The Perception and Measurement of Headphone Sound Quality: What Do Listeners Prefer?

Sean E. Olive

Headphones are the primary means through which we listen to music, movies, and other forms of infotainment. They have become an indispensable accessory for our mobile phones, providing a 24/7 connection to our entertainment, colleagues, and loved ones. This trend is reflected in the exponential growth in sales. The global market for wireless headphones alone was estimated at $15.9B in 2020 and is projected to rise to $45.7B by 2026, a compound annual growth rate of 19.1% (PRNewsWire, 2021). With this growth has come a renewed interest in improving the sound quality of headphones.

Unfortunately, headphone sound quality has not kept pace with consumers' demands and expectations. Two recent studies have measured the variance in frequency response of more than 400 headphones and found no correlation between their retail price and frequency response (Breebaart, 2017; Olive et. al., 2018a). They included the three most common types: headphones that fit around the ear (AE), on the ear (OE), and in the ear (IE). It seems that headphone designers are aiming at a target frequency response that is as random and variable as the weather.

Another telling sign that headphone sound quality has not kept pace is that headphone industry standards have not changed fundamentally since the 1990s. The International Electrotechnical Commission (IEC) 60268-7 (2010) standard specifies multiple ways to measure the frequency response of a headphone for both free-field (FF) and diffuse-field (DF) targets, with the warning: "subjective assessments are still useful because the objective methods whose results bear good relation to those from subjective assessments are under research stage" (IEC, 2010, Section 8.6.1). This does not inspire confidence.

The International Telecommunication Union Radiocommunication Assembly (ITU-R) BS.708 (1990) standard recommends that professional headphones be designed to the DF target curve to achieve best sound, but most headphone designers have rejected this suggestion and probably for good reasons. Recent psychoacoustic investigations provide evidence that listeners prefer alternative headphone targets to DF and FF target standards (Olive et al., 2013a).

The chaos that exists within the headphone industry today is reminiscent of the loudspeaker industry 30 years ago when there was insufficient knowledge on listeners' loudspeaker preferences and which loudspeaker measurements best predict them. The situation improved after Floyd Toole, an acoustician at the National Research Council of Canada, published seminal scientific papers that provided guidelines in how to measure and design loudspeakers that most listeners prefer (Toole, 1985, 1986). Later, a mathematical model was developed that could predict listeners' preference ratings of the loudspeakers based on objective measurements alone (Olive, 2004). The science provided important answers on what loudspeaker listeners prefer, design guidelines, and new measurement standards (American National Standards Institute/Consumer Technology Association [ANSI/CTA] Standard, 2015) that became widely accepted and adopted throughout the industry.

Headphone Sound Quality

In 2012, the seminal papers for headphone sound quality did not exist, and this was reflected in the headphone standards and the large variance in headphone sound quality. Skeptics argued that the variance in headphone sound was explained by a need to satisfy individual tastes in sound that vary like individual tastes in music, food, and preferred companions. If listeners could not agree on what sounds good, then a single optimal frequency response or headphone target curve could not be defined.

https://doi.org/10.1121/AT.2022.18.1.58

These same arguments were undoubtably made about loudspeakers 40 years ago and until research proved listeners largely agreed on what is a good loudspeaker.

With the lessons learned from the loudspeaker industry, the author and his colleagues embarked on a seven-year research project to improve the consistency and sound quality of headphones. There were three fundamental questions we hoped to answer.

(1) What is the preferred headphone target curve? Should the reference be a loudspeaker in a FF, a DF, or a semireflective field (SRF) found in a typical listening room?
(2) Do listeners agree on what makes a headphone sound good? To what extent does listening experience, age, gender, and geographical location influence sound quality preferences?
(3) Can listeners' subjective ratings of headphones be predicted based on an objective measurement?

These research questions were addressed for the three main headphone types, but the scope of this article is largely restricted to AE and OE headphones. The preferred target curve for IE headphones is almost identical to those for the AE and OE targets, except it has an additional 4 dB of bass (Olive et al., 2016). Each question is addressed separately, followed by conclusions.

The Search for the Preferred Headphone Target Curve

Over the past 50 years, headphone researchers have focused their attention on determining what the ideal reference sound field should be for headphone reproduction and how to measure it. Three types of reference sound fields have been proposed: a FF, a DF and a SRF that lies somewhere between the two extremes. What these sound fields are, how they are measured or derived, and psychoacoustic investigations of headphone target curves based on them are described.

Free-Field Headphone Target Curve (1970s)

The reference FF was generated by placing a loudspeaker in front of the listener in a reflection-free room. A tedious subjective loudness-matching procedure was used where a test subject would listen to narrow bands of noise at different frequencies alternately with the FF (with the headphone removed) and then with the headphone. While listening to the headphones, the levels for each band would be adjusted to match the loudness of the loudspeaker. This would be repeated for several test subjects to calculate the loudness transfer function that defined the headphone FF target curve.

Theile (1986) conducted formal listening tests and found the DF target to be preferred to the FF target, which produced an unnatural timbre and in-head localization effects. Although the FF target fell out of favor beginning in the 1980s, it remains part of the current headphone IEC (2010) standard today.

Diffuse-Field Headphone Equalizations (1980s to Present)

A DF occurs when a sound source is placed in a reverberation room with little or no absorption, so the listener receives a random and equal distribution of sounds from all directions. The headphones are calibrated to the DF using a subjective loudness procedure or alternative methods. In one method, a probe microphone is placed in the ear canals of the listener to measure and then match the transfer function of the headphone to that of the sound field (Theile, 1986).

A second approach is to substitute the listener with a head and torso simulator (HATS); this produces faster, more reproducible, and safer measurements than putting probe microphones in the listeners' ears. A third option is to use a headphone known to be DF calibrated as the reference and compare its performance with the headphone under test.

Møller et al. (1995) derived a headphone target curve based on different sound fields using a large set of head-related transfer functions (HRTFs) measured at the blocked ear canal. HRTFs define the transfer functions, both the frequency and phase responses at the entrance to the ear, for each direction and distance of a sound source. They capture both interaural time (ITD) and intensity (IID) differences and spectral cues that humans use to localize sound sources in space (Blauert, 1983). By selecting HRTFs from the appropriate directions and distances and integrating them, Møller et al. (1995) were able to derive transfer functions of reference sound fields ranging from the FF to the DF and anything in between. This method eliminated the need for a physical reference sound field, making headphone calibration more practical and reproducible. A headphone could be measured and equalized to the DF target curve using a calibrated dummy head or ear simulator.

The DF target was not seriously challenged until Lorho (2009) reported 80 listeners (25% audio engineers, 25% music students, and 50% naive listeners) on average preferred a significantly modified version of the DF target where its main feature, a wide 12 dB peak at 3 kHz, was reduced to just 3 dB. This paper sparked new interest to find better alternative headphone target curves to the ones recommended in the current headphone standards.

Semireflective Field Headphone Equalizations (2012 to Present)

Because stereo recordings are optimized for reproduction through loudspeakers in semireflective rooms, they should sound best through headphones that emulate this sound field. Sank (1980) made similar proposals three decades earlier but never conducted formal listening tests that compared these targets with the DF target.

Loudspeakers with flat on-axis and smooth off-axis frequency responses tend to produce the highest subjective ratings in formal listening tests (Toole, 2018). When placed in a typical room, they produce a uniform quality of direct, early, and late reflected sounds that in summation produce the steady-state in-room response of the loudspeaker. Due to the frequency-dependent directivity of the loudspeaker and absorption characteristics of the room, the in-room response will not be flat like the FF response nor the same as the DF response where the room absorption has been removed. Instead, the in-room response gently falls about 1 dB per octave from 20 Hz to 20 kHz.

Fleishmann et al. (2012) reported the first formal listening test results where three SRF headphone targets were evaluated. The targets were based on measurements of the steady-state in-room response of a 5.1-channel loudspeaker setup in a standard listening room and then equalized by three expert listeners to match the timbre of the speakers. Two of the SRF targets were found to be slightly preferred to the DF target, depending on the music programs. Other targets included the Lorho target, a flat target, and three unequalized headphones that generally received lower ratings than the two SRF targets. Unfortunately, no measurements or details of the loudspeakers and the three SRF targets were given. The conclusions were that the SRF targets were equal to or better than the DF target, but the Lorho target was not.

A similar study (Olive et al., 2013a) reported evidence that listeners strongly preferred headphones equalized to SRF targets to, in descending order of preference, two DF targets (Möller et al., 1995); two high-quality headphones; the Lorho target; and the FF target. The trained listeners described both DF targets as having too much emphasis in the upper midrange (2-4 kHz) and lacking bass. The Lorho target had too little energy at 2-4 kHz, which made instruments sound "muffled and dull." The FF target was strongly criticized for its strong emphasis between 2 and 4 kHz, lack of bass, and harsh and nasal colorations. Listeners described the highest rated the SRF target as having "good bass with an even spectral balance." The measured frequency responses of the headphone targets correlate to and confirm listeners' descriptions of their sound quality (see Olive et al., 2013b, Figure 2). The highest rated target curve in this study soon became known in the audio industry as the Harman target curve and is widely influencing the design, testing, and review of headphones.

Do Listeners Agree on What Makes a Headphone Sound Good?

Although the initial test results of the Harman target curve were encouraging, they were based on a small sample of 10 trained listeners. To better understand if certain demographic factors influence the acceptance of the curve, it was tested using a larger number of listeners from a broad range of ages, listening experiences, and geographic regions.

The target curve was benchmarked against three headphones considered industry references at the time in terms of sound quality or commercial sales (Olive et al., 2014). They ranged in price from $269 to $1,500 and included dynamic and magnetic planar transducer designs. A total of 283 listeners participated from four different countries (Canada, United States, Germany, and China) and included a broad range of ages, listening experiences, and genders. Most of the participants were Harman employees.

A novel virtual headphone test methodology allowed controlled, rapid, double-blind comparisons among the different headphones. Virtual versions of the different headphones were reproduced over a single high-quality replicator headphone by equalizing it to match the measured frequency response of each headphone. This removed

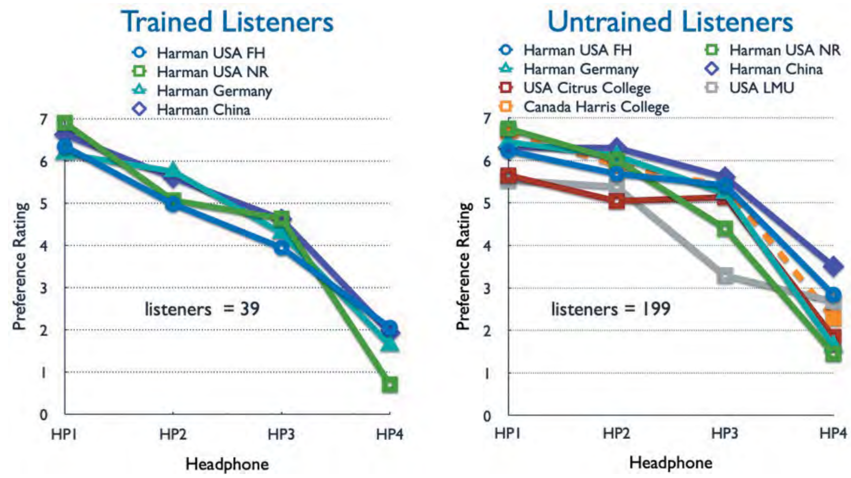


Figure 1. *The mean preference ratings are shown for 11 different groups of listeners categorized as trained (**left**) and untrained (**right**). The tests were administered in four different countries: Canada, United States, Germany, and China. HP1 is the Harman target curve and HP2 and HP3 are high-quality, high-priced headphones. HP4 was the most popular headphone in terms of sales (Olive et al., 2014).*

any potential biases related to visual (brand, model, price, design) and tactile (weight, clamping force, feel of materials) cues that might cloud their judgments of sound quality. A prior validation study confirmed that subjective ratings of virtual versus actual headphones (with the listener unaware of the headphone brand, model or appearance) had a correlation of 0.86 to 0.99 depending on the headphone type (Olive et al., 2013b). A limitation of the method is that it does not reproduce nonlinear distortions in the headphones. However, the high correlations between virtual and actual headphone comparisons and evidence from other studies indicate that these distortions are generally below masked thresholds (Temme et al., 2014).

The results show that headphone preferences were remarkably consistent across the 11 test locations for both trained and untrained listeners (**Figure 1**). As expected, the trained listeners were more discriminating and consistent than the untrained listeners.

Headphone preferences were also relatively consistent across different age groups and the four countries. The exception was listeners in the 55+-year age category who tended to prefer HP2, a brighter headphone with less bass than the Harman target curve. A possible explanation could be age related hearing loss; additional treble and less bass can help improve intelligibility. More research is needed to provide definitive answers.

Preferred Level of Bass and Treble in Headphones

The same group of listeners participated in a second experiment where they adjusted the bass and treble levels of the headphone (Olive and Welti, 2015) several times according to taste using different samples of music. The listeners' preferred levels were influenced by several factors, including the music program, as well as by the subject's age, gender, and prior listening experience (see **Figure 2**). The program interactions between preferred bass and treble levels are expected due to variability in the quality of music recordings; often they require adjustments in bass and treble on playback to restore a proper balance. Toole (2018) refers to these errors as audio's "circle of confusion." The confusion arises from not knowing the source of these errors: the recording, the loudspeaker, or its interaction with the room acoustics. The solution is a meaningful loudspeaker standard common to both the professional and consumer audio industries.

Female listeners preferred less bass and treble than their male counterparts. Younger and less experienced listeners

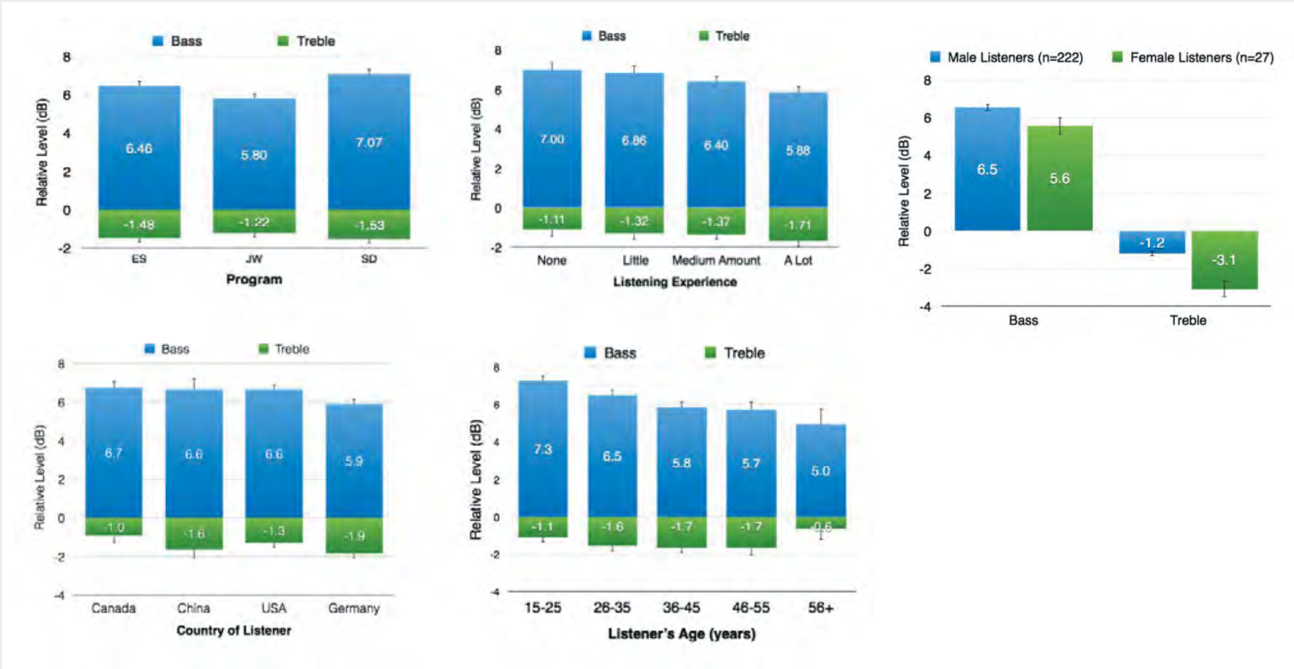


Figure 2. *The mean bass and treble levels and 95% confidence intervals for a headphone calibrated to match a flat in-room loudspeaker response. Each graph shows the interaction effect between the preferred levels and program, gender, listening experience, age, and the country of the test location (Olive and Welti, 2015).*

preferred more bass and treble than their older, more experienced counterparts. The older listeners (55+ years) were the exception here, preferring significantly more treble and less bass, consistent with their preference for headphone HP2. Altogether, these findings suggest that a single headphone target may not be sufficient to satisfy variations in the recordings, individual tastes, listening experience, and hearing loss. A simple solution for headphone personalization is to provide a simple bass and treble control that allows listeners to compensate for these variations.

Testing the Harman Target with Larger Sample of Headphones

The next goal was to test the Harman target using a larger population of headphones. A total of 31 different headphone models from 18 manufacturers were evaluated by 130 listeners, with an approximately equal number trained and untrained (Olive et al., 2018a). The headphones ranged in price from $60 to $4,000, including open and closed back designs with dynamic or magnetic planar drivers. The same virtual headphone double-blind method was used to eliminate biases from visual and tactile cues. The results establish that, on average, both trained and untrained listeners preferred the headphone equalized to the Harman target in 28 of the models tested. Four models with frequency responses close to the Harman target were equally preferred.

Segmentation of Listeners Based on Preferred Headphone Sound Profiles

Although the study established that listeners, on average, preferred the Harman target to other headphones tested, it had not explored whether segments or classes of listeners exist based on similarities in their headphone preferences and what those sound quality features or profiles are. Also, it did not identify possible underlying demographic factors that might predict membership in each class. There was already prior evidence that younger males and less experienced listeners preferred higher levels of bass and treble in their headphones compared with females, experienced, and older listeners (Olive et al., 2013a; Olive and Welti, 2015). A reasonable hypothesis was that segmentation of headphone preferences may relate to bass and treble levels, possibly predicted by these demographic factors.

A statistical method known as agglomerative hierarchical clustering exposed three different segments or classes of listeners based on similarities in their headphone preferences. By calculating the average response of the top five preferred headphones in each class, it was clear that the preferred bass level is the main feature that defines membership in a class. Class 1 includes most listeners (64%) who prefer headphones that closely comply with the Harman target. Class 2 listeners (15%) prefer the Harman target with 4-6 dB more bass. Class 3 listeners (21%) prefer the Harman target curve with 2 dB less bass.

Table 1 shows the different demographic categories and the distribution or percentage represented in each class. For example, 69% of the males in the study are members of Class 1 (Harman Target Lovers) compared with 56% of females. Class 1 has roughly equal representation from trained (70%) and untrained (65%) listeners. Class 2 (More Bass Is Better) has the fewest members overall and is represented by all categories except female; only 4% of females tested want more bass than the Harman target provides. Class 3 (Less Bass Is Better) members are disproportionately represented by females (40% of females are in this class versus 13% of males) and listeners over the age of 50 (50%). Hearing loss may be a confounding factor here. More research is needed to better understand the role it plays in headphone sound quality preferences.

The main takeaway is that the Harman target is a good design target for headphones because it satisfies the tastes of a majority of listeners (64%) over a broad range of age groups, genders, and levels of listening experience. The two smaller classes of listeners who prefer headphones with more bass or less bass can be accommodated through a simple bass tone control on the headphone or via an app on the audio device. The bass adjustment would also help compensate for inconsistences in the quality of recordings that contain either too much or not enough bass and treble. A word of caution: the research suggests adding too much bass beyond the Harman target may alienate many listeners given that the "more bass is better" segment is a small segment with little female and older listener membership. Conversely, reducing the bass too much may alienate trained and experienced listeners who are underrepresented in the "less bass is better" segment.

Predicting Listener's Headphone Sound Quality Preferences

Conducting controlled headphone listening tests is a challenging, time-consuming, and expensive proposition. An alternative solution is to model and predict listeners' headphone preference ratings using objective measurements that are relatively faster, more reproducible, and cost effective.

The 31 headphones from the same study discussed by Olive et al. (2018a) were sorted into 4 categories of sound quality based on listeners' preference scores: excellent (90-100% preference rating), good (65-76%), fair (42-54%), and poor

Table 1. *Distribution of listeners within each category.*

Distribution of Listener Categories Within Each Class (in %) Based on Preferred Headphone Sound Profile			
Category	**Class 1: Harman Target Lovers**	**Class 2: More Bass Is Better**	**Class 3: Less Bass Is Better**
Males	0.69	0.18	0.13
Females	0.56	0.04	0.40
Trained	0.70	0.30	0.00
Untrained	0.65	0.10	0.25
Age (years)			
20s	0.69	0.17	0.15
30s	0.74	0.13	0.13
40s	0.67	0.10	0.24
50+	0.30	0.20	0.50

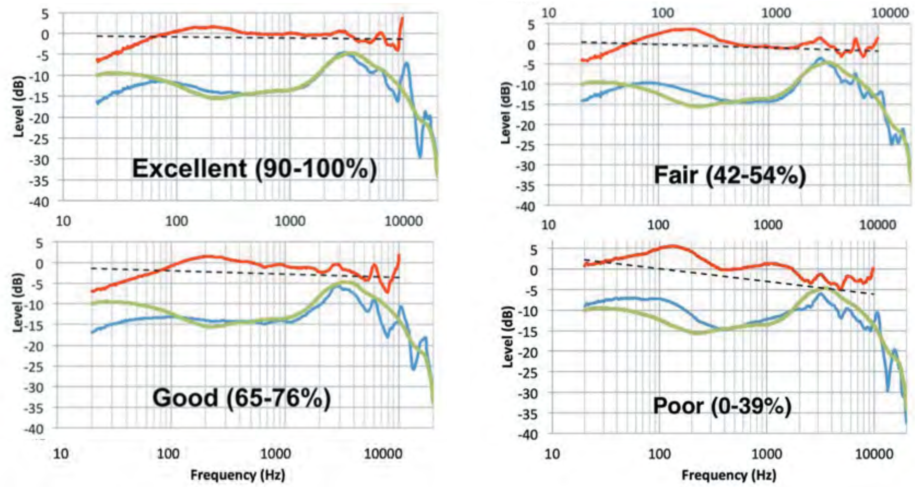
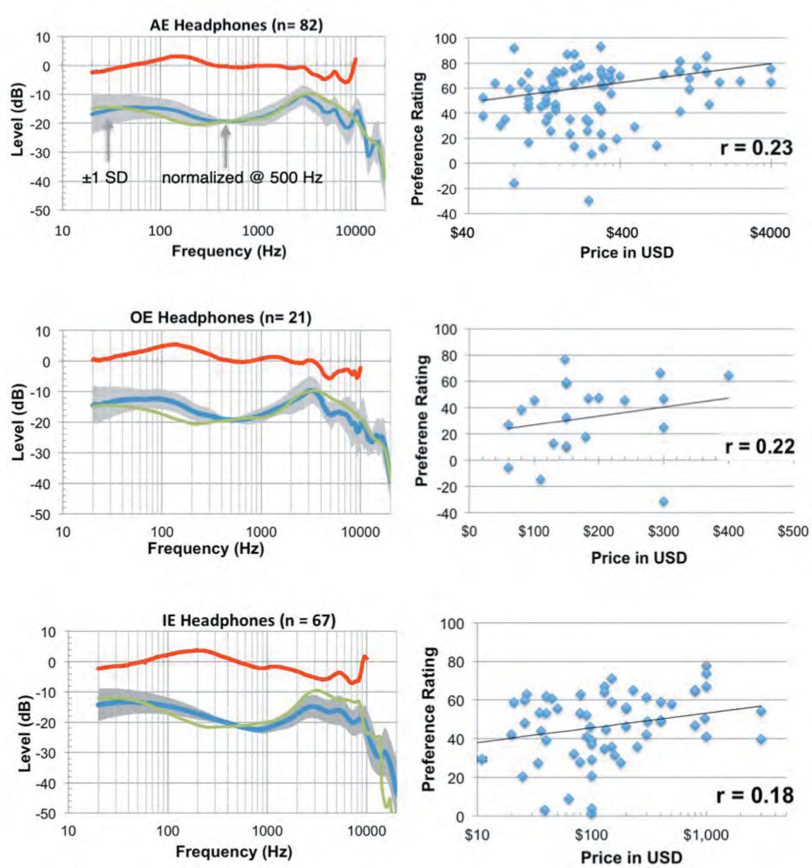


Figure 3. *The average frequency response (**blue**) and error curves (**red**) for 31 around the ear (AE) headphones assigned to four categories of sound quality based on their preference rating given in controlled listening tests. The Harman target curve used to calculate the error curve (**green**). **Dotted curve**, regression line that best fits the error curve (Olive et al., 2018a).*

Figure 4. ***Left:*** *average frequency response (**blue**), standard deviation (**gray area**), and error response curve (**red**) for AE, on the ear (OE), and in the ear (IE) headphone with the Harman target curve (**green**).* ***Right:*** *predicted preference ratings versus their retail price are plotted for the headphones with the best-fit regression line and correlation coefficient shown (Olive et al., 2018b).*



(0-39%). Frequency response measurements of the headphones were made using an ear simulator according to IEC 60318-1 (2009) equipped with a custom pinna that better simulates headphone leakage on humans.

In each category, the average frequency response for the headphones is plotted with the Harman target curve and the error curve that is the difference between the two (see **Figure 3**). **Figure 3**, *black dotted line*, is a regression line that best fits the error curve. The relationship between objective and subjective headphone measurements seems clear: the more the frequency response of a headphone deviates from the Harman target curve, the lower the listeners rated its sound quality.

A linear model was developed that predicts headphone preference ratings using two variables based on the standard deviation and the absolute slope of the error curve. The correlation between the predicted and measured ratings is 0.86, with an error of 6.7 ratings on a 100-point scale. A similar model was developed for IE headphones that produces slightly better ($r = 0.91$) predictions (Olive et al., 2016).

The two models were used to predict preference ratings for 158 headphones, including AE, OE, and IE types (Olive et al., 2018b). **Figure 4**, *left*, shows the average magnitude response, standard deviation, Harman target, and error response curve for each headphone type. **Figure 4**, *right*, plots the retail price versus the predicted preference rating for each headphone model tested. On average, the AE headphones come closest to the Harman target and produce the highest preference ratings, the OE headphones are the worst, and the IE headphones fall in between. The retail price of a headphone is not a good indicator of its sound quality based on the relatively low correlation values shown here.

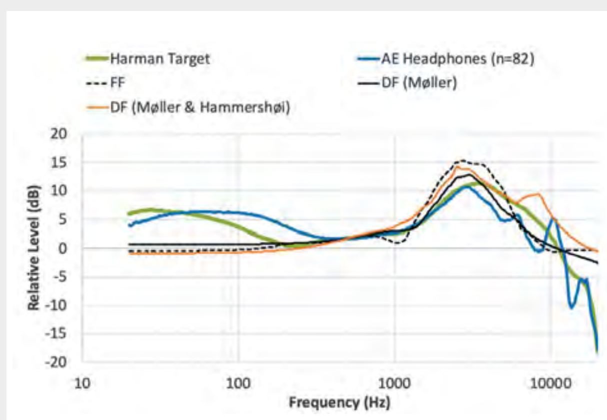
These findings are generally in agreement with those reported by Breebaart (2017). The two studies together provide evidence that headphone designers are aiming at a target curve that is closer to the Harman target than the DF or FF target curves recommended by the current headphone standards. **Figure 5** shows the average response of the 82 AE headphones in **Figure 4** compared with the Harman, DF, and FF target curves (Møller et al., 1995). Although the DF and FF targets specify a flat response below 200 Hz, the average AE headphone and Harman targets have 5-6 dB more bass, which better approximates the preferred in-room response of a full-range loudspeaker calibrated in a typical listening room.

Conclusions

Our understanding of the perception and measurement of headphone sound quality has not kept pace with consumer demand and expectations. Two independent studies measured over 400 headphones and came to similar conclusions: there is little correlation between the price of a headphone and its frequency response, the single best indicator of its sound quality. Most professional and consumer headphone designs today do not comply with the FF and DF targets recommended by current headphone standards, which warns "the objective methods whose results bear good relation to those from subjective assessments are under research stage" (see IEC 60268-7, 2010, Section 8.6.1). The research stage is largely completed, the results are in, and the headphone standards need to be updated.

Listeners largely agree on what makes a headphone sound good. For stereo reproduction, the preferred headphone target approximates the in-room response of an accurate loudspeaker calibrated in a semireflective room. This makes perfect sense because stereo recordings are intended to sound best through accurate loudspeakers in semireflective rooms. What makes a headphone sound good is the same as what makes a loudspeaker sound good.

Figure 5. *Proposed headphone target curves normalized at 500 Hz: the Harman AE headphone target (**green**), two diffuse-field (DF; **orange** and **black**) and free-field (FF; **dashed**) targets (Møller et al., 1995), and the average frequency response (**blue**) of 82 different models of AE headphones (Olive et al., 2018b).*



The Harman target curve is one example that is preferred by a majority (64%) of listeners from a broad range of ages, listening experiences, and genders. Slight adjustments in the bass and treble levels may be necessary to compensate for variance in the quality of recordings and to satisfy individual tastes. The Less Bass Is Better class (21% of listeners) includes a disproportionate percentage of females and older listeners and none of the trained listeners. The More Bass Is Better class is skewed toward males versus females by a factor of 4 to 1. There is no evidence that sound quality preferences are geographically influenced. Recognition of good sound reproduction seems to be universal.

Objective measurements of the headphones using standard ear simulators can predict how good they sound. The further the frequency response a headphone deviates from the Harman target response, the lower its perceived sound quality will be. A simple linear model based on these deviations can predict how listeners would rate it in controlled listening tests.

The reaction from the headphone industry to this new research has been largely positive. There is evidence that the Harman target curve is widely influencing the design, testing, and review of many headphones from multiple manufacturers, providing a much needed new reference or benchmark for testing and evaluating headphones. Several headphone review sites provide frequency response measurements of headphones showing the extent to which they comply with the Harman target (Vafaei, 2018; Audio Science Review, 2020); in cases where they fall short, corrective equalizations are often provided.

As expected, there are also critics whose headphone tastes in sound may not agree with the research. The Harman target is intended as a guideline and is not the last word on what makes a headphone sound good. One legitimate criticism is the limited number of headphones, programs, female listeners tested, and questions raised about the confluence of variables like hearing loss and its effect on headphone preference. Future studies will hopefully address this. Finally, I hope that this article encourages others to continue the research and improve our knowledge of the perception and measurement of headphone sound quality. Although listeners largely agree on what makes a headphone sound good, there are still many unanswered questions and more to learn.

References

American National Standards Institute/Consumer Technology Association (ANSI/CTA) Standard (2015). *Standard Method of Measurement for In-Home Loudspeakers*. Report ANSI/CTA-2034-A, ANSI/CTA, Arlington, VA.

Audio Science Review (2021). *Headphone Review and Discussions*. Available at https://tinyurl.com/cvbwmc2v.

Blauert, J. (1983). *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA.

Breebaart, J. (2017). No correlation between headphone frequency response and retail price. *The Journal of the Acoustical Society of America* 141(6), EL526-EL530. https://doi.org/10.1121/1.4984044.

Fleischmann, A., Silzle, F., and Plogsties, J. (2012). Identification and evaluation of target curves for headphones. *Proceedings of the 133rd Audio Engineering Society Convention*, San Francisco, CA, October 26-29, 2012. Available at https://www.aes.org/e-lib/browse.cfm?elib=16482.

International Electrotechnical Commission (IEC) (2009). *Electroacoustics - Simulators of Human Head and Ear - Part 1: Ear Simulator for the Measurement of Supra-Aural and Circumaural Earphones*. Report IEC 60318-1, IEC, Geneva, Switzerland.

International Electrotechnical Commission (IEC) (2010). *Sound System Equipment - Part 7: Headphones and Earphones*. Report IEC 60268-7, IEC, Geneva, Switzerland.

International Telecommunication Union Radiocommunication Assembly, (ITU-R) (1990). *Determination of the Electroacoustical Properties of Studio Monitor Headphones*. Report ITU BS.708, International Telecommunication Union, Geneva, Switzerland.

Lorho, G. (2009). Subjective evaluation of headphone target frequency responses. *Proceedings of the 126th Audio Engineering Society Convention*, Munich, Germany, May 7-10, 2009. https://www.aes.org/e-lib/browse.cfm?elib=14966.

Møller, H., Jensen, C. B., Hammershoi, D., and Sørensen, M. F. (1995). Design criteria for headphones. *Journal of the Audio Engineering Society* 43, 218-232. https://www.aes.org/e-lib/browse.cfm?elib=10274.

Olive, S. E. (2004). A multiple regression model for predicting loudspeaker preference using objective measurements: Part 2 - Development of the Model. *Proceedings of 117th Audio Engineering Society Convention*, San Francisco, CA, October 28-31, 2004. https://www.aes.org/e-lib/browse.cfm?elib=12847.

Olive, S. E., and Welti, T. (2015). Factors that influence listeners' preferred bass and treble balance in headphones. *Proceedings of the 139th Audio Engineering Society Convention*, New York, NY. https://www.aes.org/e-lib/online/browse.cfm?elib=17940.

Olive, S. E., Welti, T., and Khonsaripour, O. (2016). A statistical model that predicts listeners' preference ratings of in-ear headphones: Part 2 - Development and validation of the model. *Proceedings of 143rd Audio Engineering Society Convention*, New York, NY, October 29 to November 1, 2015. https://www.aes.org/e-lib/browse.cfm?elib=19436.

Olive, S. E., Welti, T., and Khonsaripour, O. (2018a). A statistical model that predicts listeners' preference ratings of around-ear and on-ear headphones. *Proceedings of the 144th Audio Engineering Society Convention*, Italy, May 23-26, 2018. https://www.aes.org/e-lib/browse.cfm?elib=19436.

Olive, S. E., Welti, T., and Khonsaripour, O. (2018b). A survey and analysis of consumer and professional headphones based on their objective and subjective performances. *Proceedings of the 145th Audio Engineering Society Convention*, New York, NY. https://www.aes.org/e-lib/browse.cfm?elib=19774.

Olive, S. E., Welti, T., and McMullin, E. (2013a). Listener preference for different headphone target response curves. *Proceedings of the 134th Audio Engineering Society Convention*, Rome, Italy, May 4-7, 2013. https://www.aes.org/e-lib/browse.cfm?elib=16768.

Olive, S. E., Welti, T., and McMullin, E. (2013b). A virtual headphone listening test methodology. *Proceedings of the Audio Engineering Society 51st International Conference: Loudspeakers and Headphones*, Helsinki, Finland, August 22-24, 2013. https://www.aes.org/e-lib/browse.cfm?elib=16874.

Olive, S. E., Welti, T., and McMullin, E. (2014). The influence of listeners' experience, age, and culture on headphone sound quality preferences. *Proceedings of the 137th Audio Engineering Society Convention*, Los Angeles, CA, October 9-12, 2014. https://www.aes.org/e-lib/online/browse.cfm?elib=17500.

PRNewsWire (2021), *Global Wireless Headphones Market to Reach $45.7 Billion by 2026*. Available at https://tinyurl.com/234dhxs4. Accessed November 2, 2021.

Sank, J. R. (1980). Improved real-ear tests for stereophones. *Journal of the Audio Engineering Society* 28, 206-218. https://www.aes.org/e-lib/browse.cfm?elib=3994.

Temme, S., Olive, S. E., Tatarunis, S., Welti, T., and McMullin, E. (2014). The correlation between distortion audibility and listener preference in headphones. *Proceedings of the 137th Audio Engineering Society Convention*, Los Angeles, CA. October 9-12, 2014. https://www.aes.org/e-lib/online/browse.cfm?elib=17441.

Theile, G. (1986). On the standardization of the frequency response of high-quality studio headphones. *Journal of the Audio Engineering Society*, 34, pp. 956-969. https://www.aes.org/e-lib/browse.cfm?elib=5233.

Toole, F. E. (1985). Subjective measurements of loudspeaker sound quality and listener preferences. *Journal of the Audio Engineering Society* 33, 2-31. https://www.aes.org/e-lib/online/browse.cfm?elib=4465.

Toole, F. E. (1986). Loudspeaker measurements and their relationship to listener preferences. *Journal of the Audio Engineering Society* 34, Part1, 227-235; Part. 2, . 323-348. https://www.aes.org/e-lib/browse.cfm?elib=5276.

Toole, F. E. (2018). *Sound Reproduction, the Acoustics and Psychoacoustics of Loudspeakers and Rooms*, 3rd ed. Routledge, New York, NY.

Vafaei, S. (2018). Raw frequency response. *Rtings.com*. Available at https://bit.ly/3oBUnKG. Accessed November 2, 2021.

About the Author

Sean E. Olive
sean.olive@harman.com

Harman International
8500 Balboa Boulevard
Northridge, California 91329, USA

Sean E. Olive studied music at the University of Toronto, Toronto, ON, Canada, and sound recording at McGill University, Montreal, QC, Canada, where he received master's and PhD degrees. He is a senior fellow in acoustic research at Harman International, Northridge, California, focused on research in sound quality of reproduced sound. He has published over 50 research papers for which he received Fellow of the Audio Engineering Society (AES), two AES publications awards, and the ALMA Titanium Award. Sean is the past president of AES.