

Headphone Essentials 4:

The Skinny on Headphone Frequency Response Graphs

© Dale Cotton, 2020. All rights reserved. Version 2.54 (April 2022).

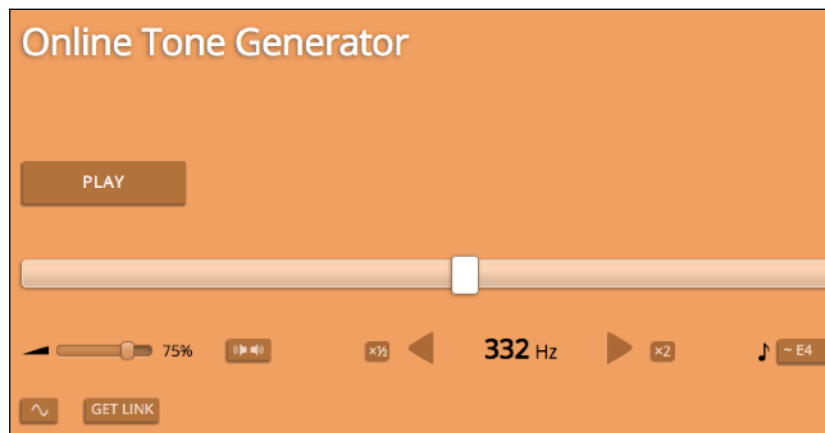
You may re-circulate this document. You may not claim authorship or copyright to it.

(Photographs are advertising copy in the public domain and copyright belongs to the original photographer.)

Note: this document is part of a instructional series. If you would like to start with more foundational material on the nature of sound, see *Basics of Musical Sound* and *Basics of Headphone Sound* at [Headphone Essentials](#).

If you are reading this document, you are likely interested in finding a headphone that matches your listening preferences. The single most important factor for many is the headphone's frequency response (FR). This is often called its tuning or tonality. Other factors, such as its dynamic responsiveness, sound staging, imaging and detail reproduction are also important considerations. But if the tuning is not agreeable those other factors are not likely to rescue it.

Getting started



To get started go to the [Online Tone Generator](#) in your browser. Make sure the volume on your computer is initially set to a fairly low level, press the space bar once to turn sound on, then move the slider fairly slowly and evenly left and right. If the sound does not change in loudness (except for tapering off at both extremes) you're hearing what's called a subjectively flat response. If it grows louder and softer as you move the slider, that's a non-flat response. If possible, you may want to try this with more than one headphone or speaker to see how they vary (remembering to start at a low loudness level with each one). (By the way, as is typical of tiny speakers, the one built into my laptop doesn't produce any sound at all to the left of about 100 hertz.)

Now let's say we want to make note of what we're hearing when sound is being produced by a particular pair of headphones. Let's say we start at the far left and hear nothing until maybe 22 hertz. Then we hear it gradually getting louder until 67 hertz, then quickly getting quieter until 83 Hz, then starting to climb again. We might draw a line like so to represent this:

The Skinny on Headphone Frequency Response Graphs

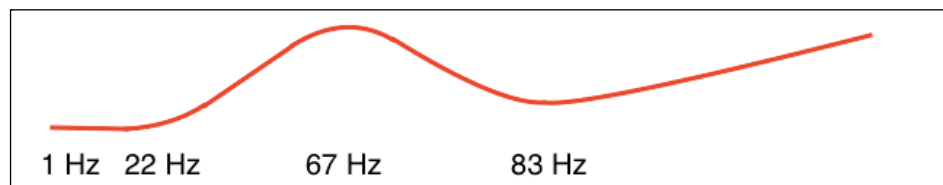


Fig. 1: freehand loudness line

What we're intuitively doing is recording the hertz numbers from left to right, just as they appear when we move the slider from left to right. Plus, we're showing the change in loudness as an up and down movement. And that's all there is to a frequency response graph:

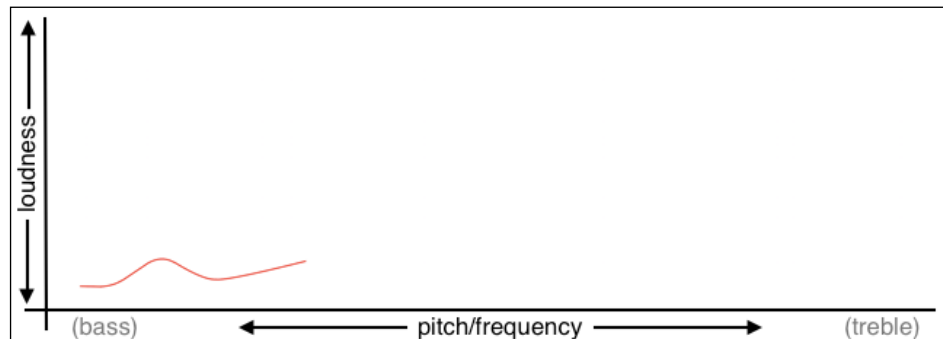


Fig. 2: frequency response graph axes

We can see that if we want to record the actual hertz numbers (abbreviated as Hz) we would need a very long graph from left to right, since the hertz numbers go from 1 to 20,154. The math trick to deal with that is to use a logarithmic scale. This simply means the distance keeps getting shorter as the numbers keep getting larger:

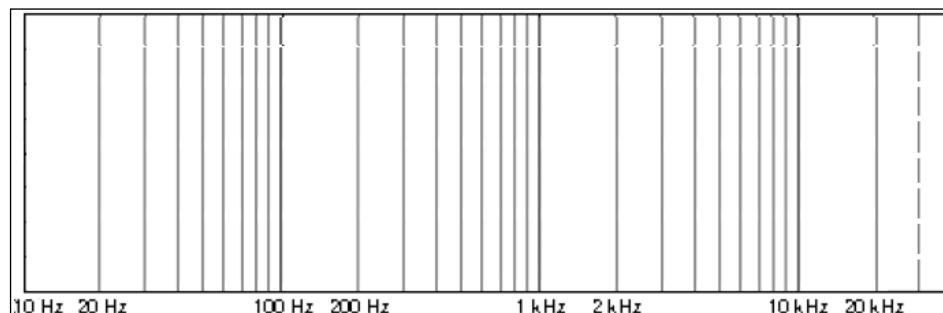


Fig. 3: frequency response graph with logarithmic hertz scale

Fig. 3 has a logarithmic scale from left to right. From 10 to 100 there are guide lines on the tens. From 100 to 1000 they're on the hundreds. From 1000 to 20,000 they're on the thousands. Notice that for convenience we use the kHz notation to get rid of the increasingly large number of zeros, so 1 kHz = 1000 Hz. The scale concentrates on the range of 20 to 20,000 Hz since this is the range of human hearing. (Keep in mind that the highest frequency you can hear decreases gradually with age. So 20,000 Hz is highly optimistic.) Also notice that the distance between 10 Hz and 100 Hz is exactly the same as the distance between 100 Hz and 1000 Hz. And both are exactly the same as the distance between 1000 Hz and 10,000 Hz.

That takes care of the left/right scale. For up/down we can use logarithms again, as shown in Fig. 4. But the units we need for loudness are normally notated as dB, which stands for *decibels*. For our purposes, the absolute values don't matter. We'll use a scale with an arbitrary starting or 0 point as a baseline and simply indicate the degree of distance from that particular loudness, either louder or quieter in decibels:

The Skinny on Headphone Frequency Response Graphs

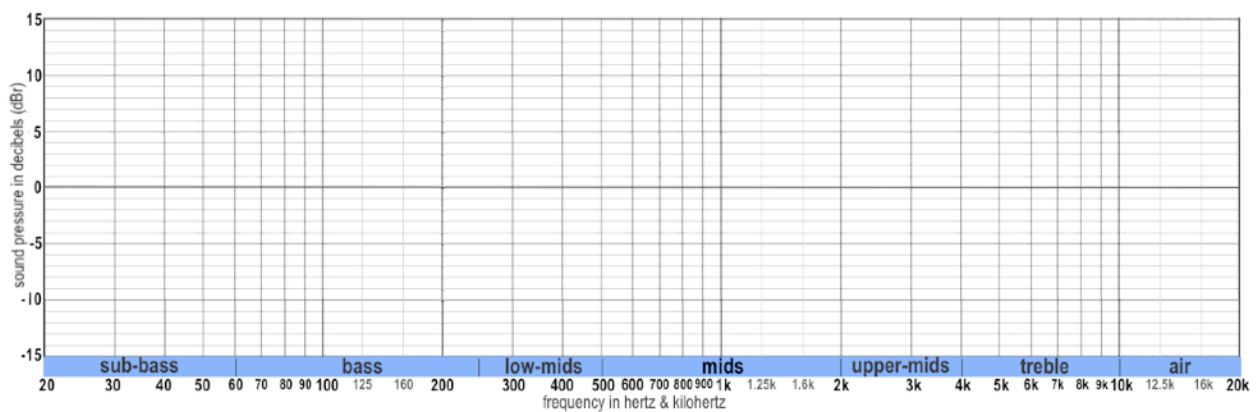


Fig. 4: frequency response graph with loudness expressed as sound pressure

Notice, however, that the vertical axis is now labeled as *sound pressure*, which is measurable, unlike the purely subjective experience of loudness. Any six decibel change going upward doubles the sound pressure. Any six decibel change going downward halves it. But neither change necessarily equates with a perceived doubling or halving of loudness. Depending on many factors it might take up to ten decibels to double or halve the subjective loudness. Going the other way, a one decibel change is about as fine a discrimination as we humans can reliably make. This new presentation might seem like an improvement until we actually try to use it. We can get exact frequency numbers from the Online Tone Generator. But how do we judge how high or low to place each vertical scale value with any degree of accuracy?

Measurements done right



For this we need some fancy, technical, and expensive measuring equipment (\$20,000+ is not uncommon) to achieve reliable results. Fortunately, there are several good samaritans doing such measurements and posting them on the web. The premiere example is acoustic engineer Oratory1990, whose measurements, made with the gold standard GRAS 45 equipment, can be found on the Headphone Database. Here's Oratory1990's measurement for the still-popular Audio-Technica ATH-M50x headphone:

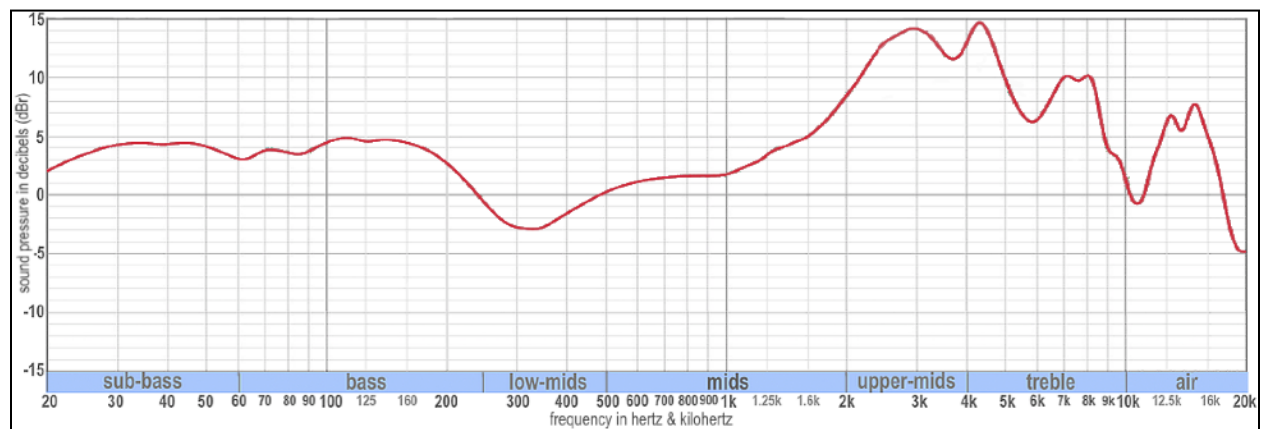


Fig. 5: Oratory1990 FR graph for the Audio-Technica ATH-M50x

The Skinny on Headphone Frequency Response Graphs

Unfortunately, there's a slight gotcha. The red line shows changes in loudness going across the human hearing range of frequencies. But, since it shows actual measurements, the line is much more of a jagged rollercoaster ride than your ears would lead you to believe. Also, our ears and brains are designed to hear the faintest possible sounds that are critical for survival. The peculiar shape of our ears are in fact natural amplifiers for high frequency sounds. When we factor in the acoustics of the enclosed spaces in which music is typically made and heard we get:

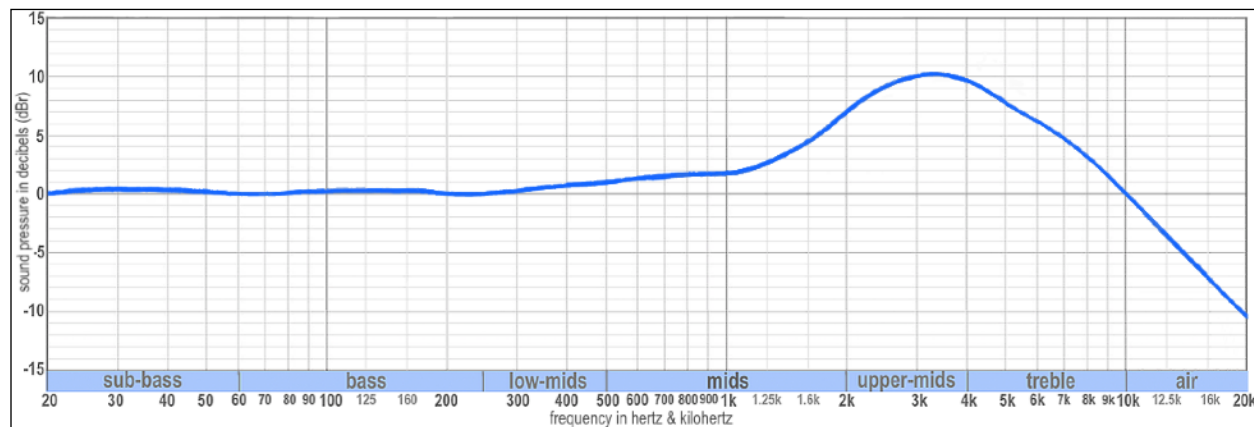


Fig. 6: Oratory Optimum HiFi over-ear headphone neutrality curve

Simplistically, a headphone that exactly re-produced this measurement curve would produce a even, or flat-line, sound if we used it to listen to the Online Tone Generator ([dig deeper here](#)). In other words, our brains would translate the curve shown in Fig. 6 into the subjective “curve” shown in Fig. 7:

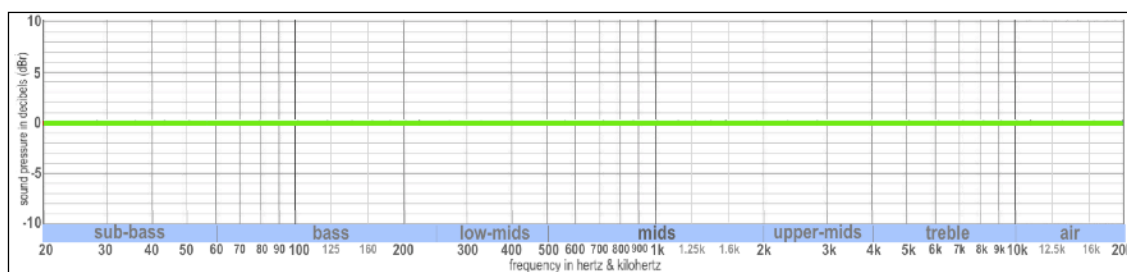


Fig. 7: subjective experience of a perfectly even loudness across the frequency range

If we overlay the curve from Fig. 6 onto the M50x measurements of Fig. 5 we get:

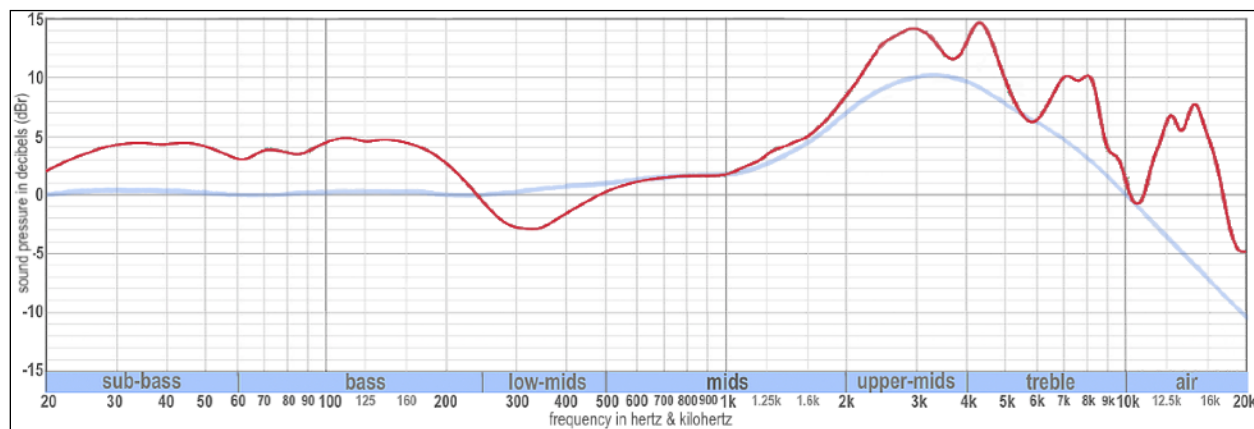


Fig. 8: ear amplification (red) vs M50x frequency response (blue)

The Skinny on Headphone Frequency Response Graphs

The sheer number of ups and downs in Fig. 8 may look like a problem. But it's actually about average for today's headphones (especially if they're a closed-back design like the M50x is). Our ears can't easily distinguish small up/down changes, and manufacturers simply don't have the technology to do better. Most often smoothing out a frequency response irregularity results in a loss of some other equally important acoustic quality, like dynamics or sound stage.

What we should hear if we use a pair of M50x to listen to the Online Tone Generator is an increase in loudness where the red line climbs above the cyan line. Plus, we should hear a decrease in loudness wherever the red line drops below the cyan line.

Another source, especially for measurements of lower-priced headphones, is [Rtings.com](https://www.rtings.com). Here's their measurement presentation for the M50x:

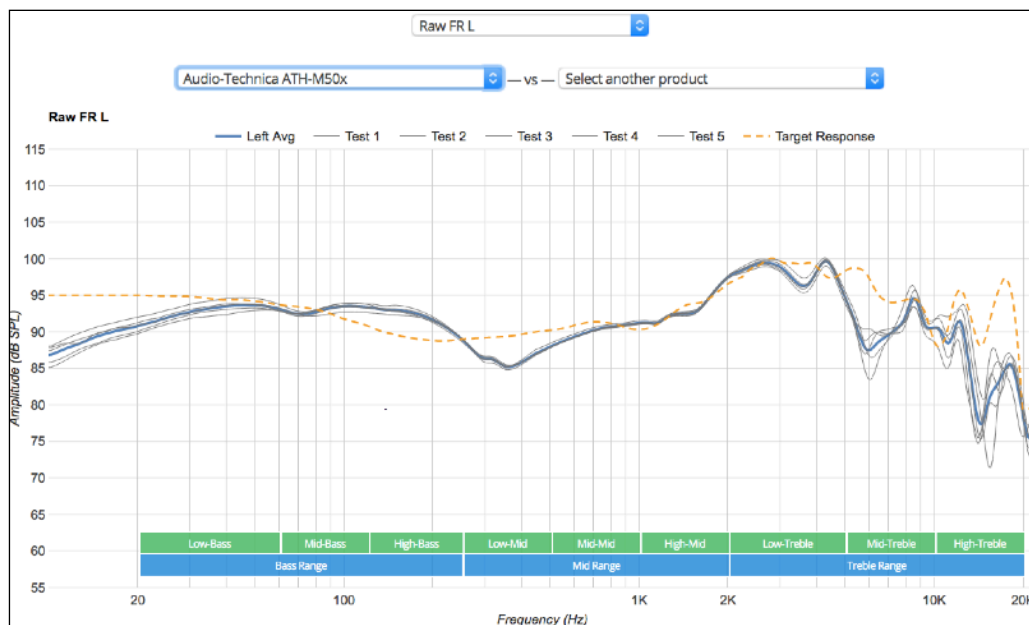


Fig. 9: [rtings.com](https://www.rtings.com) measurements for M50x

One difference is that Rtings show measurements between 10 and 20 Hz on the far left. Mostly, we would perceive these as sense of touch vibrations rather than sound. Another difference is the dotted orange line. This is their version of a common target for a headphone sound response with elevated bass (Harman 2018) as adapted to their measuring equipment.

A third site with a large collection of measurements is Crinacle's [In-ear Fidelity](https://www.crinacle.com/). Two other sites have significant collections of quality raw FR measurements: [SoundStage!Solo](https://www.soundstage!solo.com/), [Headphonetestlab](https://www.headphonetestlab.com/). Also, Resolve of the Headphone Show has recently started doing measurements which can currently be found in a forum post at [Headphones.com](https://www.headphones.com/).

Finally, there is an older site, InnerFidelity, sadly now defunct, but its extensive collection of measurements are happily available on [Stereophile](https://www.stereophile.com/). These measurements are in a PDF for each headphone and look as shown in Fig. 10. The FR graph is upper left. But confusingly it contains two sets of lines. The lower, grey lines are multiple trials of measurements for the headphone moved slightly for each one. (All competent technicians use the same multiple-trial procedure, only showing the average of the trials.) But the red and blue lines above the grey lines show the average of left and right channels separately. They are also adjusted for the properties of the measurement equipment by comparing them to a now-obsolete target response.

The Skinny on Headphone Frequency Response Graphs

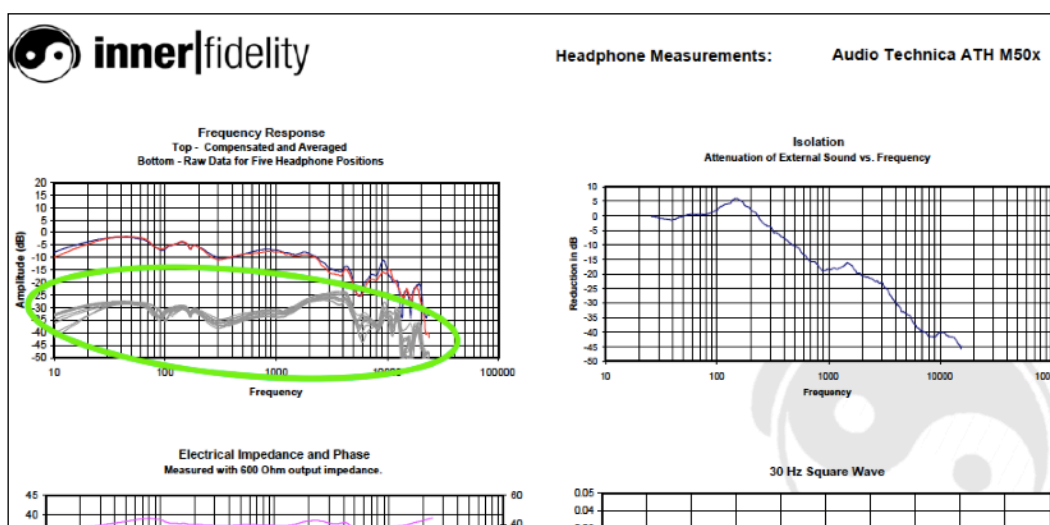


Fig. 10: innerfidelity.com measurement page for M50x, frequency response circled in green

At this time many of Oratory1990's headphone measurements have yet to be duplicated on the [Headphone Database](#) site. You can find the others on his [reddit EQ page](#) embedded in the PDF page for each headphone.

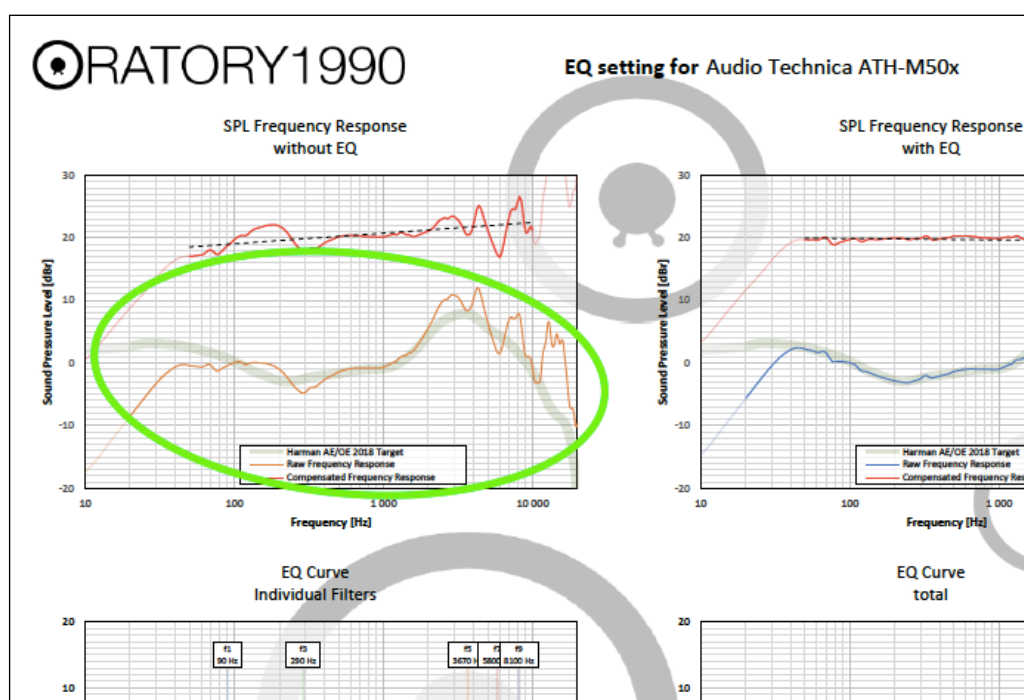


Fig. 11: Oratory1990 EQ database page for M50x with FR graph circled in green

Oratory1990 currently uses the gold standard GRAS 45BC measurement equipment. This includes not only a highly realistic simulacrum of a statistically adult average human ear. But also the ears are mounted in a mannikin-like head which facilitates achieving a consistent and repeatable signal from both ears. Crinacle, Brent Butterfield of SoundStage!Solo, Keith Howard of HeadphoneTestLab and Resolve are using some permutation of the somewhat less expensive GRAS 43 measurement rig. This employs a single copy of the same ear simulator used in the GRAS 45, requiring the technician to attempt to achieve the same clamp force, etc. when measuring both sides of a headphone. Rtings and InnerFidelity use an older HeadAcoustics set-up which gives similar results but is less accurate in the high frequencies.

The Skinny on Headphone Frequency Response Graphs



Fig. 12: random human ear photos (from Wikimedia Commons)

The GRAS mannikin ears are notable for more accurately representing what a hypothetically *average* human ear would hear, based on an international standard. Yet look at a half dozen different people's outer ears to see just how far we all depart from the norm, and ear canals vary as well. The accurate ear anatomy primarily shows up in the high frequencies. But the GRAS outer ear also has the same pliancy as a human ear, allowing a better fit of the headphone on and over the ear, resulting in more accurate bass measurements. But measurements from the other, non-GRAS sites (Rtings and InnerFidelity) are still plenty useful. The trick is to only compare the measurements on those sites to the other measurements on that same site.

A much bigger problem is that often measurements of a particular headphone won't be available on any site, let alone on the site you prefer. The headphones that do get measured tend to be ones that are of interest to headphone hobbyists. The thousands of mass market and budget headphones produced each year tend to go unmeasured. Rtings is invaluable for not ignoring this market segment.

Measurements done less right



The above are competent and useful sources of information. On user forums, however, you'll encounter measurement graphs done by amateurs using inexpensive equipment.

As a point of comparison, here are Oratory1990's measurements for the AKG K371:

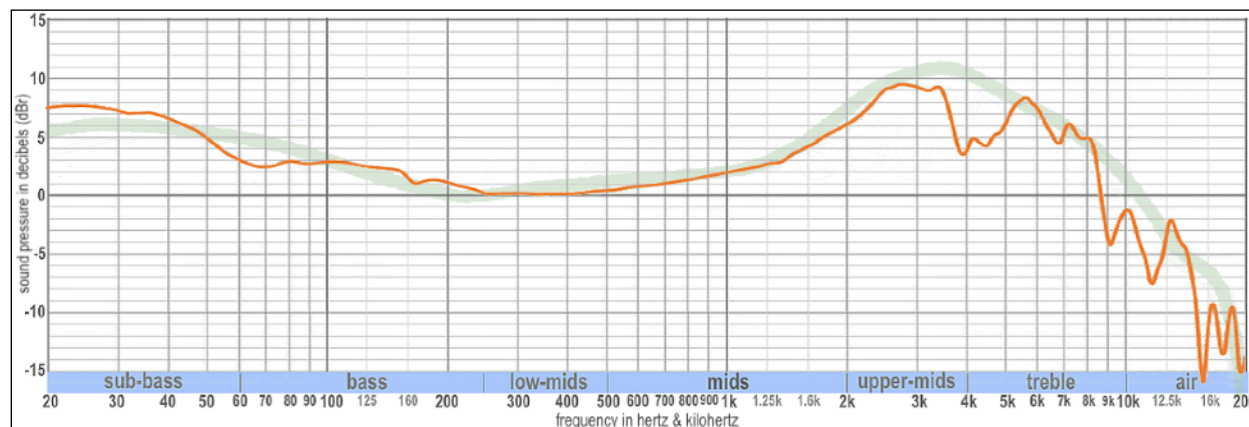


Fig. 13: Oratory1990 FR graph for AKG K371

This headphone was specifically designed to reproduce the Harman Target frequency response (pale green curve) as closely as possible. As we can see that (by headphone standards) it does so admirably. Now here is an amateur measurement of the K371 that I found by using Google:

The Skinny on Headphone Frequency Response Graphs

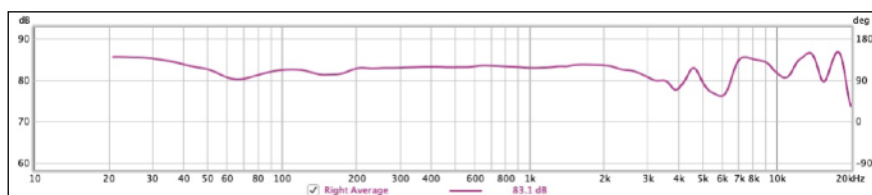


Fig. 14: MiniDSP EARS frequency response graph for AKG-K371

This was done using an inexpensive and popular tool called the MiniDSP EARS. We may scratch our heads as to how to relate the purple line in Fig. 14 to the red line in Fig. 13. They both claim to show the frequency response of the same model headphone. The first thing we need to know is that we're not seeing the actual data in Fig. 14. What we're seeing is a plot of the *difference* between the actual measurements and some unspecified target response curve. And target curves vary all over the proverbial map. The graphs we've been looking at previously are what's called raw or uncompensated. Fig. 14 is called a *compensated* FR graph. Yet it makes no mention of that fact, nor of what target it's compensated to.

Understanding compensated frequency response graphs

Let's start with a frequency response graph including measurements and a target response curve like the one by Oratory1990 of the AKG K371 in Fig. 13 above.

Now that you've dived down the frequency response rabbit hole, the odd and squiggly lines are starting to make sense. But for anyone else the graph is meaningless, not to mention off-putting, as well.

First, we'd have to explain that left to right means low-pitched to high-pitched sounds and top to bottom means loud to quiet sounds. Then we'd have to say that the red line shows how this particular headphone changes the loudness of the sounds being listened to, making sounds at some frequencies louder than they actually were and making other frequencies quieter than they actually were. Then we'd have to explain that the human brain actually expects some frequencies to be louder than others, which is represented by the green line. So we're supposed to pay attention to the differences between the green and red lines.

The green target response line is typically supposed to show how much physical loudness (sound pressure) it takes for our brains to perceive subjective equal loudness. We can make this a little easier to digest if we flatten out the green line and have the red line tag along for the ride. In other words imagine the green line is a string and we pull the string taut so the green line goes from wavy to straight. Fig. 15 is an Oratory graph showing both the raw and compensated versions of the same measurements. Notice that the relationship between the green and red lines is identical in both cases, but the green line has been flattened. I added up and down arrows to show where we'd have to press on the raw green Harman curve to morph it into the straight line of the compensated version.

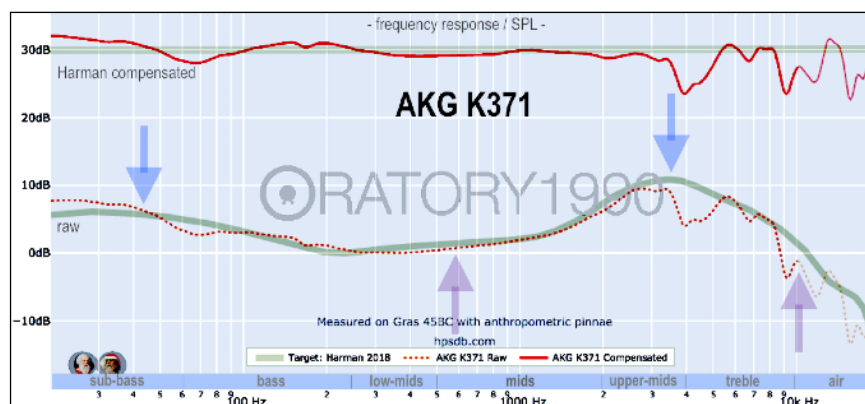


Fig. 15: AKG K371 graph morphed to straighten the Harman target curve

The Skinny on Headphone Frequency Response Graphs

Now let's return to my scathing critique on how compensated graphs are typically presented.

The MiniDSP EARS tool comes with two different target response curves. One, called HPN (headphone neutral), is somewhat similar to the diffuse field curve. The other, called HEQ (headphone equalization), is somewhat similar to the Harman Over-Ear Target. Given the relative flatness of Fig. 14 we can make an educated guess that the HEQ curve was used. But that information is rarely provided. Here I've over-laid the MiniDSP EARS graph on top of the Oratory1990 GRAS 43AG with Harman 2018 compensation graph from Fig. 15:

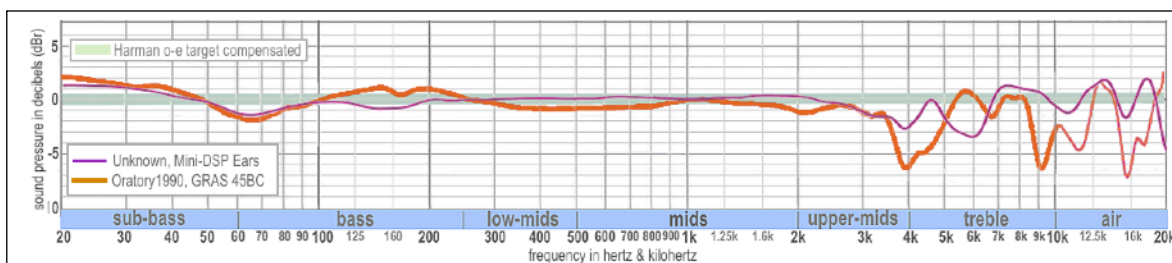


Fig. 16: compensated graph comparison: MiniDSP EARS versus GRAS 43AG for AKG-K371

As we'll see below, this is by no means a terrible match-up. The majority of the discrepancies start around 2 kilohertz, which is typical and corresponds with similar variations between human ear responses. But we still don't know how closely the MiniDSP EARS HEQ target response curve matches the Harman target curve. So the graphs might be even more closely aligned if both measurement rig results were compensated with the same target response curve.

Here's another MiniDSP EARS graph of the K371:

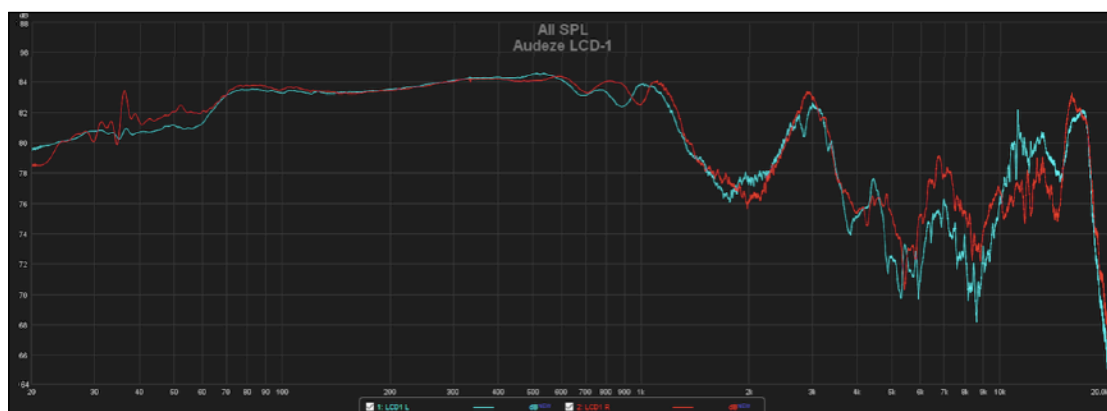


Fig. 17: [Metal571](#) MiniDSP EARS FR graph for [AKG-K371](#)

This one was done by YouTube headphone reviewer Metal571. Metal is a careful and thoughtful guy. He makes sure to tell us that the MiniDSP EARS graphs are seriously limited (especially in the region between 4 and 5 kHz). He also makes sure to tell us that he's using the HPN compensation. (Don't be confused by the dramatic black background. That's purely cosmetic.) The overall downward slope of the measurement curve is due to the increasing difference between the diffuse field ear gain curve (HPN) and the Harman Target curve (HEQ) in the ear amplification area above 1 kHz.

But another major difference between the two MiniDSP EARS graphs, Fig. 14 and Fig. 17, is the vertical scale. In Fig. 14 each horizontal reference line is 10 dB above or below the next. In Fig. 17 they're only 2 dB apart. In consequence, Fig. 14 tends to downplay the rises and falls of the measurement curve. Fig. 17 highly exaggerates them.

The Skinny on Headphone Frequency Response Graphs

This scaling effect can be used to great effect to put a headphone in a good light. The headphone forum Head-Fi.org occasionally produces FR graphs of headphones released by sponsoring companies. The Head-Fi measurement equipment used here is the same as Oratory1990's and they know how to use it.

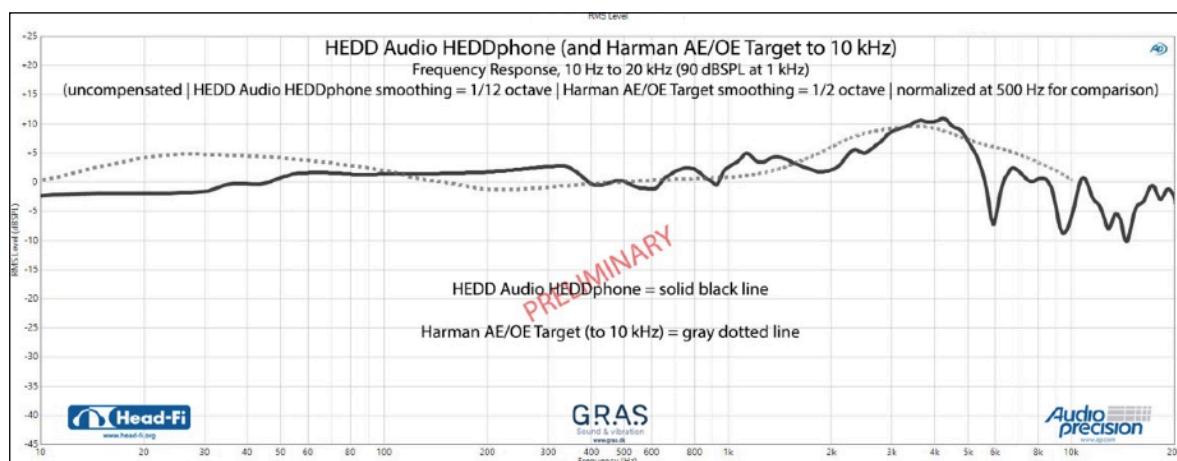


Fig. 18a: Head-fi measurements for the HEDDphone

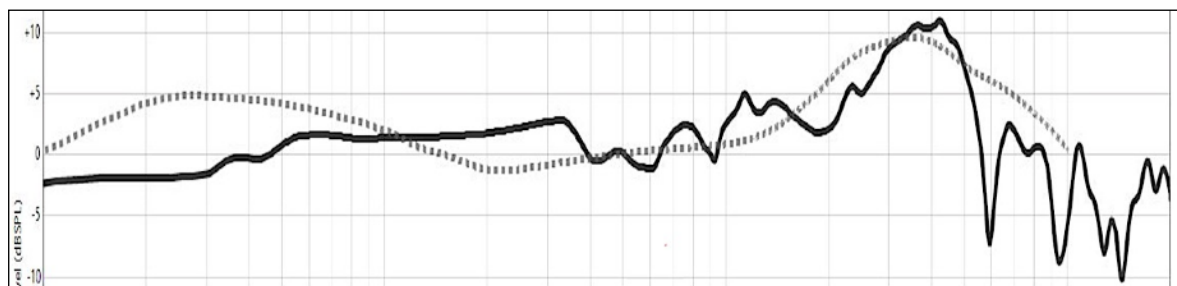


Fig. 18b: Fig. 17a vertically re-scaled

The graph in Fig. 18a is directly from the Head-Fi website. This particular headphone sells for \$1900US as I write this. The solid black line are the raw measurement data. The dotted black line is the Harman Target. Notice how shallow the Harman curve is compared to what we're used to from the green line version in Oratory1990 graphs. For the graph in Fig. 18b I've simply taken the original graph and stretched it vertically so the horizontal loudness reference lines are now a more typical distance apart. The dotted Harman Target curve now looks familiar. Same data, dramatically different presentation.

The small print in the Head-Fi graph legend says the data are presented with 1/12th octave smoothing. This was frequently done. It simply acknowledged the limitations of even the best measuring equipment and measuring techniques of the not too distant past. Applying it to GRAS results is less appropriate. We may not be able to identify small variations, but we may be able to perceive them as something like graininess. From time to time you'll see FR graphs with even smoother measurement curves, even as extreme as 1/3rd octave. When the graph line looks smooth and simple, be suspicious that more aggressive smoothing has been done.

In short, it's the wild west out there in the land of headphone FR measurements and their presentation. We have to be wary of:

The Skinny on Headphone Frequency Response Graphs

- Amateur-grade test equipment
- Unspecified compensation curves applied to the data
- Loudness scale compression applied to the data
- Excessive smoothing applied to the data.

Do good measurements agree?



Even if we choose to focus our attention on the quality measurement sites for frequency response graphs that doesn't mean their results are completely interchangeable.

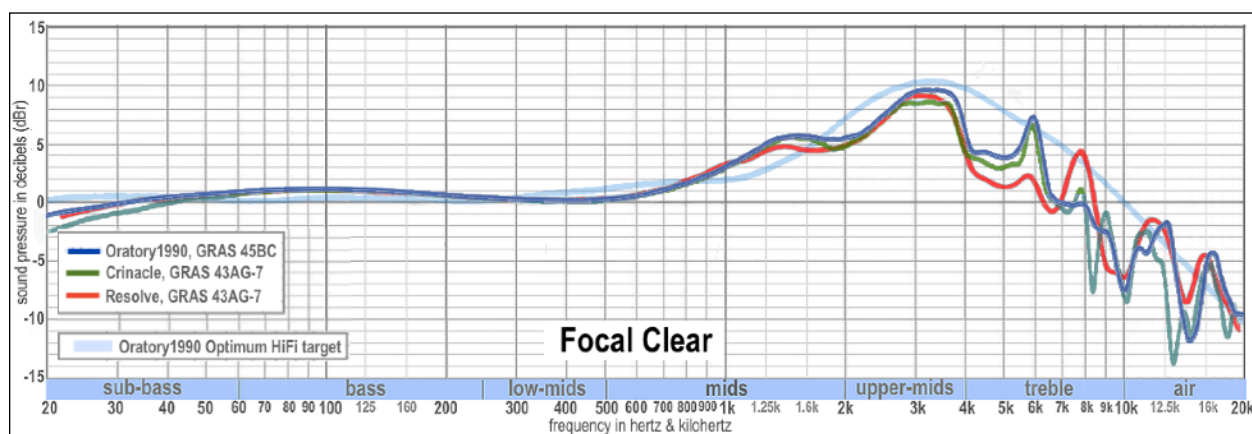


Fig. 19: three-way comparison of GRAS measurements of the Focal Clear headphone

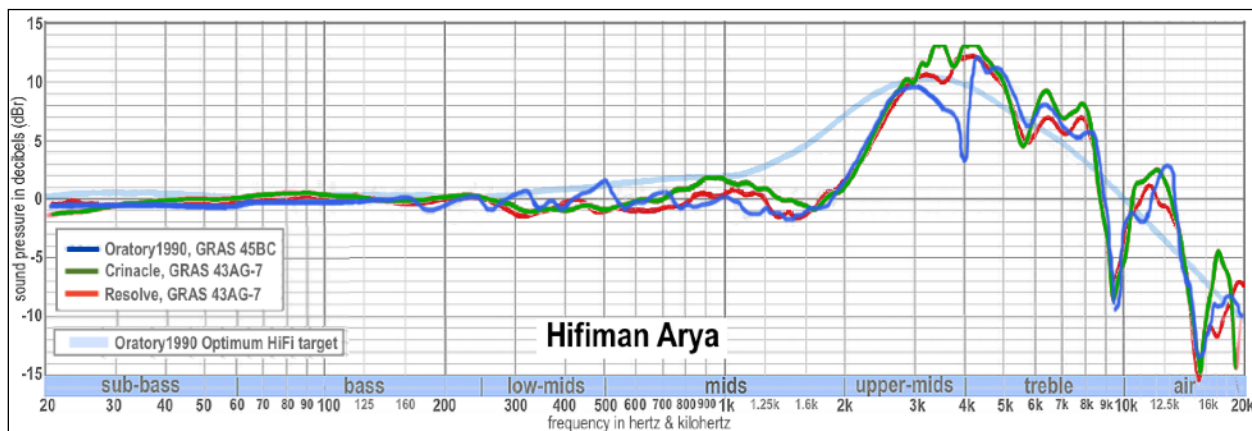


Fig. 20: three-way comparison of GRAS measurements of the Hifiman Arya headphone

Fig. 19, which compares measurements of the Focal Clear taken with three different GRAS rigs, shows remarkable overlap from 60 to 1250 hertz. There will often be disagreement in the very lowest frequencies. This is due to the difficulties getting the ear pads to rest completely on the side of the measurement head surrounding the artificial ear with no air gaps. Equally, there are typically disagreements in the high frequencies above 2 kHz. Fig. 20 of the Hifiman Arya — the first planar magnetic we've looked at — is a bit rough in the middle frequencies (I suspect

The Skinny on Headphone Frequency Response Graphs

unit variation), but overall quite consistent up to 3 kilohertz and again remarkably consistent past 4 kilohertz.

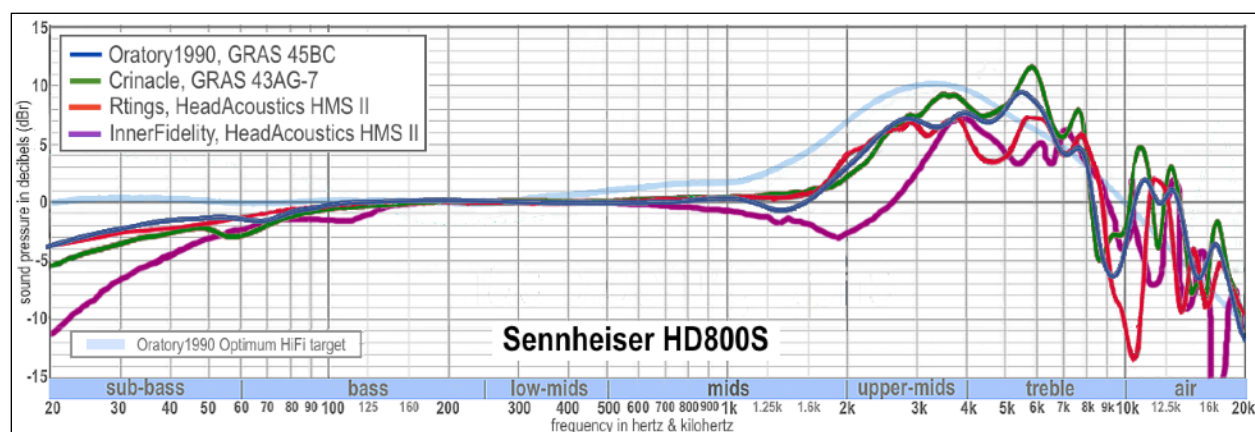


Fig. 21: four-way comparison of raw measurements of the Sennheiser HD 800S headphone

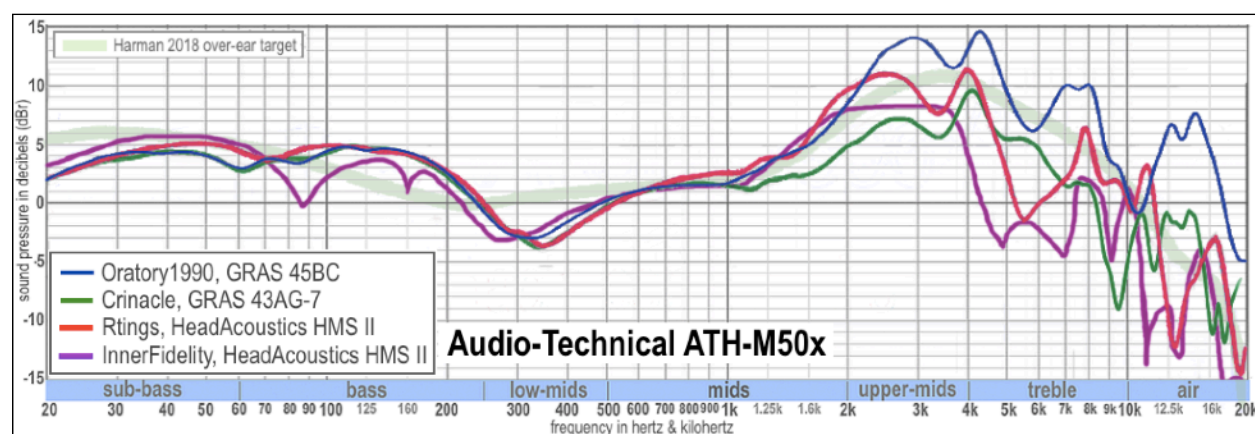


Fig. 22: four-way comparison of measurements of the Audio-Technica ATH-M50x

With Figs. 21 and 22 we're adding the two HeadAcoustics measurement rig sites (Rtings and InnerFidelity) into the mix. If these graphs are getting more and more confusing to the eye, that's pretty much the point.

In Fig. 21 we see the nice overlap in the middle as in Fig. 19. The two GRAS lines show the same so-so level of agreement we see above, while the two HeadAcoustics lines diverge more. The InnerFidelity line seems if it's from an entirely different headphone. But in Fig. 22 we're seeing differences *everywhere*. I have no idea what lies behind the M50x discrepancies, but fortunately Fig. 21 seems more typical of the kind of divergences we can expect between the measurements from these sites.

Another factor, beyond measuring equipment, is unit-to-unit product variation of the headphones themselves. Suppose you buy three units of exactly the same headphone that you find right beside each other on a store's shelves. Suppose you then send them to one of the review sites asking for measurements. Experience shows there would be measurable differences between the three units.

So when we look at Fig. 22, for example, we have no idea how much of the discrepancies between the measurement curves is due to the test equipment being used and how much is due to unit-to-unit differences between the individual M50x units that were measured. Here are the measurements for multiple units of the Sennheiser HD 650 as measured by the same technician (Oratory1990) on the same test fixture:

The Skinny on Headphone Frequency Response Graphs

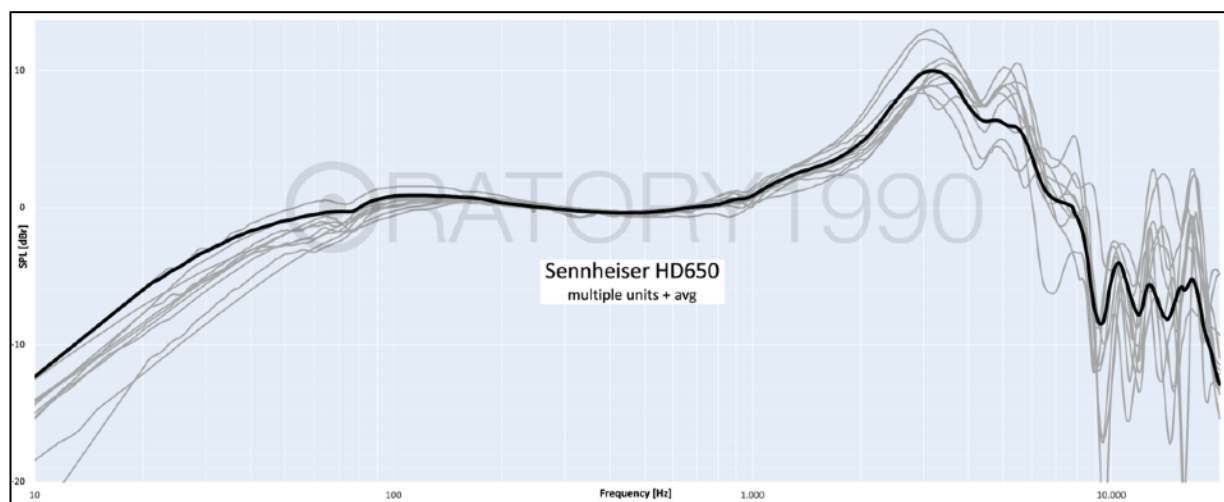


Fig. 23: different units of same headphone model (Sennheiser HD 650)

Sennheiser has a long-standing reputation for minimal unit-to-unit variation, and the HD 650 is one of its more expensive models. But each grey curve is the measurement for a different unit of the same headphone. (Note, however, that I don't know whether or not the different units had different amounts of ear pad wear.)

Yet another factor is what we call *silent revisions*. Manufacturers occasionally change internal components or assembly procedures over time in the production of a particular headphone model. A common scenario is for a manufacturer to experience a significant number of returns after releasing a new model headphone. The manufacturer will investigate what the failure was then fix the problem, if possible, in the production of that model going forward. (It's always wise to hold off for several months on purchasing an exciting but brand new product.)

Silent revisions, however, can occur at any point in a product's life cycle. Just one cause is a component supplier going out of business. Planar magnetic driver technology is still evolving as several firms continue to push the performance envelope with new materials and manufacturing techniques. Some of the more boutique of these companies make changes to existing models to incorporate new developments without changing the model designation. So any difference between Oratory1990 vs Crinacle vs Rtings, etc. measurements, for example, may be due to actual differences in the production run of the headphone model being tested. The on-going saga of the Hifiman Sundara as I write this in early 2021 with two significant revisions to the ear pads in the last two years to increase durability is a dramatic case in point. Yet another factor — the ear pads on some headphones are made of materials with desirable acoustic properties but that compress and degrade significantly over time:

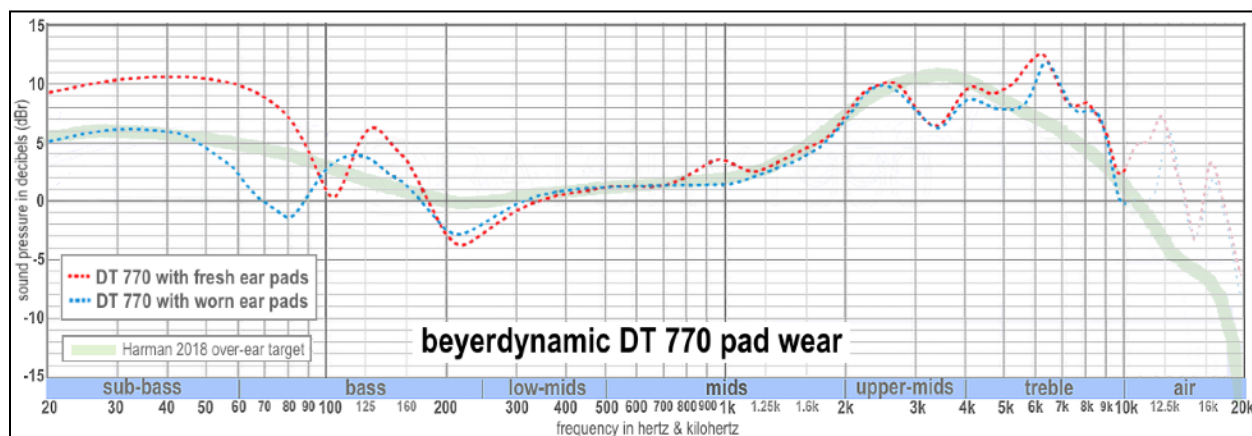


Fig. 24: effect of pad wear on the beyerdynamic DT 770 (Oratory1990 data)

The Skinny on Headphone Frequency Response Graphs

Teasing out which factors are at play in the discrepancies between graphs from different sources of the same headphone borders on impossible. So again, when details are important all we can do is restrict comparisons to headphone measurements from the same source.



And finally there is a very large fudge factor that is beyond both the manufacturer's and the measurement technician's control. I mentioned above that any competent technician will measure a given over-ear headphone multiple times in slightly different positions that result from the inner cavity of an ear pad necessarily being large enough to fit the largest ears but with room to spare for the majority of ears. Add to that different head sizes typically cause headphones to sit with more or less clamp force. This compresses the ear pads to a greater or lesser degree changing the cavity size and shape, which in turn changes frequency response. Here is a typical spread in the high frequencies from an Oratory1990 graph posted to illustrate the concept:

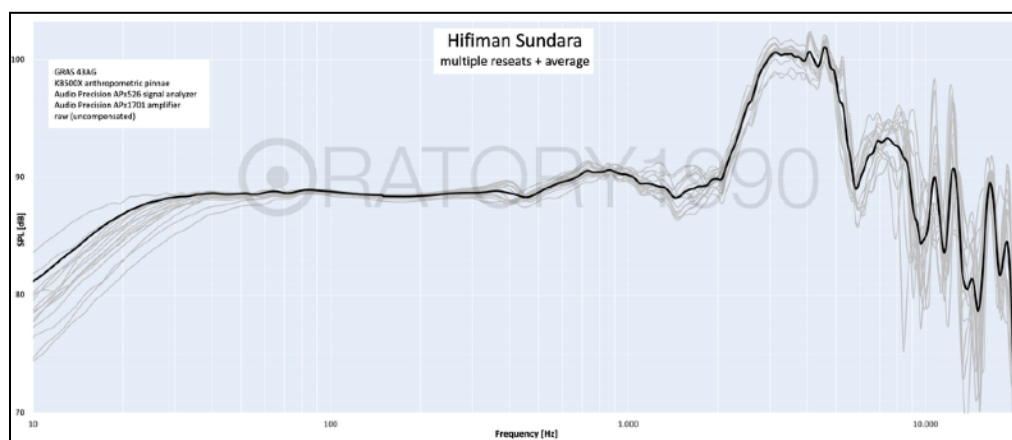


Fig. 25: Oratory1990 graph of the frequency response of a Hifiman Sundara showing multiple re-seats

The faint grey tangle is composed of at least 15 different measurement lines, each with the headphone moved slightly left, right, up or down around the ears of the measurement head. The larger spread in bass and treble versus mid range is typical. Many measurement sites, like Rtings and InnerFidelity show the measurement trial lines as well as the resulting average. While Oratory labels the black line in Fig. 25 as an average, its location near the top of the sub-bass on the far left suggests something else may be involved. Indeed, a nice thing about presenting the re-seating trials is in showing where a reduction of an unpleasant-sounding loudness or the increase in an unpleasant-sounding quietness might be achieved with an adjustment of the headphone on one's head.

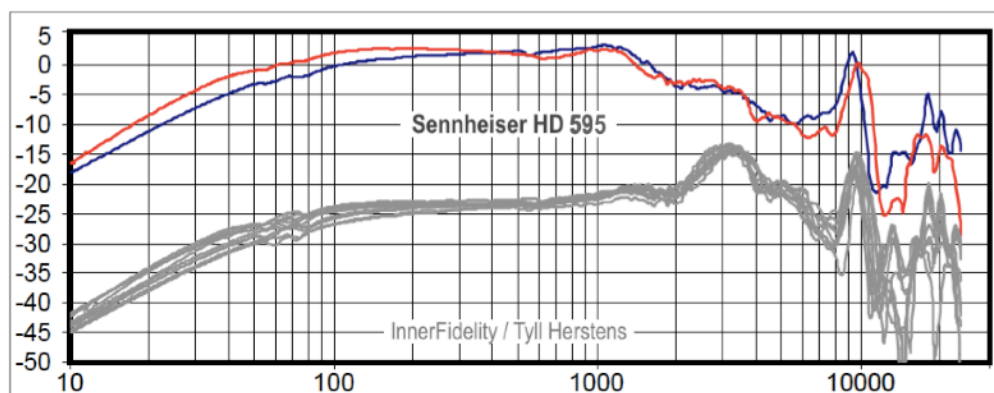


Fig. 26: InnerFidelity graph of the Sennheiser HD 595 showing multiple re-seats plus left and right ear responses

The Skinny on Headphone Frequency Response Graphs

The heart of any headphone are the (typically) two drivers/transducers, one in each ear cup. As mentioned above, just as there is unit variation of frequency response from one unit to the next of the same headphone, there is also variation between the two drivers in a single headphone. Headphone manufacturers produce or order large numbers of the drivers for a given model. They then measure and sort them to create pairs as closely matched as possible. But perfection is clearly not an option.

In Fig. 26 I also included the red and blue lines. The red line is an average of the left ear measurements and the blue, the right ear. The amount of difference between the two tends to vary from manufacturer to manufacturer. Typically, this is on the order of one or two decibels across the frequency range, like we see in Fig. 23, for the premium models of the well-established, large market share companies. But in other cases three or more decibel variance between units is not uncommon, as we see in the budget open-back HD 595 in Fig. 26.

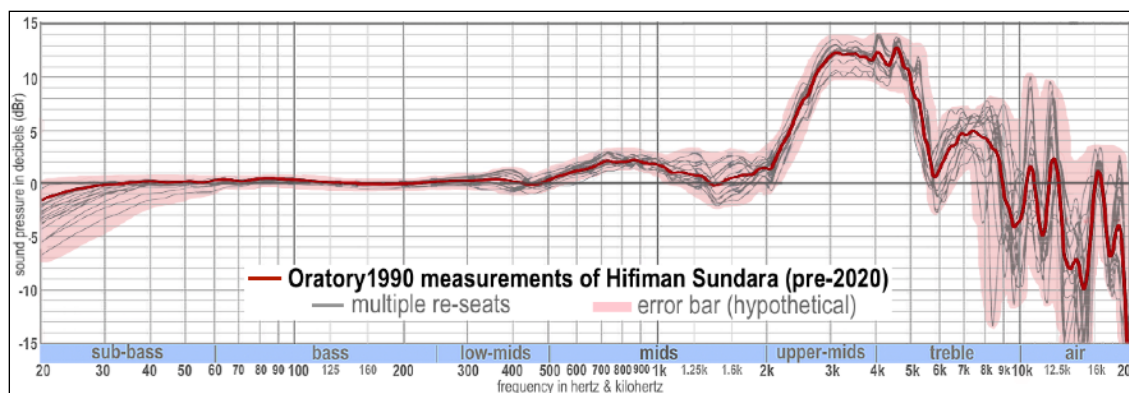


Fig. 27: Hifiman Sundara graph from Fig. 25 showing re-seating trials and (fictitious) error band

So — do good measurements agree? Nope. The graphs and data plots in science research papers include error bars to show visually how much variance or room for error there is in each measurement. The evidence presented in this section suggest sizeable error bars (or more practically, an error band) are missing from the neat, tidy, precise-looking graphs we often see posted in great numbers on-line. When even a GRAS 45 and a GRAS 43, both fitted with the expensive anthropomorphic ear analogue option, fail to quite agree with each other — when unit variation and silent revisions on the manufacturer's end play equal or greater havoc — when the headphones give different responses simply due to play in positioning on the listener's head — we need to have some idea what fudge factor to apply. Fig. 27 shows what a hypothetical headphone frequency response graph with an error band might look like. Of course, error band incorporation isn't going to happen. But when we see multiple re-seats shown, it's a great stand-in for what the error band might look like.

Are good measurements useful?

Despite all the above, good measurements are far from useless. I would look at GRAS measurements done by a skilled like those highlighted above as most likely coming closest to an accurate portrayal of the average-ear frequency response of a given headphone. But beyond that, quality measurement graphs still serve to answer questions like: "does this headphone have a reasonably accurate frequency response?" "Does this bass elevation on this headphone bleed well into the mid-range?" "Is the treble on this headphone likely to stab at my eardrums?" Questions like these are the proverbial meat and potatoes of endless forum arguments and endless buying decision deliberations. I have four headphones that have GRAS measurements available on-line. Using those measurements to create EQs that correct them to the same neutral reference target (Fig. 6), they all sound very nearly identical in frequency response. Switching from one to the other they eerily sound so similar I have to pay close attention to other sound qualities like sound stage and level of detail to tell them apart.

The Skinny on Headphone Frequency Response Graphs

A given measurement rig may not be completely accurate in delivering the correct frequency response for the hypothetical anatomically average human ear. But since few of us sport a pair of those average ears, a particular measurement rig may actually give a more realistic result for your hearing than a theoretically more correct rig does. One point of variance between sites is the sub-bass. But this is due to the intrinsic problem that ear cups can be seated in a fairly wide range of positions on an adult head. Presumably each person will tend to optimize positioning. From there, the range of frequencies from about 80 to 1200 hertz tends to be pretty consistent between the better measurement rigs. But higher frequencies start to diverge again, paralleling the wide variability in human ear anatomy.

A critical aspect of this are ear canal resonances. Every tube has a natural resonance frequency. For the ear canal, depending on its length, cross-section shape, friction, etc., this frequency is *on average* 2.5 kilohertz. But multiples of this frequency are also resonant:

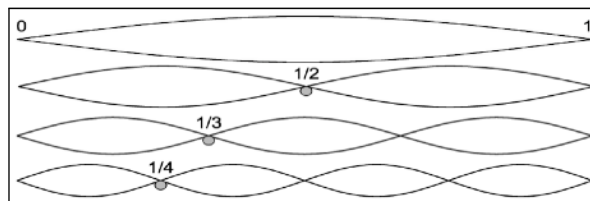


Fig. 28: first four vibrational modes

So that yields the series 2.5, 5, 7.5, 10, 12.5 and 15 kilohertz within the range of human hearing. But for someone whose ear canal has a resonant frequency of 2.7 kHz, the series becomes 2.7, 5.4, 8.1, 10.8, 13.5 and 16.2 kilohertz. The upshot is each of these frequencies gets a signal boost courtesy of the ear canal. If your particular ear canals generate a signal boost at 8 kHz and the headphone you're listening to has an intrinsic loudness spike at 8 kilohertz, then whenever the music or speech you're listening to peaks at 8 kilohertz you've got a problem. The straight tube ear canal of a HeadAcoustics rig may not match the hypothetical average ear canal — which includes both male and female ear measurements — but it may be a closer match to one or both of your own ear canals. The same basic idea applies to variance in outer ear shape and even to the boundary gain imparted by the human head and torso.

Unfortunately, as we've seen, other variables, such as unit variation and ear cup seating, are likely to obscure any match between your ears and a particular measurement rig. Whether researching a new purchase or creating an EQ curve for a headphone you own, you cannot count on any given measurement being truly accurate to a particular unit of a given model of headphone at any specific frequency. But understanding how to weight all the variables we've discussed will take you a long way to bridging the gap.



And that's it. You're now far more knowledgeable about headphone frequency response graphs than 90% of the Internet's self-appointed expert forum commentators. By the way, in this unit we only looked at measurement graphs for over-ear headphones. But most concepts apply equally to the separate world of in-ear monitors (also known as earphones). The main difference is that a different set of target response curves are typically used for in-ears. Also, all the headphones we looked at, excepting the two Hifimans and the HEDDphone, use dynamic (cone) drivers. Again, all concepts apply to planar, electrostatic and, yes, even AMT driver headphones. But, especially with planars, the amount of unit variation, plus the frequency of silent revisions, makes it difficult to treat any individual unit's measurement as representative of a given model.

In this episode we looked at what the frequency response of headphones actually looks like when properly presented, why they sometimes don't and how much they can be trusted even when no trickiness is involved.

So now be sure to proceed on to the next exciting instalment in my must-read Headphone Essentials series: [HE5: The quest for a reference headphone](#). In that episode we'll delve into

The Skinny on Headphone Frequency Response Graphs

what headphones *ought* to sound like but don't. We'll do this while exploring the mysterious world of headphone target response curves that we used but didn't explain in this unit.

Appendix A: measurement systems



Here's an old, fuzzy, but still relevant walk-through of the headphone frequency response measurement process in two parts. This is Tyll Herstens of InnerFidelity using a HeadAcoustic measurement head hooked up to an Audio Precision signal analyzer: [Part 1](#), [Part 2](#).

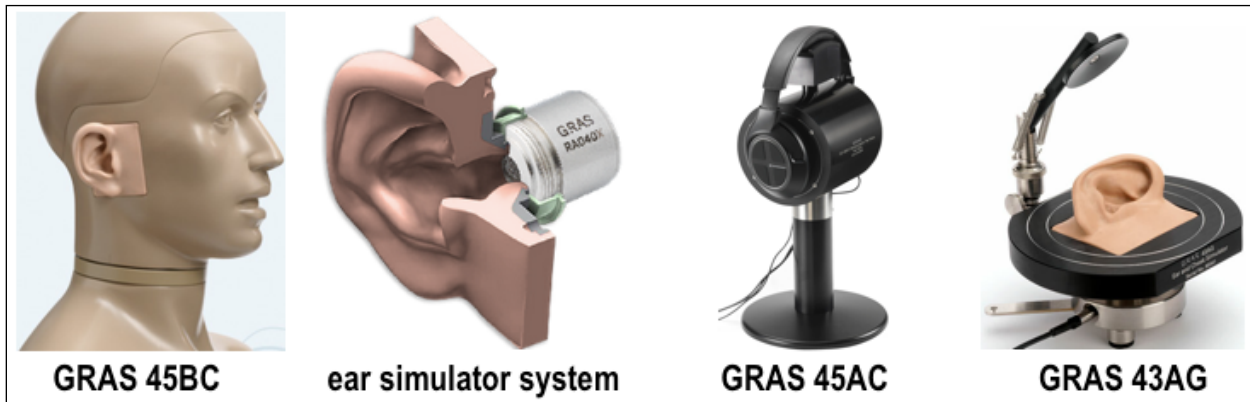


Fig. A1: current GRAS measurement systems

The industry standard relevant to headphone frequency response measurements is IEC60318-4. At this time GRAS Acoustics is the go-to source for equipment meeting this spec. An ideal, price-no-object GRAS 45BC set-up includes a mannikin-like head and torso simulator with very precisely moulded, flexible human ear replicas plus a microphone capsule for each ear that precisely replicates the acoustic impedance of the human eardrum. The 45CA forgoes the human-like head and torso. The GRAS 43AG is a minimum system incorporating one each of the same human outer ear (pinna) simulator, RA0402, and eardrum simulator, KB5000, that are used in both GRAS 45 systems.



Fig. A2: two alternatives

The older, no longer in production, measurement system used by InnerFidelity and Rtings is the HeadAcoustics HMS II, which also conforms to IEC60318-4, but nevertheless produced somewhat different results, particularly above 2 kilohertz. Its hard plastic outer ear also made

The Skinny on Headphone Frequency Response Graphs

getting a good seal, and therefore accurate low bass measurements, problematic on some headphones.

The inexpensive, non-standards-compliant, MiniDSP EARS is owned and used by many technically inclined headphone enthusiasts — especially those in the sub-hobby of headphone modding. It's results look to be surprisingly repeatable and not *that* far off the IEC60318-4 standard. What's missing for our needs is any way to compare EARS results with GRAS results.



Fig. A3: very recent addition to the field

At the opposite extreme is the new and ultra-expensive B&K 5128 system. Its selling point is that it produces results accurate to 20 kilohertz (IEC60318-7), while the GRAS system only claims accuracy out to 10 kilohertz (IEC60318-4). This is a moot point, given that each individual's high frequency ear response differs so dramatically from that of the IEC standard ear, that seeing how the standard ear measures in that range is of little practical significance. But even below 10 kilohertz the 5128's results differ so greatly from GRAS results that they will need their own neutrality reference curve. Anecdotally, GRAS results simply come closer to most people's subjective experience than the 5128 results do. Most obviously, relatively few people have a resonance spike at 8 kilohertz.

Delving into this a bit further, here is a comparison of the same headphone measured on both the GRAS 45BC and the B&K 5128:

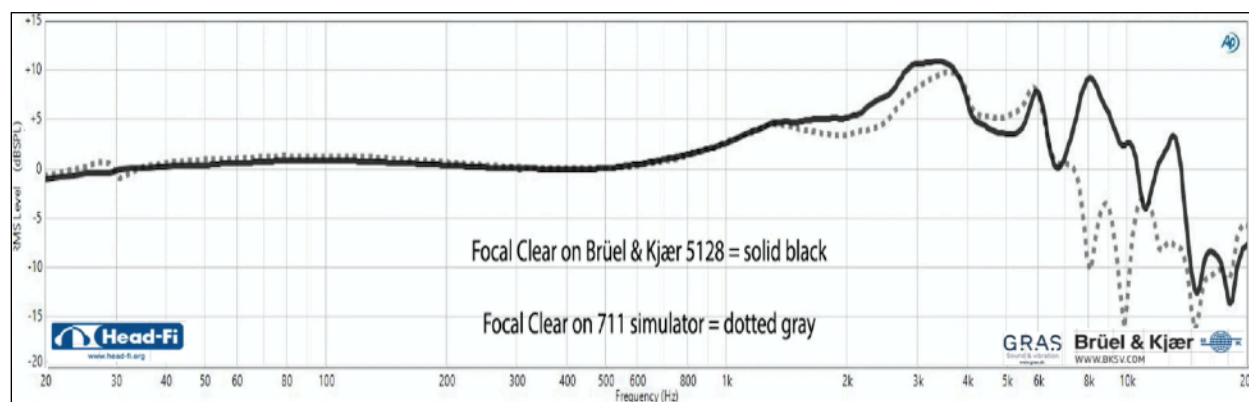


Fig. A4: Focal Clear (open-back dynamic driver) measured on 5128 (solid) and GRAS 45 (dotted)

(Because Fig. A4 is a Head-Fi graph, I vertically stretched it to a more standard ratio.) Below, are six more Head-Fi (Jude Mansilla) 5128 measurements, this time compared to Oratory1990 GRAS 45. Both are extremely competent technicians, as evidenced by the near-identical bass and mid-range regions for most of the headphones:

The Skinny on Headphone Frequency Response Graphs

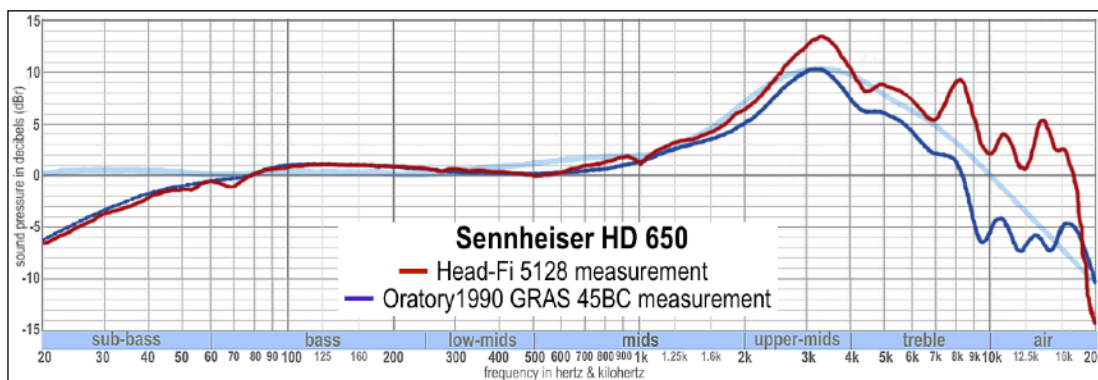


Fig. A5: Sennheiser HD 650 (open-back dynamic driver)

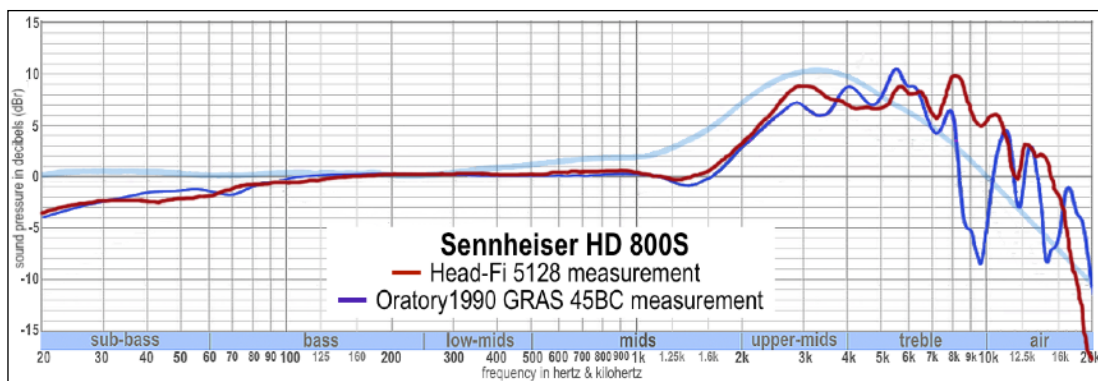


Fig. A6: Sennheiser HD 800S (open-back dynamic driver)

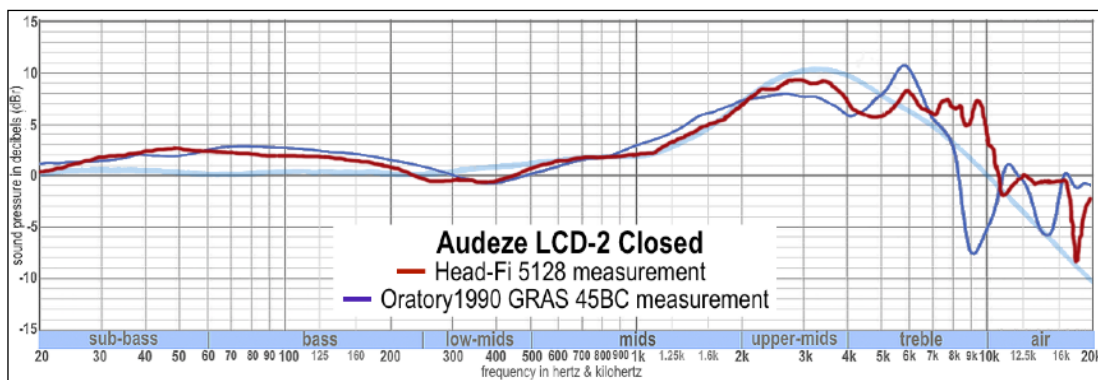


Fig. A7: Audeze LCD-2 Closed (closed-back planar magnetic driver)

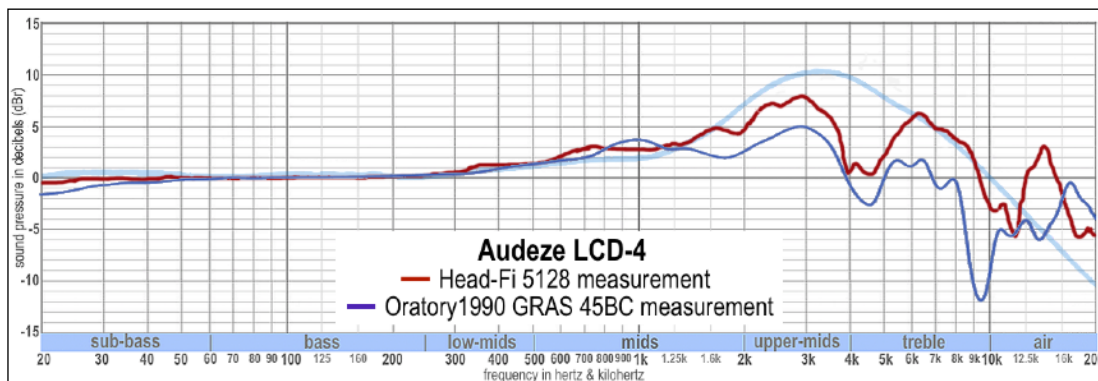


Fig. A8: Audeze LCD-4 (open-back planar magnetic driver)

The Skinny on Headphone Frequency Response Graphs

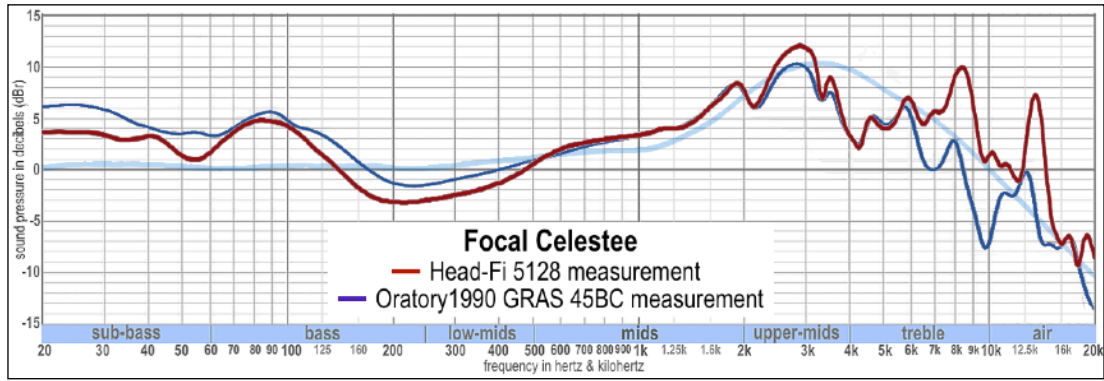


Fig. A9: Focal Celestee (closed-back dynamic driver)

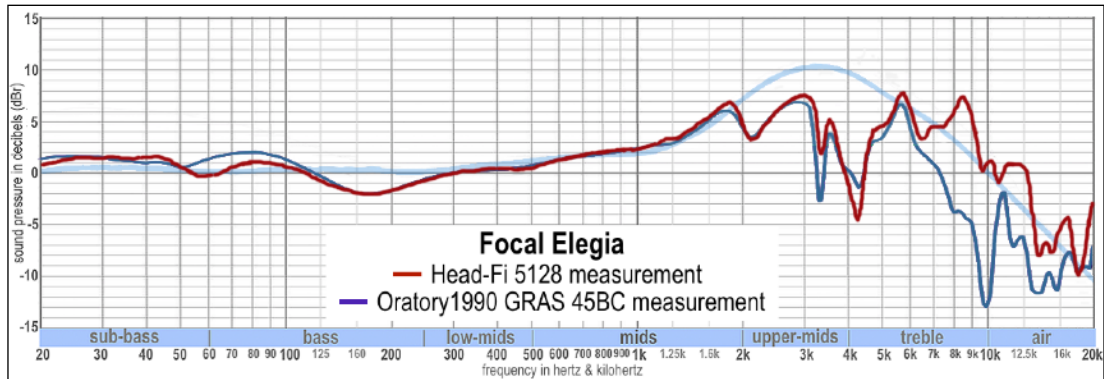


Fig. A10: Focal Elegia (closed-back dynamic driver)

We see the repeated elevation of the red line over the blue line at 8 kilohertz. Yet in each case I'm not aware that any anecdotal evidence for such an artifact exists. There is also a fairly consistent 5128 elevation between 2 and 3 kilohertz above the GRAS measurements.

The research behind the 5128 was exhaustive and heroic. Its focus on ear canal morphology seems to be especially important for in-ear headphone measurement. I can't speak to in-ears. And undoubtedly, more information will surface in time that may well totally change my thinking. But at this time for on-ear and over-ear I'm not seeing any reason to abandon the large and useful body of GRAS measurements we already have access to.

Dig deeper into the 5128: YouTube [search results](#) for Brüel & Kjær 5128.