

Factors Of A Good System

What makes the difference?

By Bob Thurmond

How many sound systems have been built and are in use? Many millions, for sure, and they're found in all types of venues and for all kinds of programs. So one would think we'd know exactly how to do it by now. But there seems to be plenty of examples to prove that we don't.

Why should this be? What is it we don't yet understand? Do we even know enough to know what we don't know?

Perhaps we should start by trying to define the characteristics of a good system. Not just "it sounds good" but - exactly - what makes the difference between "good" sound and not so good. Then we might be able to quan-

tify how good each characteristic needs to be and how to judge whether it's good enough or not.

After nearly 40 years spent designing and testing sound systems, I've finally come up with a list of the factors that I feel make up what we could call quality in a system, and why. In this installment, I'm going to confine my list and discussion to systems for speech reinforcement only. Next time, we'll look at factors for music systems.

RELIABILITY

The most important quality factor has to be reliability. No matter how good the performance of a system may be, if it fails to work, it is useless.

Reliability is largely an engineering matter, involving component selection, configuration design, as well as assembly and installation correctness. But any system can be abused to the point of failure.

Significantly, failure may not be abrupt and catastrophic, but instead may take the form of performance decline due to damage. One particular, and common, example of damage-induced deterioration can be found in commonly-used transducers for higher audio frequencies, the horn and compression driver combination.

Drivers have a severe amplitude limit; if overdriven, the driver diaphragm will impact the phasing plug, an essential part of the structure. If the diaphragm material is metallic, it can fracture and fail.

Some diaphragms, however, are made of a resin-impregnated fabric, which is much less brittle and, therefore, more able to survive a collision with the phasing plug. Repeated collisions, however, still cause progressive deformation (or warping) of the diaphragm, resulting in eventual failure and therefore, progressive decline of the driver's performance characteristics.

Predicting and detecting this impending failure, however, is not easy to do. The audible change in performance is fairly subtle and can be detected reliably only by careful comparison of the sound of a single questionable driver with that of a known good one. In the field, such a comparison is usually impractical. Further, a driver that has been used heavily for some time will also exhibit some performance deterioration, even though it has never been overdriven into diaphragm collision.



Topping the quality system list: reliability of components, individually and collectively.

Think Tank

Figure 1 illustrates these performance differences. The frequency response (amplitude versus frequency) of three drivers of the same model (with an impregnated-fabric diaphragm), one new, one well-used but apparently undamaged, and one with observable damage.

It can be seen that the response at higher frequencies changes with use or abuse. The differences between the upper two measurements are slight, while the third one is significantly different. There seems to be a good relationship between the measured and (subjectively) observed performances in cases like these, but no real study of this relationship has been performed.

So it would seem that a response measurement could be a valid substitute for a listening test. In fact, such a relationship has been established under certain circumstances, but not definitively in a sound reinforcement context. An investigation of this relationship would certainly be worthwhile.

However, there is another measurement that is easy to make, even though it's seldom done. The bottom three curves on **Figure 1** represent the measured electrical impedance at the input terminals of each of the three drivers. Such a measurement is usually quite easy to make, even on a driver installed in a system.

It's apparent that these curves separate the characteristics of the three drivers as well as any other common measurement does, especially in the case of the damaged unit, and much

more easily. In fact, automated tests of this type have been designed into integrated systems as performance and reliability checks, with good results.

Thus it appears that different types of tests on the same items can yield corresponding results. In fact, experience has shown that such relationships hold in some cases but not in others, and that it may be difficult to predict which is which. And in many cases, no acceptable substitute for a listening test has yet been found. Worse, some widely accepted tests might prove inadequate.

LOUDNESS

It's obvious that any sound system must provide enough sound level at the audience locations to ensure a satisfactory listening experience. Defining what this level actually should be is less obvious, and use of a valid measurement technique is not obvious at all. Subjective opinions on appropriate sound levels often vary widely as well, depending on a host of factors. (Investigating this matter alone could become a major research project!)

In fact, the correct sound level may not be just a matter of loudness. How well speech is understood (intelligibility) is often the overriding concern, and this is the result of many factors other than just loudness. In some cases, the loudness may need to be set other than as would normally be expected, because of adverse acoustical or system functional characteristics. It may also be found that the audience prefers a sound level different from

that which exists near the performer.

Other acoustical factors may also be highly significant. The level of the reinforced sound must be sufficiently higher than that of any background noise so that speech intelligibility or program enjoyment is maintained. Some guidelines in this regard have been established empirically, and they may be adequate for most situations.

A common and complicating factor is that background noise level may vary significantly, rapidly and unpredictably. Further, since adequate performance in this area may be a matter of life safety, accuracy can be quite important.

It's often the case that the desired sound level is greater than that which the system is capable of producing without difficulty. This difficulty is the result of one or more components overloading, which results in an audible distortion of the sound.

Distortion may take various forms, depending on the type of component that is overloaded, the magnitude of the overload, and the nature of the program material, among other factors. Therefore, the audibility of the distortion may vary greatly with the situation, and each type of distortion must be evaluated individually.

Many listeners even believe that certain types of distortion are desirable, such as that typically produced by vacuum tube amplifiers. This usually applies to music playback systems in small rooms, however, so it's unclear if such an effect is valid in a larger sound reinforcement situation.

Some devices are available that deliberately introduce controlled distortion, specifically for pro audio applications. Many have noticed that a limited amount of distortion adds to the apparent loudness of amplified sound, and without being objectionable. If anyone has actually studied this effect, the results remain obscure

TIMBRE

The overall timbre, or tonal balance, of a sound system undoubtedly has the strongest influence on the overall perceived quality. This characteristic is easy to measure, both subjectively and objectively, and there is a very good correlation between the two in a small-room configuration.

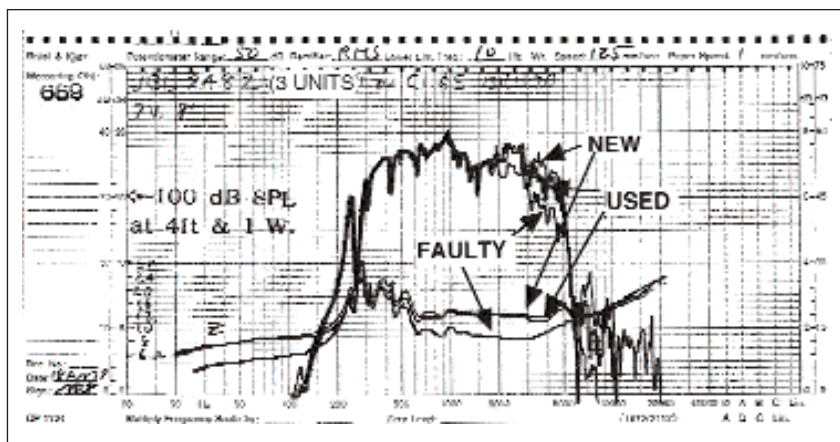


Figure 1: Measurement of output acoustical and input electrical impedance characteristics of three high-frequency horn drivers of identical model but different usage histories.

Think Tank

In a large-room sound reinforcement situation, however, this correlation does not hold. If the system has an overall response that is measurably flat (has nearly the same input-to-output level ratio at all frequencies), it will sound too bright, with the high frequencies being too loud. A system which sounds subjectively flat, so that the reproduced sound is perceived as being a close duplicate of the source, will have a measured response which rolls down at high frequencies.

Should the analysis be done with a swept filter, which yields more information, or is a stepped filter technique acceptable? What amplitude smoothing or averaging is appropriate? If measurements are taken at single, discrete frequencies, as are commonly done with contemporary techniques, how many measurement points are needed and at what spacing? This could be a major source of misleading data, especially at lower frequencies.

Whatever the technique – how many measurement locations should be taken, and where should they be located? And exactly how should the individual measurements be averaged to yield the overall system response? Also, how much variation between individual measurements is acceptable, and what should be done if the variation exceeds this tolerance?

Schulein documented this discrepancy in 1975 in an elegant experiment and offered a plausible explanation. He noted that in all rooms, the listener receives sound directly from the source and also reflected from the room surfaces. (See *Sidebar*)

Schulein's Experiment

Two identical loudspeakers are arranged in a room so that the listener receives the direct sound from both, one close and one distant. The bandpass filter is first set at 1000 Hz and the attenuator adjusted so that the sound level appears to be the same when the noise is switched from one loudspeaker to the other.

Then the filter is changed to another band and the corresponding band on the equalizer is adjusted until the apparent loudness is the same from each loudspeaker. This is continued through all bands, and repeated as necessary until no further adjustments are necessary.

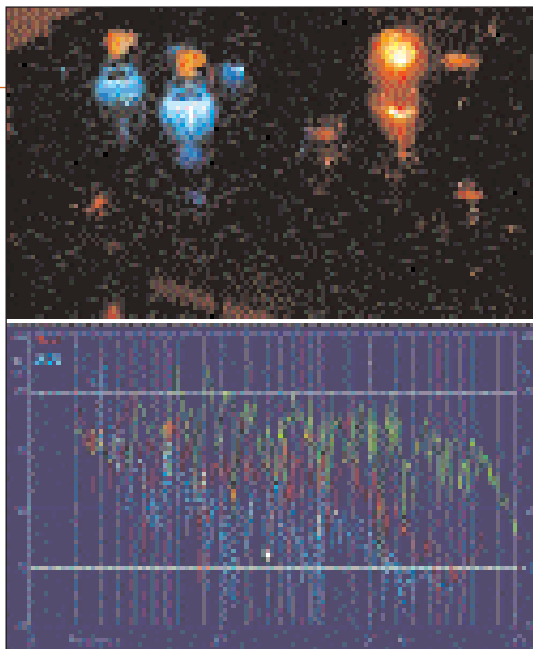
Finally, the overall frequency response of the equalizer is measured, this being the correction necessary to make the distant loudspeaker sound like the near one in this room.

In a small room, the level of the direct sound is almost always higher than that of the reflected sound and, therefore, dominates in the perception process. Because of directional characteristics of human hearing at high frequencies, largely due to head shadowing effects, less total sound energy enters the ears at high frequencies than at lower. This imbalance is perceived as normal.

In a large room with typical acoustics, however, the opposite is true; the level of the reflected, or reverberant, sound is significantly higher than that of the direct at most listener locations. Since this reverberant sound arrives at the listener from all directions rather than just one, more of it enters the ears at high frequencies. Thus the highs are perceived as being louder.

A simple experiment tends to confirm this theory. A loudspeaker is located at head level in a relatively non-reverberant environment and fed with broadband noise. A listener stands one to two meters (about three to six feet) in front of the loudspeaker and slowly turns around while listening to the tonal character of the noise. Typically, the overall tonal balance will change little, if at all, with head direction.

However, if two identical loudspeakers are placed two or three meters apart facing each other and



Some desire the added distortion that can result from tube-based components. But is this a bad thing?

both are fed the same broadband noise, a listener between them, turning around as before, will hear the high frequencies more loudly when his ears are toward the loudspeakers than when he is facing one or the other loudspeaker.

The measured response (and perceived timbre) of a loudspeaker in a room deviates significantly from its performance in an anechoic environment, in ways that are complex and quite difficult to predict. Also, these deviations are different at each location in the room. Therefore, the only practical solution is to measure the actual response of the completed system and correct it as needed with additional circuitry.

This turns out to be a bit trickier than one might expect, however. If a pure tone, slowly swept in frequency, is fed over a sound system and the resulting level is measured at a point in the audience area, it will be found to consist of strong peaks and valleys, tens of decibels in amplitude, and spaced at intervals of about 1 Hz, caused by room resonances.

It's almost impossible to get meaningful information from such readings. Besides, we don't perceive these variations because they are averaged by our hearing process in ways that are only partly understood. The measurements must incorporate averaging which simulates the hearing process.

Think Tank

However, this presents us with a shopping list of unanswered questions pertaining to the measurement techniques. What frequency resolution (bandwidth) is needed? A first assumption might be to use a bandwidth similar to that of the auditory (critical bandwidth) filters, but system measurements are typically done with third-octave filters, which are considerably wider than critical over much of the spectrum.

Should the analysis be done with a swept filter, which yields more information, or is a stepped filter technique acceptable? What amplitude smoothing or averaging is appropriate? How many measurement locations should be taken, and where should they be located? And exactly how should the individual measurements be averaged to yield the overall system response?

Despite countless practical field experiments in this area, beginning at least 65 years ago, little critical research has been carried out. As a result, there

exist only a few de facto standards, and the actual results of these procedures vary considerably in quality.

In addition to these considerations, it might be expected that non-linear distortion in any of the system's components, especially the loudspeakers, would significantly affect its timbre, but such does not seem to be the case. The distortion levels of modern components, properly used, are low enough to be unnoticeable in a reinforcement situation.

INTELLIGIBILITY

As the name suggests, intelligibility is the measure of how easy or difficult it is to understand speech over a system. It's ultimately measured subjectively and directly, typically using rhyming words as the test signal.

The execution of this test is tedious and time-consuming with only one test subject, which is quite inadequate. Different subjects will render somewhat different results even under apparently identical conditions, and conditions vary significantly with location, program sound levels, room noise, hearing acuity, and many other factors.

The typically broad variance of test results makes it difficult to determine whether a system is actually performing acceptably or not. It hardly seems worth the rather considerable effort required to execute such a test, but there may be little choice.

Because of these difficulties, a lot of effort has gone into devising an objective test regime, with several products resulting. All these involve dedicated gear and techniques, which, while not simple, are quite preferable to subjective tests.

These objective

tests have been demonstrated to produce results comparable to those obtained subjectively in some, but not all, conditions. Unfortunately, the worst correlations tend to occur in conditions that produce low scores, exactly where accurate results are most desired. In fact, after extensive experience with all the commonly used objective techniques, Mapp has concluded that all are inadequate.

It gets worse. Low intelligibility scores, which indicate serious problems, usually provide little or no information on the nature of these problems. Sometimes one or more physical problems are apparent in such cases, but are these really the causes of the poor performance?

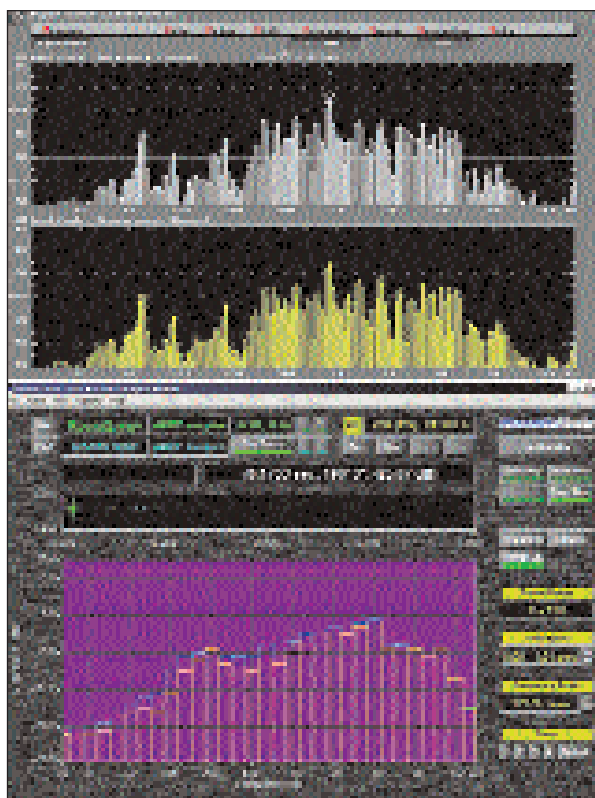
Often, the only way to be sure is to correct the problems and see if that improves the scores. Of course, this may be completely impractical, and in fact, there may be multiple problems, some masking others, so that correcting the most obvious might accomplish nothing useful.

A much more practical approach might be to identify exactly which physical factors adversely affect speech intelligibility, and how, and calibrate physical measurements to subjective effects. If this were accomplished, then not only would meaningful test methods be available, but effective design criteria could be established to predict results and avoid problems in the design stage. Some significant work has already been done in this area, with results pointing to the ratio of direct to reflected (or reverberant) sound being the most important factor.

Think this over, work on your own list, and next time, we'll look at key quality factors of music reinforcement systems.

Editor's Note: *This paper was originally presented at the 146th Meeting of the Acoustical Society of America, Session 4aAAA. For a full bibliography of research materials referenced, see ProSoundWeb.com. ■*

Bob Thurmond began his professional audio career in 1959 as a recording technician. He became a consultant in 1965, specializing in sound reinforcement, and formed his own company, G. R. Thurmond and Associates, in 1971. He may be reached at thurmond@io.com.



Tools like SIA SMAART and Meyer SIM can be of great help along with an entire understanding of what's happening with a system and room.