

Excerpts from *Good References*, by Brian Knave, in *Electronic Musician*

JB: Read for an understanding of the following:

Close-Field or Near-Field, Yamaha NS-10, Minimizing Room Effects, Characteristics of Home Stereo Speakers, Frequency Response (on and off-axis), Frequency Response Specifier, Sweet Spot, Transient Response, Distortion, THD, Woofer Tweeter Midrange Drivers, Speaker Timing Problems, Passive and Active, Soundstage, Sound Pressure Level or SPL, "Practicing with Monitors"

PRE ROLL

Speakers used in recording studios are called monitors and generally fall into two categories: main monitors and compact or close-field reference monitors. Mains, as they are called, are mostly found in the control rooms of large commercial studios, often flush-mounted in a "false" wall (called a soffit); close-field reference monitors are freestanding and usually sit atop the console bridge or on stands directly behind the console.

Most personal studios don't have the space or funds for main monitors, so this article will focus on the compact reference monitor — a relatively recent studio tool. The first "compact" monitor to see widespread use in recording studios was the JBL 4311, a 3-way design introduced in the late 1960s. The 4311 was quite large, however (it had a 12-inch woofer, a 5-inch midrange speaker, and a 1.4-inch tweeter), and today would qualify more as a mid-field monitor.

As engineers increasingly realized the importance of hearing how their mixes sounded on car and television speakers, smaller reference monitors gained in popularity. One of the earliest favorites (around the mid-1970s) was the Auratone "cube," which had a single 5-inch speaker.

Car and home-stereo speakers kept improving, of course, so engineers were always on the lookout for better close-fields. One compact model that caught on big was the Yamaha NS-10M. A bookshelf-type speaker introduced in 1978 for home use, the NS-10M soon became a familiar sight in commercial studios, and it remains popular — or at least ubiquitous — to this day.

ENVIRONMENTAL ISSUES

Curiously, because close-field reference monitors have become increasingly accurate during the course of time, the original rationale for using them — to generate a good indication of how mixes will translate to low-cost car and home-stereo speakers — has waned. But there are also other good reasons close-field monitors have become all but indispensable in music production. For one, professional mix engineers are typically hired on a project-by-project basis, which means they may end up in a different studio from one day to the next. Close-field monitors, because they are portable enough to be carted from studio to studio, make

for an ideal solution and guarantee, at the minimum, some level of sonic consistency, regardless of the room.

But don't the monitors sound different in different rooms? To a degree, they do. But another advantage of close-field monitors is that they can partially *mitigate the effect of the room on what you hear*. As their name makes clear, they are meant to be used in the “*near field*,” typically about three feet from the engineer's ears. At that distance, assuming the monitors are well positioned and used correctly, the sound can pass to the ears largely unaffected by surface reflections (from the walls, ceiling, console, and so forth) and the various sonic ills they can wreak.

For the same reason, close-field monitoring is also a good solution for the personal studio, where sonic anomalies are the norm. As engineer, consultant, and all-around acoustics wizard Bob Hodas has so well demonstrated, however, it's foolhardy to think close-field monitors entirely spare you from the effects of room acoustics. “near-field monitors can be accurate,” explains Hodas, “only if care is taken in the placement of the speakers and room issues are not ignored.” (Find more information at www.bobhodas.com/pub1.html.)

DIFFERENT WORLDS

A common misconception among those new to music production is that home-stereo speakers are adequate for monitoring. That is, in fact, not the case. The problem is one of purpose: whereas manufacturers design reference monitors to reproduce signals accurately, home-stereo speakers are specifically designed to make recordings sound “better.” Typically, that perceived improvement is accomplished by boosting low and high frequencies. Although it may sound like an enhancement to the average listener, such “hype” is really a move away from accuracy.

Home-stereo speakers may also be engineered to de-emphasize midrange frequencies so as to mask problems in this critical range. That makes it difficult to hear what's going on in the midrange, which can tempt mixers to overcompensate with EQ. It can also lead to fatigue because the ear must strain to hear the mids.

Yet another reason home-stereo speakers are inappropriate for monitoring is that they are meant to be listened to in the far field, where much of the sound is reflected. But as we've seen, close-field monitors are designed to be used in the near field, in order to help minimize the effects of room acoustics. Of course, it's important not to sit too close to near fields. Rather, they should be positioned far enough back to allow the sound from the speakers to blend into an apparent point source and stereo soundstage. As you move in closer than three feet or so, the sound from each speaker becomes distinguishable separately, which is not what you want.

ELUSIVE BULL'S-EYE

Everyone can agree that reference monitors are meant to reproduce signals accurately. But what is accuracy? For our purposes, there are three objective tests that can be performed to help quantify accuracy in reference monitors. The tests measure frequency response, transient or impulse response, and lastly, distortion.

Frequency response is a measure of the changes in output level that occur as a monitor is fed a full spectrum of constant-level input frequencies. The output levels can be plotted as a line on a graph — called a frequency response plot — in relation to a nominal level represented as a median line typically marked 0 dB (see Fig. 2). The monitor is said to have a “*flat*” or linear frequency response when that line corresponds closely to the median line — that is, does not fluctuate much above or below from one frequency to the next.

When they are written out, frequency-response specifications first designate a frequency range, which is typically somewhere between 40 and 60 Hz on the low end and 18 to 22 kHz on the high end. To complete the specification, the frequency range is followed by a range *specifier*, which is a plus/minus figure indicating, in decibels, the range of output fluctuation. For example, the spec “50 Hz — 20 kHz (± 1 dB)” means that frequencies produced by the monitor between 50 Hz and 20 kHz will vary no more than 1 dB up or down (louder or quieter) from the input signal. (That spec would suggest a very flat monitor, by the way!) Note that the range specifier may also be expressed as two numbers, for example “+1/-2 dB,” which is useful when the response varies more one direction than the other.

Primary frequency-response measurements are made on-axis, that is, with the test mic directly facing the monitor, often at a distance of one meter. Also helpful are off-axis frequency response plots (measured with the mic at a 30-degree angle to the monitor, for example), which give an indication of how accurate the response will be — or how much it might change — as you reach for controls or gear located outside of the “sweet spot.” (The *sweet spot* is the ideal position to sit at in relation to the monitors; it is calculated by distance, angle, and listening.)

Transient or impulse response is a measure of the speaker's ability to reproduce the fast rise of a transient and the time it takes for the speaker to settle or stop moving after reproduction of the transient. Obviously, the first characteristic is critical to accurate reproduction of instrument dynamics and transients (such as the attack of a drum hit or a string pluck). The second is important because a speaker that is still in motion from a previous waveform will mask the following waveform and thus muddle the sound.

Distortion refers to undesirable components of a signal, which is to say, anything added to the signal that was not there in the first place. For monitors it can be divided into two categories: harmonic distortion and intermodulation distortion (IM). Harmonic distortion is any distortion related in some way to the original input signal. It includes second- and third-harmonic distortion, total harmonic distortion (THD), and noise (which are the types most commonly measured; see Fig. 4), as well

as higher harmonic distortions (fifth, seventh, ninth, and so on). Intermodulation distortion is a form of “self-noise” that is generated by the speaker system in response to being excited by a dynamic, multifrequency signal; typically, it is more audible and more annoying than harmonic distortion.

Frequency response, impulse response, and distortion levels should all be taken into account to get an idea of a monitor's accuracy. However, frequency response is often the only measure mentioned in product literature and reviews, and even it gets short shrift on occasion. (In many instances, I have seen frequency specs given with no range specifier — and of course, without it the specification is meaningless). Few manufacturers provide an impulse response graph (even assuming they have measured impulse response), and often the only distortion specification given is “THD + noise.” In fact, the lack of established and agreed-upon standards for monitor (and for microphone) specifications — for both measuring them and reporting them — is a long-standing industry issue. Though it is true that specs don't tell the entire story, they are useful for corroborating what our ears tell us, and as such they can help educate us so that we can more exactly listen.

MIRROR IMAGE

Now that we've established the *raison d'être* of the close-field monitor, let's take a look at its anatomy. We'll start with the internal components and work our way outward to the enclosure. Understanding how monitors are put together will help you know what to look for when deciding which best suit your needs.

In a loudspeaker, the coil of wire is known as the voice coil. As the electric current (audio signal) fluctuates in the wire, it generates an oscillating magnetic field that pushes and pulls against the magnet, causing the voice coil and attached diaphragm (in this case, the speaker cone; see Fig. 5) to vibrate. In turn, the vibrating speaker cone agitates nearby air molecules, creating the sound waves that reach our ears.

DRIVING LESSONS

A loudspeaker's magnet, voice coil, and diaphragm form, collectively, an assembly called a driver. Close-field monitors usually contain either two or three drivers, and thus are designated 2-way or 3-way, respectively. Standard 2-way monitors contain a *woofer and tweeter*; standard 3-ways contain a woofer, a tweeter, and a *midrange* driver. The woofer, of course, reproduces lower frequencies and the tweeter, the higher frequencies.

Cones and domes are the two most common types of diaphragms used in monitor drivers. Most moving-coil tweeters use a small dome, typically measuring one inch in diameter. One advantage of a small dome is that it exhibits fast transient response and a wide dispersion pattern, both of which are critical to the reproduction of upper frequencies.

When monitors employ separate drivers, as 2-way and 3-way monitors do, the design is termed discrete. In discrete designs, the drivers are usually mounted on the front face of the enclosure as close together as possible, which helps the sound blend into a coherent point source at the sweet spot. Depending on the monitors, the sound can change dramatically as you move away from the sweet spot.

IT'S ABOUT TIME

Some companies, for example Tannoy, employ an alternative driver design in some of their monitors in which the tweeter is mounted in the center of the woofer cone. Though more expensive, this coaxial design is naturally more time coherent than discrete designs because the drivers are positioned on the same axis (as well as closer together). Indeed, the coaxial driver arrangement is one of the design elements (among others) that manufacturers have used to meet Ed Long's Time Align specification, mentioned before.

Before we can understand how time alignment can improve a monitor's accuracy, we must first understand the *timing problems* inherent in conventional monitor designs. Discrete loudspeakers cause minute delays that spread sounds out in time, resulting in lost detail and a blurred or smeared sound. Specifically, sound from the woofer is delayed more than sound from the tweeter. This problem has two main sources, one structural, the other electronic. In a discrete monitor with a flat-face enclosure, the woofer voice coil is naturally set back further than the tweeter voice coil because of the extra depth of the cone in relation to the dome. The tweeter is therefore closer to your ears, causing the high frequencies to arrive slightly ahead of the lows.

The problem is compounded by the *crossover*, an electronic circuit that splits the incoming signal into separate frequency bands and directs each band to the appropriate driver (more on crossovers momentarily). As it happens, crossovers also tend to delay low frequencies more than highs.

With his Time Align scheme, Long was the first to specify corrections for these problems, including physically lining up the drivers and adjusting driver and crossover delay parameters. When correctly implemented, Time Alignment ensures that the time relationships of the fundamentals and overtones of sounds are the same when they reach the listener as they were in the electrical signal at the input terminals of the monitor.

WHEN I CROSS OVER

As mentioned, the crossover's job is to divide the incoming signal into separate bands and then send each band to the appropriate driver. In inexpensive monitors, this is typically accomplished using simple lowpass and highpass filters that split the signal coming from the power amp. This is called a *passive* crossover. In more sophisticated systems, an active crossover splits the line-level signal before it gets to

the power amp. This requires each driver to have its own power amp, and is called biamping in 2-way monitor, triamping in a 3-way, and so on.

Typically, monitors that have active crossovers incorporate internal power amps. These are called powered monitors. The terms active and powered, though often used interchangeably, actually refer to different things: active refers to the crossover, and powered to the fact that the amplifiers are part of the package. In other words, although active monitors are almost always powered, not all powered monitors are active.

In addition to giving a more exacting crossover performance, powered, active monitors offer other advantages over passive designs. Perhaps most importantly, because the amps and electronics are specifically designed to match the drivers and enclosure, powered monitors eliminate the guesswork and the potential pitfalls of matching an external amp to your monitors. This means *reduced risk of blowing the drivers* and virtually no risk of overtaxing the amps. In addition, the internal wiring is much shorter, which cuts down on frequency loss, noise induction, and other gremlins attributable to long cable runs. The upshot is that a power, active system provides a more reliable reference — no matter where you take the monitors, you can be sure the only variable is room acoustics.

BOX SET

The enclosure is a critical part of any reference monitor design. Compact monitors present a particular challenge to designers because diminutive enclosures do not support low frequencies well. For many small monitors, the lowest practical frequency is around 60 Hz. However, certain techniques allow manufacturers to extend the low-frequency response of their boxes.

A common solution is to vent or port the enclosure. Ports tend to be round, ovular, or slit-shaped, and usually are located on either the front or rear panel of compact monitors. Rear ports allow for a smaller front face, and therefore a more compact monitor, but they can also lead to sonic imbalances — the main one being excessive bass — in cases where the monitor is mounted too close to a wall or corner. Front ports help avoid this problem, but require a larger front face on the enclosure.

PERFORMANCE ISSUES

As we've seen, the technical recipe for accuracy has three basic ingredients: accurate frequency response, accurate impulse response, and low distortion. Superior monitors boast a very flat frequency response, typically within ± 3 dB of a nominal level. In addition, the frequency response should roll off smoothly at either end of the spectrum, as well as fall off evenly as you move away or off axis from the monitor.

Also critical is a monitor's impulse (transient) response. Ideally, this should be a direct analog to changes in air pressure in response to transient electrical signals; a superior monitor keeps all the "time domain" qualities of a signal intact, reproducing them in exactly the same time relation as they appear at the monitor's input terminals. In addition, in a superior monitor the frequencies issuing from discrete drivers are time aligned so as to compensate for the time misalignment inherent in discrete designs, as described earlier. That way, the highs, mids, and lows reach the listener's ear simultaneously.

Both impulse response and time alignment (among other things) figure prominently into two other critical sonic qualities of a reference monitor: *soundstage* and *imaging*. Soundstage refers to the imaginary stage that forms between two speakers (including width and depth), and imaging refers to how well the monitors can localize individual instruments on the soundstage. Obviously, a good soundstage and precise imaging are necessary for accurate positioning of instruments within the stereo field.

Distortion levels vary considerably from system to system. Whereas home-stereo speakers typically exhibit as much as 1 percent distortion above bass frequencies, some high-quality reference monitors may deliver as little as 0.1 percent. Though a low distortion spec is always desirable, some monitors with less-than-spectacular distortion specs still excel thanks to superiority by other measures. The human ear, however, is very sensitive to distortion, especially in the midrange (distortion is often a major contributor to ear fatigue).

Another helpful specification is speaker sensitivity or efficiency, which shows the monitor's output *sound pressure level* (in dB SPL) at a distance of 1 meter with an input signal of 1W. All things being equal (which they rarely are), speaker sensitivity has no determining effect on sound quality. However, if you are doing an A/B comparison of two or more sets of passive monitors and running them from the same power amp through a switching box, it is important to be aware of differing sensitivities. Our ears can readily perceive even slight differences in SPL, and our *brains naturally perceive louder sources as sounding better*. If you fail to compensate for any sensitivity differences — that is, to ensure that each monitor is playing back at the same level — you are more prone to reach incorrect assessments of monitors while comparing them.

FAITHFUL TRANSLATOR

As helpful as specs may be, they are not really an indicator of how a monitor sounds; two similar monitors with near-identical specs can sound very different, for example. Therefore, as in all things audio, careful listening must be the final measure. After all, monitoring is inherently subjective.

But even if monitoring weren't subjective and reliable standards for accuracy could be decided on and agreed upon, the problem of wide-ranging sonic differences

among playback systems would still persist. More important than accuracy is knowing how your mixes will translate to other speakers in other environments. That's the real bottom line. And the only way to gain that certainty is from experience. As they say, practice makes perfect — and it's no different with reference monitors than with musical instruments. After all, a monitor is a musical instrument of sorts. *Thus the need to spend many hours, many days, many months working with a set of monitors, “practicing” on them, listening to your results on countless playback systems, always fine tuning, adjusting, figuring out what the quirks are, where the bumps and holes are, and how every little thing translates, until you reach a level of familiarity that allows you to work undaunted, confident that the mix you dial in will bear a strong resemblance to what the end-user ultimately hears. Regardless of what monitors you use, until you are intimately familiar with them, mixing will remain something of a guessing game.*

CAN OF WORMS

This brings us to a can of worms I'd just as soon not open — but open it we must if we're to inquire seriously into the nature of reference monitoring. Anyone who has searched for the “perfect” monitor has run smack into this dilemma, which is best summed up by these questions: Who, ultimately, are you mixing for? The snooty audiophile with speakers that cost more than most folks' cars? Or the masses who listen to music on cheap systems?

Lord-Alge's (who has used Yamaha NS-10s his whole career) answer is enlightening: “Ninety-five percent of people listen to music in their car or on a cheap home stereo; 5 percent may have better systems; and maybe 1 percent have a \$20,000 stereo. So if it doesn't sound good on something small, what's the point? You can mix in front of these huge, beautiful, pristine, \$10,000 powered monitors all you want. But no one else has those monitors, so you're more likely to end up with a translation problem.”

Similarly, I learned a few years ago that John Leventhal, who was one of my heroes at the time, did the bulk of his mixing on a pair of small Radio Shack speakers. (Leventhal, a New York City-based guitarist, songwriter, and engineer, made his mark by producing Shawn Colvin's acclaimed 1989 record, *Steady On*.) Leventhal owns both a pair of Yamaha NS-10Ms and a pair of Radio Shack Optimus 7s. But he prefers the latter.