THE AMP/SPEAKER INTERFACE

BY E. BRAD MEYER

ARE YOUR LOUDSPEAKERS TURNING YOUR AMPLIFIER INTO A TONE CONTROL?

If you've been reading STEREO REVIEW *for a while, you've probably read comments by Julian Hirsch and others to the effect that well-designed modern amplifiers normally sound pretty much the same. (You may even have seen "Do All Amplifiers Sound the Same?" in our January 1987 issue, which reported on controlled listening comparisons that found no statistically significant differences in the sound of five quite different amplifiers.) Although there are some provisos associated with this claim—that the amplifiers have adequately low noise and distortion, that they not be driven beyond their power limits into overload, that they have flat (or simply identical) frequency response into the loudspeakers they are driving—these don't seem like very difficult restrictions. And for the most part, they probably aren't. But E. Brad Meyer has discovered that there can be situations in which the last condition is violated, creating audible differences between amplifiers that one would ordinarily expect to sound the same. He has detailed his discovery in the form of a dialogue between audiophiles. —Ed.*

ILLUSTRATION BY ANTHONY RUSSO

SOMETIMES you can hear differences between power amplifiers, even in a carefully controlled test.

Why did you say that? And what exactly did you say?

I said, sometimes you can tell the difference between two power amplifiers by the sound, even in a controlled test. And I said it because there's been an argument going on for a long time about whether you can hear the difference, and I've just conducted a test that proves you can. But remember, I said "sometimes." That's important.

But why only sometimes? If amplifiers didn't sound different, everyone would just buy the cheapest one. Some amplifiers cost thousands of dollars, and people buy them. There must be a reason.

I'll talk about the reasons for buying an expensive amplifier later. First of all, people don't just buy the cheapest amplifier because some amps are more powerful than others, and power costs money. One of the premises of the controlled test is that neither amplifier is being driven into serious distortion. If you're comparing a $6,000 high-power amplifier with a cheap receiver, you can just turn both of them up until the receiver distorts, and the difference will be obvious. The longstanding argument is about whether you can hear any difference between well-designed amplifiers operating within their power limits.

Well, I still don't see what's so hard about that. Last month I was over at the house of a friend who had taken home two amps for a weekend so he could decide which to buy. The salesman predicted that one of them would sound more musical, and he was right. We listened to first one and then the other for a whole evening, and it was no contest.

I can't say for sure what you did or didn't hear, but for a lot of reasons that wasn't a controlled test. The first problem is that the salesman told you what you should be hearing. Second, you probably didn't make sure that the two amps were playing at exactly the same level. And third, you had to take at least a couple of minutes to change from one amp to the other, during which time your auditory memory would have faded.

Hold on. Are you saying that one amp sounded better in the same way to both of us just because the salesman told my friend it would? I didn't even talk to the salesman, and I heard the same things my friend did.

Yes, people can and do hear things just because they expect to. And you didn't have to talk to the salesman yourself for your friend to communicate those expectations to you. You probably don't even remember how it happened, and your friend probably didn't mean to do it, but it can happen anyway.

I still don't see how I could be hearing something just because of something the salesman told my friend. What was that about the levels? We listened at about the same volume the whole time.

If you compare two components at slightly different levels, the louder will tend to sound better, and if the difference in level is small it will masquerade as something else—greater transparency, more detail, more depth in the stereo image, or whatever. The effect is even stronger if you use a switch box to compare the two components quickly.

I still don't understand why all that trouble and extra equipment are necessary.

The methods I'm talking about were arrived at by audiophiles and audio engineers who heard differences between all kinds of equipment, just as you do. They set out to identify the causes of the differences between electronic components, and they built switch boxes to make the job easier. Then they discovered that differences in overall level could always be heard unless the two components were within about 0.1 dB of each other.

They found that the ear is sensitive to level differences even if they occur over only part of the spectrum—an octave or two, perhaps. In other words, if the two devices had different frequency responses, they would sound provably different. As with differences in overall level, small variations in frequency response typically would sound like differences in detail or presence or warmth, or something like that.

So they tried using equalizers to eliminate the often tiny response differences, to enable them to concentrate on properties like overall musicality, the amount of depth in the stereo image, and so on. That's when the trouble really started.

What trouble? What happened?

When they finished equalizing the two components to within 0.1 dB of each other, not just in the midrange but all across the spectrum, they stopped being able to identify the components in their blind tests. And to this day no one has been able to do it, except under special conditions.

What do you mean by "special conditions?" What about differences in things like distortion and noise?

The audibility of noise and distortion depends on what you're listening to. With most music you can't hear ordinary harmonic distortion unless it's well over 1 percent. The ear is more sensitive to distortion with a pure tone, a single-frequency sine wave, but most people don't listen to those very much. The same is true of noise, only the most sensitive test condition for noise is no signal at all: You can hear hiss or hum most easily if there's no music playing to cover it up or distract you. The other exception is that with some music played on some loudspeakers you can hear differences in polarity (absolute phase)—that is, whether the speaker diaphragms move toward or away from you when a positive-going signal is applied.

But what about all the writers in audiophile magazines who go on for pages characterizing the sounds of preamps and power amps? Surely some people must hear better than others. Maybe the ones with the golden ears haven't been tested.

Most of the subjectivist audio writ-

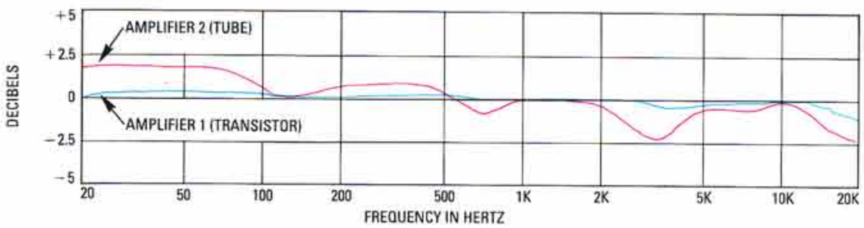

Figure 1. Frequency-response plots of a typical solid-state power amplifier and a tube amplifier driving the midrange and tweeter sections of Speaker A. The difference between them was audible with both pink noise and music.

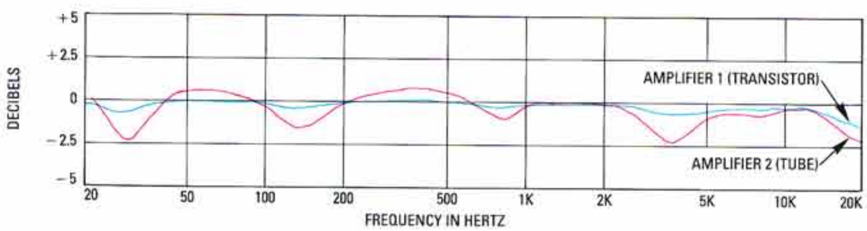

Figure 2. Frequency responses of the same amplifiers driving the entire Speaker A system, including the woofer. The upper-midrange difference was audible with pink noise, but the more similar overall balance failed to reveal the difference with music.

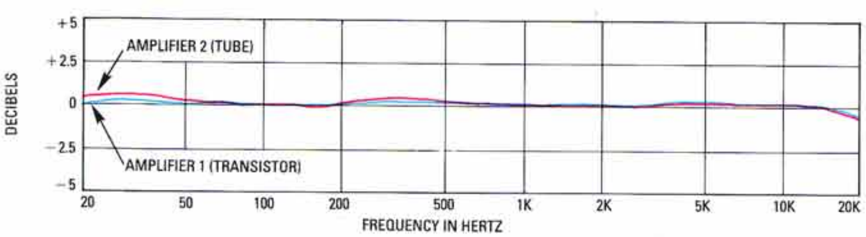

Figure 3. Responses of the same two amplifiers into Speaker B, a three-way system presenting a less difficult load than Speaker A. The slight measured difference in the midrange was barely audible with pink noise and not audible at all with music.

THE LISTENING TESTS

A proper double-blind test requires proper equipment and much care in setting it up. Levels were matched using a 1,000-Hz test tone. Polarity and levels were controlled with a custom-built line-level module from DB Systems of Jaffrey Center, New Hampshire. Its response is within ±0.05 dB from 5 to 80,000 Hz, and its maximum distortion is less than 0.0003 percent. The module has itself proved to be inaudible in double-blind tests. The comparator, made by the ABX Company of Troy, Michigan, switches the amplifiers' inputs and outputs simultaneously.

Each trial represents many comparisons between A, B, and X or between X for the current trial and X for the next higher or lower trial. In the pink-noise tests a rapid switching sequence worked best, but with music, which is constantly changing, many strategies were used, including rapid switching during sustained notes, listening to entire passages on A or B and then to the same passages on X, and switching between similar phrases in popular music. Occasional periods of silence, including breaks for food or coffee, were helpful in maintaining concentration and aural acuity.

At first glance, the graphs in Figures 1 and 2 may appear too similar to account for the results of the music tests, which show a very strong probability of audible differences when the speakers were biamplified and none when they weren't. Closer examination of Figure 1 shows an upward tilt below 500 Hz that gave the tube amplifier a warmer overall tonal balance despite the rolloff below 100 Hz in the loudspeaker's crossover.

The much smaller differences on my own speaker system, plotted in Figure 3, were still faintly audible with pink noise: By concentrating on an apparent difference in the vocal range, corresponding to the 0.25-dB rise between 300 and 500 Hz, I got a score of twelve out of fifteen choices correct, representing a confidence level (probability that the results could not be ascribed to chance alone) of 98.2 percent. With music, however, I could not hear any difference.

Because of the relatively small number of trials in these tests, the results should be taken primarily as an indication of where to look and how to proceed with more thorough tests in the future. But for the biamplified configuration, the results have very high confidence levels, and the measurements, together with the results of previous work (especially that of Floyd Toole and others on the audibility of resonances), fully support the conclusion that with these speakers, with these amplifiers, the difference was audible.

Some audiophiles maintain that the rapid switching and accompanying tension of double-blind testing is somehow unfair, or at least unlike their usual listening conditions and states of mind. Give us time, they say, and let us relax, and we can identify not just amplifiers, but speaker cables, interconnects, or cryogenically treated CD's. Such claims are, at any rate, impossible to disprove.

Double-blind testing is *frustrating. Even after our scores showed we were just guessing, we still heard what we expected to hear. Was this merely convincing illusion? I think so, and I prefer to spend my money on things I know I can hear.*

Setup	Signal Source	Correct Choices	Confidence Level
1	pink noise	25 of 25 (100%)	99.999+%
1	Bach: *St. John Passion*	13 of 15 (87%)	99.6%
1	Cowboy Junkies	13 of 15 (87%)	99.6%
1	combined music trials	26 of 30 (87%)	99.98%
2	pink noise	15 of 15 (100%)	99.9%
2	Bach: *St. John Passion*	9 of 15 (60%)	70%
2	Cowboy Junkies	13 of 30 (43%)	18%
2	combined music trials	21 of 45 (47%)	34%
3	pink noise	12 of 15 (80%)	98.2%

Setups 1, 2, and 3 correspond to Figures 1, 2, and 3. The confidence level expresses the probability that the correct identifications of the two amplifiers in the ABX trials could not be ascribed to chance alone. A confidence level above 95% is considered statistically significant.

ers publicly refuse to participate in double-blind tests, even if some of them have taken them in private. And of those who do take such tests and write about them, most deny that the negative results prove anything. They give a lot of reasons, and the argument has been going back and forth for years. It's often called "The Great Debate."

I'm not saying I agree with you, but if what you say is true, why did you claim that differences between power amps were audible?

Because I found a pair of amplifiers that, with a certain loudspeaker system, provably sound different. One of the amplifiers is a solid-state unit from a company that makes low- to medium-priced equipment, and the other is a vacuum-tube model from a well-known high-end manufacturer. The first lists for less than $900, and the second is priced at more than three times as much. The loudspeaker is a large, expensive multidriver system that presents a complicated load to the amplifiers.

Look at Figure 1 and you'll see what I mean. The frequency-response curve labeled Amplifier 1 is for the medium-priced solid-state amp; Amplifier 2 is the expensive tube amp.

Wait a minute. The expensive amp's frequency response isn't flat, and not just by a fraction of a decibel, either. It varies by almost 3 dB. Is something the matter with it? Are you sure it was working right? What's going on here?

I wondered the same thing myself, but I tried two different examples of the same model and they behaved almost identically. You never see graphs like this in equipment reviews because no one tests amplifiers with speakers attached; they use a simple load resistor instead.

But this speaker system's impedance varies widely with frequency. In the parts of the spectrum where the impedance dips, the speaker draws more current from the amplifiers, so their output voltage tends to fall. In general, transistor amplifiers have lower internal resistance (higher damping factor) than many tube designs, so their response doesn't change as much when they're presented with a load like this one.

You mentioned some kind of scientific test. What did you do?

We used a double-blind comparator—a switch box with three positions, A, B, and X. A and B are the two amplifiers, and X is one of the two, but the box decides which, and it doesn't tell you until after you've made your guess.

But you have a 50 percent chance of guessing right even if you can't hear any difference.

That's right. So you do a number of trials, and if you're correct all, or almost all, of the time, you probably are hearing a difference. I did two sets of trials, one with a steady signal called pink noise that makes it easy to hear response differences and another with music. With the pink noise I got fifteen out of fifteen choices correct in about 5 minutes; with music, I got thirteen out of fifteen correct in about 20 minutes. The audiophile at whose house we did the tests got ten out of ten choices correct with pink noise and thirteen out of fifteen in his own independent music test. All these results satisfy what is loosely called the 95-percent criterion, meaning that we could expect to score that well by guesswork alone fewer than one time in twenty. In fact, the "confidence level" is almost 100 percent for the pink-noise trials and 99.9 percent for the music trials.

Now I'm really puzzled. It looks like we should buy the transistor amplifier because its response is flatter. But the subjectivist writers don't agree with that.

Well, the situation is more complicated than it looks. Remember, the graphs show the amplifiers' electrical outputs, not the sound in the room. From where we sat, the tube amplifier *sounded better* with these speakers. The slight rise below 500 Hz added a warmth that was pleasing with most recordings we tried, and the dip between 3,000 and 4,000 Hz softened a slight upper-midrange hardness and made vocals more natural sounding and easier to listen to.

But can't you accomplish the same thing with an inexpensive equalizer and save yourself a lot of money?

Practically speaking, no. I've tried to duplicate curves like this one with equalizers, and although it's possible if you have the right equipment, it takes a long time and many tiny adjustments. You can't do it accurately without expensive measuring equipment to check your work. You might come up with something that sounded as good or better by trial and error, but it would still take a lot of work and almost certainly would not be exactly the same.

It looks as though the amplifier has errors that happen to compensate for the speaker's errors. Wouldn't it be simpler if the amplifier had flat response and the speaker were designed to sound best with the ideal amplifier?

In a perfect world, yes. In the real world, people buy speakers like these for their other fine qualities and then buy the amplifier that sounds best with them. It's even possible that these speakers were designed using an amplifier like our tube model, which is why the combination sounds good.

How strongly you feel about this depends on your philosophy of equipment design. If you buy this tube amp with these speakers, then for your money you get a slightly mellower and very musical sound, a beautifully massive physical package, a slightly higher electrical bill, a slightly lower heating bill, and the pride of owning a hand-assembled piece of American craftsmanship. That combination holds a lot of appeal for some people.

There's another complication I haven't mentioned. In the comparison where the differences were obvious with both pink noise and music, the system was biamplified: The amplifiers we tested were driving the system from the lower midrange up, while a separate amplifier drove the woofers. With the same amps driving the entire system, the response difference decreased, as you can see in Figure 2. Under those conditions we could hear which amp was which with pink noise, but we failed to identify them using music.

What about your own system? What kind of amplifier do you use when you have to put your money where your mouth is?

I'm currently using a large solid-state amplifier. But when I compare the same two amplifiers on my speakers, both of them sound fine. My speakers don't present as difficult a load, so the two amps actually sound more alike on them, as you would expect from looking at Figure 3. The speaker systems used for some previous controlled amplifier tests were electrically more like mine, which may be one reason the results were mostly negative. And in many cases, both the amplifiers compared were more like Amplifier 1 than like Amplifier 2, and, as we've seen, Amplifier 1 maintains almost perfectly flat frequency response even into the more difficult loudspeaker load.

The reason we succeeded in our tests wasn't because we hear better than other people; we don't. If we had just listened, instead of doing double-blind tests and making measurements of the amplifiers with the speakers attached, we would have said that the biamplified system was the most revealing of the differences in power amps, and my speakers the least revealing. But the controlled tests and measurements show that we heard differences on the biamplified system because the speaker actually *produced* them.

What you seem to be saying is that the amplifier and speaker form a system and that to talk about the sound of an amplifier without reference to a specific speaker is meaningless.

I couldn't have said it better myself.

Then what does this say about other combinations of amplifiers and speakers? You got very different results for your three systems.

Most speakers are probably more like the one that produced the curves in Figure 3, but we need more tests like these to answer that question. What we've found so far suggests (1) that most good solid-state amplifiers probably sound identical, or at least very much alike, within their power limits; (2) that tube amplifiers (and solid-state amps designed deliberately to behave like tube amps) will tend to behave differently with different speakers; and (3) that speakers with strongly varying impedance curves tend to make the two types of amplifiers sound different. □

E. Brad Meyer *works as an audio consultant, recording engineer, and producer for Point One Audio, Inc., of Lincoln, Massachusetts, and is president of the Boston Audio Society.* *He wishes to thank Audio Vision of New England, in Arlington, Massachusetts, for the generous loan of equipment and especially audiophile Jean-Marc Matteini, for both his gracious hospitality and his intellectual courage.*