tests were done in a free field anechoic room with a loudspeaker in front of the test subject. In 1980s, several research papers reported that ISO 226 contained a large error. (Suzuki et al. 2003.) ISO 226 standard was revised in 2003, and new ISO 226-2003 version was based of 12 studies carried out between 1983 to 2001. Comparison with previous contours is shown in Figure 1.

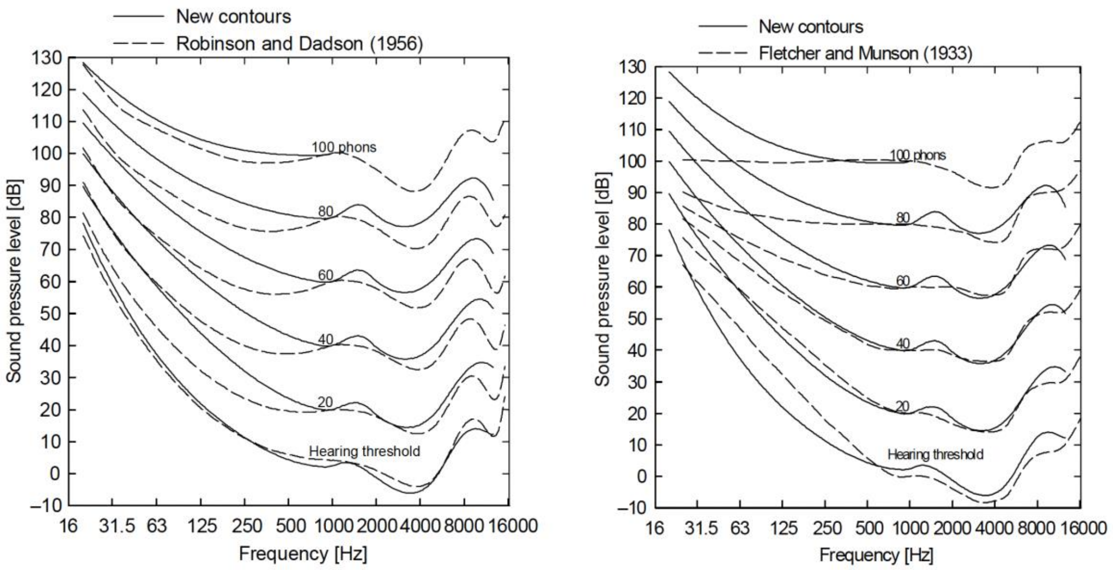

FIGURE 1. Comparison of the equal loudness contours (Suzuki et al. 2003)

Hearing is most linear at the 80 dB loudness. In Figure 2, the equal-loudness contours have been normalized to 80 dB, so we can see how the perception of lower frequencies depends very much on the loudness of the sound. This is important to understand when making frequency balance decisions by ear.

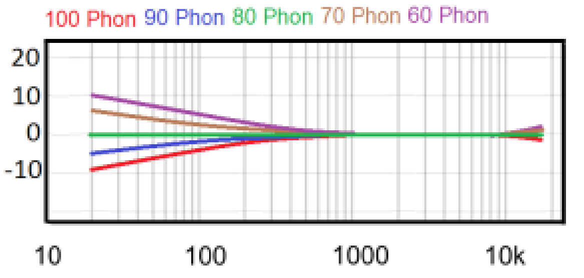

FIGURE 2. Equal-loudness contours normalized to 80 dB average SPL (Diy-audio-heaven 2019)

2.1.2 Anatomy of the ear

The ear consists of three main parts that we can see in Figure 3. The outer ear comprises the visible part of the ear called auricle or pinna. In the middle of the pinna, there is a horn-like cavity called the concha that leads to the ear canal. Pinna and concha together direct the sound to the ear canal, and their shape plays an important role in how a person hears the sound and from which direction the sound is perceived to come. The ear canal, which leads to the eardrum, belongs also to the outer ear.

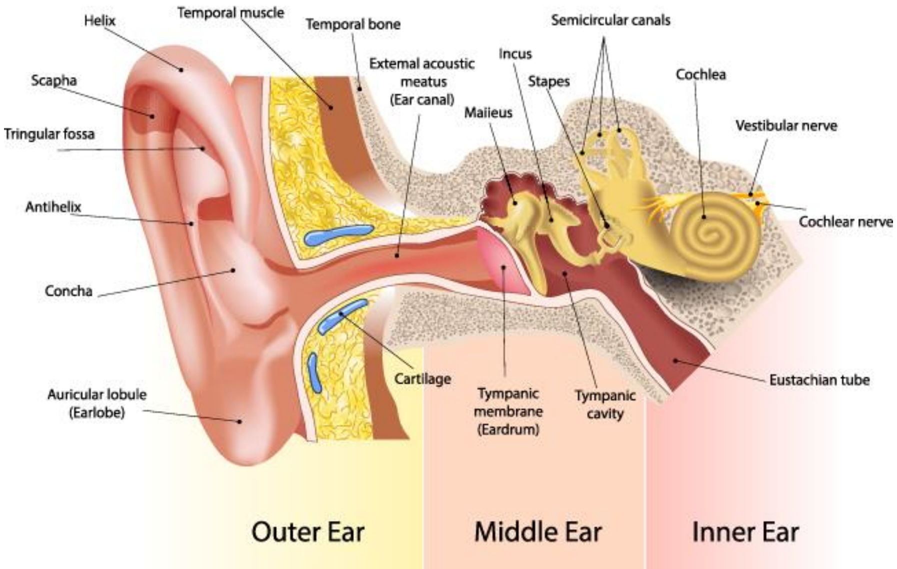

FIGURE 3. Anatomy of the ear (Shutterstock 2022)

Eardrum, also called tympanic membrane, separates the outer ear from the middle ear. The middle ear consists of the eustachian tube, and three small bones called ossicles. The eustachian tube links to the back of the nose to equalize the pressure needed for the proper transfer of sound waves. Ossicles are named malleus, incus and stapes, and these transmit and amplify the vibration of the eardrum to the inner ear. The inner ear consists of cochlea, vestibule and semicircular canals. Spiral-shaped hollow bone cochlea converts sounds into nerve impulses that are interpreted by the brain. (Stanford Children´s Health n.d.)

In addition to being a hearing organ, the ear is part of the sense of balance. The proprioceptive system, which indicates the position of different parts of the body in relation to each other, forms the sense of balance, together with the vestibular system and the visual system. Vestibular system is located in the inner ear. It consists of three semi-circular fluids containing canals that respond to rotation and two otolith organs that sense linear acceleration. (Cullen 2008.)

If we look at the outer ears of different people, we can see that there are many different shapes and sizes. The result of this is that people’s hearing and ear resonances are not the same between different people. In Emeritus professor

Abraham Tamir's study, he examined a total of 2 425 different ears and divided them into 36 different categories based on the shape of the concha (Figure 4). The most common shape was number 29, which was found in a total of 393 copies (16.2%). (Tamir 2017, 3.)

FIGURE 4. Different shapes of concha (Tamir 2017, 3)

There are also many studies in which the ears and concha have been modified with, e.g., silicone mold and the test subjects locational hearing has adapted to the changes after a while. The brain plays a big role in this, and the human hearing can adapt to modified spectral cues. (Trapeau & Schönwiesner 2018.)

2.1.3 Ear resonances

The ear forms various resonances due to the structures of the ear. One of the largest resonances occurs in the ear canal, which is tubular in structure, averages 26 mm in length and 7 mm in diameter and acts as a quarter wavelength resonator (Hiipakka 2008, 6).

Resonance frequency can be calculated by quarter wavelength formula

$$f = \frac{\mathrm{v}}{4L}$$

in which *f* is the root resonant frequency, *v* is the speed of sound (340,000 mm/s) and *L* is the length in mm.

There are also odd-numbered resonances with multiples of 3 / 5 / 7 of the root resonance. With the previous formula, the ¼ wavelength resonant frequency of the 26 mm long ear canal is theoretically about 3,3 kHz, with odd numbered multiples in 10 kHz, 16,5 kHz and so on.

As seen in Figure 5, the resonance of the ear canal is about 10 dB at 2 500 Hz. The exact frequency is determined by, among other things, the length of the ear canal, and it is the most characteristic resonant of the HRTF (Head-Related Transfer Function). Another large resonance is generated in the concha, the horn-like structure of which collects sound and directs it towards the ear canal. This resonance is of the order of about 9 dB and settles at 5 kHz. Another influencing factor is the slight accentuation due to the reflections of the pinna at 3 kHz. At higher frequencies the reflections begin to form a slight comb filter effect. The head, body and other parts of the body also contribute to resonances. All these together form an ear resonance of about 17 dB at the center frequency of 2700 Hz, where real-ear audiological measurements generally place the ear resonance (Staab 2014).

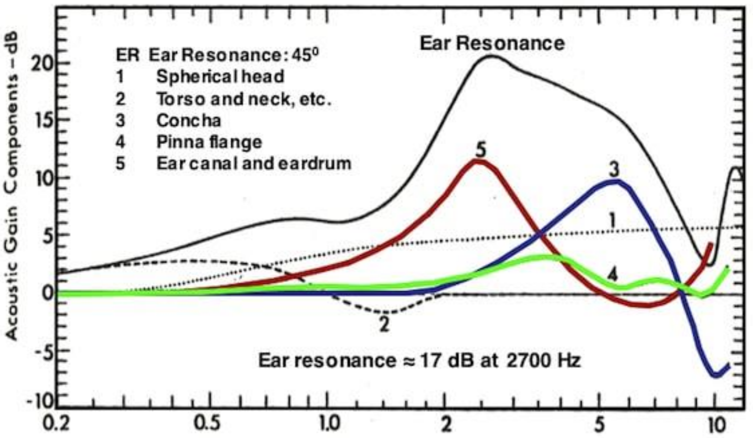

FIGURE 5. Combination of different resonances caused by the ear and the human body at 45 azimuth, x -axis in kHz (Staab 2014)

Attenuation in frequency response called “Pinna Notch" can be seen usually around 10 kHz but depending on the individual it can vary. At these high frequencies, the reflected signal from pinna is out of phase with the direct signal, and destructive interference occurs. It also has a greater effect on the sound coming from above than the sound coming directly from the front, and this can help to localize vertical sound. (University of California Davis 2011.)

2.1.4 Spatial hearing

When we talk about spatial hearing, we mean the ability of the human auditory system to interpret and exploit the different spatial pathways along which sounds travel towards the listener. The modern study of the location of sound sources can be said to have begun in the late 19th century. John William Strutt, also known as Lord Rayleigh, conducted a 'garden experiment' in which his students and staff gathered in circle around him to play pure-tone stimuli with vibrating tuning forks. He reasoned that a binaural ratio of sound level at each ear could explain his ability to identify the location of people in the garden (Yost 2017).

In his Duplex theory from 1907, he states that at frequencies above 500 Hz, the interaural level difference (ILD), due to acoustic shadowing by head, gives a clue to the lateral position of the source. At lower frequencies, the sound wavelength is greater than the head diameter, so ILD differences are insignificant. In contrast, humans are sensitive to interaural time differences in the ongoing phase of low frequencies, so ITDs can give us clues about the lateral location of lower frequencies. (Macpherson & Middlebrooks 2002.) Binaural cues ITD and ILD are illustrated in Figure 6.

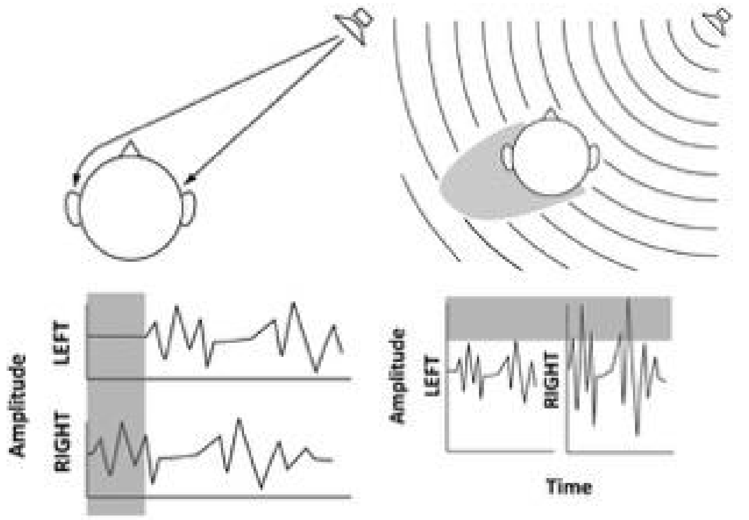

FIGURE 6. Interaural Time Delay (ITD) and Interaural Level Differences (ILD) (Sun, Zhong & Yost 2015)

Rayleigh could not discriminate the front-versus-back locations of pure-tone stimuli, but Jens Blauert conducted an interesting study in 1969 examining sound localization in the median plane. He found out that certain frequency ranges are localized either front, back, or on the top of the listener, regardless of the actual location of the audio source. As seen in Figure 7 narrow band sounds in 300 Hz and 3 kHz localized to front, 1 kHz and 10 kHz to back and 8 kHz to the top of the listener.

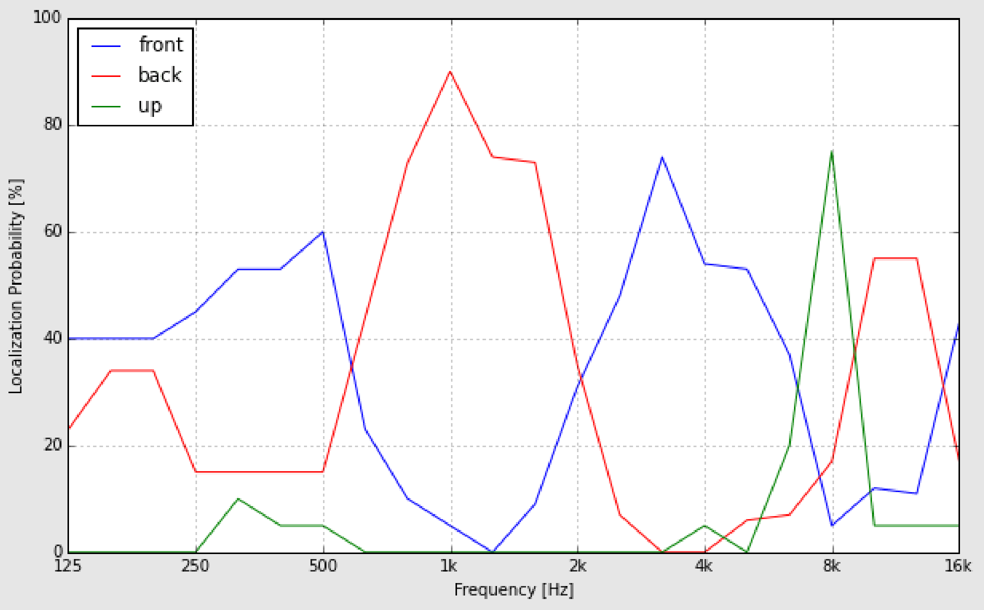

FIGURE 7. Localization of narrow band sounds in the median plane (Luther 2022, reproduced from Blauert 1969)

In his doctoral thesis, Klaus Genuit proposed (1984) for the first time that the HRTF be presented as a combination of different anatomical structures, e.g., head and pinnae would have their own filter structures. The Lopez-Poveda and Meddis study (1996) modeled the concha more accurately as a spinal cylinder and explained the "pinna notch" behavior. The Shaw (1997) deepened the pinna model and provided more detailed information on resonances and high-frequency HRTF behavior. (Algazi, Duba, Morrison & Thompson 2001, 103.)

Human torso, head and pinna reflect and scatter sound waves on their way to the eardrum differently depending on the angle from which the sound arrives. These head-related transfer functions (HRTF) help us to determine the direction of the sound. HRTF consist of two components that are unique to each listener. The Directional Transfer Function (DTF) consist of directional cues caused by the pinnae and shoulders, and the Common Transfer Function (CTF) consist of sound coloration cues caused mainly by the longitudinal resonance of the ear canal. (Bomhardt 2017, 12.)

According to PhD David Griesinger (2009), human hearing uses mapping features to identify elevation and azimuth. The frequency maps used for the identification of elevation and azimuth as well as timbre correction are fixed. Because a human must be able to identify the direction from sounds as short as 1 ms, these must be fixed. After the match to a particular spectral cue has been found, neural processing corrects the timbre. A simple model of this can be seen in Figure 8.

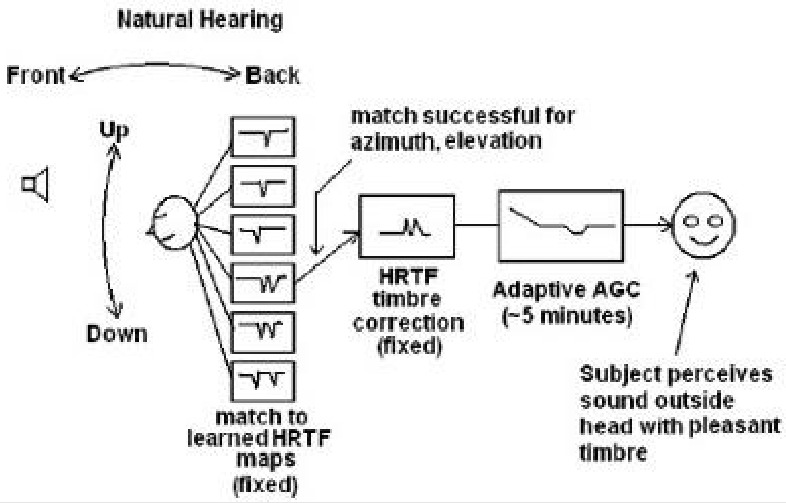

FIGURE 8. Simple model of human hearing (Griesinger 2009, 2)

After HRTF timbre correction there is adaptive gain control, with a time constant of several minutes. After a few minutes of listening, automatic gain control (AGC) tends to make the perceived loudness of the frequency spectrum constant, and even large errors in frequency response of for example headphones are reduced. (Griesinger 2009, 2–3.)

HRTF measurements can be taken at many different points. The most common points are the drum measurement point (DRP) and the ear canal entrance point (EEP) shown in Figure 9. Measurements can be done also between the points. As the different points give different results, this must be considered when examining the results. This also makes it difficult to compare measurements taken from different points.

The measurement requires a miniature microphone to be placed at the desired point in the human ear. Both directional cues and ear resonances can be detected by measuring at the DRP. Directional cues can be detected at EEP, as Wiener & Ross noticed already in 1946 when they made the first studies about linear distortions caused by head, pinnae, and ear canal (Blauert 1969). Due to the measurement deep in the ear canal and close to the eardrum, the DRP measurement may be slightly uncomfortable and may cause damage to the ear if caution is not exercised.

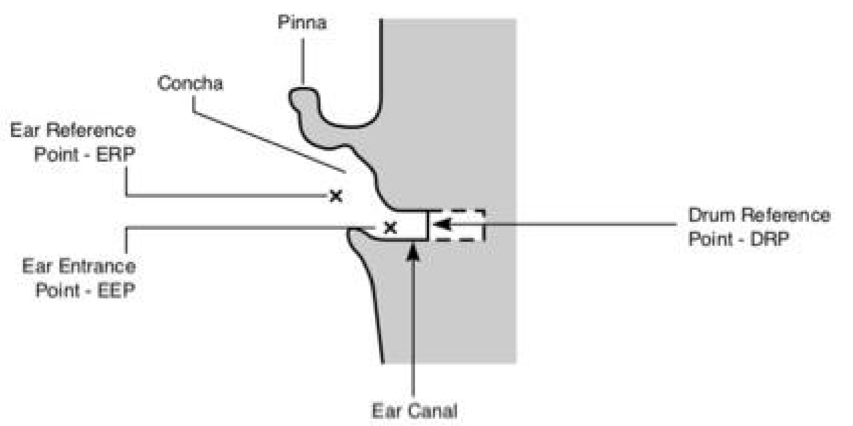

FIGURE 9. Ear measurements reference points (Jønsson, Matthisson & Borg 1997, 2)

Møller (1992) modeled the sound transmission within the ear canal. This was valid up to 10 kHz, and the conclusion was that all localization cues can be measured from blocked ear entrance point. Later Hammershøi and Møller (1996) verified this concept and further refined it so that the acoustic transfer pathing from the sound source to the eardrum can be divided into three parts: Transition from sound source to blocked entrance of the ear canal, transition from blocked to open entrance of ear canal and transition along the ear canal. For this reason, the recommendation is that binaural recording should be done in blocked ear canal entrance as it contains all localization cues but minimum individual information. (Hammershøi & Møller 1996, 408.) Algazi, Avendano and Thompson (1999) later tested the method and found it to be successful. Thus, measurement performed in a blocked ear canal is the generally accepted method of performing HRTF measurements today. (Li & Peissig 2020, 12.)

To produce spatial sound for example in virtual reality (VR) headsets, devices typically use generic HRTF profiles. Generic profiles do not work as well as individual HRTF profiles, because everyone has a unique profile. This has slowed down the breakthrough in VR systems as the problem of individual differences in HRTFs has not yet been solved (Iida 2019, 5).

The Finnish pro audio manufacturer Genelec has made customized HRTF profiles for a few years with their Aural ID service. After sending them a 360 -degree video of your head and torso, they analyze it and produce two SOFA files (Spatifially Oriented Format for Acoustic) that you can use in binaural rendering plugins. (Corbett 2020). One version includes only DTF effects, and it can be used preferably with diffuse field equalized headphones. Full HRTF version is for headphones that are equalized flat response on-head, because Aural ID does not include ear canal resonance. (Genelec 2020.)

2.2. Headphones

Headphones are a device that contains a pair of small loudspeaker drivers that are worn in the ears, on the ears or around the ears. Headphones can be used to listen audio sources more privately, unlike loudspeakers which play the audio also to others nearby.

2.2.1 History of headphones

The roots of headphones lead to the telephone technology invented by Alexander Graham Bell in 1876. The earliest incarnations of headphones could be found in 1880 when inventor Ezra Gilliland, who worked also with Bell Telephone Company and Thomas Edison, combined a Blakes telephone transmitter and associated receiver into a “Gilliland harness”. This was the first time the receiver was not held to the ear but was on the shoulders of telephone operators and attached to the waist. Although this device (Picture 1) let operator work in handsfree, the device was bulky and weighted almost three kilograms (Beardsley 1929, 206).

PICTURE 1. Gilliland harness (Beardsley 1929, 204)

In 1891, French engineer Ernest Mercadier patented (U.S. Patent No. 454,138) stethoscope-like in-ear headphones called bi-telephone. They were also designed for telephone operators, but hey weighed only a fraction of the previous designs. They allowed the user to wear the headphones while moving, and they had rubber covers that fitted securely into user’s ear canal. This invention (Figure 10) can therefore be seen as the ancestor of current earbuds.

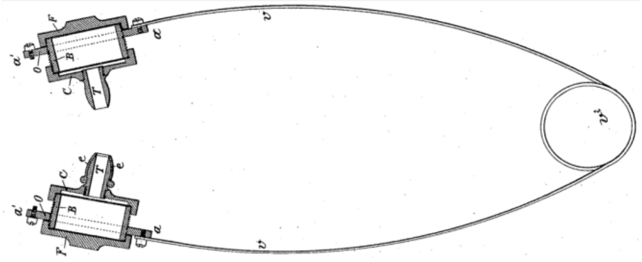

FIGURE 10. Bi-telephone drawing in patent application (Mercadier 1891, 1)

The first consumer headphone inventions date back to 1890, when the British Electrophone was invented. The device looked like a stick with stethoscope-like receivers at one end, which the user had to place over their ears. (Picture 2). With Electrophone, you could call a switchboard operator, who would then connect you to live performances or Sunday church services, e.g., for an annual subscription fee of around £5. In 1896 Electrophone had about 50 subscribers, and by the 1906 there were fourteen different theatres and fifteen churches which you could connect. Its subscriber base peaked in 1923 with over 2 000 subscribers, but when wireless radio receivers started come to market by 1924, Electrophone lost its subscribers and went out of business in 1925. (Loeffler 2020.) So, while one might think that Apple Music or Spotify have invented streaming platforms, Electrophone has pioneered the industry, at least to some extent.

PICTURE 2. Electrophone receivers (Britishtelephones.com)

The first device resembling modern headphones was made by inventor Nathaniel Baldwin in 1910s (Figure 11). These headphones consisted of two super-sensitive telephone receivers attached to the operator's headband. Each receiver had a mile of fine copper wire and a mica diagram. Baldwin was inspired to develop this sensitive receiver in a general conference of the LDS church, where he had difficulties in hearing the speaker because of the quiet sound system. He set out to develop a voice amplification device, which led to the development of this sensitive telephone receiver. (Singer 1979, 46–47.)

Baldwin patented the telephone receivers (US patent no. 957,403) but had trouble selling the device until he sent it to U.S. Navy for testing. The test request was not taken seriously at first because of the clumsy looking device, but eventually they tested it and informed Baldwin that the electrical properties were exceptional. They were twice as sensitive as any other devices they had tested before. Thus, they wanted to order more of these if Baldwin could make changes to the impractical head harness design while keeping the price the same. Baldwin accepted this, but because he made headphones at home on his kitchen table, he made contracts in delivering headphones as batches of ten. After several contracts, the Navy asked if he could further fine-tune the headband. Baldwin developed a new simpler headband consisting of two leather-covered spring wire rods, which allowed the spring-loaded driver units to be moved up and down. This lightweight and practical design was admired, and the Navy urged Baldwin to patent it, but he replied that it was a trivial invention and not worth the trouble. (Howeth 1963, 149-150.)

However, it seems that later in 1915 Baldwin also patented this headband design for his headphones (U.S. Patent No. 1,127,161). The design is identical to that of modern headphones, so it can be said that the current design of headphones comes from Baldwin’s kitchen table due to the collaboration between him and the U.S. Navy.

FIGURE 11. Nathaniel Baldwin receivers type C “Navy standard” (McMahon 1973, 192)

Modern consumer type of headphones originates to 1930s. Eugen Beyer, who was born in Sweden but lived his early life in Russia, was on holiday in Berlin when the World War I started. Eugen and his family were unable to get back to Russia, so they had to leave everything behind and get to Sweden as refugees. In 1921, Eugen settled to Berlin where his ancestral roots were. Eugen’s passion was to add sounds to motion pictures, and he started his company Elektrotechnische Fabrik Eugen Beyer in 1924. (Beyerdynamic 2022.) The company developed rapidly and later changed its name to Beyerdynamics. They built sound systems to the cinemas, and Eugen convinced all major film distributors in German that silent cinema has no future. In 1931, there were no cinemas left without sound system (Sudonull 2017).

With his good instincts and innovator mind, Eugen started to develop headphones to the market. In 1937, Beyerdynamics unveiled the world’s first dynamic driver headphones, the DT48. DT stands for “Dynamic Telephone”, and the model remained in production until 2013 with very few modifications (Figure 12). It was closed supra-aural (on-the-ear) design and mono in the early versions, and it was popular by sound engineers and reporters. The DT letter combination have continued to be used in their PRO series headphones even today. (Beyerdynamic 2022.)

FIGURE 12. Beyerdynamics DT48 evolution

At the same time across the border in Austria, two Viennese men, Dr. Rudolf Görike and engineer Ernst Pless, had also been supplying technical equipment to cinemas. In 1947 they formed Akustische und Kino-Geräte Gesellschaft m.b.H., better known by its abbreviation AKG. In 1949, they began to produce their first headphone model, named k120 dyn. (AKG 2022.) The model became popular, and the company focused later purely on pro audio.

After the first commercial stereo long-play record by Audio Fidelity Records hit the markets in 1957, it forced the other record labels to follow their lead. There were no devices on the market that could play stereo records yet, but they started to appear soon after records appeared. In 1958, American jazz musician John Koss teamed up with engineer Martin Lange Jr. to develop portable stereo phonograph player Koss Model 390. It included stereo loudspeakers but also a unique “privacy switch”. This allowed the listener to connect their Koss SP-3 headphones to the player and enjoy audio “silently”. (Koss 2022.) Koss SP-3 were the first stereo headphones, and the design was closed circumaural (around-the-ear) with 3” drivers and foam earpads to comfortable wear.

Headphones had dynamic drivers until 1959, when Japanese Stax produced the first electrostatic headphone model, the SR-1. Stax called its headphones "Ear speakers", because this level of Ultra-High Fidelity could only be compared to the best no-cost loudspeakers (Stax 2022). At this time, the headphones had been quite bulky in design. In 1968, the German company Sennheiser released the first lightweight headphone model, the HD 414. These were supra-aural headphones with new “open-aire” technology, featuring exciting bright yellow foam pads. More than 10 000 000 units were sold, and the headphones were at least in 2012 still the best-selling headphones of all time (Sennheiser 2012).

In 1972, the next breakthrough in the industry was seen when the British company Wharfedale produced the first isodynamic headphone model, the ID1. A few years later, the Fostex T50v0 and Yamaha HP-1 headphones were also released. Yamaha described their models as orthodynamic, while Fostex called their models as "regular-phase" headphones. However, it can be said that all these three concepts represent the same technology. Today, the technology is typically referred to as planar magnetic.

The industry evolved and the portability of products became a bigger priority. The Japanese company Sony introduced the Walkman brand of portable audio players in 1979, which included very small and portable headphones. They featured Sennheiser's licensed “open-aire” technology and were the first headphones with a 3.5 mm plug. As headphones became smaller, in-ear and earbud headphones also appeared on the market in the 1990s, peaking in the 2000s with the launch of Apple's iPod.

Although Active Noise Cancelling (ANC) headphones have been on the market for a relatively short time, they have a long history. The first ANC patent was granted to Paul Lueg, a German Doctor of Philosophy and Medicine, in 1933. Although the technology did not yet allow its implementation, he had already recognised the need for it. The patent (U.S. Patent No. 2,043,416) got him into serious trouble with the Wehrmacht, the unified armed forces of Nazi Germany. That eventually forced him to give up his research. (Guicking 1990). The first ANC headphones were developed in the 1950s by Dr. Lawrence Jerome Fogel. They were designed to reduce the noise experienced by helicopter pilots and thus improve the quality of communication. It was not until 2000 that Bose released the first consumer ANC headphones, the Bose QuietComfort QC1. Until then, ANC headphones were designed mainly for the aviation industry.

Wireless headphones were also seen as early as the 1960s, when "radio headphones" allowed you to listen directly to the radio. However, the great revolution in wireless headphones began in the late 2000s with brands like Bose and Beats. In 2020, nearly 550 million headphones were sold (Figure 13). This number has increasing despite the impact of COVID-19, and headphones have become the most popular commercial audio product today.

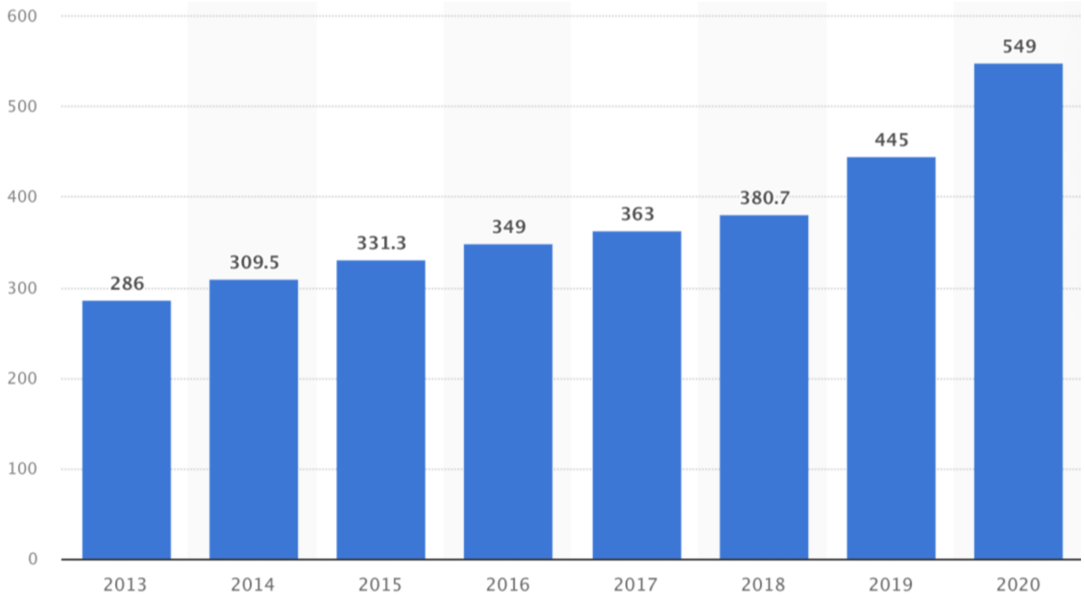

FIGURE 13. Unit shipments of headphones worldwide from 2013 to 2020 (in millions) (Statista 2021)

2.2.2 Types of headphone design

Headphones can be roughly divided into four categories: circumaural, supra-aural, intra-concha and insert headphones. Circumaural headphones are commonly called over-the-ear or around-the-ear headphones because the ears fit completely inside the pads of the headphones. Supra-aural headphones are known as on-the-ear headphones because they rest on the earlobes. Their lightness and small size make them well suited for portable use, unlike the larger circumaural headphones. Intra-concha headphones or, more commonly, earbuds are intended to be placed in the concha cavity, while insert headphones (in-ear-monitor / IEM) are inserted directly into the ear canal. (ITU-R 2021, 2–4).

As mentioned in the previous section, there are different types of headphones in terms of how well they isolate external sound. Closed headphones isolate sounds from the outside, and they do not let out the material you are listening to. These are popular for recording use, but also when a little privacy is desired, e.g., when walking in public. Open headphones, on the other hand, do not isolate the sound, so the sound can be heard by others. This also ensures that the listener can better hear what is going on around them, which can be desirable in certain situations. The closed headphone design often makes low frequency reproduction easier to achieve, but also causes reflections inside the headphones that can slightly interfere with the reproduction. Semi-open headphones are again a model between the two, with the outward sound slightly attenuated.

Headphones can be divided into three main categories according to the type of transducer that converts electrical signals into sound. Dynamic headphones have small cone or dome shaped drivers with moving voice coils. Dynamic drivers are used in the majority of headphone designs due to the low cost. Electrostatic headphones, on the other hand, are generally more expensive to manufacture. They do not have a moving voice coil like in dynamic headphones, but a very thin film membrane between two perforated and conductive push-pull structured electrodes (Poldy 2001, 609). The membrane has a continuous electrical charge, hence the name electrostatic.

Planar magnetic headphones, also called Isodynamic or orthodynamic headphones, are like a hybrid version of the previous two. They have a large membrane, like electrostatic headphones, and a system of magnetic rods on one or either side of the membrane. Magnets create very even magnetic field, called isodynamic magnetic field. However, they work on the dynamic principle of electromagnetic induction, like dynamic headphones. The coil is replaced by a serpentine path attached to the diaphragm (Kahrs 2009, 6). A current passing through this wire pattern embedded in the membrane creates a varying magnetic field that reacts with the constant field of the magnetic rod system to produce sound.

2.2.3 General design criteria for headphones

Unlike loudspeakers, which produce a sound field in a larger space around the listener, the sound field of the headphones is produced in a partially leaky and relatively small pressure chamber with a volume of up to about 30 cm^3. Since the ear is essentially a pressure detector, this difference has little relevance in terms of subjective sound image. (Poldy 2001, 587.)

However, when listening with headphones, the usual crosstalk and phasing in loudspeaker listening is still absent. With digital signal processing (DSP), artificial crosstalk can be made. Crosstalk means that right channels signal is not heard only with the right ear, but also with the left ear and vice versa. Also, the reflection from the head, torso and partly also from the pinna does not filter the sound as much as when listening to the loudspeakers. Crosstalk and the absence of reflections from the listening room mean that when listening with headphones, the sound is usually localized inside the head (Møller 1992,199). As Dr. Floyd Toole puts it, we do not hear real stereo with headphones, only spatially distorted, but still entertaining rendering (Toole 2018, 403).

Headphones cannot create the same temporal reproduction as a loudspeaker in a room. Therefore, the requirements for headphone design are reduced to ensuring that the frequency response of the headphones is reproduced with the same timbre as that of the loudspeakers. ITU-R BS.708 recommendation states: “The frequency response of studio monitor headphones should provide the same sound-color neutrality as required for loudspeaker monitoring in control rooms and high-quality listening rooms” (ITU-R 1990). So, the amplitude of the frequency response should be the same in the headphones as the sound produced by the loudspeakers. This interpretation is the general design criterion for headphones (Møller, Jensen, Hammershøi & Sørensen 1995, 208–209).

As most music material is optimized for listening through loudspeakers, it is recommended to aim for a similar timbre in headphones. To take the sound reproduced by the loudspeakers as a reference, we should first determine what the desired target curve is when listening to the loudspeakers. However, the lack of standards in the field makes this difficult. Indeed, Dr. Toole has mentioned the term "circle of confusion" (Figure 14) that plagues the industry.

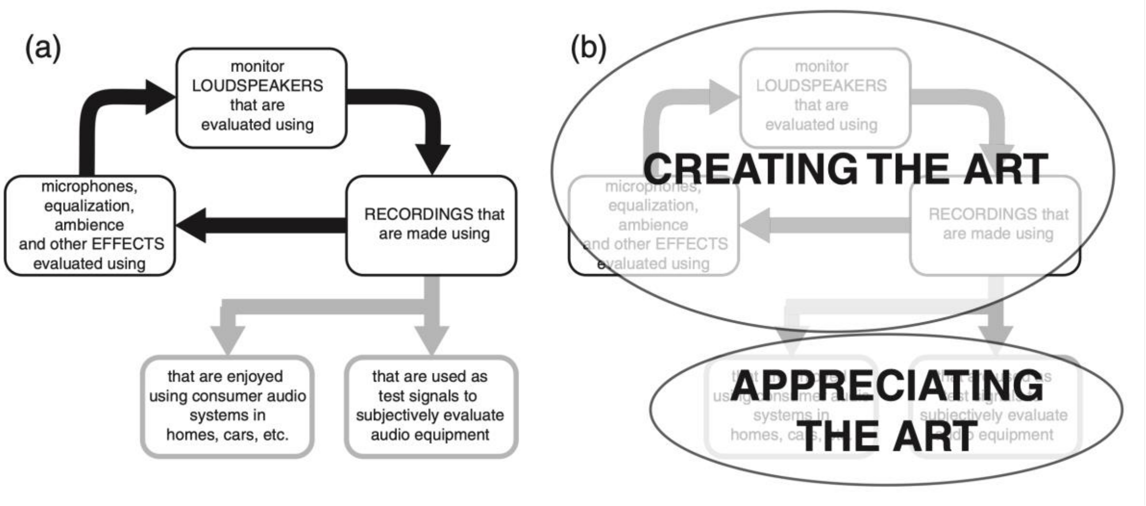

FIGURE 14. Circle of confusion (Toole 2015, 513)

The term describes the problem of evaluating the quality of sound equipment and loudspeakers without meaningful standards. Recordings are produced by listening to them through loudspeakers, which in turn are designed to listen to the recordings. Without precise standards for how loudspeakers and rooms should sound together and how they should be calibrated, the quality of sound recordings and loudspeakers will remain highly variable. Toole also mentions that if the listening equipment and environment are very different between consumers and professionals, does the art created in the studio translate well enough to the consumers for whom the art is eventually intended? (Toole 2018, 9.)

3 MEASURING HEADPHONES

The development of devices for measuring headphones dates back to the first half of the 20th century. Measurement devices have evolved greatly over time, and the latest Head and Torso Simulators (HATS) include relatively convincing technology that model the human ear and the hearing system. Equipment that meets industry standards is expensive due to the difficulty of modelling the human auditory system, but there are also devices designed for consumer use.

In this chapter we will review which organizations set standards in the industry and how standards and measuring instruments have evolved. We also look at the target curves that have been used to design headphones and the latest research on them.

3.1. International standards-making bodies

There are few international standards-making bodies. IEC (International Electotechnical Commission) is concerned with international standardization in electrotechnology. Electroacoustics is within the term of reference of the IEC, and close liaison is maintained with ISO on matters of joint interest. ISO (International Organization for Standardization) is concerned with standardization in fields other than electrotechnology, including mechanical and purely acoustic technology. ITU-R (International Telecommunication Union Radio Communications Bureau) would deny that it is a standards-making body, but Recommendations and Reports that it produces in the field of broadcasting technology have become de facto standards in many countries. (Woodgate 2001, 694).

There are also multiple regional, supranational, and national bodies such as CENELEC (European Committee for Electrotechnical Standardization), ANSI (American National Standards Institute) and DIN (German Institute for Standardization). Whereas the actual members of the international bodies like ISO and IEC are national standards committees, which are governmental or quasi-autonomous national bodies, The Audio Engineering Society (AES) members are individual engineers and scientists mostly from industry and academic institutions.

The AES is audio industry´s forum of choice for the discussion of important matters, especially concerning new technology. AES has adopted standards committee procedures and published a considerable number of standards. (Woodgate 2001, 695).

3.2. Headphone measurement device standards

Headphone measurement devices can be divided into a few different categories. Firstly, a small probe measurement microphone can be inserted into person’s ear canal and headphones can be then measured by inserting headphones to person’s head and running test signals. This technique is called MIRE, the microphone-in-real-ear technique (ISO 2018).

Another way is to use ear simulators or head and torso simulator (HATS) instead of real human head (ISO 2021). Head and torso simulator (or Manikin) models the human head and torso and in the latest models there are silicon made anthropometric pinnae and anatomically shaped concha and ear canal. HATS also include ear simulator that emulate impedance on human ear and microphone that finally captures the sound in the DRP.

3.2.1 Standards on using the MIRE technique

ITU-R BS.708 recommendation defines the electro-acoustical properties of studio monitor headphones. It states that the frequency response of studio monitor headphones should provide same neutrality as loudspeaker monitoring in control rooms and high-quality listening rooms. Measurements are made with a small probe microphone which is inserted at least 4 mm away from the beginning of the auditory canal. The reference sound field is the diffuse field, and the headphone under test is compared to the diffuse field of a reverberation chamber with filtered pink noises. Measurements should be done with at least 16 persons. Difference of left and right earphone should not exceed 1 dB in the frequency range 100 Hz - 8 kHz and 2 dB in the frequency range 10 kHz - 16 kHz. Tolerance for the diffuse field frequency response of the studio monitor headphones should meet the specs found in Figure 15 (ITU-R 1990, 2).

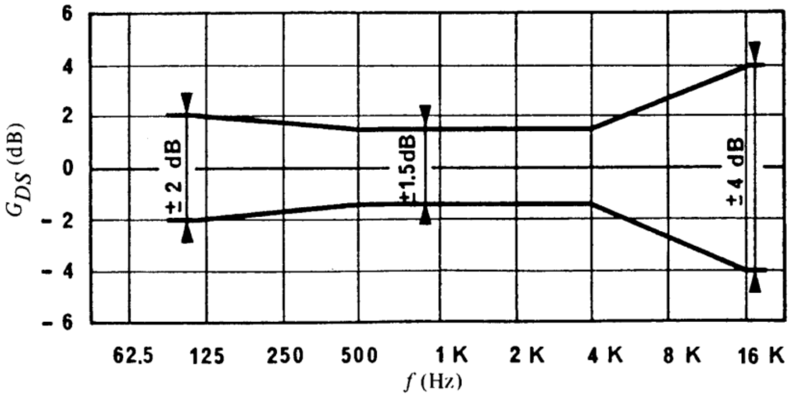

FIGURE 15. Tolerance for the diffuse field frequency response (ITU-R 1990, 2)

The standard of MIRE technique is specified in ISO 11904-1:2002 (R2018): "Acoustics — Determination of sound immission from sound sources placed close to the ear — Part 1: Technique using a microphone in a real ear (MIRE technique)".

3.2.2 Standards on ear simulators and HATS

The main standards for head- and earphones test equipment are IEC 60318 series and its parts 1, 4 & 7 (Audio Precision 2020, 1). IEC 60318-1:2009 specifies the ear simulators used for making measurements of the supra-aural and circumaural earphones mainly in use of audiometry and telephonometry. IEC 60318-4:2010 describes the occluded-ear simulator used for measuring earphones coupled to ear by ear insert. It has replaced the original IEC-60711 and specifies the simulation of the impedance of human ear canal.

IEC 60318-7:2017 specifies the head and torso simulator or manikins. The document is based on commonly used types of manikins and pinnae and specifies the free- and diffuse field frequency response characteristics of a manikin. What comes to testing procedures itself, IEC 60268-7:2010+AMD1:2020 describes electroacoustic measurements, needed sound system equipment and relevant methods of measurement. (Audio Precision 2020).

ITU-T P57 (06/21) covers the specifications on the artificial ears. Type 1 is preferred the IEC 60318-1 simulator, type 2 as a 60318-4 occluded-ear simulator and type 3.3 as an artificial ear simulator that consist of 60318-4 occluded-ear simulator and anatomically shaped pinna simulation. Type 3.3 is recommended for all types of devices. Type 4.3 is the latest version which adds fully anatomically shaped ear canal allowing a smoother and more naturally transition from concha bottom to the ear canal. (ITU-R 2021, 5–20).

ISO 11904-2:2021: “Acoustics — Determination of sound immission from sound sources placed close to the ear — Part 2: Technique using a manikin”, specifies the basic framework measurement methods for sound immission from sound sources placed close to the ear using with a manikin. (ISO 2021).

3.3. Hearing simulators and other measuring devices

The hearing simulators can be incorporated into few different categories when it comes to measuring headphones. There are head and torso simulators (HATS) that use a realistic human torso and head where microphones and ear simulators are located. These devices are used more widely in acoustical measurements where it is needed to get a simulation on whole human upper body. The second category is smaller ear simulators that include some fixture that mimic the acoustic properties of average human skull, pinna, and ear canal. These devices have been designed to test hearing aids or telephone handset transmission quality, but also to test headphones and earphones.

3.3.1 Evolution of hearing simulators

The origins of audio simulators date back to the first half of the 20th century, when the quality of telecommunications began to be more closely monitored. Initially, quality monitoring was done by having a person at one end reading text and the person in the other end listening and observing the quality of the connection. In order not to rely entirely on subjective judgements, there was a need to develop equipment to measure this transmission quality.

The first hearing simulators made by CNET and Swiss P.T.T consisted only of a cavity of a reference volume, but their accuracy did not match the human ear and they were difficult to mass-produce at that time (Rasmussen 2006, 2). However, in 1970 this development led to the IEC-60318 standard. The standard is still in use today, mainly in measuring telephones and headsets. IEC-60318-1 ears (Figure 16) are considered to have the advantage of measurement repeatability, and they cover the frequency range up to 8 kHz. They are very accurate in the range 100 – 4 000 Hz, which covers most part of the audio communication frequency band. Measurement is taken from ear reference point (ERP), so it omits the simulation of the ear canal.

In 1981, the IEC-60711 standard was published (IEC 1981). The 711 coupler that simulates the acoustical transfer impedance of the human ear canal was developed, and it extended the audio frequency range of the measurements up to 10 kHz, as the hearing aid industry considered that 8 kHz was adequate. Adapters were also developed to allow, for example, in-ear headphones or hearing aids to be fitted and measured. Later in 2010, IEC-60711 was cancelled and replaced by IEC 60318-4 standard, and it extended the usable upper limit to 16 kHz, although the standard tolerances are still only defined up to 10 kHz.

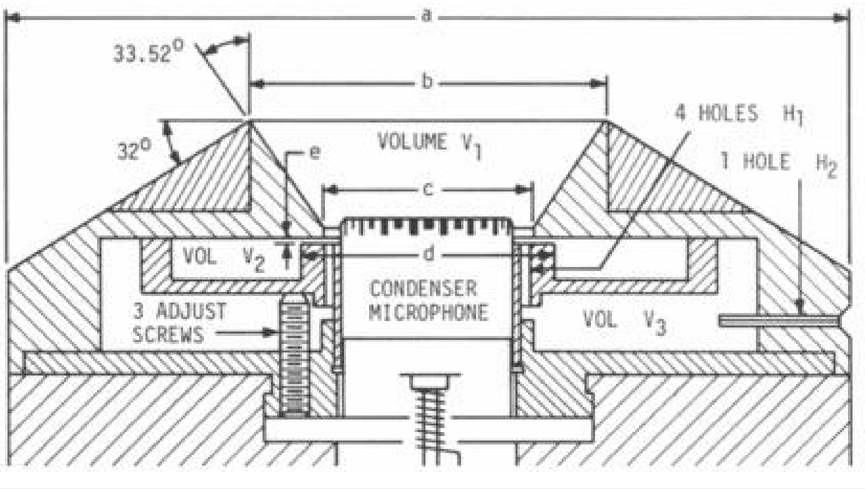

FIGURE 16. IEC-60318-1 ear. Dimensions and cavity of 2500mm^3 is important to correct high frequency performance (Rasmussen 2006, 3)

The interest in binaural sound and sound localization in spatial environment led to the need for more accurate simulations to model how the human head, shoulders and torso affect frequency response in free- and diffuse field. One of the first and most documented HATS-style devices is KEMAR (Knowles Electronics Manikin for Acoustic Research), which was launched in 1972. KEMAR was the first head and torso simulator designed for acoustic research and use in hearing aid laboratories. Earlier mentioned Klaus Genuit founded his company HEAD acoustics GmbH in 1986, which company build their first HATS HMS II in 1988. In 1987 Brüel & Kjær launched its own HATS, type 4128, which soon became industry standard for testing worn audio devices. Later in 2005, KEMAR was taken over by GRAS and they have continued to develop it under their brand (Kemar 2021). These three common HATS can be seen in Figure 17.

High-frequency HATS type 5128 launched by Brüel & Kjær in 2017 included type 4620 ear simulator that allows human ear impedance loading up to 20 kHz. This type 4620 ear simulator has the latest type 4.3 artificial ear, which includes both anatomically shaped pinna simulation and fully human shaped ear canal (Figure 18). The price for this state-of-art HATS is 40 000 – 50 000 €, depending on the configuration.

FIGURE 17. From left to right, Brüel & Kjær Type 4128-C, Head Acoustics HMS II.3 and G.R.A.S. KEMAR Type 45BM (Snaidero, Jacobsen & Buchholz 2011,30)

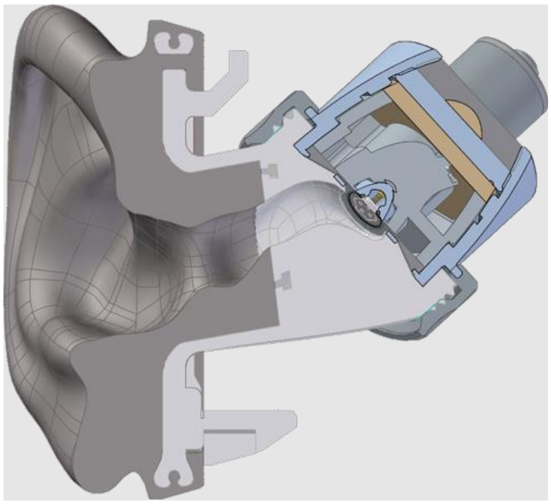

FIGURE 18. Type 4620 ear simulator by Brüel & Kjær (Brüel & Kjær n.d.)

Common binaural dummy heads such as the Neumann KU80 (free field), KU83 (diffuse field) and KU100 (diffuse field) should not be confused with the HATS devices used for measurements: These binaural dummy heads have a built-in free or diffuse equalization filter (Figure 19).

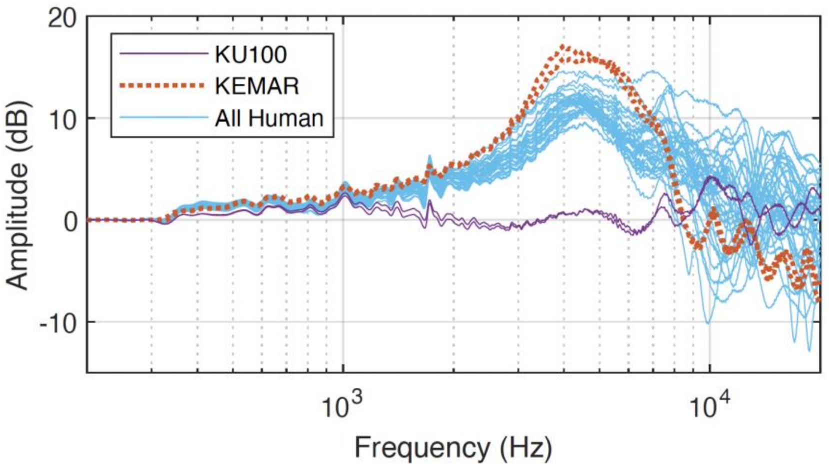

FIGURE 19. Neumann KU100 pre calibrated with diffuse field equalization filter (Armstrong, Thresh, Murphy & Kearney 2018, 9)

3.3.2 Common simulators for testing headphones

The Head Acoustics HMS II.3 is head simulator that is designed to act like the average human auditory system. It uses IEC-60318-4 compliant ear simulator and anatomically shaped pinnae type 3.3 according ITU-T P.57.

HMS II.3 has been used on the well-known website Innerfidelity.com, which has recently been taken over by Stereophile.com. Tyl Hertsen, editor of the original Innerfidelity website, has created an extensive database of over a thousand headphones, and all his measurements can now be found on the Stereophile website. However, no new measurements are forthcoming (Hertsens, 2018).

Another famous review website rtings.com uses HMS II.3 in their testing. Rtings.com uses a head simulator for measuring mid and treble frequencies, but for over/on-ear headphones they measure the low frequencies with five real humans with calibrated blocked ear canal microphones. The crossover between HATS and real human measurement is between 200 – 450 Hz. The reason for this hybrid measurement is that in their opinion, HMS simulator does not give reliable and consistent results with bass range due to the lack of proper sealing, because the simulator has larger and stiffer ears than humans. (Rtings.com 2018).

Harman International has done a lot of research on target curves for headphones, and they have also derived their own Harman target curve from the research. In their 2013 and 2015 studies, they have used the GRAS 45CA-6 configure that includes the 45CA test figure, two RA0045 ear simulators and KB0070/71 pinnae. In 2018 study, Harman used custom pinnae, because the standard pinnae were too stiff and did not work like a human earlobe. The retail price for GRAS 45CA-6 configuration is around 16 000 €. GRAS 43AG Ear & Cheek simulator complies to same IEC-60318-4 standard, but as a single channel measurement setup price is starting at 6 000 € with same pinna and ear simulator. (Klippel 2020a). However, these configurations include outdated pinnae and ear simulators.

The new anthropometric pinna for these models (KB50XX) has anatomically shaped concha and ear canal and has more flexibility so that it collapses more naturally when the headphones is set in place. Up-to-date configurations include also the RA0402 high-frequency ear simulator, which can extend the usable range to 20 kHz. 45CA-10 configuration including these updates costs over 18 000 € and 43AG-7 around 8 000 € (Klippel 2020b). With 43AG you only get right pinna (KB5000), so in case you want measure both headphone sides of the headphone, also the left pinna (KB5001) is needed, and it raises the total cost closer to 9 000 €.

The website www.headphonedatabase.com contains more than 600 measured headphones from more than 150 different manufacturers. The site has compiled the measurements of Reddit user Oratory1990 and made a useful site where you can compare the measurements with each other or how they compare to different target curves. Oratory uses nowadays mainly a GRAS 45BC KEMAR head and torso simulator with KB5000/5001 anthropometric pinnae. Price for the latest GRAS 45BC-14 KEMAR configuration that complies to IEC-60318-7 standard is over 31 000 euros (Klippel 2020b).

Interestingly, headphone manufacturer Ollo audio uses GRAS 45CC test figure + RA0039 coupler that complies to IEC-60318-1 standard. This setup is stereo, but does not have a pinna, as discussed in the previous section. The measurement is made from the ERP point, and thus it does not include ear/pinna gain. Retails price for 45CC-4 configuration is little bit over 11 000 euros (Klippel 2020b).

3.4. Target curves for headphones

Industry standards define two types of reference fields: free field (FF) and diffuse field (DF), which have been used to create target curves for headphones (Lorho 2009, 3). A target curve refers to the “ideal” frequency response curve. The preferred target curve for optimal sound quality is a controversial topic, with little consensus among headphone manufacturers (Olive, Welti & McMullin 2013a, 10). In addition to the free field and diffuse field targets, the Harman target from the Olive et al. study is commonly known in the field (Figure 20).

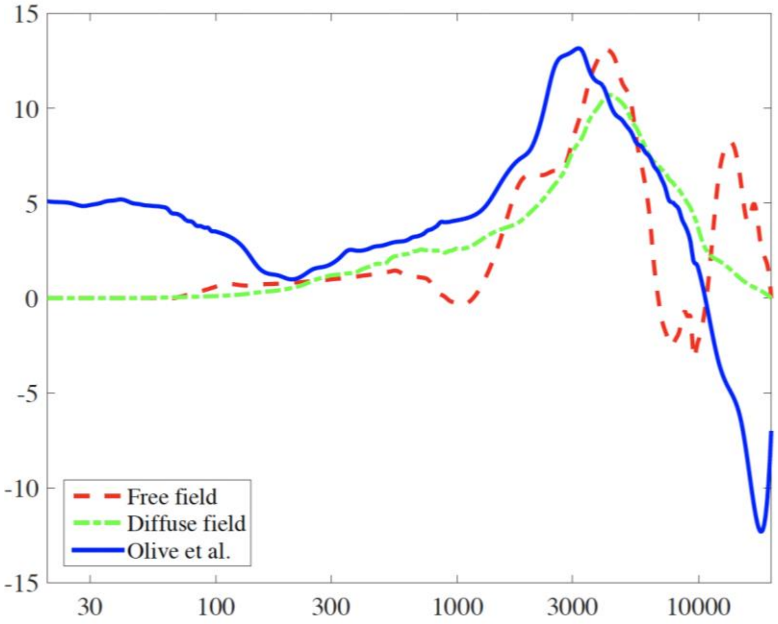

FIGURE 20 Different target curves. Olive et al. curve from 2013 study. (Välimäki & Reiss 2016, 30)

3.4.1 Free field response

Free field measurements are made in anechoic free field conditions, with the loudspeaker placed directly in front of the listener and a flat probe measurement mic inserted at the DRP of the ear. Using human measurements, many measurements and persons should be used. Alternatively, the measurements can be made using standardized HATS as a replacement for human subjects. In Figure 21 we see that while the measurement microphone shows a flat frequency response, the human ear does not, due to the resonances discussed earlier. So, when we talk about a free field target curve, the frequency response of the headphone is modelled to match that of a HATS placed in an anechoic room with a single loudspeaker directly in front of the HATS.

When measured directly from the front, near 10kHz there is a dip due to comb-filtering reflection from pinna. Since the resonance of the 3λ/4 ear canal is about the same as a 10 dB dip at around 10 kHz, there is still a good chance of hearing such sounds despite the inherent attenuation. Also, at around 1kHz there is a small dip due to reflections from the shoulders. (Poldy 2006, 17.)

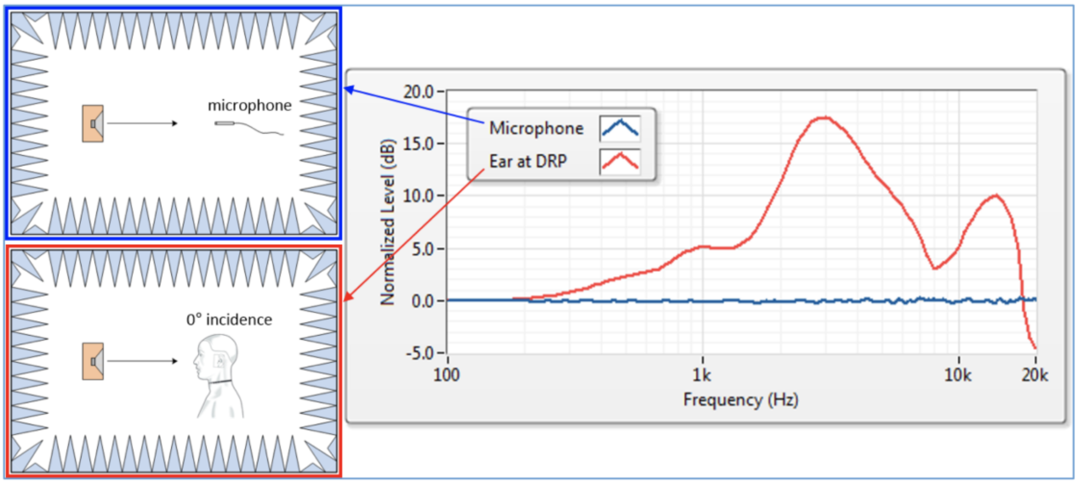

FIGURE 21. Flat loudspeaker in anechoic chamber vs in-ear measurement in DRP (free field response) (Audio Precision 2016, 4)

In 1985, Günther Theile from Institut für Rundfunktechnik Gmbh (IRT) presented in his AES paper 2207, that free field equalization applied to headphones cannot produce good results because of mono source from single direction (Theile 1985, 22). According to Theile, the free field frequency response of the headphones would be only accurate if the stereo image were perceived to be in front of the listener, as it is when played through the loudspeakers. However, this is not the case, and the sound image is placed inside the head.

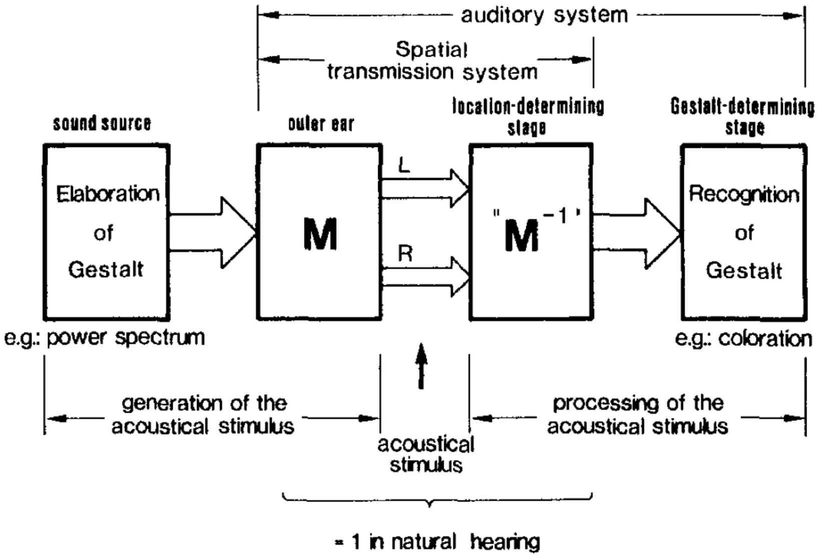

FIGURE 22. Spatial transmission system (Thiele 1985, 26)

Thiele's diagram in Figure 22 shows how sound is exposed to spatial information *M* as it passes by the outer ear. The brain uses this information to help determine the location of the sound source and then applies inverse filtering M^{-1} to prevent the tone of the sound from being distorted. If M is not the inverse of M^{-1}, as Theile argued for the response of free field headphones, the cancellation is incomplete, and the result of perceived sound is colored. When listening with headphones, the localization process does not receive localization cues and the sound image is localized inside the head because the simulated free field response is not accurate enough. The ideal headphone frequency response should be the average value of the transfer functions of the outer ear in all different directions, as in diffuse field (Theile 1985, 22). Although the FF target has lost its popularity since the 1980s, it is still part of the current standard (IEC 2010) for headphones (Olive 2022, 59).

3.4.2 Diffuse field response

Diffuse field frequency response replicates the SPL in the ear of a listener for sound impinging from all directions (Poldy 2001, 586). The diffuse target curve can be measured by setting the HATS in an echo chamber where the sound source produces flat sound to all directions and is far enough away from the HATS that it is no longer in a direct field where the direct signal would dominate. Since the measurement is done in a highly reverberant echo chamber, the sound is theoretically reflecting from all directions to the HATS and the diffuse field target curve is the average of these sounds. Specifications of the diffuse field can be found in IEC 60268-7:2010 recommendation for calibrating headphones (IEC 2010).

In Figure 23 we can see three common HATS diffuse field measurements and the tolerances on diffuse field standards. In (Snaidero et al. 2011), the authors found that the current standard that specifies the diffuse field conditions lacks a few details, for example the number of the spatial locations to be used.

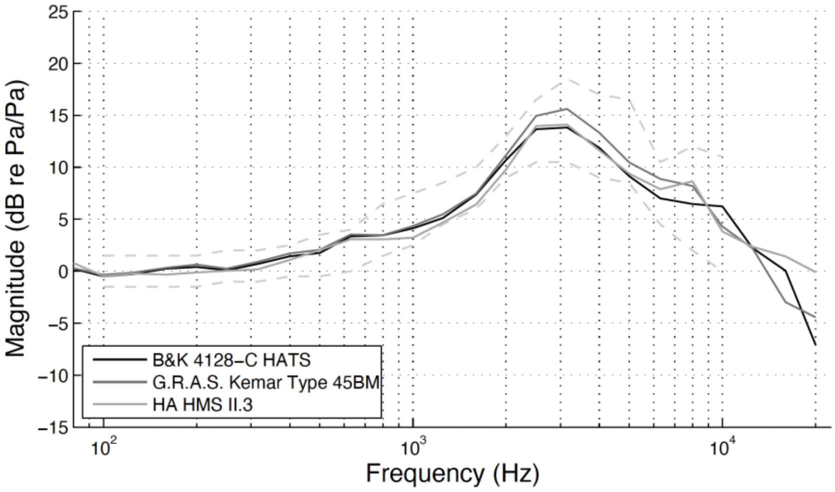

FIGURE 23. Three common HATS compared in same diffuse field HRTF setup (Snaidero et al. 2011, 23)

Christopher Struck mentions in his AES paper, that both diffuse field response, as free field response, represents an unrealistic listening situation, because a normal listening room is never completely anechoic, nor completely diffuse. According to him, these responses represent more a kind of boundary condition of the listening conditions. (Struck 2013, 4.) This topic was also addressed in study by Fleischmann, Silzle and Plogsties (2012). They measured a 5.1 surround sound system in a standard listening room, and a group of three experts equalized three different headphones to match the timbre of the speakers. This semi reflective field (SRF) was thought to better match normal listening room conditions. The conclusion was that in listening tests these new target curves were equal or better than the DF targets. (Olive 2022, 60.)

3.4.3 Harman target curves

In 2013, Harman International studied listener preferences for different headphone target response curves. In this study, ten trained listeners from Harman evaluated Sennheiser HD518 and Audeze LCD-2 headphones unequalised and equalized to six different target curves. The target curves were free and diffuse field curves from ISO 11904-1 (2003), one older diffuse field target proposed by Møller et al (1995), one modified diffuse field target by Lorho (2009) and two

Harman target curves based on the equalized in-room response of a loudspeaker. Listeners preferred the Harman target RR1 compared to the other DF and FF target curves or unequalised headphones.

Harman target curve RR1 curve originates from in-room study, where Harman compared different control room target curves and room correction products. Listeners preferred target curve that was downward tilted -9 dB from 20 Hz to 20 kHz (Figure 24). In this study, the two worst selected target curves pursued a flat frequency response, so the conclusion was that a flat frequency response is not necessarily desirable (Olive, Jackson, Devantier, Hunt & Hess 2009, 13).

2013 targets were made using in-room target responses of the loudspeaker system in the Harman listening room as a reference. The listening room contained a seven-channel loudspeaker system, and each channel was measured with a calibrated measurement microphone at six different locations in the listening location to obtain a spatial average. Each channel was then equalized to a flat frequency response and then adjusted to match the target curve using shelving filters.

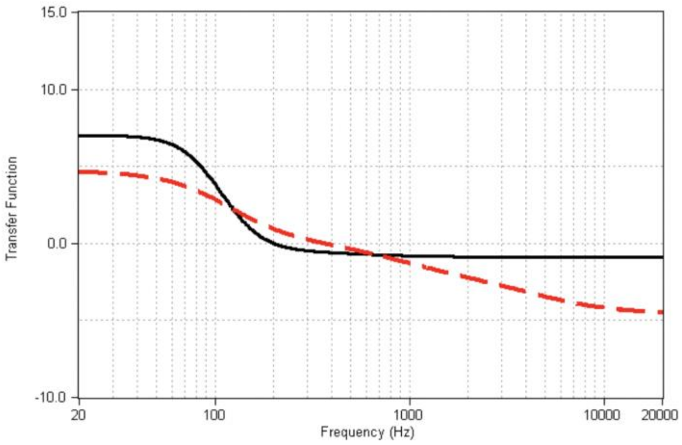

FIGURE 24. RR (solid line) and modified RR1 (dotted line) in-room target responses (Olive et al. 2013a, 4)

A GRAS 43AG simulator was then installed in the head of Styrofoam manikin, which was inserted in primary seat of the listening room. A total of 21 measurement were taken: three different rotations (-30, 0 and 30 degrees in horizontal plane) for every channel and spatial average was made and smoothed to generate RR and RR1 target curves. Headphones were then measured with the simulator and equalized to the targets.

The study concluded that the current industry DF and FF standards for calibrating headphones are flawed and sound too thin and bright. The most neutral target curve in the study turned out to be a new Harman target that simulates high-quality loudspeakers in an acoustically treated listening room. Since in the earlier study the listeners could not adjust the bass and treble levels themselves and the target curves were derived empirically, the authors decided to conduct a new study. In this study (Olive, Welti & McMullin 2013b), conducted the same year, Harman examined both headphone and loudspeaker listening. The loudspeakers were equalized to a flat frequency response, and listeners were allowed to adjust the bass and treble controls to their liking as they listened the test material. A similar test was performed on the headphones, where the headphones were first equalized to match the flattened frequency response of the loudspeaker response. Preferred curves can be seen in Figure 25.

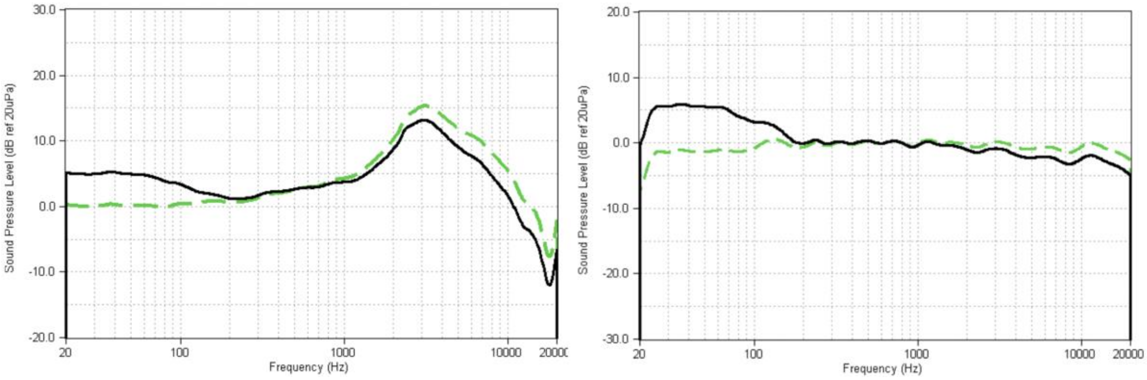

FIGURE 25. Preferred curves for headphones (left) and loudspeakers (right) as solid line. Green dashed line as a reference of flat in-room response (Olive et al. 2013b, 13)

The result of the study was quite like what they found in the previous study, with the exception that listeners preferred the bass and treble levels in headphones to be about 2 dB lower than in loudspeakers. The Harman 2013 target was published as a result of this study. Since the 2013 test had only eleven test subjects, eight of whom were experienced listeners and three of whom were novices, it was decided to conduct a larger study in 2015. In a 2013 study, three less experienced listeners preferred a 3.6 dB louder bass and treble level in headphone listening compared to more experienced listeners. Now it was decided to study only headphone listening with a larger test group to see if the age and experience of the test subjects had a big impact on the results. 249 test subjects from four countries took part in the 2015 study. The test design was otherwise the same except that the headphone model changed. The study found that the music material used, the listeners' age, gender and amount of listening experience all influenced the results. Younger and less experienced listeners preferred more bass and treble than older and more experienced listeners. On average, women also preferred less bass than men. Nationality did not seem to play a major role, or at least it was difficult to prove. However, a new Harman 2015 target curve was created based on this study. (Olive & Welti 2015.)

Compared to the 2013 target, the new target curve did raise in both bass and treble (Figure 26). The researchers concluded that the relatively small test group in the original study and the lower average age and larger number of inexperienced subjects in the new test led to this result. The rise in both bass and treble levels meant that both the headphone and loudspeaker listening curves converged even further.

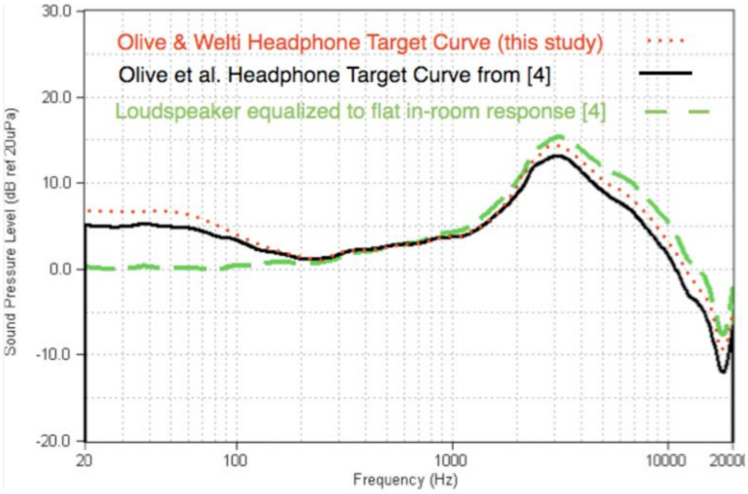

FIGURE 26. Harman 2013 (black) and 2015 (red) target curves compared (Olive & Welti 2015, 10)

Because the volume did not remain constant when using the bass and treble controls, the researchers speculated that some of the subjects' reactions may have been because as the volume increased, the sound may have felt better. (Olive & Welti 2015, 11). Later in 2016, in a similar study done with in-ear headphones, tests used a bass shelf filter that automatically compensated for the volume loss or gain of the material and researchers found that without compensation the bass was increased by 2 dB more than if the compensation was on (Olive, Welti & Khonsaripour 2016, 8). Further study done in 2017 found that when comparing listening volumes and bass levels in headphone and loudspeaker listening, test subject liked 1 dB more bass with loudspeakers and they listened 2 dB louder also with loudspeakers (McMullin 2017, 1).

The latest version of Harman’s target curve was published in 2018, when they compared 31 different target curves (Olive, Welti & Khonsaripour 2018a). In the study, the frequency responses of 29 different commercial headphone models in the range of 60 - 4000 dollars were measured. Authors used a virtual headphone method in which the frequency response and phase response were modelled. Each test subject used then only one pair of headphones, but the person was able to change the frequency response from the control panel. This study also introduced a slightly modified new target curve with a slight decrease in frequency at 3 kHz (Figure 27). The study involved 130 test subjects which were both trained and untrained. They tested the headphones in batches of eight, each batch containing a hidden high anchor (Harman's new curve) and a so-called low anchor headphone.

Experienced testers rated the new target better than the other 31 models, and more inexperienced testers rated the new target better than the other 27 and four others as equally good. The study also concluded that, given this study and two previous studies (Breebaart 2017; Olive, Welti & Khonsaripour 2017b), there is a little correlation between headphone price and sound quality based on listener preferences and sound quality does not improve much beyond 300 dollars (Olive et al. 2018a, 9).

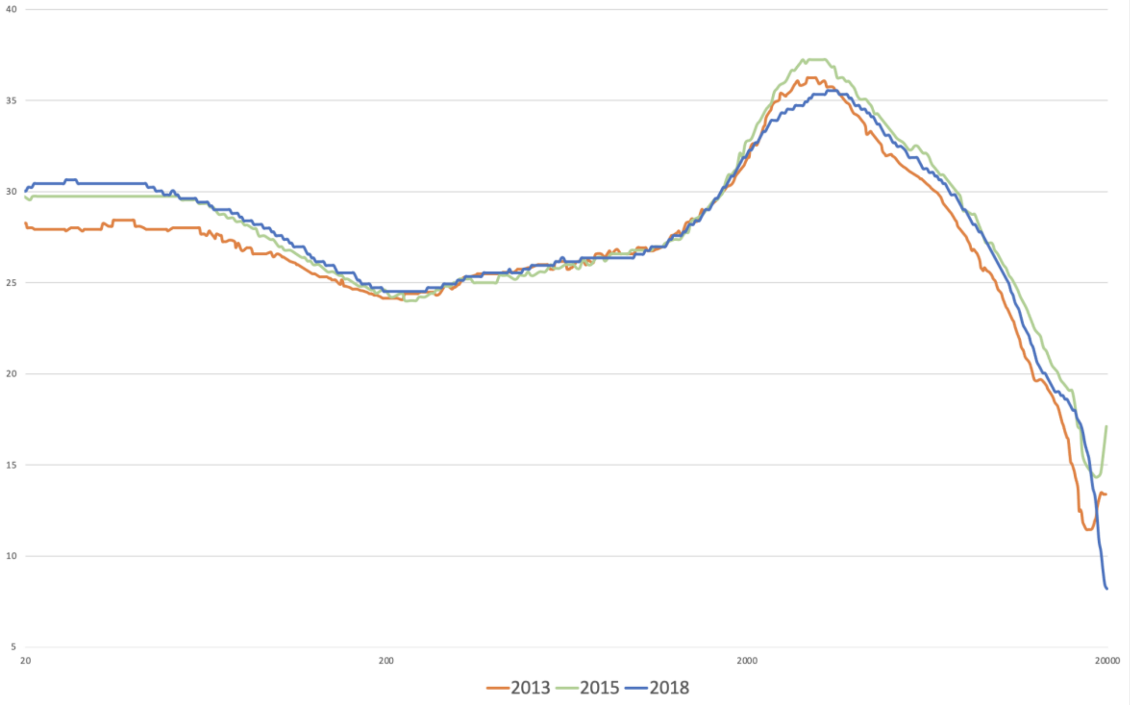

FIGURE 27. Comparison of Harman over-ear headphone targets (preproduced from Olive & Welti 2015 and Olive et al. 2018a)

4 PRODUCT DEVELOPMENT PROJECT

At MojoLab, we have tried out numerous different hifi and pro audio headphones during the mastering process. In most cases, the timbre of the headphones has not been similar to what you would hear in the studio with loudspeakers. This has caused more confusion than help. In general, headphones designed for studio and hi-fi use have often sounded too bright and lacked low-frequency extension. In contrast, headphones designed for consumer use are often hyped especially at the low end of the frequency response. This was also noted by Olive et al. (2018b) in their study comparing consumer and professional headphones. Because of all these shortcomings, headphones have mostly been used at MojoLab only to check clicks and other low-level noises. These things are often easier to hear with headphones than with loudspeakers.

We have also tested a few different pieces of software designed to correct headphones frequency response with DSP. However, the experience of using them has not been entirely convincing. Also, as these kinds of software are usually machine specific, it is not practical if you listen to music from different devices and those all do not have that piece of software installed. Since nowadays not all work is necessarily done in just one studio, and you do not want to carry around expensive headphones with imperfections in the sound, the idea was born to start modifying the headphones.

4.1. Equipment

Before any equipment was purchased, modifications were made based solely on listening observations, but it soon became apparent that it was difficult to rely only on one's own listening observations. The ear gets quickly tired of various experiments, and it is also challenging to compare changes to the previous ones, especially if those have been done on a different day. It was also noticed that when you were concentrating on listening to a particular change, you might miss something else that was happening elsewhere.

Because industry standard measuring devices are very expensive, starting around 15 000 € with all the needed equipment, it was decided to test one consumer grade device to see if it would be helpful in supporting listening observations.

4.1.1 Measurement device

The main measurement device in the project was MiniDSP EARS (Picture 3) which is affordable consumer type of headphone measurement device. EARS (the Earphone Audio Response System) costs about 200 €, and it works by USB -connection with Mac, Windows and Linux. EARS have molded silicone pinnae that that mimic the human outer ear, but the ear canals are just short cylindrical canals. The device also does not include an ear simulator to mimic the input and output impedance of the human ear. EARS provides you with calibration files containing the correction data for the microphones placed at the end of the "ear canals".

PICTURE 3. MiniDSP EARS

Compensation files are also included, which simulate a few different target curves. HEQ compensation proved to be the most useful in this project. The purpose of HEQ compensation is that measurement result of the neutral headphones is approximately flat. It is based of Harman target curves but adapted for EARS (MiniDSP n.d.). Depending on listener preference, for low frequencies the frequency response may either increase or decrease slightly and for frequencies above about 1 kHz the frequency response may decrease slightly, as we can see in Figure 28.

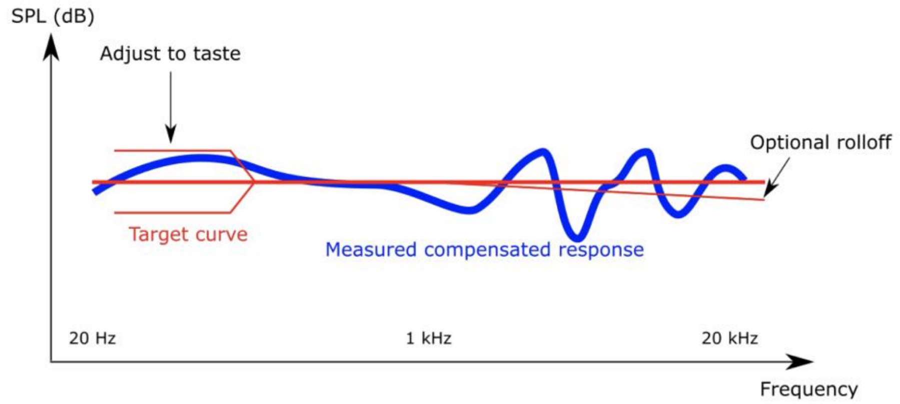

FIGURE 28: MiniDSP EARS HEQ compensation and it´s optional high and low varitions (MiniDSP 2021, 25)

4.1.2 Software

REW is a free room acoustics, loudspeaker, and audio device measurement software (Mulcahy 2022). Because we had used the REW in the past to make room acoustic measurements, it was natural to use it also for headphone measurements in our project. REW also supported the MiniDSP measuring device as well as the necessary compensation files. The software is easy to use, and the developer regularly update the software and its features.

In addition to the REW program, a few other applications were used for measuring and fine-tuning the headphones. The University of New South Wales have produced a simple java program that play different sine wave tones at different loudness (Wolfe 2022). Comparing different frequencies with each other using for example 1 kHz frequency as a reference, you could quickly create a curve that shows how you perceive different frequencies of tested headphone or set of loudspeakers.

In Figure 29, you can see one frequency response measurement made in studio by listening observations. The curve should be viewed so that the frequency with the higher value have been raised so that it has sounded as loud as the reference frequency. Similarly, if the frequency was louder than the reference frequency, you have selected quieter sound of the same frequency below. So, if this would be compared to the frequency response measurement, the curve should be viewed vertically as a mirror image.

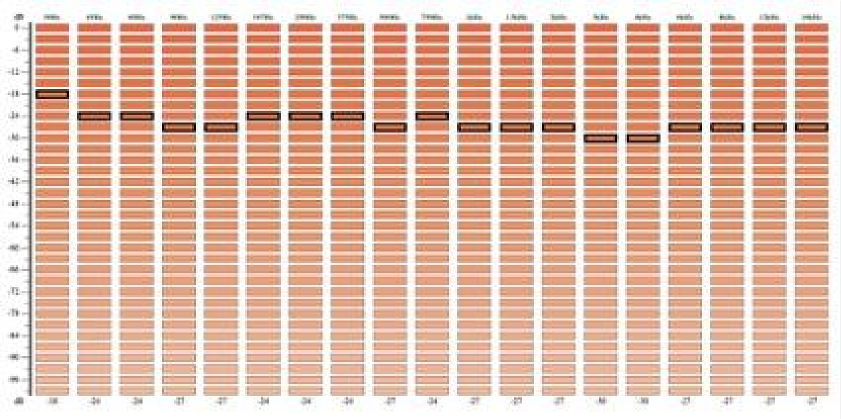

FIGURE 29. Measured response curve by listening observations in one studio

This program has proved to be useful in many situations. A measurement signal can be played directly from the laptop and thus form even some kind of measurement to support listening observation if actual measuring devices are not available. Since the measurement results cannot be easily saved and the jumps between the different volume options are also quite large (3 dB), the program is only suitable for quick testing on the road.

When the benefits of measurements made by ear and the shortcomings of the above program were noted, searches were started to find a more advanced program. Since no other useful program was found at that time, a Pro Tools multitrack session was made, where filtered 1/3 -octave band pink noise signals were imported to the tracks. When listening the different tracks soloed, you can compare audio frequency bands rather than just one pure sine wave frequency. By adjusting the levels of the tracks, you can balance the different bands with each other and thus form a more accurate equal loudness curve. After comparing different bands, levels from all tracks were written to Excel. By multiplied the values with -1, you could get the vertically mirrored result, which forms a frequency response curve.

During this Pro Tools experiment, the DgSonicFocus software developed by ex-Harman researcher David Griesinger was found on the internet. He had developed an application that used filtered 1/3 -octave pink noise bands to match the frequency response of the listener's eardrums to the flat front loudspeaker response. His app works in such a way that you first make acoustic measurement in front of loudspeaker at the distance of 18 inches and with the application you equalize the loudspeaker signal to flat. The listener sits then in the measurement position, compares the equalized filtered bands to the 500 Hz reference band and adjust the bands that all play same loudness. This way you get your own frontal “free field response”. The next step is to put headphones on and make same loudness comparison with headphones. The app then compares the test with the headphones to the test with the equalized loudspeakers and forms a curve with how you should equalize the headphones to match the loudspeakers. (Griesinger 2022).

4.2. Background study of headphones and in-room responses

Since acquiring the MiniDSP measurement device, over thirty headphones have been measured, and experiences on how the device works and how measurements correlate with listening observations have been gathered. Some problems with the device were also identified. Because silicone ears are much stiffer material than the human ear, the headphones do not always seal perfectly around the ears and the device. Leakage problems have a big impact on the low-frequency reproduction of headphones. Poor sealing can cause headphone response deviations of up to 20 dB or more. (Welti 2015, 1.) With MiniDSP, it was difficult to measure supra-aural headphones, and future measurements ended up focusing more on circumaural headphones.

Also, some circumaural headphones, which had a shallow air cavity for the ear, seemed to suffer from the same kind problems. Pressing the headphones slightly against the device sometimes helped to seal the headphones better and thus produce more accurate results at low frequencies. However, in some cases too much pressure caused the whole ear, rather than just the earlobe, to flatten, and this began to show up as a change in upper mid-frequencies.

Another pair of silicone pinnae sets were also purchased for the project, which were modified to make pinnae a little more flexible. However, the result was not better, and the pinnae were eventually cut off completely. All that remained were the holes in the silicone base where the microphone was embedded. This allowed the MiniDSP to also be used for flat plate measurements, by swapping pinnae set with flat silicone base.

Figure 30 shows measurements from 25 circumaural headphones aligned at 600 Hz. The measurements have been 1/6 -octave smoothed to make the picture a little clearer to examine. When the headphones that gave clear error measurements were excluded from this background study, it was found that the measurements were somehow consistent with the auditory findings.

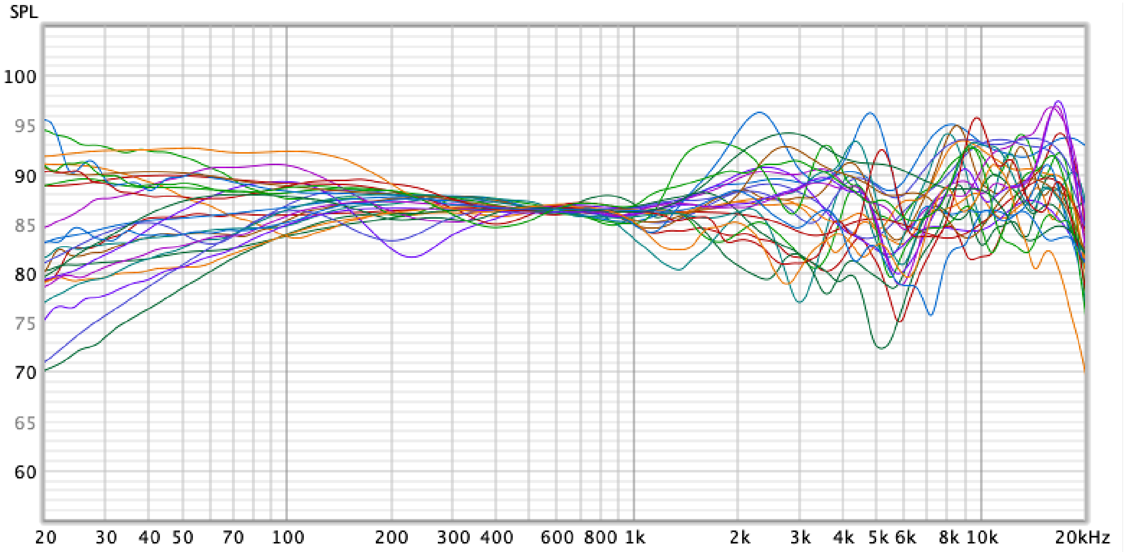

FIGURE 30. Measurement of 25 different commercial headphones

The MiniDSP´s Harman based HEQ compensation was used for the measurements. Harman target seemed a good starting point as previous studies had shown its superiority over other target curves. The device proved to be a good tool to investigate different modifications and to support auditory observations, but the measurement results did not always fully correspond to the auditory findings. Especially at frequencies above 6 kHz, auditory perception found differences that were not necessarily indicated by the meter. One reason for this is certainly that the meter does not include any ear simulator to model the acoustical impedance and transfer impedance of the human ear, so the accuracy of the highest frequencies is limited.

The inaccuracy of the highest frequencies is also reflected in HATS devices that meet industry standards. Thomas Snaidero from the Technical University of Denmark studied three well-known HATS devices. In their study, the devices were examined under both free- and diffuse fields conditions. In Figure 31, we can see a free field comparison of the three common HATS devices in an anechoic chamber with 0 azimuth. While the results show similarities, there are over 5 - 10 dB differences between the devices in certain frequencies. Other azimuths showed similar differences in other frequency bands. The same models from the same manufacturer gave the same results, but there are big differences in the measurement results between models from different manufacturers. Thus, the conclusion is that it is not necessarily worth comparing measurements done with different manufacturers' measurement devices, even if they all meet the standards.

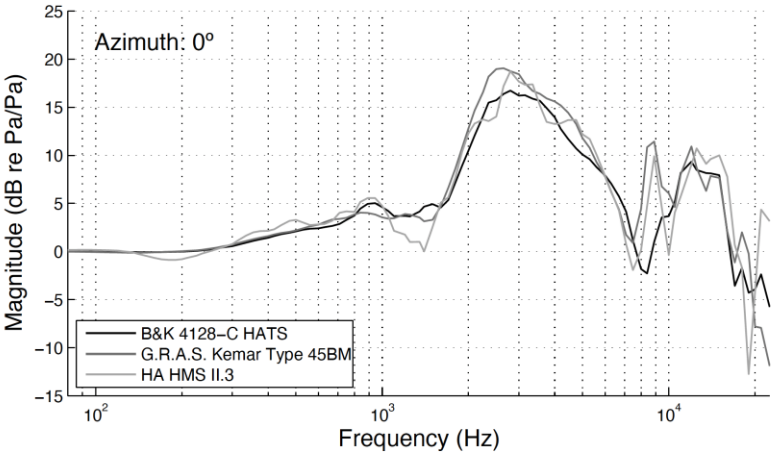

FIGURE 31 Three common HATS compared in free field HRTF at 0 azimuth (Snaidero et al. 2011, 27)

4.2.1 Comparison of measuring devices and methods

A few headphones were selected for more detailed testing. The selected headphones are widely used by pro audio users and stand out to some extent in terms of their sound, as well as the measurement results. Each of these headphones also sounded somewhat right with their own shortcomings. AKG k371 is one of the few headphones that have had the Harman curve as their design target, largely because AKG is a Harman's own brand. Sennheiser HD600 is prolonged favorite of many pro audio and hifi users, and its roots are in a diffuse field target. Audeze is respected brand that is oriented more in the higher price range headphone models. Their LCD-3f headphones measure quite differently from many other headphones, but the sound is pleasing, and it has found its users in many pro audio and hifi users.

In Figure 32, we see how these three headphones measured with the similar GRAS device that Harman has also used in their studies. In Figure 33 we can see same headphones measured with MiniDSP EARS. Of course, the headphones have not been the same individuals in these tests, but we can assume that there are no major individual differences, and we will see at least some comparison between the devices. We can see that the measurements are surprisingly similar when you look at the measurements in broad terms. Whether it is worth looking too closely at the results of any of the devices is not certain, but this comparison showed that the consumer measuring device does give good enough results to support our listening observations, at least with these headphones.

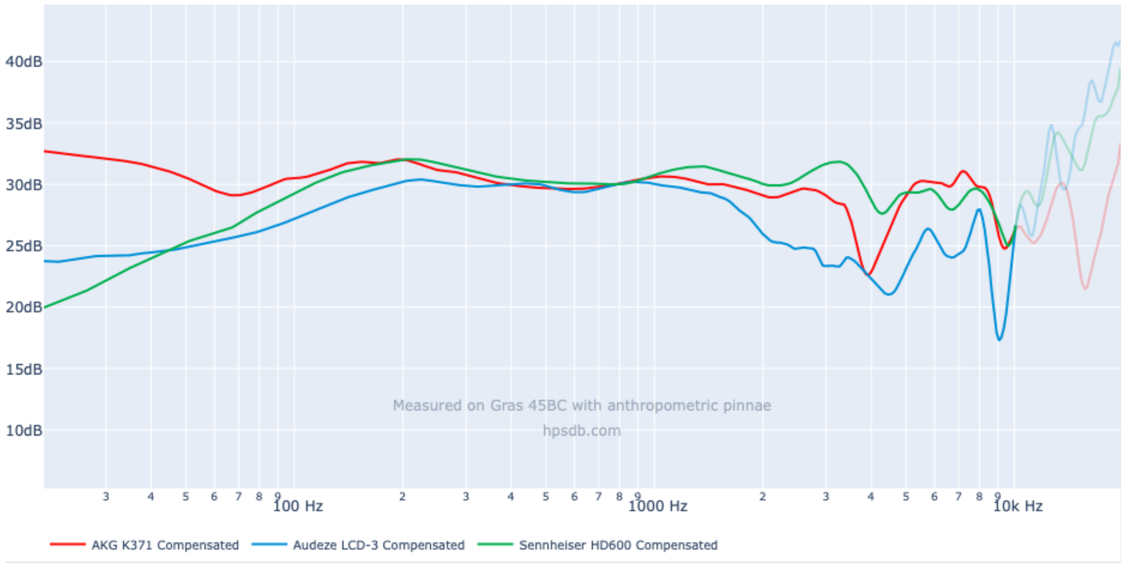

FIGURE 32. Gras 45bc setup; red AKG 371, blue Audeze LCD-3f & green Sennheiser HD600

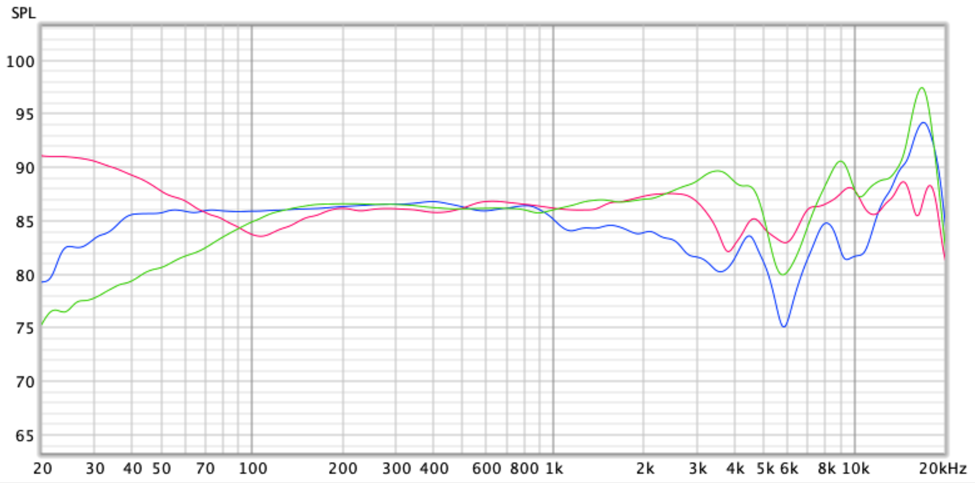

FIGURE 33. MiniDSP EARS setup; red AKG 371, blue Audeze LCD-3f & green Sennheiser HD600

The headphones frequency responses were also tested with the DGSonicFocus software, using auditory findings. This was done to see if a simple equal-loudness hearing test could provide reliable measurement results. If measurements could be made with the ears that are used for critical sound work anyway, many of the errors caused by the measurement equipment could be eliminated. Even the most sophisticated measuring device will hardly never be able to fully model the human ear and brain, especially since the human ear and hearing are so individual. Thus, the goal was to find a measurement method that would allow headphones to be easily and maybe more reliably measured by auditory perception. DGSonicFocus seemed to be a useful program for this purpose.

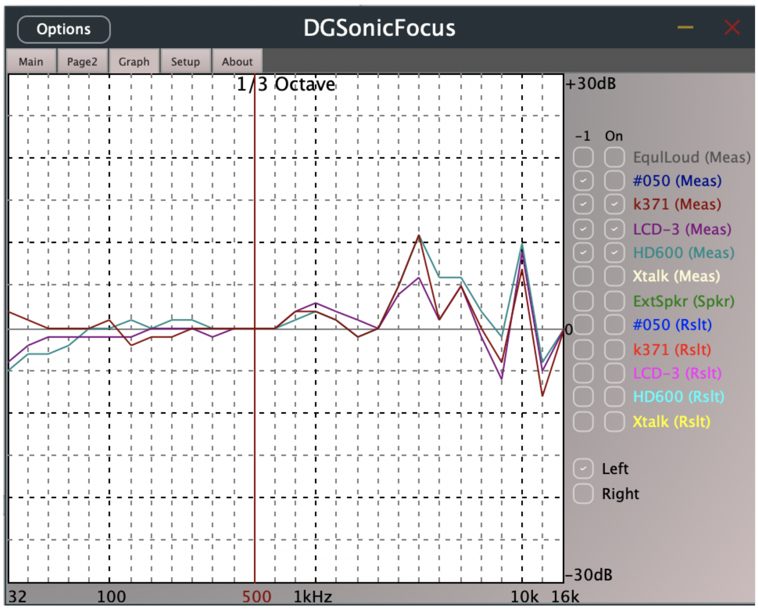

FIGURE 34. Headphones measured by listening and comparing different filtered noise bands. Red AKG 371, purple Audeze LCD-3f & green Sennheiser HD600

In Figure 34, we see the same three headphones measured by auditory perception by comparing the bandpass noise with the 500 Hz reference band. The differences between the headphones are somewhat similar to the differences measured by the measuring devices. We can also see how the ear is most sensitive in the 3 kHz range, due to the previously discussed ear resonances. The measurements are more time-consuming when done this way, and although it is a simple test comparing the volume of two different bands, it always leaves room for interpretation as well. Listening volume also plays a role; when listening at low volume, especially low frequencies are reproduced differently than when listening louder.

It is worth bearing in mind that the ear is more sensitive to different resonances when listening to a steady noise signal than when listening to music with temporally and spectrally varying content. In Figure 35 we see that with noise signals we can distinguish much smaller resonances than we can when listening to music. The differences may be smaller when listening to music, so it may not be worth paying so much attention to the smallest differences in measurement results. The Q-value, or sharpness of the bandwidth, also affects how sensitive we hear different resonances.

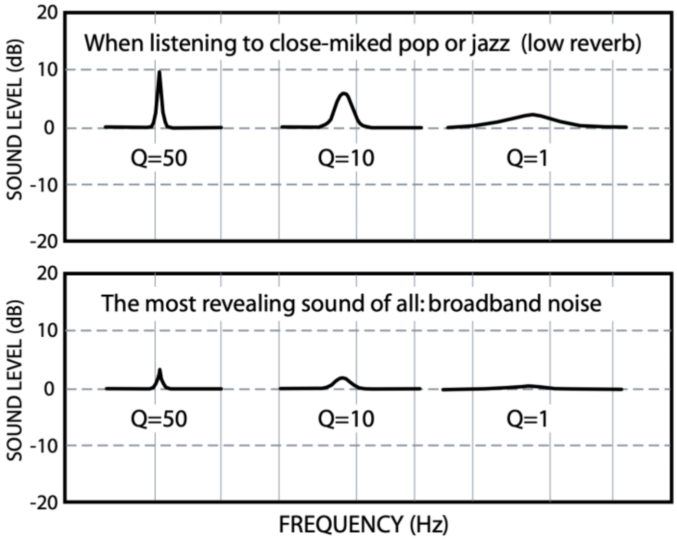

FIGURE 35. Detection thresholds on different Q-value resonances and type on material (Toole 2004, 17)

This test method gave rise to the idea that if headphone listening should be similar in sound quality to loudspeaker listening, why not simplify the whole thing by conducting the listening test in a room with good acoustics and loudspeaker system and taking these results as the target curve for headphones. The headphones are then fine-tuned so that the listening test with the headphones matches the test with the loudspeakers.

4.2.2 Preference in-room responses

As mentioned earlier under "General design criteria for headphones" section, the amplitude of the frequency response of headphones should be the same as the sound produced by the loudspeakers. In our project, two studios were used as reference for the project. One was a mixing control room equipped with Genelec 8341a nearfield monitors and the other was a MojoLab´s mastering room with custom-made full-range loudspeakers. Both listening rooms were measured with a calibrated Earthwork m23 measurement microphone from the listening position. Genelecs had been calibrated with their own DSP to their flat target curve. In MojoLab, loudspeakers had been calibrated with Acourate software to correct amplitude and phase with its finite impulse response (FIR) filters to the desired target curve. Reverberation times, first reflection levels and frequency responses comply with the EBU and ITU standards in both control rooms. In Figure 36 we see the steady state frequency response of both rooms with recommend 1/3 -octave smoothing.

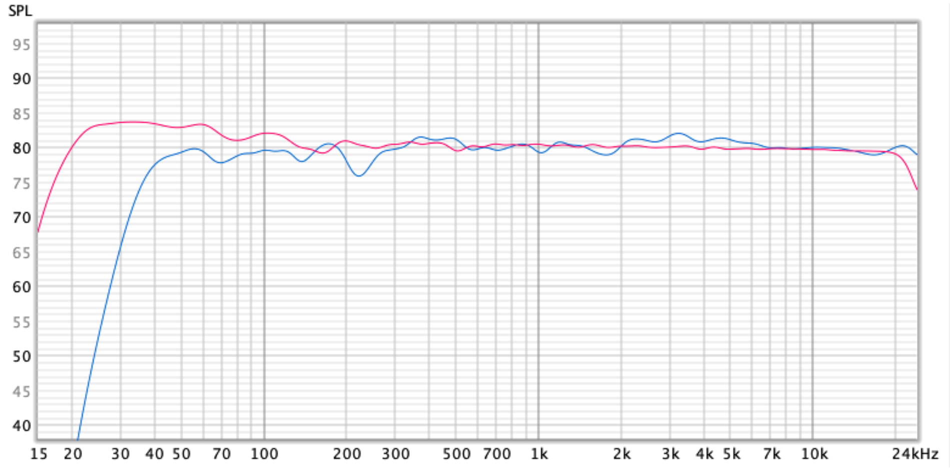

FIGURE 36. Mixing room with Genelec 8341a (blue) / Mastering room in MojoLab (red) (1/3 -octave smoothing)

In Figure 37, the same rooms and setups are measured with the DgSonicFocus software. This shows that a listening room measured as flat is not flat according to the listening observations, but the frequency response has peaks and dips due to HRTF. Generally, in loudspeaker listening, the loudspeakers are at about a 30 -degree angle to the listener and looking at the Figure 38, we can see that our measurement corresponds quite well to HRTF for 30 -degree azimuth.

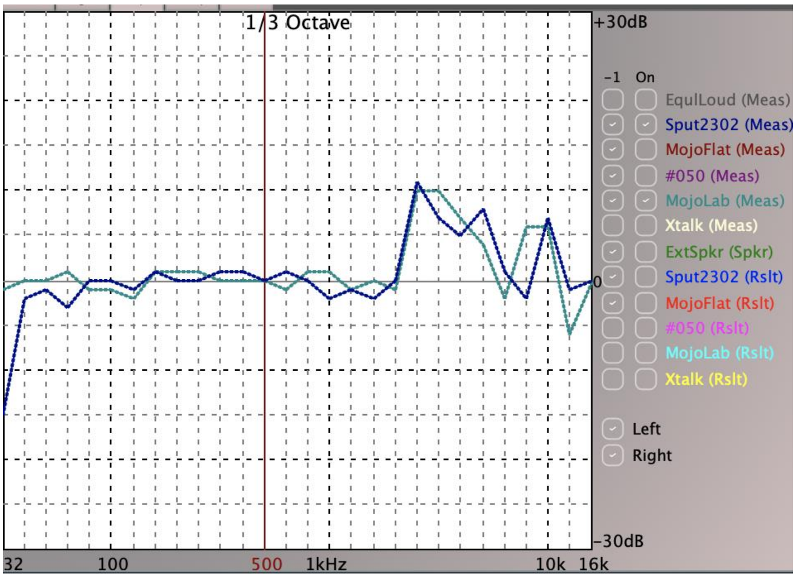

FIGURE 37. Two control rooms measured with DgSonicFocus. Mixing room: blue & mastering room: green

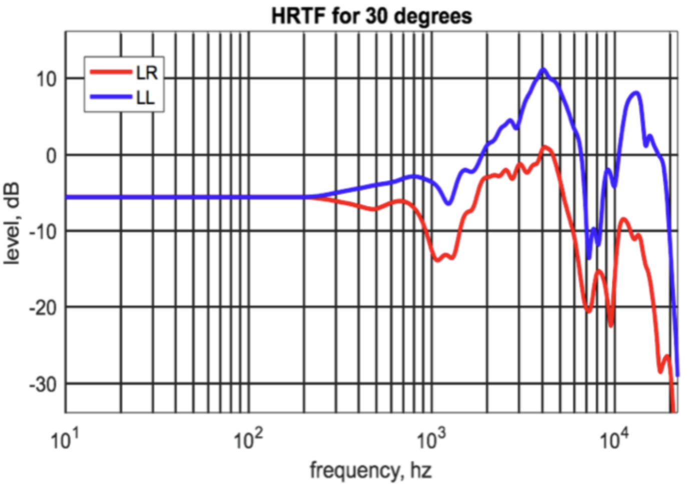

FIGURE 38. HRTF 30 -degree azimuth. LR=left loudspeaker, right ear. LL=left loudspeaker, left ear. (Gunnarsson 2016)

A fully flat in-room frequency response has not been used in our mastering room, because it always sounded a little bit too bright and underpowered at low frequencies. In a comparison of different room processors and their target curves, Olive et al. (2009) concluded that a fully flat response is not desirable. Listening tests showed that the preferred response was a frequency response that steadily declines towards high frequencies. This was also found in a 2017 study in which twenty test listeners from Samsung adjusted the frequency response to their liking using a shelving filter. The preferred target curve was +3.9 dB for the second-order low shelf at 105 Hz and -3 dB for the first-order high shelf at 3 kHz. The loudness was not compensated for in this study when the filters were adjusted, so it may have had a subtle effect. Previous studies had shown that without compensation, test subjects did increase low frequencies about 2 dB more than when using loudness normalization. (McMullin 2017.)

Both Harman's and Samsung's studies have tended to be a little more targeted towards consumer objectives. But there has also been a lot of research into frequency response on the professional audio side. In 2001, the Finnish studio monitor manufacturer Genelec tested a total of 372 of their loudspeakers in 164 professional studios after calibration (Mäkivirta & Anet, 2001, 3). A slight tilt can also be seen in the median, derived from the frequency responses of professional studios (Figure 39).

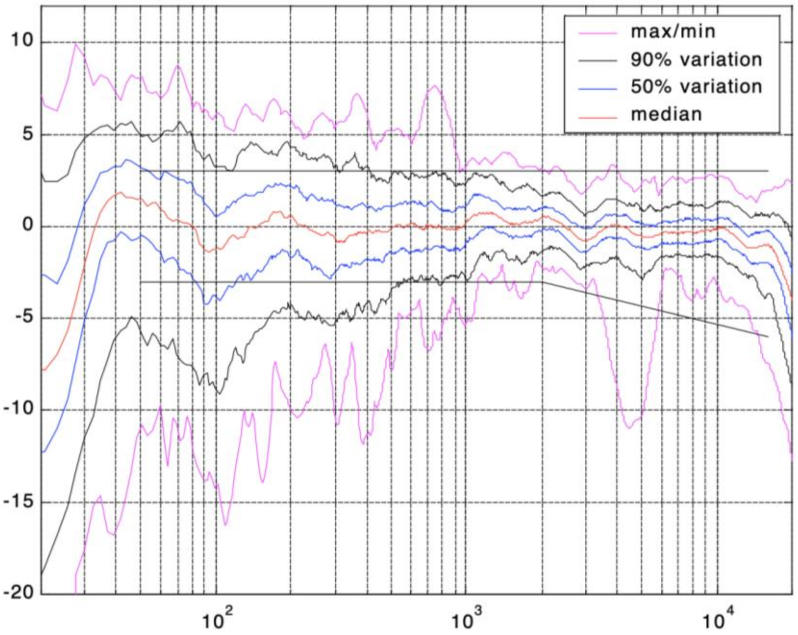

FIGURE 39. Genelec monitors (n=250) in professional studios after calibration (Mäkivirta & Anet, 2001, 3)

Dr. Toole has shown in his studies that flat on-axis frequency response of a loudspeaker in anechoic chamber is not perfectly flat in normal rooms. Depending on the acoustics of the room, the directivity characteristics of the loudspeaker and the listening distance, the desired response tends to slightly decrease from the low frequencies to the upper frequencies. Figure 40 shows that when a loudspeaker with a flat on axis frequency response and a uniformly variable directivity index is placed in rooms with different acoustics, the frequency response is also different. As most mixing and mastering rooms have been acoustically quite treated and the reverberation times are short, the target curve is usually more likely to be quite flat.

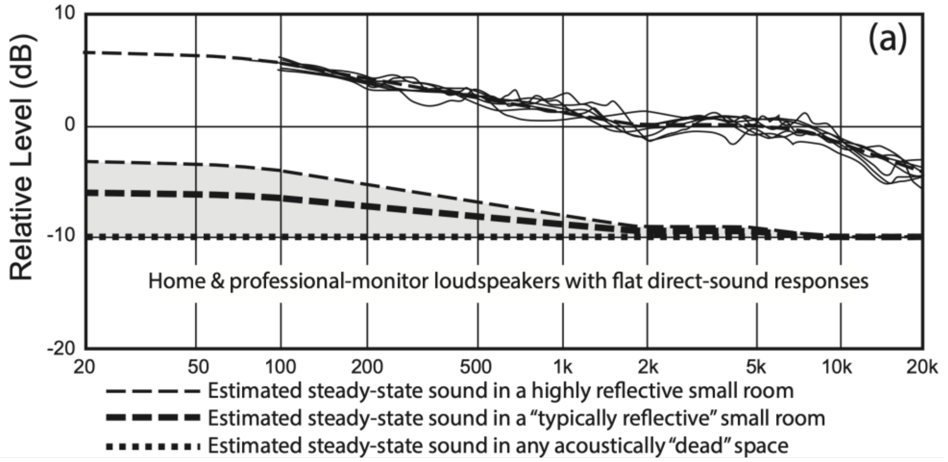

FIGURE 40. Room curves in different acoustical spaces (Toole 2015, 527)

Technical specifications for two-channel studio listening rooms can be found in technical papers EBU 3276 (1998) and ITU-R BS 1116-3 (2015). The specifications between electroacoustic and acoustic parameters are basically the same in these documents. The latest standard (ITU-R 2015) for room response curve is very broad (Figure 41). With the 1/3 -octave smoothing used, the tolerance window allows practically any shape of target curve to meet the standard. It does not guarantee quality sound, only that the loudspeakers will make sound (Olive et al. 2013b, 14). Smoothing means averaging certain frequency bands. This is often used to make the frequency response curves better match the frequency discrimination accuracy of the human auditory system (Laukkanen 2014, 20). One argument for 1/3 -octave spectral analysis has been that it is a rough approximation to the "critical bands" that can be detected. The concept of critical bands was already introduced by Fletcher (1933). According to Toole, the claim that this is the resolution of our hearing is simplistic and misleading. Another concept related to the auditory filter is the equivalent rectangular bandwidth (ERB), which are about 1/6 or 1/4 -octave wide at mid and high frequencies. There are situations where smoothed measurements are useful to describe general trends in spectral balance, such as room curves, but a full understanding of loudspeaker performance requires measurements with a resolution greater than the critical band or ERB. (Toole 2018, 87–89.)

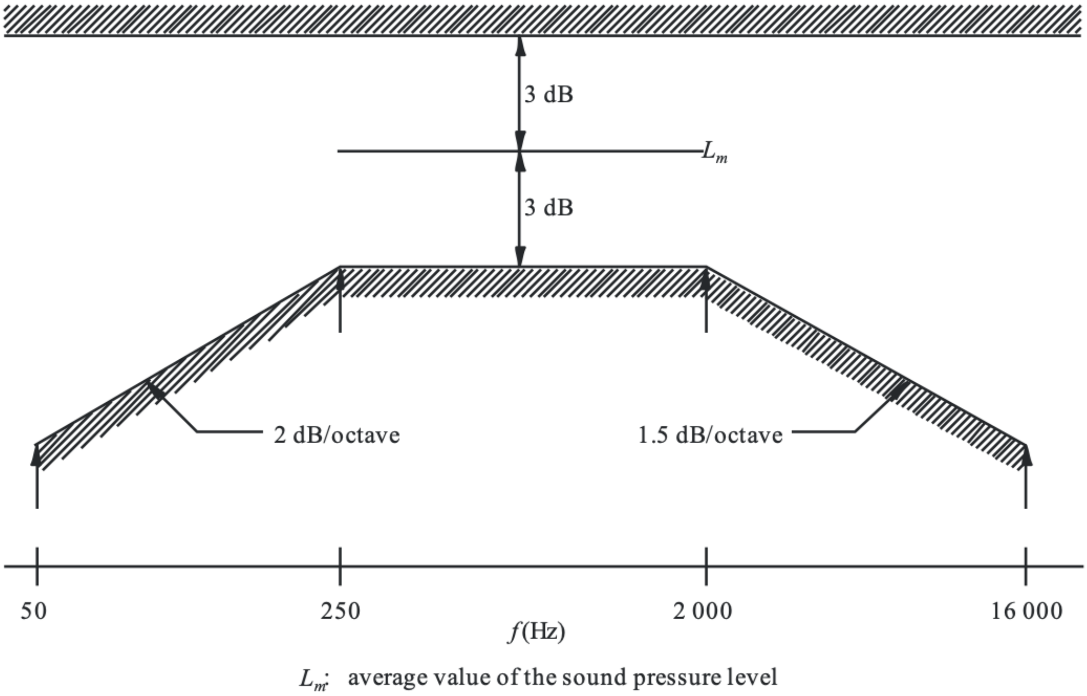

FIGURE 41. Tolerance limits of room response (ITU-R 2015, 16)

Listening tests have also shown that EBU Tech 3276 compliant control rooms do not guarantee a uniform and neutral sound image (Thiele 2016, 2). The standards and their measurement methods are outdated, so they are of no use when looking for factors that contribute to neutral sound quality.

At MojoLab, the target response of the listening system has changed over the years due to different rooms, changing acoustic solutions and the loudspeakers used. Also, the listening preferences have evolved as the experience has grown and it has been noticed, e.g., how the material translates to other listening rooms and systems. Many types of target curves have also been tested, including the completely flat, Bruel & Klaer target curve (Møller, 1974), Toole's idealized room curve (Toole 2018, 344) and all the revisions of Harman and Samsung in-room curves (2009, 2013, 2017). In addition, the target curve (Katz 2017) favored by renowned mastering engineer Bob Katz has been tested. All these curves (Figure 42) fits to the standards, but none of them have been immediately proven to be perfect in our listening system.

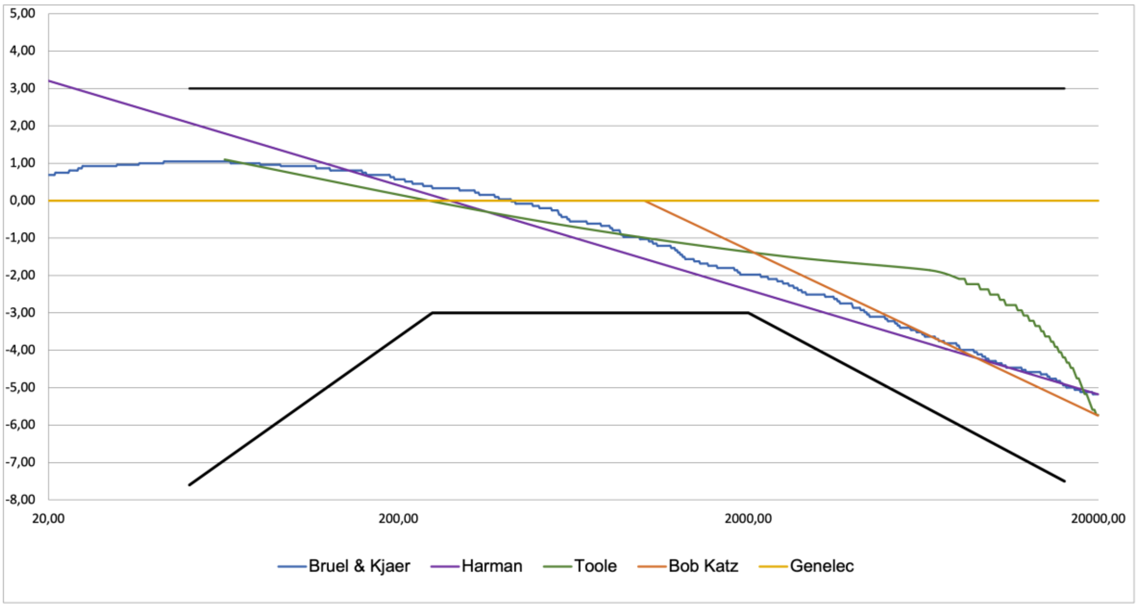

FIGURE 42. Different steady-state room curves

As Toole has mentioned, the acoustics of the room, listening distance and available loudspeakers, among other things, all play a role. There are no perfect one size fits all solution. In my experience, the best results are obtained by equalizing the direct signal to flat at the upper frequencies. At low frequencies, a small increase may be desirable below 100 Hz or so, where the human ear is less sensitive and where normal room gain occurs. Room gain means that each room emphasizes low frequencies depending on its acoustic characteristics. Pedersen & Morten (2007a) and Pederson & El-Azm (2007b) from Lyngdorf Audio argued in their studies that this natural phenomenon should not be removed during room correction to preserve the natural timbre of the room.

Steady state measurements include both the direct signal and the first and lateral reflections of the room. By using gated measurement windows, we can examine the direct signal. By keeping the time window open only a few milliseconds after the first sound waves have reached the microphone, room reflections or reverberations can be excluded from the measurement. This also reveals any nonlinearities in the speakers and allows them to be corrected if necessary. The EBU standard recommends using an equalizer only for low frequencies ($f < 300$ Hz). According to email conversations with Dr. Brüggemann, the developer of the Acourate software, he considers this information to be outdated. In his opinion, speakers are never perfect at the highest frequencies and a room correction should be done for the whole frequency band (Brüggemann 2018).

Because the loudspeakers at MojoLab are very directional specially in the high frequencies, and the lateral reflections are more dampened that flat direct sound, the steady-state in-room response is decreasing 1 dB in the high-end. The room is not anechoic chamber, so the low frequencies are slightly accentuated. They have been adjusted so that they sound as loud as the midrange frequencies based on auditory perception. The frequency response is +/- 2 dB between 20 Hz – 20 kHz with 1/3 -octave smoothing. The in-room response even meets EBU/ITU measurement standards for reference speakers in an anechoic chamber, despite the emphasis on low frequencies. The target curve can be switched to other stored curves quickly if needed to quickly test what e.g., a flat curve sounds like. In-room response and for comparison, the ITU tolerances and Tooles recommendations for different acoustically treated control rooms are shown in the Figure 43.

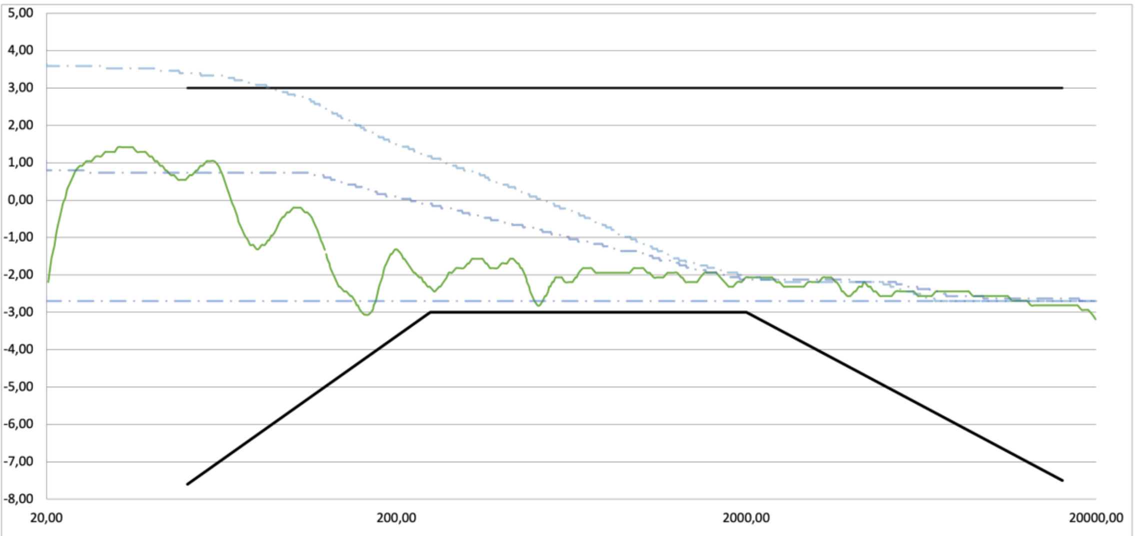

FIGURE 43. In-room frequency responses at MojoLab against the ITU-R BS-1116 standard and Toole's proposals

Since the headphones should sound similar in timbre to high-quality loudspeakers used in a professional studio, we decided to measure our mastering room with DGSonicFocus software and use it as one of the targets in our headphone project (Figure 44).

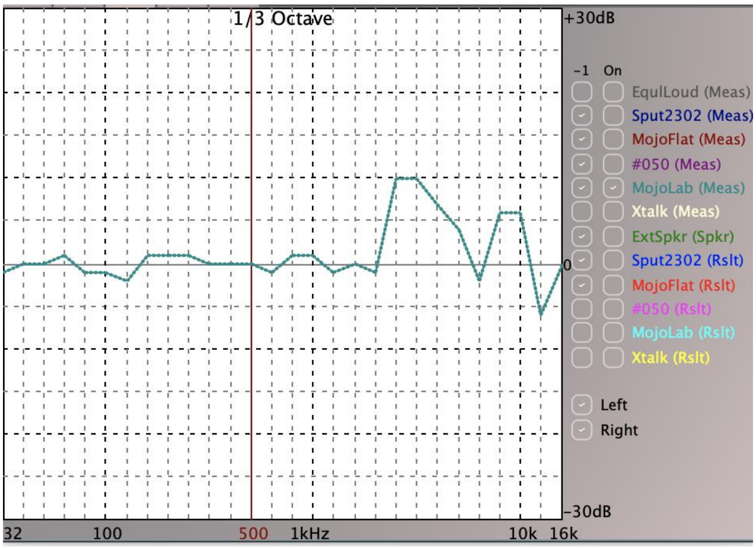

FIGURE 44. Mastering room measured with DgSonicFocus

4.3. Headphone modification project

Since the perfect headphones had not been found in testing over 30 different commercial headphones, we decided to start modifying one headphone model to better meet our needs. In our background studies, the inexpensive MiniDSP

EARS provided somewhat usable data that also correlated with the listening observations. It would also allow us to get feedback on how the headphones correlate with the Harman target. This was chosen as a starting point in our project, as it had been shown to be the most recommended target curve in studies. In addition to using the Harman target to modify the headphone, we also wanted to use listening observations to test how the headphone compares to different professionally used mixing and mastering control rooms.

The model of the headphone and further details of the modifications are not discussed because they are beyond the scope of the thesis. Headphone also proved to be commercially successful, so these things has been kept ourselves. All adjustments were made by adjusting the acoustics of the headphones, with no electronic correction or replacement of electronic components in the headphones.

4.3.1 Revisions of the modified headphones

After more than thirty different headphones where measured and tested by listening test, one specific model was selected for the modification project. The model had workable measurement results and was inexpensive, so it was good for the project, because we did not have to be so careful with the different experiments. The headphone also had a relatively natural midrange reproduction. Low and high frequencies were pronounced, but from a modifying perspective, it was easier to reduce excessive low or high frequencies than to try to create something that the element cannot produce. Figure 45 shows the frequency response of the fifty stock headphones used in the modification project, measured with MiniDSP EARS and its HEQ target.

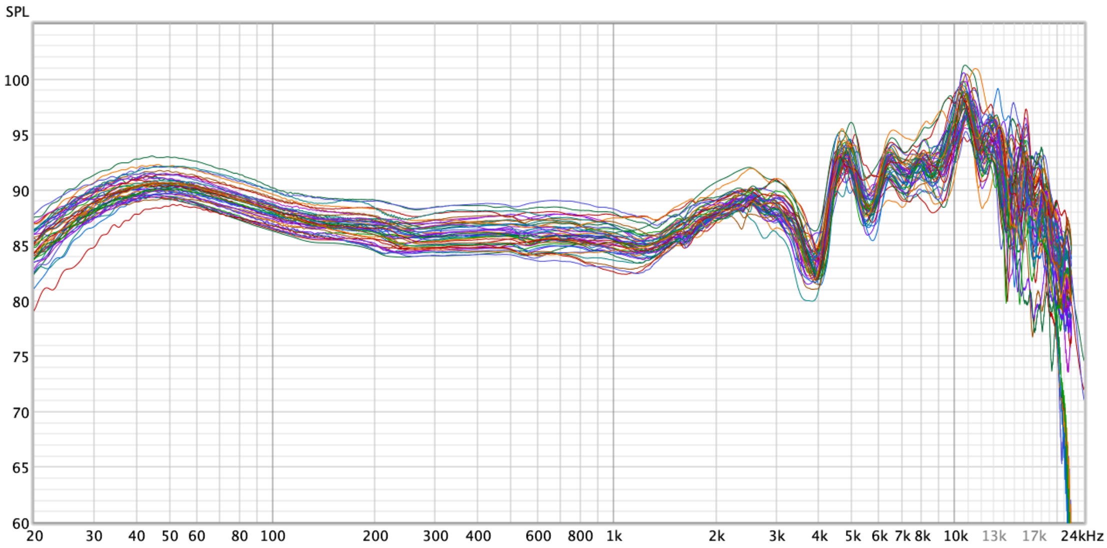

FIGURE 45. Fifty stock headphones before modification (no smoothing)

During the project, fifty modified headphones were done, and the headphones underwent four major revisions. Each buyer of the headphones was able to influence the result by selecting the headphones they liked best from the slightly differently modified headphones. A few experienced audio professionals were also consulted in more detail, and their feedback was used to help shape future versions. The fit and comfort of the headphones were also constantly improved as the project progressed.

At low frequencies, already the first versions worked well. In the 5 kHz range, there was a slight accentuation in the headphones that made the modifications difficult. If this problem area could be made perfect, the higher frequencies above it were too attenuated, so compromises had to be made. Also, with the 3 kHz range, a lot of effort was made to find a solution that was a suitable compromise for that range. Due to slightly different ear canal lengths, different people have different accentuation in this range. In the latest version, a good solution was found for the 5 kHz accent, so this frequency could be kept under control, without attenuating the higher frequencies too much. All fifty modified versions of the headphones are grouped by revision in Figure 46, with the first revisions at the top, newer revisions below it and the latest revisions at the bottom.

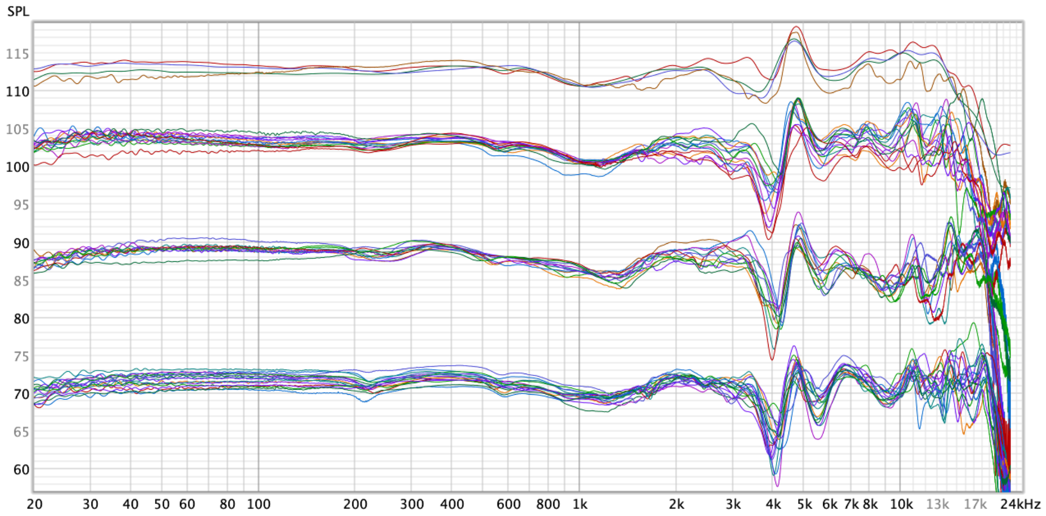

FIGURE 46. Fifty modified headphones (no smoothing). The four revisions are divided into own groups. Revisions from the oldest to the latest, sorted from top to bottom.

In Figure 47 we can see how the headphone modification process progressed. The green line represents the average of the fifty unmodified headphones. The other colors represent the different revisions. Compared to the original unmodified headphones used as the base model for the project, low and high frequencies have been attenuated and are now more balanced and neutral. The midrange reproduction has been kept natural, but the upper midrange frequencies have been smoothed out. The comfort of the headset has also been improved with softer and less sweaty pads.

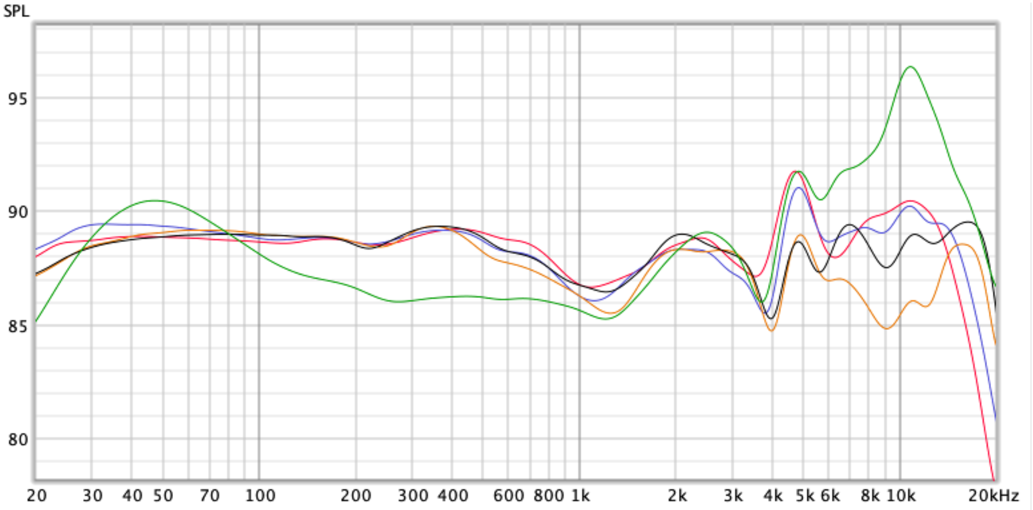

FIGURE 47. Fifty headphones averaged and grouped by unmodified and modified revisions with 1/3 -octave smoothing. Unmodified = green, rev1 = red, rev2 = blue, rev3 = yellow & rev4 = black.

4.3.2 Results

The HEQ target curve of the MiniDSP EARS device, based on the Harman headphone target, turned out to be quite useful. At the beginning of the project, the headphones were adjusted mainly based on measurements taken with this device. Towards the end of the project, the fine tuning was based more on listening observations. The goal was also to get the headphones as close as possible to the timbre of the MojoLab studio loudspeakers, and I think it succeeded quite well. So well, in fact, that you could no longer tell which was the reference for the other.

In Figure 48 we can see the #050 of the headphones compared to the loudspeakers in the MojoLab. We can see that the headphones match quite well to our mastering room and the timbre is very similar. The 2.5 - 3 kHz range could not be brought up to the same level without compromises. It could be raised, but at the cost of lower midrange frequencies being lowered. On the other hand, the level is now correct according to the MiniDSP and its Harman target. Also, e.g., in HD600 this area is more boosted, and it have been always dislike in our opinion. We have also received feedback that the MojoPhones in this area are more balanced than the HD600.

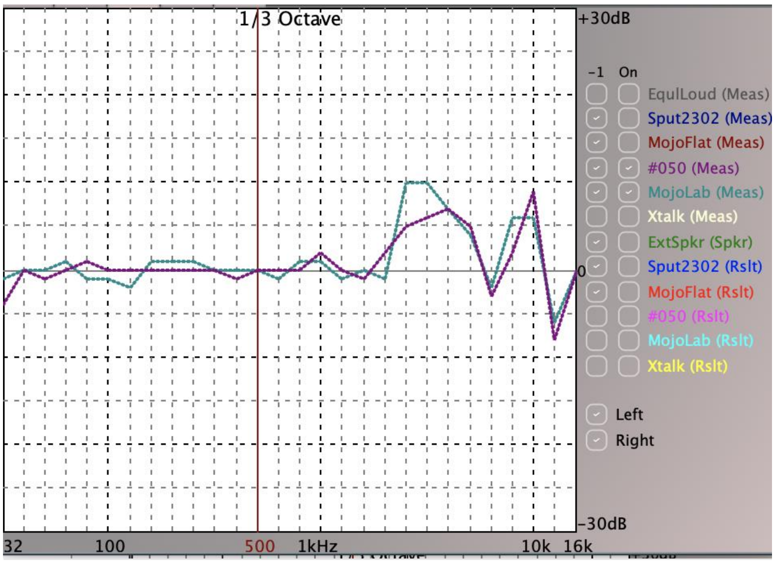

FIGURE 48. Modified headphone #050 (purple) versus MojoLab room (cyan)

After investigating this issue, we found that many loudspeakers and systems have dip in this frequency range. A dip in this area could be either deliberate or the result of poor speaker design. In poorly designed loudspeakers on-axis response might be flat, but directivity mismatch in midrange and tweeter drivers cause sound power response issues in this crossover range. This small notch in the 2 - 3 kHz range is highlighted in Holman & Green's extensive study of 1375 home theatre speakers in 275 different rooms. In the pre-equalized measurements, the dip was already noticeable, but this bump was wanted to be further amplified in the equalized results. (Holman & Green 2010, 18.) This phenomenon is also evident in Toole's loudspeaker studies (Toole 2018, 344).

Sometimes this notch is done on purpose. A dip at this frequency is often associated with the term BBC dip. BBC engineer and man behind legendary British Harbeth loudspeakers, H.D. Harwood wrote in 1974 that flat response sounds unnatural. Orchestral music in particular sounds too close and much better sense of perspective is obtained when slight dip in 1 - 3 kHz region is applied. Few dB is enough to provide more natural perspective without destroying the sound quality. (Harwood 1976, 50.) Siegfried Linkwitz, the late well-known loudspeaker designer, author in the field, and co-inventor of the Linkwitz-Riley filter mentions that he often uses a 4 dB psychoacoustic dip around 3 kHz to add greater realism and allow louder playback levels (Linkwitz 2022).

Audyssey, a leading provider of research-based room correction solutions, uses by a default "Midrange compensation" in their products that makes a dip of about four decibels in the 1.5 to 3 kHz range (Audyssey 2021, 25–26). Another manufacturer of room correction products, Finnish DSPeaker, also recommends trying -4.2dB “psychoacoustic cut” at 3kHz with a bandwidth of around 1000Hz. According to DSPeaker, the idea behind the 3 kHz cut is that human hearing is insensitive to diffuse fields in this range. Because of the flat response of the microphones, many recordings have too much energy in the 3 kHz range compared to the original listening situation. To compensate for this, the user must judge for himself whether the cut is beneficial. (DSPeaker 2014, 18.)

Our speaker system theoretically reproduces this area correctly, on-axis response of the loudspeakers is flat, and the power response is smooth and evenly descending. With our system DSP it would be easy to equalize that range, but this has not yet been tested to see what effect this would have on the perceived sound quality. The most important thing in mastering is not whether the material sounds as pleasant as possible in the mastering room. What matters is that the mastering engineer can trust what he hears and, based on that, make the right decisions that will translate as widely as possible on different listening devices and listening environments. However, this is an interesting topic, and our further research needs to look more closely at what is the correct level in this frequency range in our loudspeaker system.

In Figure 49 we can see latest version of the modified headphones (MojoPhones #050) and other reference headphones used in our project measured with MiniDSP EARS. Since the flat measurement result with HEQ compensation corresponds Harman's target, we can say that the modified headphones match it very well. The measurements of the MojoPhones are the closest to the Harman target of all the headphones measured in the project. Even Harman's own-brand AKG k371, considered one of the best headphones to meet Harman's target, pales in comparison. MojoPhones has more balanced low frequencies than the HD600 and k371, which are the weak points of these headphones. The weak point of the Audeze is its slightly soft upper midrange frequencies, which make everything sound a little too good and forgiving. In the MojoPhones, these frequencies are more pronounced, bringing the sound more in your face. The frequency response of the original headphones to HEQ compensation was within +/- 5,5 dB and the latest revisions are within +/- 2 dB with 1/3 octave smoothing.

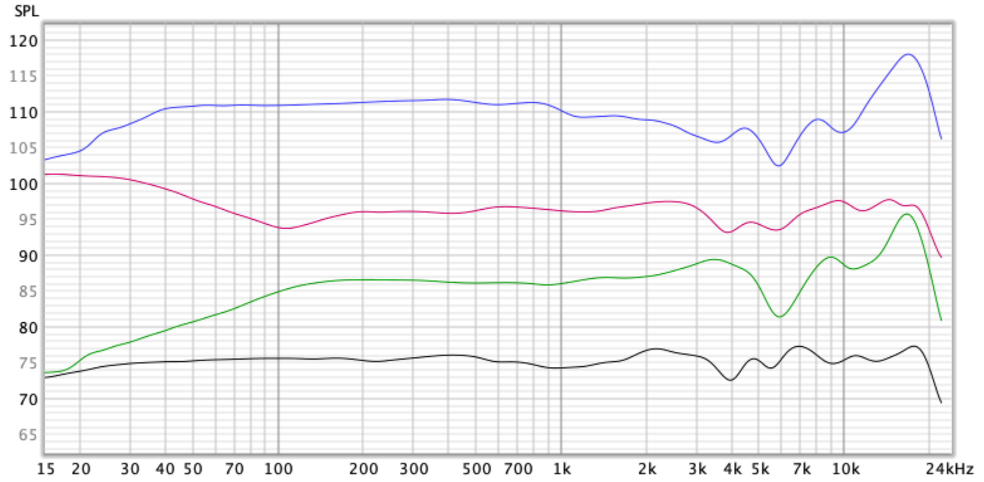

FIGURE 49. Modified headphones (MojoPhones #050) versus three highly respected headphones. Blue: Audeze LCD-3f, red: AKG k371, green: Sennheiser HD600 & black: MojoPhones

Feedback on the headphones has been very positive, and everyone who has used them has said they are one the best headphones they have used. Other common comments have been that the headphones reproduce the entire sound range evenly, right down to the lowest frequencies. You can mix with the headphones, and you do not have to equalize them at all. You can hear the balances of he mixes very well, and they have the same kind of resolution and reproduction as, e.g., Genelec or Amphion that the users had used in their studios. These headphones are great tools of the trade, and the overall sound is reminiscent of vintage ribbon mics - those who are tired of the "hi-fi" top end of modern headphones will enjoy these.

Early versions of the headphones put a little pressure on the ears after long sessions. This has been corrected in later versions, and now there has been also positive feedback on the comfort of use. MojoPhones (Picture 4) are available to order and can be delivered anywhere in the world.

PICTURE 4. MojoPhones promo shoot session in relaxing Thailand. Now there is a pleasant reason to be working even on holiday

5 CONCLUSIONS AND DISCUSSION

The project lasted about three years and was very rewarding in terms of what was learned about the subject. At times the project was really demanding because of the complexity of the whole thing and the fact that there was a lot to learn and even new things to be invented. The subject was sometimes complicated by the fact that, e.g., many psychoacoustic topics were difficult to understand. There was a lot of information and research data on the topic, but sometimes it felt challenging to narrow down what made sense to cover within the scope of the thesis.

MiniDSP EARS proved to be very useful, even though it is not industry standards compliance research tool. As experience gained of using it, it became clear that the measurement results could not be fully trusted. However, the reason was not always the device, but the way the measurement results were interpreted. When looking at the measurement results, it is important to understand that they are only indicative. You need to know which things in the results are relevant and audible. In general, the ear is more sensitive to peaks than dips (Johansen 2006, 3). The width of the peak plays also a role. With certain musical material, the ear may not detect a really narrow peak, because the music signal may not be hitting that frequency accurately or playing long enough for the resonance to be noticed. Therefore, a much lower peak with a wider footprint can be much more audible than a narrow, very high peak (Toole 2004, 17).

It is important to evaluate headphones with a wide range of music material, rather than listening to just one or two songs. In this way, we cover a wider range of resonances and frequencies. Studies have been done on which tracks are suitable for assessing differences in headphone sound quality. Olive, Welti & Khonsaripour (2017a) investigated this topic and looked at ten different music tracks and their suitability for headphone assessment. Songs such as "Bird on a Wire" by Jennifer Warnes and "Cousin Dupree" by Steely Dan, which have been used in many studies, were found to give good results in terms of discrimination and reliability. Key factors for a good track included the bandwidth of the spectral content and how familiar the material was to the reviewer. It is also desirable to use only short loops of songs which spectral and temporal characteristics do not vary greatly over the duration of the loop. This will help listeners to evaluate the headphones according to ITU recommendations (ITU-R 2015).

The thesis also included measurements based on listening observations, but these were not without problems either. It was found that the measurement results were not always repeatable, and many things influenced the results. When listening to test signals, hearing becomes fatigued relatively quickly. Care must be taken to maintain a fixed volume level during the different test periods. According to Poldy, SPL about 70dB would be suitable for bandpass loudness comparisons (Poldy 2001, 670). Dr. Ulrich Horbach notes in his AES paper (Horbach 2015) that the single 500 Hz reference noise band used in the Griesinger´s program might not be the optimal. With bands that are wide apart, it is difficult to adjust loudness differences with sufficient accuracy, making the method unreliable and hardly repeatable. His solution was to divide the audio band into five different overlapping sub-bands, each with its own center reference band. Poldy also states that when measuring frequency response using the subjective loudness balancing method, at frequencies below 250 Hz the test signal should be changed from filtered band noise to sine wave. This is because of the stochastic nature of filtered band noise. (Poldy 2001, 671.)

These two problems were also found in our own experiments. In an email conversation with Horbach, he said that in the end, his method did not produce completely reliable results either. However, he had continued his research in this area but could not yet reveal his findings. (Horbach 2022.) One of our own future interests is to explore better ways to implement loudness balancing methods. By comparing the frequency response of the headphones to the loudspeakers in the room using listening observations, we can more straightforwardly compare how the two differ. This can be done without the need for expensive measuring equipment that does not fully match the human hearing, at least not with individual accuracy. This is also the approach that Genelec has started to use in the recently released Aural ID headphone technology. They recommend a listening distance of 20 inches when comparing the timbre of the headphones to these ultra-near field calibrated speakers. When I asked them why the headphones should be calibrated in an ultra-near field, rather than at the normal optimized listening position in the studio, Aki Mäkivirata and Thomas Lund from Genelec answered that either way works. In the ultra-near field, the room effect can be minimized, but if you want the headphone to match the interaction between the speakers and the room as closely as possible, comparison from the listening position is the better way to go. (Mäkivirta, Lund 2022.) The latter approach was used in our project. The acoustics of control rooms and how well the loudspeakers are set up to work together with the room is an important element in neutral sound reproduction. In the future, I will continue to explore these issues and increase my knowledge on the subject. After all, an acoustically neutral control room with well-tuned, high-quality loudspeakers is what the target curve for headphones should be based on.

During the headphone modification phase, it was found that there was quite a lot of individual variation in the modified headphones. A modification that worked in one of the headphones did not necessarily work perfectly in another. This required the ability to make many different types and levels of modifications to accommodate individual differences in both the headphones and the user's preferences. A lot of trial and error had to be done to find new ways to modify the headphones. A variety of different parts and materials were also tested that would change the sound in the desired way. As other people became interested in headphones, the project was no longer just about making headphones for yourself. This increased the pressure, because in addition to your own goals, you also had to be able to meet other people's wishes.

Not all standards in the industry are free and many of them are relatively expensive, e.g., the IEC 60268-7 standard cost 320 € and IEC 60318-7 190 €. In the context of this thesis, I did not have access to all these paid standards, so I only dealt with them superficially. Measuring equipment that meets industry standards are expensive, ranging from 15 000 € to 50 000 € with two-channel configuration. During the thesis I was not able to test any of these personally, so I had to rely on the judgements of others. However, I did find out that in Finland at least Aalto University have GRAS 45BC KEMAR and 45CA simulators for rent. They rent their facilities and equipment for a minimum duration of one day, so it did not make financial sense for me to use their services at the end of the thesis just to measure the final version of the headphones. Especially when the 45CA standard pinna set (KB0070/KB0071) may not provide reliable headphone measurement results, as Welti found in his 2015 study. Pinna is too rigid because it was originally designed for testing Behind The Ear (BTE) hearing aids and has problems sealing some headphones properly (Welti 2015, 5).

A telephone conversation with Aleksi Öyry, research engineer at Aalto Acoustics Lab, revealed also that KEMAR is not without problems when it comes to measuring headphones. According to Öyry, because the surface of the removable pinnae is not perfectly flush with the shape of the head, some circumaural type of headphones have problems with sealing the headphones (Öyry 2022). So even the industry-standard devices have shortcomings, and the technology is still somewhat lacking because it is very difficult to fully model the human auditory system. Product development will certainly continue in the field, and consumer grade equipment will also improve as the equipment used in scientific research improves. Unbranded test jigs, couplers, and pinnae are already available from China that, at least according to them, meet industry standards. It would be interesting to test these in the future and build a measurement device at a fraction of the cost of the well-known brands.

The IEC/ITU/EBU headphone standards currently do not serve the industry, as the recommendations produce sub-optimal results and are ignored by most manufacturers (Olive et al. 2018b, 10). It is to be hoped that manufacturers would invest more in quality headphone research, as people are increasingly listening with headphones. As Olive et al. states: “While most loudspeaker manufacturers today aim to achieve a flat frequency response on-axis, headphone manufacturers seem to be aiming at a target response that is as variable and random as the weather.” (Olive et al. 2013a, 2). According to current research, the Harman target is the most preferable for non-spatial content such as stereo music. However, when listening to spatial audio content, a flat headphone response should be preferred over the Harman target. (Engel et al. 2022, 278.) I look forward to seeing what new research we will see in this area. Especially when the studies are obtained using the latest versions of measuring devices, that include anthropometric pinnae, fully anatomically shaped ear canals and ear simulators that allows human ear impedance loading up to 20 kHz.

During the thesis, the question arose that could the Harman target be better. The target was formed by placing the HATS at a listening point in the reference room and taking the spatial average by measuring each seven surround loudspeakers in three different HATS horizontal angles (-30, 0 and +30 degrees). It would be interesting to see how the target would differ if only stereo loudspeakers had been used. The target could have a pinna notch in the 8 kHz range, which is observed in the 30 -degree azimuth HRTF measurements. This phenomenon was also seen in this thesis, when rooms with two loudspeakers were measured by listening observations. In an email conversation with Harman's Todd Welti, he specified that the Harman Target curve itself was derived using a spatial average of all seven speakers and three different HATS angles. However, in the actual Method of Adjust tests, only the left and right front speakers were used. Welti speculated that perhaps the reflections from the semi-diffuse room used in the tests, as well as the average derived from the HATS at three different head positions have filled in the blanks whose absence I wonder about.

In my view, this is an area for further research. Diffuse field and Harman target do not possess this natural phenomenon seen in our studies, because the target is the sum of equally strong sounds coming from many different angles. This may be also one of the reasons why headphones in general are unnecessarily bright. But as Griesinger mentioned about modern headphones: "They sell better this way. Accurate is not always perceived as best." (Griesinger 2013, 9).

My future goal is to transfer the knowledge I learned in this project to building my own headphone brand. I believe that by building an entire product from scratch, I will be able to make fewer compromises and the result will be better than these modified headphones. But now is a good time to distance yourself from the project for a while and just enjoy a good pair of headphones. The thesis is now finished, and the objectives were achieved, and expectations were even exceeded. It is time to look ahead to the other challenges and objectives that have been emerged or neglected during this thesis.

Thanks are due to Tampere University of Applied Sciences, its teachers, students, and all the people who have been involved in helping with this project. Special thanks to Tipi Tuovinen, who was the first person who wanted to buy the headphones and provided important feedback specially in the early stages of the project. I am also grateful to you for the promotional work you have done for the headphones. Thanks to Petri Levonen, whom I could always rely on when I needed confirmation of what I was hearing. We did not always hear the same things, but your analyses were always informative and gave me a new perspective on things. Thanks to the thesis supervisor Jukka Holm, who pushed me to start writing and was attentive from the beginning to ensure that the written part of the thesis was also of highest quality. There may still be some errors, but the author takes full responsibility for them.

Big thanks also to Leena Mäkelä who made the thesis possible, Timo Kivikangas who urged me to apply for these studies and Klas Granqvist who was examiner of the thesis. A big hug to my family and apologies for me being so focused on the project that I have not always been as present as I should have been.

REFERENCES

Akg. 2022. Our Story. Read on 13.4.2022. https://www.akg.com/history.html

Algazi, V., Avendano, C., & Thompson, D. 1999. Dependence of Subject and Measurement Position in Binaural Signal Acquisition. Journal of the Audio Engineering Society, vol. 47, no. 11, 937–947.

Algazi, V., Duba, R., Morrison, R. & Thompson, D. 2001. Structural Composition and Decomposition of HRTF's. UC Davis: Institute for Data Analysis and Visualization.

Armstrong, C., Thresh, L., Murphy, D. & Kearney. G. 2018. A Perceptual Evaluation of Individual and Non-Individual HRTFs: A Case Study of the SADIE II Database. October 2018. Applied Sciences 8(11):2029.

Audio Precision. 2020. Headphone Electroacoustic Measurements, Technote TN143. Read on 3.9.2021. https://www.ap.com/wp-content/uploads/2020/11/Audio-Precision-TN143-Headphone-EA-Measurements.pdf

Audyssey Laboratories. 2021. MultEQ-X User Guide v1.0.

Audio Precision. 2016. Application Note: Headphone Electroacoustic Measurements. Read on 14.9.2021. http://opt-in.ap.com/headphonetest-Offerbased2018-en/images/AudioPrecision-AppNote-Headphone-EA-Measurements-0617.pdf

Beardsley, H. 1929. Operators' Transmitters and Receivers. USA: Bell Laboratorie Record Vol. 7 January 1929.

Beyerdynamic. 2022. Story. Read on 13.4.2022. https://global.beyerdynamic.com/company/about-us/story

Blauert, J. 1969. Sound Localization in the Median Plane. Acta Acustica united with Acustica, Volume 22, Number 4, 1969, pp. 205-213(9). Germany: S. Hirzel Verlag.

Breebaart, J. 2017. No correlation between headphone frequency response and retail price. The Journal of the Acoustical Society of America 141, EL526 (2017).

Brüel & Kjær. N.d. TYPE 5128 HIGH-FREQUENCY HEAD AND TORSO SIMULATOR. Read on 22.3.2022. https://www.bksv.com/en/transducers/simulators/head-and-torso/hats-type-5128

Brüggemann, U. Owner of the AudioVero 2022. Re: Web contact AudioVero. Email message. Read on 29.3.2022.

Bomhardt, R. 2017. Anthropometric Individualization of Head-Related Transfer Functions Analysis and Modeling. Germany: Logos Verlag Berlin GmbH.

Corbett, I. 2020. Mic It!: Microphones, Microphone Techniques, and Their Impact on the Final Mix. 1st edition. UK: Routledge.

DSPeaker. 2014. Anti-Mode 2.0 Dual Core Owner's Manual Rev. 1.1.

Diy-Audio-Heaven. 2019. Fletcher-Munson. Read on 10.9.2021. https://diyaudioheaven.wordpress.com/tutorials/fletcher-munson/

Engel, I., Alon, D., Scheumann, K., Crukley, J. & Mehra, R. 2022. On the Differences in Preferred Headphone Response for Spatial and Stereo Content. Journal of the Audio Engineering Society, vol. 70, no. 4, 271-283.

European Broadcasting Union (EBU). 1998. EBU Tech. 3276 – 2nd edition - Listening conditions for the assessment of sound programme material: monophonic and two–channel stereophonic.

Fleischmann, F., Silzle, A. & Plogsties, J. 2012. Identification and Evaluation of Target Curves for Headphones. Audio Engineering Society Convention Paper 8740.

Fletcher, H. & Munson, W. A. 1933. Loudness, its definition, measurement and calculation. The Journal of the Acoustical Society of America 5, 82.

Genelec. 2020. Aural ID. Support page. Read on 10.9.2021 https://support.genelec.com/hc/en-us/articles/360015520800-What-sort-of-equalization-I-need-to-apply-to-my-headphones-when-listening-to-audio-that-is-binauralized-using-my-Aural-ID-

Genuit, K. 1984. Ein Modell zur Beschreibung von Außenohrubertragungseigenschaften. Doctoral dissertation. Germany: Rheinisch-Westfalischen Technichen Hoch-schule Aachen.

Griesinger, D. 2009. Frequency response adaptation in binaural hearing. Audio Engineering Society Convention Paper 7768.

Griesinger, D. 2013. Binaural Hearing, Ear Canals, and Headphone Equalization. Powerpoint presentation. Read on 30.3.2022. http://www.davidgriesinger.com/Binaural_hearing_and_headphones.ppt

Griesinger, D. 2022. Welcome to DGSonicFocus and David Griesinger’s home page! Read on 26.2.2022. http://www.davidgriesinger.com

Guicking, D. 1990. On the invention of active noise control by Paul Lueg. The Journal of the Acoustical Society of America 87, 2251.

Gunnarsson, V. 2016. How to Make Headphones Stereo-Compatible. Read on 30.3.2022. https://www.dirac.com/how-to-make-headphones-stereo-compatible/

Hammershøi, D. & Møller, H. 1996. Sound transmission to and within the human ear canal. The Journal of Acoustical Society of America. Jul;100(1):408-27.

Harwood, H. 1974. Some factors in loudspeakers quality. Wireless World, May 1976.

Hertsens, T. 2018. InnerFidelity Headphone Measurements. Read on 23.10.2021. https://www.stereophile.com/content/innerfidelity-headphone-measurements

Horbach, U. 2015. Characterizing the Frequency Response of Headphones – a new Paradigm. Audio Engineering Society Convention Paper 9274.

Horbach, U. Engineering Fellow 2022. VST plugin in your aes paper 9274. Email message. Read on 29.3.2022.

Howeth, L. 1963. History of communications-electronics in the United States Navy. U.S. Government Printing Office.

Iida, K. 2019. Head-Related Transfer Function and Acoustic Virtual Reality. Singapore: Springer.

International Electrotechnical Commission (IEC). 2010. IEC 60268-7:2010 Sound system equipment – Part 7: Headphones and earphones, Edition 3.0. International Standard.

International Organization for Standardization. 2018. ISO 11904-1:2002 (R2018), Acoustics — Determination of sound immission from sound sources placed close to the ear — Part 1: Technique using a microphone in a real ear (MIRE technique).

International Organization for Standardization. 2021. ISO 11904-2:2021 Acoustics — Determination of sound immission from sound sources placed close to the ear — Part 2: Technique using a manikin.

ITU Radiocommunication Sector (ITU-R). 1990. Rec. ITU-R BS.708-0 (06/90) – Determination of the electro-acoustical properties of studio monitor headphones. Recommendation paper for Broadcasting Services (sound).

ITU Radiocommunication Sector (ITU-R). 2015. Recommendation ITU-R BS.1534-3: Method for the subjective assessment of intermediate quality level of audio systems. (10/2015)

ITU Telecommunication Standardization Sector (ITU-T). 2021. Recommendation ITU-T P.57: Artificial ears. (06/21).

Johansen, L. 2006. Psychoacoustics and Audibility – Fundamental Aspects of the Human Hearing. Lecture note for the course TI-EAKU. Read on 14.4.2022. https://citeseerx.ist.psu.edu/viewdoc/dowload?doi=10.1.1.580.3611&rep=rep1&type=pdf

Jønsson, S., Matthisson, A. & Borg, C. Application Note: Improving Telephone Handset Performance – On the Use of Ear Simulators ITU-T P.57 Type 1 and Type 3.2 for Telephonometric Measurements 1997. Read on 8.3.2022. https://www.bksv.com/media/doc/bo0455.pdf

Kahrs, M. 2009. SPICE simulation of headphones and earphones. Audio Engineering Society Convention Paper 7843.

Katz, B. 2017. Katz's Corner Episode 17: Perfecting Perfection. Read on 28.4.2022. https://www.stereophile.com/content/katzs-corner-episode-17-perfecting-perfection

Kemar. 2021. The Life of KEMAR. Read on 23.10.2021. http://kemar.us

Klippel. 2020. KLIPPEL Analyzer System Price-List 15.3.2020. Read on 22.2.2022. https://www.klippel.de/fileadmin/klippel/Bilder/allgemein/Preisliste%20aktuell/Klippel%20Analyzer%20System%20Pricelist_March_15_2020.pdf

Klippel. 2020. KLIPPEL Analyzer System Price-List 1.9.2020. Read on 23.2.2022. https://www.klippel.de/fileadmin/klippel/Bilder/allgemein/Preisliste%20aktuell/KLIPPEL%20Analyzer%20System%20Price%20List_September_01_2020.pdf

Koss. 2022. Koss History. Read on 13.4.2022. https://koss.com/pages/koss-history

Cullen, K. & Sadeghi, S. 2008. Vestibular system. Scholarpedia, 3(1):3013. Read on 29.3.2022. doi:10.4249/scholarpedia.3013

Laukkanen, P. 2014. Evaluation of Studio Control Room Acoustics with Spatial Impulse Responses and Auralization. Master´s Thesis. Aalto University. School of Electrical Engineering.

Linkwitz, S. 2022. Models for a dipole loudspeaker design - Psycho-acoustic 3 kHz dip. Read on 30.4.2022. https://www.linkwitzlab.com/models.htm#H

Loeffler, J. 2020. Headphones: A History of the Original Werable Tech. Read on 12.8.2021 https://interestingengineering.com/headphones-a-history-of-the-original-wearable-tech

Lopez-Poveda, E. & Meddis, R. 1996. A Physical Model of Sound Diffraction and Reflections in the Human Concha," The Journal of Acoustical Society of America, Vol. 100, 3248-3259, 1996.

Lorho, G. 2009. Subjective Evaluation of Headphone Target Frequency Responses. Audio Engineering Society Convention Paper 7770.

Luther, C., 2022. About JENS: Sound Localization In The Median Plane. Read on 18.3.2022. https://science-of-sound.net/2016/01/about-jens-sound-localization-in-the-median-plane/

Macpherson, E. & Middlebrooks, J. 2002. Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. The Journal of the Acoustical Society of America 111(5 Pt 1):2219-36.

Mercadier, E. 1891. Bi-telephone. U.S. Patent No. 454,138. Read on 14.4.2022. https://patents.google.com/patent/US454138A/en

McMahon, M. 1973. Vintage Radio. Second edition. USA: Vintage Radio.

McMullin, E. 2017. A Study of Listener Bass and Loudness Preferences over Loudspeakers and Headphones. Audio Engineering Society Convention Paper 9855.

MiniDSP. 2021. EARS User Manual. Hong Kong: MiniDSP.

MiniDSP. N.d. Headphone EQ with EARS and REW. Read on 26.2.2022. https://www.minidsp.com/applications/headphone-equalization/headphone-eq-with-rew-ears

MojoLab. 2022. MojoLab – Mastering & Acoustics. Read on 13.4.2022. https://www.mojolab.fi

Mulcahy, J. 2022. Room Acoustics Software. Read on 26.2.2022. https://www.roomeqwizard.com

Møller, H. 1974. Hi-Fi Tests-with 1/3 Octave Pink Weighted Random Noise. Audio Engineering Society Convention Paper A-5.

Møller, H. 1992. Fundamentals of Binaural Technology. Applied Acoustics 36, 171-218. Denmark: Aalborg University.

Møller, H., Jensen, C., Hammershøi, D. & Sørensen, M. 1995. Design Criteria for Headphones. Journal of the Audio Engineering Society, vol. 43, no. 4, 218-232.

Møller, H. & Pedersen, C.S. 2004. Hearing at low and infrasonic frequencies. Article in Noise & Health 2004, 6;23, 37-57. India: Wolters Kluwer.

Mäkivirta, A. & Lund, T. Genelec R&D director & Senior Technologist 2022. DIscussion in G*Connect Webinar: Getting started with Aural ID, held on 28.4.2022.

NIDCD: National Institute on Deafness and Other Communication Disorders. 2016. U.S.Department of Health and Human Services – National Institutes of Health. NIH Pub. No. 97-4235. Read on 14.4.2022. https://www.nidcd.nih.gov/sites/default/files/Documents/health/hearing/AgeRelatedHearingLoss.pdf

Olive, S., Jackson, J., Devantier, A., Hunt, D. & Hess, S. 2009. The Subjective and Objective Evaluation of Room Correction Products. Audio Engineering Society Convention Paper 7960.

Olive, S., Welti, T. & McMullin, E. 2013a. Listener Preference for Different Headphone Target Response Curves. Audio Engineering Society Convention Paper 8867.

Olive, S., Welti, T. & McMullin, E. 2013b. Listener Preferences for In-Room Loudspeaker and Headphone Target Responses. Audio Engineering Society Convention Paper 8994.

Olive, S. & Welti, T. 2015. Factors that Influence Listeners' Preferred Bass and Treble Balance in Headphones. Audio Engineering Society Convention Paper 9382.

Olive, S., Welti. T. & Khonsaripour, O. 2016. The Preferred Low Frequency Response of In-Ear Headphones. Audio Engineering Society Conference Paper 6-1.

Olive, S., Welti. T. & Khonsaripour, O. 2017a. The Influence of Program Material on Sound Quality Ratings of In-Ear Headphones. Audio Engineering Society Convention Paper 9778.

Olive, S., Welti. T. & Khonsaripour, O. 2017b. A Statistical Model That Predicts Listeners' Preference Ratings of In-Ear Headphones: Part 1 – Listening Test Results and Acoustic Measurements. Audio Engineering Society Convention Paper 9840.

Olive, S., Welti, T. & Khonsaripour, O. 2018a. A Statistical Model that Predicts Listerers' Preference Ratings of Around-Ear and On-Ear Headphones. Audio Engineering Society Convention Paper 9919.

Olive, S., Welti, T. & Khonsaripour, O. 2018b. A Survey and Analysis of Consumer and Professional Headphones Based on Their Objective and Subjective Performances. Audio Engineering Society Convention Paper 10048.

Olive, S. 2022. The Perception and Measurement of Headphone Sound Quality: What Do Listeners Prefer? Acoustics Today. Volume 18, issue 1, Spring 2022.

Pedersen, J. & Mortensen, H. 2007a. Natural Timbre in Room Correction Systems. Audio Engineering Society Convention Paper 7112.

Pedersen, J. & El-Azm, F. 2007b. Natural Timbre in Room Correction Systems (Part II). Audio Engineering Society Conference Paper 5.

Poldy, C. 2001. Headphones, chapter 14 in Loudspeaker and Headphone Handbook (ed. J. Borwick). Third edition. UK: Focal Press.

Poldy, C. 2006. Headphone Fundamentals. Audio Engineering Society Convention Tutorial Material.

Rasmussen, G. 2006. The Artificial Ear Dilemma: The Challenges of Modern Handset Testing. ITU-T Workshop on ”End-toEnd QoE/Qos”; June 14-16; Switzerland, Geneva. Read on 3.9.2021 https://www.itu.int/ITU-T/worksem/qos/200606/presentations/s3paper-rasmussen.pdf

Robinson, D. W. & Dadson, R. S. 1956. A re-determination of the equal-loudness relations for pure tones. British Journal of Applied Physics, 7, 166-181. American Physical Society.

Rtings.com. 2018. Raw Frequency Response. Read on 13.8.2021. https://www.rtings.com/headphones/tests/sound-quality/raw-frequency-response

Sennheiser. 2012. A history of sound in perfection: An insight into Sennheiser´s high-end world. Rean on 22.4.2022. https://en-us.sennheiser.com/news-a-history-of-sound-in-perfection-an-insight-into-sennheisers-high-end-world-160412

Shaw, E. 1997. Acoustical Features of the Human External Ear, Part I, chapter 2 in Binaural and Spatial Hearing in Real and Virtual Environments (ed. R. H. Gilkey and T. A. Anderson) USA: Lawrence Erlbaum Associates.

Singer, M. 1979. Nathaniel Baldwin: Utah inventor and patron of the fundamentalist movement. Utah Historical Quarterly. vol. 47, no. 1 (winter 1979). USA: Utah State Historical Society.

Snaidero, T., Jacobsen, F., & Buchholz, J. 2011. Measuring HRTFs of Brüel & Kjær Type 4128-C, G.R.A.S. KEMAR Type 45BM, and Head Acoustics HMS II.3 Head and Torso Simulators. Technical University of Denmark, Department of Electrical Engineering.

Staab, W. 2014. Human Ear Canal – VIII. Read on 12.8.2021. http://hearinghealthmatters.org/waynesworld/2014/human-ear-canal-viii/

Stanford Children´s Health. N.d. Anatomy and Physiology of the Ear. Read on 29.3.2022. https://www.stanfordchildrens.org/en/topic/default?id=anatomy-and-physiology-of-the-ear-90-P02025

Statista. 2021. Unit shipments of headphones worldwide from 2013 to 2020 (in millions). Read on 13.4.2022.

https://www.statista.com/statistics/236075/revenue-of-headphone-shipments-in-the-united-states/#professional

Stax. 2022. Company Profile. Read on 13.4.2022. https://staxheadphones.com/

Struck, C. 2013. Free Plus Diffuse Sound Field Target Earphone Response Derived From Classical Room Acoustics Theory. Audio Engineering Society Convention Paper 8993.

Sudonull. 2017. Personality and Sound: Eugen Beyer is a man who has always looked to the future. Read on 13.4.2022. https://sudonull.com/post/18063-Personality-and-sound-Eugen-Beyer-a-man-who-has-always-looked-to-the-future-Pultru-Blog

Sun, L., Zhong, X., Yost, W. 2015. Dynamic binaural sound source localization with interaural time difference cues: Artificial listeners. The Journal of the Acoustical Society of America.

Suzuki, Y., Mellert, V., Richter, U., Møller, H., Nielsen, L., Hellman, R., Ashihara, K., Ozawa, K. & Takeshima, H. 2003. Precise and Full-range Determination of Two-dimensional Equal Loudness Contours. Technical Report, Japan: Tohuku University.

Suzuki, Y. & Takeshima, H. 2004. Equal-loudness-level contours for pure tones. The Journal of the Acoustical Society of America 116, 918 (2004).

Tamir, A. 2017. Numerical Survey of the Different Shapes of Human's Ear Concha. SM Otolaryngol. 2017; 1(1): 1002. Read on 12.9.2021: https://dx.doi.org/10.36876/smotol.1002

Thiele, G. 1985. On the standardization of the frequency response of high quality studio headphones. Audio Engineering Society Convention Paper 2207.

Thiele, G. 2016. Equalization of studio monitor headphones. Audio Engineering Society Conference Paper 4-3, (2016 August)

Toole, F. 2004. Audio – Science in the service of Art. Harman International Industries, Inc. Read on 28.3.2022: https://www.harman.com/documents/audioscience.pdf

Toole, F. 2015. The Measurement and Calibration of Sound Reproducing Systems. Journal of the Audio Engineering Society, vol. 63, no. 7/8, 512-541.

Toole, F. 2018. Sound Reproduction. The Acoustics and Psychoacoustics of Loudspeakers and Rooms. Third Edition. USA: Routledge.

Trapeau, R. & Schönwiesner, M. 2018. The Encoding of Sound Source Elevation in the Human Auditory Cortex. The Journal of Neuroscience 28 March 2018, 38

(13) 3252-3264. Read on 15.9.2021 https://doi.org/10.1523/JNEUROSCI.2530-17.2018

University of California Davis. 2011. Psychoacoustics of Spatial Hearing - Elevation Cues. Read on 29.3.2022. http://interface.cipic.ucdavis.edu/sound/tutorial/psych.html

Yost, W. 2017. History of sound source localization: 1850-1950. Proceedings of Meetings on Acoustics, 30, 050002.

Välimäki, V., Reiss. J. 2016. All About Audio Equalization: Solutions and Frontiers. Applied Sciences 6, no. 5: 129. MDPI. Read on 12.9.2021. https://doi.org/10.3390/app6050129

Welti, T. 2015. Improved Measurement of Leakage Effects for Circum-aural and Supra-aural Headphones. Audio Engineering Society Convention Paper 9275.

Woodgate, J. 2001. International standards, chapter 15 in Loudspeaker and Headphone Handbook (ed. J. Borwick). Third edition. UK: Focal Press.

Öyry, A. Research engineer at Aalto Acoustics Lab 2022. Telephone and email conversation.