Private Investigations

PRIVATE HEADPHONE LISTENING HAS BECOME HUGELY POPULAR, AND NOW ITS IDEAL FREQUENCY RESPONSE IS BEING RE-EVALUATED. KEITH HOWARD EXPLORES, AND ENVISAGES FURTHER DEVELOPMENTS

You may never have tested an item of hi-fi equipment in your life but will surely know that the measured frequency response of a pre-amplifier (from a line input) or a power amplifier will usually be flat to within tight limits across the 20Hz to 20kHz audio spectrum. Something similar should occur when measuring the on-axis pressure response of a loudspeaker, although the departures from a flat response will usually be larger, especially at low frequencies. In all audio equipment, 'flat is right' (at least in conventional thinking), albeit with two obvious exceptions. One is the phono pre-amplifier, which should have a frequency response that is the inverse of the RIAA/IEC replay curve to give a neutral and 'flat' response.

However, the second exception of headphones is a very different proposition. If you measure a headphone's frequency response (specifically, the response at the eardrum, properly termed the drum reference point, DRP) it will be anything but flat, and will have a marked peak around 2-3kHz. But in this case the correction to apply (that is, the correction which will generate a response that represents the headphone's perceived tonal balance) is altogether less certain. The 'target' response that creates a perception of neutral tonal balance in a headphone has been the subject of debate for decades, and recent research has reignited the controversy.

For a headphone intended to reproduce conventional stereo recordings – binaural recordings are a different case, not considered here – it might seem obvious that the frequency response at the eardrum should match that of a loudspeaker placed in the classic equilateral triangle layout for speakers and listener, 30 degrees off the median plane (the vertical plane that bisects the head from front to back). In headphone parlance this is termed the free-field response assumption, and it held sway until the mid-1980s when it came under challenge from research conducted by Günther Theile at IRT – the Institut für Rundfunktechnik (a Broadcast Technology Institute) in Munich [1].

Drawing on principles of Gestalt psychology (Gestalt refers to a theory of perceptual organisation in which the whole amounts to more than the sum of its constituent parts), Theile argued that the free-field response would only be appropriate were the stereo image from headphones to be perceived in the same way as that from loudspeakers, *ie* externalised in front of the listener. In fact the headphone image is normally perceived quite differently, as being inside or very close to the head.

This matters because of perceptual constancy. If in an open space (to obviate room effects) someone walks slowly around you while talking, their voice remains the same despite the frequency response at each eardrum varying according to their position. What the brain does is exploit the frequency response changes to help localise the person talking, then compensates for them so that the vocal quality appears unchanged.

If the headphone image does not match that from a pair of loudspeakers, it follows that the free-field response assumption must be wrong: a headphone with free-field frequency response will colour the sound because the brain applies the wrong correction to it. Theile went on to demonstrate this in experiments, from which he concluded that the correct headphone target response is not the free-field but the diffuse-field response – that experienced when sound impinges on the head with equal intensity from all directions. The diffuse-field target response was subsequently adopted by many headphone manufacturers, although some continued to prefer responses that approximated the free-field alternative.

Theile's work went substantially unchallenged for over 25 years until the issue was taken up by Sean Olive and colleagues at Harman International

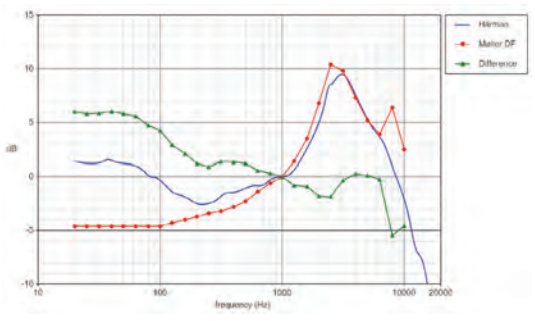


Figure 1. Harman's headphone target response (blue trace) compared to the diffuse-field response (red trace). The difference between them (green trace) is particularly marked at low frequencies

in Northridge, California [2,3]. In the course of a series of experiments involving multiple listeners – some trained, some not – Harman has identified a new headphone target response that is significantly different from the diffuse-field response particularly at low frequencies. As *Figure 1* shows, the two responses are not that disparate through the area of the 2-3kHz peak, but below 1kHz the Harman target response (compared to the diffuse-field response) shows first a gentle sloping upwards of output, followed by a steeper upwards slope beginning at about 200Hz, and culminating in a +6dB shelf below 50Hz.

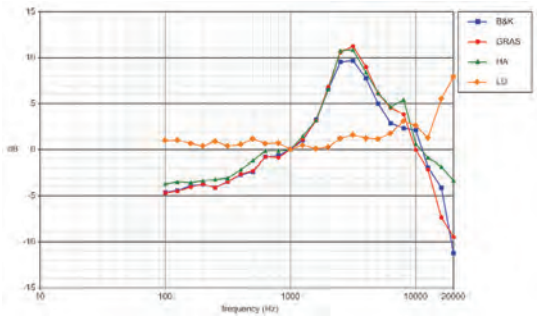


Figure 2. Diffuse-field response as measured using three different head and torso simulators: the B&K Type 4128-C (blue trace), GRAS Type 45BM (red trace) and Head Acoustics HMS II.3 (green trace), with the orange trace depicting the largest disparity at each third-octave test frequency (from [4])

(At this juncture I ought to point out that the diffuse-field response is not so well-defined as the above might suggest. The IEC, IEEE and ISO all publish diffuse-field responses in different standards; all, while broadly similar, are different. As *Figure 2* shows, so too are the diffuse-field responses recorded using different artificial heads in the form of head and torso simulators (HATS) [4]. The diffuse-field response I use here, and apply as a correction to the headphone measurements I make, using key elements from a GRAS *43AG* ear and cheek simulator, is due to Hammershøi and Møller [5]. The 'head' I use – built from bonded sheets of MDF over which a headphone is placed as in normal use – is pictured in *Figure 3*.)

The shelving up of bass output in the Harman headphone target response broadly matches the in-room response of a loudspeaker in an acoustically well behaved room. Given that a headphone is not subject to room gain, like a loudspeaker, this might seem odd. One speculative explanation is that room gain in mastering studios causes bass level to be equalised down, requiring a headphone (and loudspeakers) to boost it again.

Independently, similar research has been undertaken at the National Research Council in

Figure 3. The artificial head used by the author for measuring headphones, which incorporates key elements of the GRAS 43AG ear and cheek simulator. Left and right artificial pinnae are used as appropriate

Canada – particularly well-known for its audio research in the 25 years that Floyd Toole was resident before moving to Harman – and the results are incorporated in the *RoomFeel* headphones that have appeared from PSB and NAD. In this case, though, the research results have not been published. While it is clear from measurements on *RoomFeel* headphones (I show an example later) that shelved-up bass is again a feature, more than this we don't know about NRC's target response and how closely it resembles Harman's. (Would that I could overlay the two on a graph for you, but I can't.)

Identification of the Harman target response (which applies, please note, to circum- and supra-aural headphones only; insert earphones, also known as 'earbuds', are the subject of further research) is a significant event in the history of headphone science. Already many headphone manufacturers seem to be preferring it to the diffuse-field response. But I consider that Harman's findings regarding individual variability are even more important.

The Harman target response is, of course, an average: the mean preferred response of the subjects tested. But those subjects varied widely in their preference, a finding which particularly intrigues me because the Harman target response doesn't align with my own – although neither does the diffuse-field (DF) response. My favourite headphones, I've found, have a DF-corrected response that shows a presence band dip of a few dB, and a maximum bass lift of about 3dB (both relative to 1kHz). Any

more low frequency lift than that and I find the bass/lower-mid exaggerated and the overall sound dulled and muddied as a result.

What Harman's research uncovered were variations in listener preference that were much larger than those I've just described in myself. In the second Harman paper of 2013 [3], user-adjustable linear-phase bass and treble shelf filters with the both-maximum and both-minimum frequency responses shown in *Figure 4* were used to determine the preferred bass/treble balance of 11 listeners, using both loudspeakers and headphones. Eight were male and three female; eight of the listeners had scored highly in Harman's listener training whereas three external listeners had no formal training. The amplitude response of the headphone (a Sennheiser HD 800) was first equalised to match the in-room target response of the loudspeaker measured at the DRP. The three tracks used were Jennifer Warnes' *Bird On A Wire*, Steely Dan's *Cousin Dupree* and Estelle's *American Boy*. (Harman has not used classical programme in any of its published headphone target response research.)

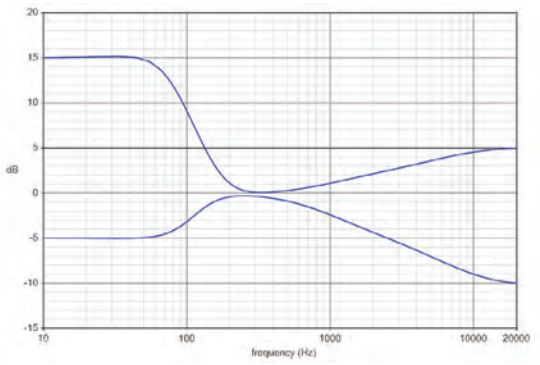


Figure 4. Frequency responses of the Harman shelf filters at their maximum and minimum bass/treble settings (+15dB/+5dB and -5dB/-10dB)

Shocking is not too strong a word, I think, for the variation in bass/treble filter settings exhibited by this small group of listeners. For headphones the bass filter settings ranged from about +13.5dB to about -1.0dB (I say 'about' because I don't have tabulated results; I'm scaling from the published graph), and treble filter settings from about +0.5dB to about -8.5dB. I need hardly tell you that these variations are enormous – in fact they dwarf the response variations found between many headphones.

We shouldn't be entirely surprised at this: it has been known for at least 20 years that headphone frequency response measured on human ears [6] give a large variation between individuals at high frequencies. But why there should be such large disparities in preferred bass level is something of a puzzle. Guessing, it might be an attempt by some listeners to compensate for headphones' lack of visceral bass impact, and/or it might also reflect expectations born of habituation to bass excess. (It's worth noting, however, that the variations in filter setting were even larger when using loudspeakers!)

Does this mean that the concept of a single defined headphone target response has had its day? No. Headphone manufacturers have to have some frequency response to aim at, and hitting the mean preferred response – assuming a normal distribution of preferences – will ensure that the greatest number of listeners will find the tonal balance acceptable.

But what of listeners whose preferred response is sufficiently different to make the target response is unacceptable? It appears there will be significant numbers in this position, and widespread adherence to any target response – whatever it may be – must inevitably marginalise them. Hope is not a strategy, so we can't simply suppose that sufficient 'non-compliant' products will continue to exist for those who fall towards either extreme of the bell curve to find a headphone which meets their needs. This is why I've been opining for some time – verbally if not in print – that headphone amplifiers incorporating user-adjustable equalisation are well overdue. It's easy to adopt a sniffy audiophile 'no tone controls' attitude to this suggestion, but if the evidence suggests that individualised headphone EQ is needed, the debate should not be about whether to have it but how best to realise it – shelf filters like Harman's or third-octave equalisation, say?

Apps like Golden Ears' Accudio which equalise supported headphones to a set target response are no answer to this. By further imposing a universal target function they ignore the variety in listener preferences, and may make the perceived tonal balance worse rather than better for some users.

What does this mean for headphone reviews? Is there any point in me promulgating my reactions to a headphone's tonal balance when yours may be quite different? I think there is *provided that the review is accompanied by frequency response measurements*. It also helps if the reviewer is as upfront as possible about his/her personal taste, and describes the tonal balance as they perceive it – you'd be surprised at how many headphone reviews make little or no explicit mention of this. Then individual readers have a fighting chance of identifying headphones that may suit them, whatever the opinion offered in the review.

If I criticise a headphone for a dull tonal balance, for instance, but you hear it and disagree, the combination of my subjective impressions and the measured frequency response can help identify other

References

[1] Theile, G, 'On the standardization of the frequency response of high-quality studio headphones', *J. Audio Eng. Soc.*, vol 34, no 12 (December 1986)

[2] Olive, S; Welti, T; McMullin, E, 'Listener Preference For Different Headphone Target Response Curves', Convention Paper 8867, Audio Eng Soc 134th Convention (May 2013)

[3] Olive, S, Welti, T; McMullin, E, 'Listener Preferences for In-Room Loudspeaker and Headphone Target Responses', Convention Paper 8994, Audio Eng Soc 135th Convention (October 2013)

[4] Snaidero, Thomas; Jacobsen, Finn; Buchholz, Jörg, '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

[5] Hammershøi, D; Møller, H, 'Noise from sources close to the ears', *Acustica United with Acta Acustica*, vol 94, no 12, p114-129 (2008)

[6] Møller, H; Hammershøi, D; Jensen, C B; Sørensen, M F, 'Transfer Characteristics of Headphones Measured on Human Ears', *J. Audio Eng. Soc.*, vol 44, no 4 (April 1995)

headphones likely to make your tail wag, even if they don't mine. Without a measured frequency response – and how many of those do you see in headphone reviews? – this is harder to achieve.

What doesn't make sense, to me, is a committee approach to reviewing. Individual reactions hidden behind a homogenised verdict are no more useful, on the basis just described, than the judgment of a single reviewer. In fact a committee verdict is likely to be worse because the preference of a group is less easy to identify, and not even constant if the members of the committee keep changing. Neither is it any use for a single reviewer to attempt to adopt others' perceptions. It is one thing to understand that your opinions won't be shared by everyone; it's quite another to walk in their shoes and listen with their ears. Manifestly you can't. You can be careful to say that listeners who prefer a different tonal balance may like/dislike a particular headphone more than you do, but it's never possible to enthuse or castigate on the basis of anyone's perception but your own.

Not long ago headphones were an almost disregarded backwater of the hi-fi industry. Today, because mobile listening has burgeoned, and headphone tonal balance has become a hot topic for research, this has become a vibrant mainstream market with sufficient research effort and attached controversy to whet the audiophile appetite. Much has changed in the last five years; the next 10 will surely bring more revelations.

Sincere thanks to Sean Olive and Todd Welti of Harman International for providing the data for Figures 1 *and* 4.

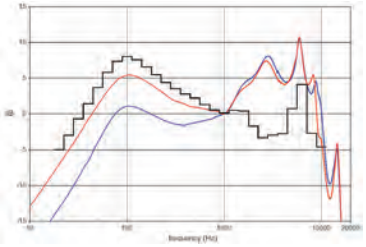
Beyerdynamic T1 gen 2

Reviews

I measured four disparate headphones to illustrate the range of frequency responses currently on offer – although more extreme examples than these do exist. In each case the blue and red traces in the graphs represent the uncorrected left and right channel responses respectively, averaged from 10 separate measurements. The black third-octave trace is a consequence of applying diffuse-field correction to the mean response of the two capsules.

Apple Ear Pod

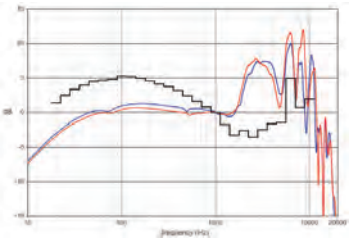
I included the Apple *Ear Pod* because it has clearly been thoroughly engineered but costs just £25. Below 1kHz the response from the two channels was very different, suggesting an issue with sealing the right pod to the artificial ear. Even with a good seal, though, the bass response falls away below a peak at about 100Hz, and the DF-corrected response shows a mild dip in the presence band followed by a peak at 6kHz. Subjectively I just didn't get on with this design. Realign the pods as I might I couldn't achieve a sufficiently weighty tonal balance and the stereo image seemed flat and somewhat constrained.

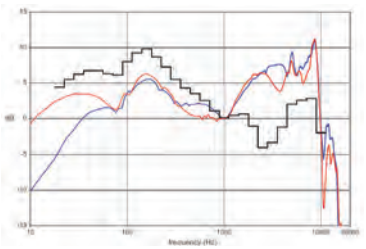


[EDITOR'S NOTE: A hitherto unresolved difference of opinion has emerged between Keith Howard and Martin Colloms over the correct means of mounting and measuring the Apple *Ear Pod*.]

Beyerdynamic T1 gen 2

I really liked Beyerdynamic's original *T1*. With DF correction applied its response had a mild lift at 100Hz of about 3dB and a mild suckout immediately above 1kHz to the tune of 1-1.5dB. I thought its subjective bass performance excellent and also appreciated its powers of transparency and analysis. With the generation 2 model, Beyerdynamic has raised the bass shelf in line with the Harman target response and presence band output is reduced by 1-2dB – and I now find the *T1*'s sound distinctly less engaging.





NAD Viso HP50

NAD's *Viso HP50* is one of the models to incorporate *RoomFeel*, based on the target response identified at the Canadian NRC. The lower-midrange lift in its DF-corrected response reaches almost 10dB, and would have been better maintained at bass frequencies but for an apparent sealing problem with the left capsule. Together with the dip centred on 3kHz this makes the *HP50* sound much too sluggish and warm-hued for my taste.

Oppo PM-2

Typical of planar magnetic headphones, the Oppo *PM-2* has very well maintained bass extension, but the DF-corrected response declines from 800Hz to 5kHz which again makes its sound too thickened for me. Still, it has been well reviewed elsewhere.