

# No correlation between headphone frequency response and retail price

Jeroen Breebaart<sup>a)</sup>  
Sydney, Australia  
[jeroen@jeroenbreebaart.com](mailto:jeroen@jeroenbreebaart.com)

**Abstract:** This study quantifies variability of measured headphone response patterns and aims to uncover any correlations between headphone type, retail price, and frequency response. For this purpose, the mean, variance, and covariance of the frequency magnitude responses were analyzed and correlated with headphone type and retail value. The results indicate that neither the measured response nor an attempt to objectively quantify perceived quality is related to price. On average, in-ear headphones have a slightly higher measured bass response than circumaural and supra-aural headphones. Furthermore, in-ear and circumaural headphones have a slightly lower deviation from an assumed target curve than supra-ear models. Ninety percent of the variance across all headphone measurements can be described by a set of six basis functions. The first basis function is similar to published target responses, while the second basis function represents a spectral tilt.

© 2017 Acoustical Society of America  
[CCC]

**Date Received:** March 8, 2017    **Date Accepted:** May 8, 2017

## 1. Introduction

The market for headphones and earphones has been growing steadily over the last decade, both in value and volume. Market reports indicate a world-wide aftermarket reaching over  $11 \times 10^9$  units sold per year (Iyer and Jelisejeva, 2016), with growth rates of up to 55% in the 25–49 USD segment. Research suggests that factors influencing consumers' choice as to which model to purchase are mostly based on wireless functionality (Iyer and Jelisejeva, 2016) and attributes such as shape, design, and comfort (Jensen *et al.*, 2016). Interestingly, sound quality does not seem to be a major attribute for purchase decisions.

Objective assessments and subjective metrics for sound quality on headphones have been a subject of research, in particular over the last 5 years. Work by various authors has indicated that the subjective quality is mostly correlated with linear (spectral) attributes instead of non-linear (distortion) metrics (cf. Temme *et al.*, 2014; Fleischmann *et al.*, 2014). In particular, research suggests that the frequency (magnitude) response is a major factor in listener preference scores (Olive and Welti, 2012; Fleischmann *et al.*, 2012; Olive *et al.*, 2013), and one headphone can effectively be transformed into another one by means of headphone equalization (Welti *et al.*, 2016). The preferred response however seems to be listener, content, and headphone dependent (Olive and Welti, 2015; Olive *et al.*, 2016). Nevertheless, if one aims for a one-fits-all target response, diffuse field or free-field responses seem to be less preferred than a response based on measurements of a calibrated loudspeaker system in a listening room (Fleischmann *et al.*, 2012; Olive *et al.*, 2013).

The body of research work covering headphone responses and target curves is typically based on a relatively small set of headphones. Therefore, variance patterns of headphone responses and their dependencies on headphone type and retail price are difficult to uncover. The aim of this paper is to reveal such patterns across a large set of headphones and to investigate whether and how these response patterns vary with price and headphone type.

## 2. Method

Headphone spectral magnitude responses were measured on a HMS II.3 artificial head (HEAD Acoustics, Germany) equipped with a type 3.3, open ear canal artificial ear (ITU-T P.380 2003). Each headphone was resealed at least five times. A 5.46-s log sine sweep covering a frequency range between 20 Hz and 20 kHz was converted to

---

<sup>a)</sup> Author to whom correspondence should be addressed.

electrical signals by a Fireface UC (RME, Germany) sound card using the headphone output connector. Log sine sweeps rather than linear sine sweeps were employed to allow verification that non-linear distortion components were virtually absent. The responses for both the left and right channel were measured with a sampling rate of 48 kHz. The 10+ measurements were averaged in the spectral power density domain using a discrete Fourier transform. Subsequently, the averaged response was converted to the log domain, and resampled to a perceptually-relevant frequency distribution between 20 Hz and 19 kHz with 0.1 equivalent rectangular bandwidth resolution [cf. Glasberg and Moore (1990)] using linear interpolation. After interpolation, the mean response across frequency was subtracted. In total, measurements for 283 headphones [in-ear (IE), supra-aural (SA), and circumaural (CA)] were acquired. Headphone retail prices were determined using Google shopping in Australia and converted to USD using a fixed exchange rate of 0.75. All headphones were categorized according to retail price quantile (with cutoffs of 25% and 75%) and headphone type (CA, SA, and IE).

### 3. Results

#### 3.1 Mean and variance

The mean and variance of the headphone response curves as a function of frequency are shown in Fig. 1. The solid black line within the gray area represents the arithmetic mean, and the gray area denotes the standard deviation. The top panels group the data into quantiles according to retail price; the bottom panels represent CA, IE, and SA headphones, for the left, middle, and right panel, respectively. The variable  $N$  in the panel titles presents the number of headphones in each category. The text inset in each panel denotes the root-mean-square deviation (RMSD) for each category across all frequencies, and for frequencies below 100 Hz only.

Across all groups, the averaged responses demonstrate a resonant peak at around 3.5 kHz, a secondary resonance at 10 kHz, and a general decrease in response toward 19 kHz. The average response does not seem to differ much across categories, with the exception of the IE headphones demonstrating an amplified low-frequency response and a wider resonance around 2–5 kHz (bottom-middle panel of Fig. 1). The higher low-frequency magnitude response for IE headphones could be resulting from an acoustic seal between an insert headset and an artificial ear greater than that typically occurring for humans. In other words, listeners may not necessarily perceive IE headphones as on average having a greater bass response.

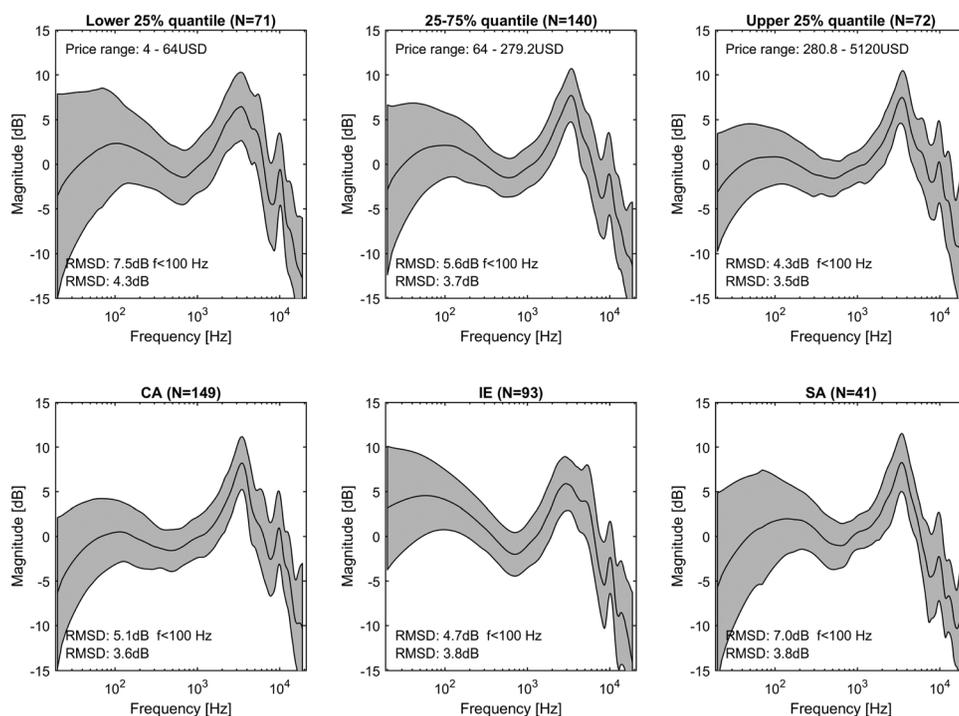


Fig. 1. Mean and variance of headphone response curves as a function of frequency. Top panels: lower, middle, and upper price quantiles. Bottom panels: CA, IE, and SA headphones. RMSD denotes the root-mean-square deviation across frequency (full frequency axis and below 100 Hz only).

The variance and the wide-band deviation across headphones (RMSD) decrease slightly with increasing price (RMSD of 4.3 and 3.5 dB for the lower and upper 25% quantile, respectively). The largest decrease in variance with price is observed for frequencies below 100 Hz, with the highest 25% price quantile having less than half the RMSD of the lowest 25% price quantile.

### 3.2 PCA

A principal component analysis (PCA) was performed across the headphone responses to investigate how response patterns correlate across frequency. The first six eigenvectors are shown in the left panel of Fig. 2, along with the cumulative percentage of variance accounted for. The right panels show the corresponding eigenvalue vs price for each headphone. Different symbols represent different headphone types (see the legend for Fig. 2).

The first eigenvector is identical to the mean response across headphones. The corresponding eigenvalues are all positive while their magnitude deviates considerably across headphones. The second eigenvector seems to reflect an overall spectral tilt, with eigenvalues roughly centered around zero. The third to sixth eigenvectors seem to mostly model the position and magnitude of high-frequency peaks and notches. The corresponding eigenvalues are scattered around zero. Most interestingly, none of the eigenvalues seem to correlate strongly with the retail price, nor does there seem to exist any clear dependencies on headphone type. The results also indicate that 90% of the headphone response variance can be accounted for by six eigenvectors. This result supports the notion that headphone equalization may be realized by means of a small set of basis functions (Fleischmann *et al.*, 2013).

### 3.3 Deviation from a target curve

Root-mean square errors (RMSEs) were calculated across frequency for each headphone with respect to an assumed target curve to assess an objective quality metric. Two target curves were used: the overall mean curve across all measured headphones, and the target curve suggested by Olive and Welti (2015). The two target curves are shown in the top panel of Fig. 3; the curve from Olive and Welti (2015) was lowered by 6.5 dB to facilitate easier comparison between the two. When comparing the two

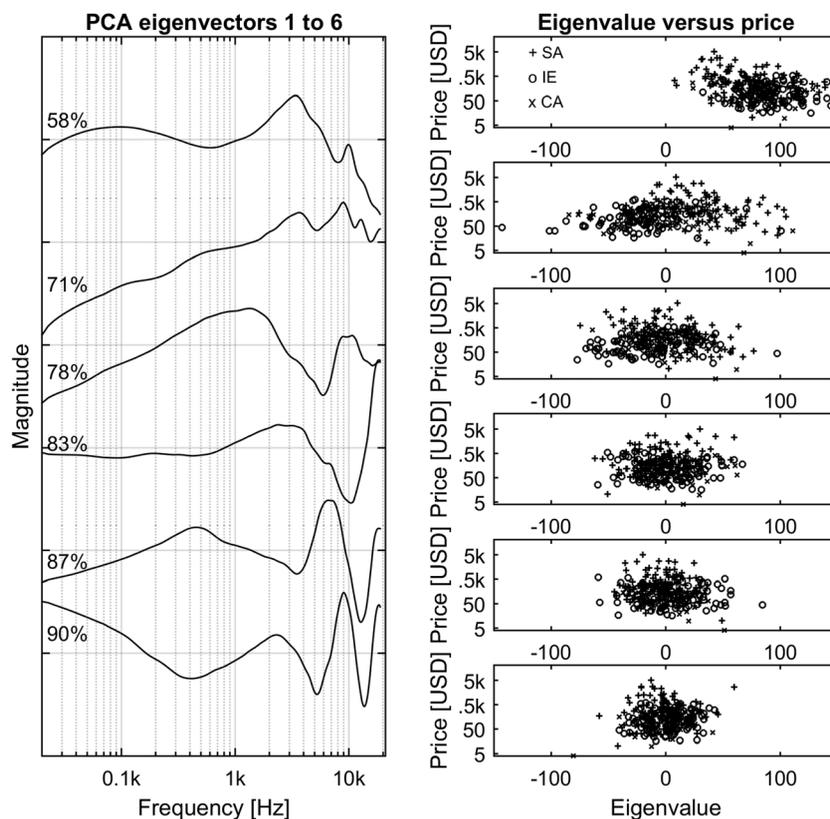


Fig. 2. Left panel: first six principal component eigenvectors and accumulative variance accounted for. Right panels: eigenvalue vs retail price for each headphone under test; top to bottom panels represent eigenvector one to six, respectively. Different symbols represent different headphone types (SA, IE, CA).

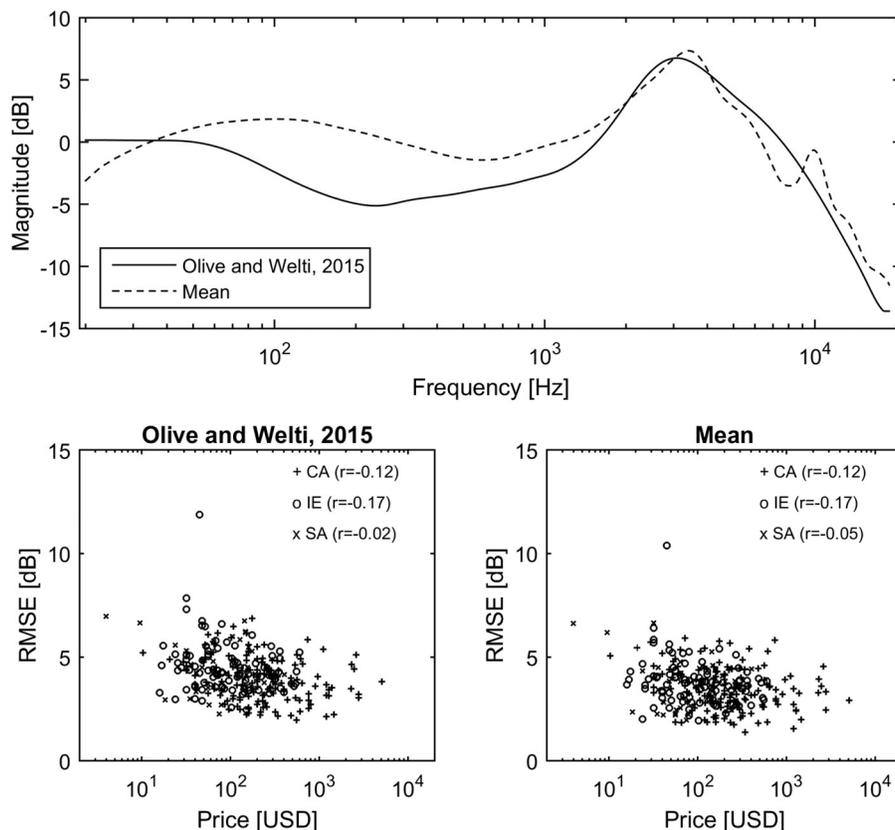


Fig. 3. Top panel: target curve adopted from Olive and Welti (2015) (solid line) and average headphone response (dashed line). Bottom-left panel: RMSE as a function of price for all headphones using Olive and Welti (2015) as target function. Bottom-right panel: same as bottom-left panel, based on mean headphone response as target curve. Different symbols represent different headphone types (SA, IE, CA). Numerical values in the legend represent the Pearson correlation coefficient between RMSE and the retail price for each headphone type.

target curves, the largest difference between the two seems to exist between 50 Hz and 2 kHz, for which the average headphone response is up to 5 dB higher than the target curve suggested by Olive and Welti (2015), in particular around 100–200 Hz. The scatter plots of RMSE against retail price (lower panels of Fig. 3) are nevertheless very similar in the sense that none of them demonstrates a high correlation between price and deviation from the target curve. The numerical value between brackets in the inset of the lower panels denotes the Pearson correlation coefficient between RMSE and price. The IE headphones demonstrate the largest (absolute) correlation but its magnitude is nevertheless very small.

To further establish the significance of headphone type and RMSE, a two-way analysis of variance (ANOVA) was conducted with headphone type and RMSE as independent factors, and price as a dependent variable. These tests were based on a single target curve for all three headphone types (Olive and Welti, 2015). The effect of headphone type was shown to be significant [ANOVA F statistic  $F(279, 2) = 21.05$ ; significance level or probability of falsely rejecting the Null hypothesis  $p < 0.0001$ ] while the effect of RMSE was not [ $F(279, 1) = 2.59$ ;  $p > 0.10$ ]. A *post hoc* comparison of mean RMSEs across headphone types was performed using a non-parametric permutation test and the median as location statistic. Bonferroni-adjusted  $p$  values indicated that the RMSEs for CA and IE headphones are on average significantly smaller than those for SA headphones ( $p < 0.025$ ).

#### 4. Conclusions

- Based on the evaluation of the mean, variance, PCA, and mean square error with respect to a target function, no correlation could be observed between the measured magnitude response and retail price of headphones. However, the variance in low-frequency response seems to decrease with increasing price, indicating an improved bass response measurement consistency across headphones in the higher price range. It is however unclear whether this improved consistency with a higher retail price is the

result of better headphones or better repeatability of measurements with more expensive models. Nevertheless, assuming that the perceived audio quality is largely determined by the spectral magnitude response of headphones, there are plenty of relatively cheap models that match the assumed target function, as well as very expensive ones that deviate significantly from an assumed ideal response.

- On average, IE headphone measurements demonstrate slightly more bass than CA and SA headphones. The difference amounts to approximately 4 dB at 100 Hz. This finding is in line with subjective preference ratings reported by Olive *et al.* (2016). The observed higher bass response could be resulting from a greater seal between headphones and artificial ear compared to the seal for human ears (ITU-T P.380, 2003).
- CA and IE headphones have a response that (on average) more closely mimics the target curve proposed by Olive and Welti (2015) than SA headphones.
- The target function suggested by Olive and Welti (2015) is fairly similar to the average headphone response found in this study, with the exception of a deviation of up to about 5 dB for frequencies between 50 Hz and 2 kHz.
- PCA can account for 90% of the variance across all measured headphones with six eigenvectors. The first eigenvector is similar to published target responses, while the second eigenvector represents a global spectral tilt.

## 5. Limitations

First, all analyses were based on magnitude responses only, discarding phase or non-linear attributes. Research has suggested that non-linear properties of headphone responses have little effect and no significant traces of non-linear behavior were observed in this study. However the author is not aware of any work on the perceptual implications of the phase response of headphones. A second limitation is that a specific target curve was assumed, irrespective of the headphone type, audio content, listener demographics, or influence of environmental sounds. Third, it is well known that dummy-head measurements for headphones can be inaccurate at low and high frequencies and deviate from human ear canal measurements (ITU-T P.380, 2003; Christensen *et al.*, 2013). Repeated measurements for resealed headphones should partly alleviate this problem. However, it could be that measurement inaccuracies have some implications for the results and conclusions drawn in this paper.

## References and links

- Christensen, A. T., Hess, W., Silzle, A., and Hammershøi, D. (2013). "Magnitude and phase response measurement of headphones at the eardrum," in *Proceedings of the 51st AES International Conference*, Helsinki, Finland.
- Fleischmann, A., Silzle, F., and Plogsties, J. (2012). "Identification and evaluation of target curves for headphones," in *Proceedings of the 133rd AES Convention*, Convention Paper 8740, San Francisco, CA.
- Fleischmann, F., Estrella, J., and Plogsties, J. (2014). "A method for comparison of non-linearities of consumer earphones using equalized stimuli," in *Proceedings of the 136th AES Convention*, Convention e-brief 141, Berlin, Germany.
- Fleischmann, F., Plogsties, J., and Neugebauer, B. (2013). "Design of a headphone equalizer control based on principal component analysis," in *Proceedings of the 134th AES Convention*, Convention Paper 8869, Rome, Italy.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- ITU-T P. 380. (2003). "Electro-acoustic measurements on headsets," International Telecommunications.
- Iyer, R., and Jelisejeva, Z. (2016). Headphones Market Report—Worldwide Outlook.
- Jensen, R., Lauridsen, N., Poulsen, A., Tofte, C., and Christensen, F. (2016). "Analysis of subjective evaluation of user experience with headphones," in *Proceedings of the AES Conference on Headphone Technology*, Aalborg, Denmark.
- Olive, S. E., and Welti, T. (2012). "The relationship between perception and measurement of headphone sound quality," in *Proceedings of the 133rd AES Convention*, Convention Paper 8744, San Francisco, CA.
- Olive, S. E., and Welti, T. (2015). "Factors that influence listeners' preferred bass and treble balance in headphones," in *Proceedings of the 139th AES Convention*, Convention Paper 9382, New York, NY.
- Olive, S. E., Welti, T., and Khonsaripour, O. (2016). "The preferred low frequency response of in-ear headphones," in *Proceedings of the AES Conference on Headphone Technology*, Aalborg, Denmark.
- Olive, S. E., Welti, T., and McMullin, E. (2013). "Listener preference for different headphone target response curves," in *Proceedings of the 134th AES Convention*, Convention Paper 8867, Rome, Italy.
- Temme, S., Olive, S. E., Tatarunis, S., Welti, T., and McMullin, E. (2014). "The correlation between distortion audibility and listener preference in headphones," in *Proceedings of the 137rd AES Convention*, Convention Paper 9118, Los Angeles, CA.
- Welti, T., Olive, S. E., and Khonsaripour, O. (2016). "Validation of a virtual in-ear headphone listening test method," in *Proceedings of the 141st AES Convention*, Convention Paper 9658, Los Angeles, CA.