LETTERS TO THE EDITOR

COMMENTS ON "IMPROVING THE SIGNAL-TO-NOISE RATIO AND COVERAGE OF FM STEREOPHONIC BROADCASTS"*[1]

I found it somewhat ironic that CBS's FMX stereo broadcasting scheme seems to be justified primarily on the basis of increased coverage area for FM stations (translation: higher advertising rates and larger profits), with the prospect of a badly needed improvement in audio quality relegated to somewhat secondary status.

With pop FM stations already applying virtually infinite compression to their signals in an attempt to squeeze every last listener out of each watt of radiated power, the last thing we need is another gimmick they can abuse to wreak further indignities on the music. In the FM that I listen to, the most objectionable background noise comes not from the channel, but from the pumping and breathing of the compressors working to process all the life out of the program. Of course, FMX cannot address this problem.

FMX proponents will say, "But FMX will reduce the need for compression to maximize station coverage area." A naive theory at best. For if *uncompressed* FMX covers well, then can we not extend coverage a little further by precompressing the source?

I would be in favor of FMX on one condition: stations using it would be required by the FCC to reduce preprocessing of their audio to the absolute minimum, that is, to assure that clippers would be operating no more than a tiny fraction of the time, and that the time-averaged gain reduction applied by any compressors could never exceed some small value. When broadcasting digital source material, a station might be allowed to use gentle decilinear compression, say, with a slope of no less than 0.9 dB/dB. Special feedback control circuits (with *very long* release time constants) should be specified to automatically reduce the audio level if these maximum allowed compression limits are exceeded.

Of course, adhering to these requirements would require some effort by broadcast engineers to determine appropriate gain settings for different source programs—rather than simply cranking the knob to "maximum" and letting the limiters take over. Besides, it should be possible to digitally encode dynamic range and peak level information in the lead-in "grooves" of Compact Discs to allow automating the level-setting process.

Artists—and the music-listening public—should have the right to hear recordings more or less as they were intended, and not the "osterized" travesties that stations spew out in the pursuit of ratings. If FMX could help more people enjoy *quality* (that is, faithful) FM sound, then it would be a godsend indeed.

Rob Lewis

* Manuscript received 1986 February 28.

[1] E. L. Torick and T. B. Keller, *J. Audio Eng. Soc.*, vol. 33, pp. 938–943 (1985 Dec.).

COMMENTS ON "STEREO TELEVISION: MARKET FORCES AND ISSUES"**[2]

I would like to clear up some historical points in the above extensive and informative article. The first station in the United States to broadcast BTSC stereo sound was WNBC-TV, New York, when *The Tonight Show* was broadcast in stereo thereon and fed in stereo to the Skypath™ satellite network on Thursday, 1984 July 26.

The first program to be regularly broadcast in stereo was *Friday Night Videos*, beginning on Friday, 1985 July 19, followed by *The Tonight Show* and *Late Night with David Letterman*, which both began regular stereo telecasting on Tuesday, 1985 July 23.

Randy Hoffner
Senior Staff Engineer—Technical Development
National Broadcasting Co.
New York, NY 10020

** Manuscript received 1986 April 14.

[2] M. Polon, *J. Audio Eng. Soc. (Feature)*, vol. 34, pp. 188–204 (1986 Mar.).

COMMENTS ON "SUBJECTIVE AND PREDICTIVE MEASUREMENTS OF SPEECH INTELLIGIBILITY—THE ROLE OF LOUDSPEAKER DIRECTIVITY"†

In the above paper[3] Mr. Jacob apparently misunderstands the concept of loudspeaker Q and the way in which it affects speech intelligibility in reverberant spaces. The generally specified value for the Q of a loudspeaker is an *axial Q*, that is, the value of Q with

† Manuscript received 1986 February 3.

[3] K. D. Jacob, *J. Audio Eng. Soc.*, vol. 33, pp. 950–955 (1985 Dec.).

reference to the loudspeaker's axis. The Q as perceived by a listener off axis is given by

$$Q_o = Q_a \times 10^{dB/10}$$

where Q_o is the perceived Q off axis, Q_a is the axial Q, and dB is the sensitivity of the loudspeaker at the off-axis angle relative to on axis (e.g., -6 dB for 20° off axis of a 40° horn). For most loudspeakers, those whose sensitivity falls off axis, the perceived Q off axis is lower than that on axis. Therefore as a result of this lowered perceived Q, and all other things being equal, the intelligibility is reduced off axis.

The effects on intelligibility of this lowered Q may be compensated for by the inverse square law. This is evident from the variable D^2, the source-to-listener distance, in the equation for $\%AL_{cons}$ [Eq. (1) in Mr. Jacob's paper]. Therefore if a listener were on the -6-dB angle of the loudspeaker *and* twice as close as a listener on axis, the effects of lowered Q and the inverse square law would cancel each other, and both listeners would enjoy the same intelligibility. This is the reason for all of the emphasis on achieving even direct sound coverage in reverberant spaces.

It is assumed that Mr. Jacob used *axial* Q to predict the $\%AL_{cons}$ at each listening location. (We are not told where the Q values for the three loudspeakers come from.) If the direct sound coverage was not even, and this appears to be the case, then errors of the type experienced by Mr. Jacob are to be expected. It has been my experience that when uniform direct sound coverage has been achieved, the intelligibility will be uniform and in good agreement with that predicted by the $\%AL_{cons}$ method. Had Mr. Jacob determined the Q at each listener position and used these individual values when predicting the $\%AL_{cons}$ at the individual locations, his predictions would have matched the intelligibility tests much more closely.

Also, on page 950, Mr. Jacob states that ". . . two listening locations were chosen, roughly in the middle and at the rear of the auditorium floor." He then states on page 951 that ". . . two listener locations were chosen to coincide with 1) the critical distance of the high Q source, and 2) the 'intelligibility distance' of the high Q source. . . ." It is unlikely that both of these statements will be satisfied at the same time in one room, let alone *five*. As the critical and limiting (intelligibility) distances are Q dependent, they will vary off axis in the same manner as the Q. I doubt that *any* listener was at the limiting distance, even for the low Q loudspeaker, in any of the five rooms.

In addition, I would like to point out that by deriving the intelligibility from the measured impulse response of the room in the cases of signal-to-noise procedure and modulation transfer function, Mr. Jacob has in effect measured the intelligibility. His Fig. 4 therefore compares not *predicted* but *measured* intelligibility with actual, the actual being the intelligibility measured in yet another way. The $\%AL_{cons}$ is the only value actually predicted, and then with erroneous input. Quick and easy methods do exist to measure $\%AL_{cons}$ directly—TEF, for example.

There seem to have been a great many variables in Mr. Jacob's experiments: Q varying with frequency, listeners at various off-axis angles from the loudspeakers, and so on. In order to limit better the variables to Q alone, I suggest, as an experimental setup, a sound system consisting of a bass cabinet and a high-frequency horn at one end of the test room. Aim the loudspeakers with their axes parallel to the floor and at ear height such that a listener may move fore and aft along the axis to any desired distance. Several manufacturers offer "families" of constant-directivity high-frequency horns. By substituting low, medium, and high Q horns from the same family into the sound system, the Q may be varied in a known way. Under these conditions, the axial Q is the Q perceived by the listeners, and a more accurate test of the effects of Q on intelligibility may be made.

FARREL M. BECKER
Audio Artistry
Kensington, MD 20795, USA

ADDITIONAL COMMENTS‡

It was with regret that we noted the failure of the *Journal* reviewers to detect and correct the many errors in the above paper.[3] For example, Eq. (3) leaves out the dimensional constant 0.4 normally associated with that equation when SI parameters are used. There is a nondimensional form of the same equation that is much more widely used:

$$D_c = 0.141\sqrt{QS\bar{a}} \ .$$

(See Klein and Davis.[4]). The use of C_d for critical distance when *Journal* usage for the past 15 years has been D_c and the use of "intelligibility distance" in place of the agreed-upon "limiting distance" D_L would seem to indicate unfamiliarity with the literature by both the author of this paper and the *Journal* reviewers. A further irritation is the failure to provide dimensional labels in Eq. (2).

Neither Jacob nor the reviewers recognize that the audible difference to be expected from going from $Q = 1.0$ to $Q = 7.5$ is

$$10 \log \left(\frac{7.5}{1}\right) = 8.7 \text{ dB}$$

‡ Manuscript received 1986 February 10.
[4] W. Klein and D. Davis, "Formulas for Distributed Loudspeaker Systems," *J. Audio Eng. Soc.* (*Letters to the Editor*), vol. 20, pp. 401–402 (1972 June).

whereas the difference in level between $Q = 7.5$ and $Q = 17$ is only

$$10 \log \left(\frac{17}{7.5}\right) = 3.5 \text{ dB} .$$

Perhaps the reason Jacob heard little difference between his "medium Q" array and his "high Q" loudspeaker is that it was the difference between a "medium Q" device and a "medium Q" device.

To have been consistent, he should have gone from $Q = 7.5$ to $Q = 7.5 \times 7.5 = 56.3$ for another

$$10 \log \left(\frac{56.3}{7.5}\right) = 8.7 \text{ dB} .$$

Since loudspeakers listed in today's catalogs have Q values up to 80+, it is a pity that Jacob did not try a true "high Q" device.

This writer feels that arrays do not have a Q, but rather an N (see Davis[5]). My personal experience has been that because of "lobing" in arrays, the calculation of their Q values becomes a highly subjective matter. Therefore it would have been useful to have seen the polar data used in Jacob's calculations.

Q is always at a point and not an area. Therefore any listener not sitting on the axis of the device is not experiencing the indicated Q value. The relative Q at other positions can be calculated from the polar response by

$$Q_{\text{rel}} = Q_{\text{axis}}(10^{\pm \text{dB}_{\text{CL}}/10}) .$$

In order to evaluate the accuracy of Jacob's data, we would require:

1) The seating plan of the listeners relative to the on-axis line of the loudspeakers
2) The distance to the listener
3) The variation in level with angle to each listener.

Other authors, notably Houtgast and Steeneken, have shown the close correlation between V. M. A. Peutz's $\%AL_{\text{cons}}$ and intelligibility calculated from the MTF technique as validation of MTF (see Jacob's Ref. [3]). Jacob uses the same techniques to do the reverse. The truly remarkable work of Peutz during the past 14 years was completely overlooked by both Jacob and the reviewer.

As a user of all of the techniques named and a number not referred to in the measurement of speech intelligibility in real systems in real installations, may I be permitted the comment that rarely are signal-to-noise ratios or direct-to-reverberant sound ratios the culprit when poor intelligibility is encountered. Rather, poor coverage, misalignment of alike devices, improper equalization, and discrete high-level late reflections are the causes found and corrected.

The question raised, namely, "What is the relationship of the Q rating of a loudspeaker to the intelligibility experienced?" is a worthy one, but it must be explored with the other variables under control and documented.

DON DAVIS
Synergetic Audio Concepts
San Juan Capistrano, CA 92693, USA

AUTHOR'S REPLY[6]

In Response to F. M. Becker's Concerns

The author regrets not having stated more clearly the relationship between the listeners and the major axes of the loudspeakers. In all cases, listeners were within 10° of being on axis of the loudspeakers. (See Fig. A and Fig. 2 in the paper.) This being the case, axial Q was used as input to the $\%AL_{\text{cons}}$ predictive formula. This clarification should also relieve concerns about the listening locations and their relation to "critical distance" and "intelligibility distance."

Readers should note that every axis is a "major" axis in the case of the omnidirectional loudspeaker. Note too that the largest errors in the $\%AL_{\text{cons}}$ formula occurred not with the $Q = 17$ device, but rather with the $Q = 7.5$ and $Q = 1$ devices.

Speech intelligibility is a subjective parameter. By definition, any test paradigm for measuring intelligibility must include listeners. The measurement of a system's impulse response—or any other objective measurement for that matter—cannot, therefore, be used to measure speech intelligibility.

In Response to D. Davis's Concerns

It is not clear which audible differences are to be expected from changes in the directivity index (= 10 log Q). If, by audible differences, differences in actual subjective speech intelligibility is meant, it is unknown whether the writer is referring to published or private data relating the two. The predictive formula shown in the paper to be least accurate uses as an input parameter loudspeaker directivity Q, rather than the directivity index (= 10 log Q),

$$\%AL_{\text{cons}} = \frac{200D^2T^2}{QV} \tag{1}$$

(units are SI metric). Therefore the loudspeaker Q values in this experiment were chosen in part to reflect the importance placed on them by this predictive formula, as was stated in the body of the paper.

It is not true that "Houtgast and Steeneken have

[5] D. Davis, "The Modified Hopkins–Stryker Equation," *J. Audio Eng. Soc.*, vol. 32, pp. 862–867 (1984 Nov.).

[6] Manuscript received 1986 April 28.

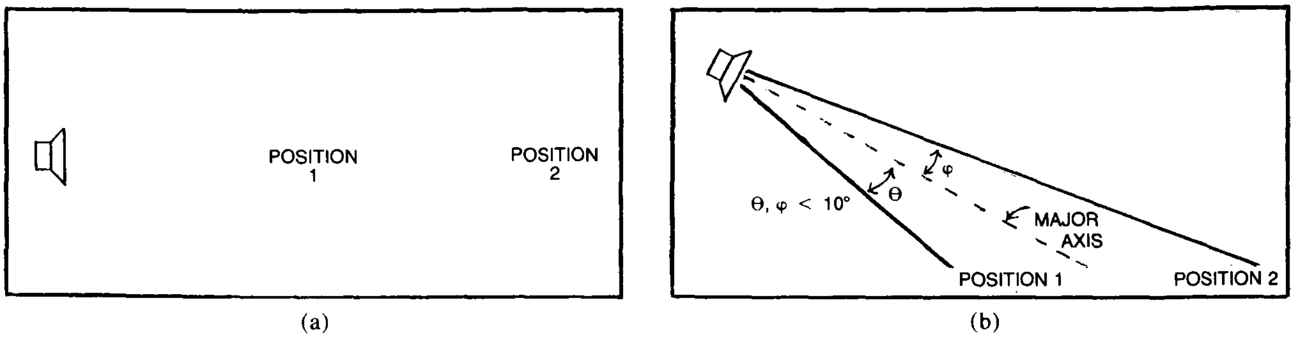

Fig. A. Relationship between listening positions and loudspeaker major axis. (a) Plan view. (b) Elevation.

shown the close correlation between . . . %AL_{cons} and intelligibility calculated from the MTF technique." What has been stated by Mr. Steeneken (private communication and [3]), and what was stated in the body of the paper, is that intelligibility as predicted by the two techniques will agree only in the hypothetical case of an exponentially decaying squared room impulse response.

While Q is unambiguously defined for an array loudspeaker, it is true that these loudspeakers tend to exhibit lobes in certain frequency bands due to complex pressure summation from individual array elements. The effect of lobing on intelligibility, if any, should be treated in a separate experiment.

Readers should be aware that other algorithms have been developed by Peutz for predicting %AL_{cons} since the one used here. While preliminary results indicate that Eq. (1) can be inaccurate, this experiment was in no way intended to discredit the work of Dr. Peutz or his associates.

KENNETH D. JACOB
Bose Corporation
Framingham, MA 01701, USA

CORRECTION

We regret the publication of two errors in the computer code of "Design of Optimized Loudspeaker Crossover Networks Using a Personal Computer," by Peter L. Schuck (vol. 34, no. 3, 1986 March). The author's corrections are as follows: on p. 139—

```
  1570 FOR I=1 TO NV: VAR(I)=X0(I): NEXT I: REM set the best variable values
should read
  1570 FOR I=0 TO NV: VAR(I)=X0(I): NEXT I: REM set the best variable values
```

and on p. 141—

```
     2620 VR(I) = RHSR(I): VI(I) = RHSR(I) : REM save the node voltages
   should read
     2620 VR(I) = RHSR(I): VI(I) = RHSI(I) : REM save the node voltages
```