Psycho-Acoustic Aspects of Higher Quality Reproduction

C. J. LeBEL*

The author emphasizes the commercial advantages of better reproduction.

THIS IS A DELIBERATELY CHALLENGING PAPER, for wide attention to better sound quality in the home is at least ten years overdue. It will stress the beneficial commercial effect of improved sound quality, because that most important feature has been generally overlooked.

The radio field has been badly confused over the public's reaction to higher quality. The incorrect idea that extended frequency range and high quality are synonymous has caused much of the confusion. A further source has been the fact that many of the factors lie in the semi-subjective field of psycho-acoustics. The typical set designer or chief engineer is generally a radio-frequency engineer with no feel for the subjective, and hence is easily misled. As Ben Olney has often said, the average set designer thinks of it only as a device to join a signal generator to an output meter. Such few organizations as have maintained a staff of audio engineers, *with adequate authority*, have found the result profitable.

Listener Fatigue

That listening may cause fatigue is a matter which has had little engineering attention. There has been a limited amount of unpublished work, and one or two magazine articles, on the subject, but it has never been discussed before a learned society.

The fatigue effect results after a listening period of a few minutes to a dozen hours. It is a fatigue of the understanding centers of the brain, not of the cochlea in the ear, and ranges from the mildest weariness to one so extreme as to force an immediate respite. There has been some effort to blame such weariness wholly on program character, but experiment leaves no doubt that system characteristics are equally effective.

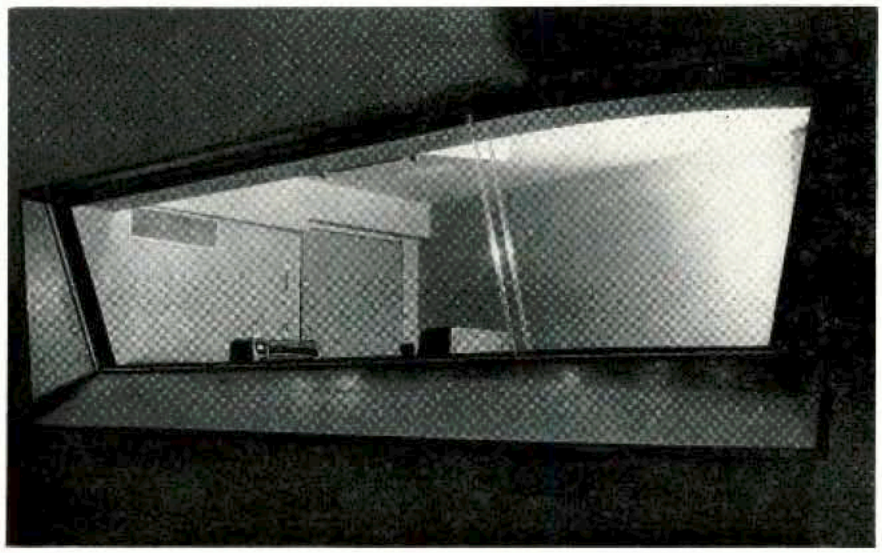
Reduced-Fatigue Designs

In prewar home radios the low, medium and high priced fields each had one manufacturer whose designs were definitely less fatiguing than those of his competitors in the same price class. All three makers were very successful. Much of their success was due to aggressive selling, but we have the word of an experienced merchandiser that their easier listening led to repeat business over a period of more than fifteen years. The fatigue effect of most radio sets cuts listening time which reduces the business of radio stations and tube manufacturers. In fact, it has been proposed that NAB and the tube manufacturers grant a subsidy to makers of sets which have exceptionally low fatigue effect in their respective price classes.

In the hearing-aid field we have interesting observations which are applicable to other fields as well, for approximately one-half of the hard of hearing have hearing characteristics very much like the normal except for a sensitivity reduction which is essentially uniform with respect to frequency. It has been demonstrated that a drastic reduction in fatigue effect, with no visible change in the instrument, may double sales within a period of months. In one case, ten thousand dollars' worth of engineering was as effective as several hundred thousand dollars' worth of advertising. Such a high ratio of yield is probably due to the fact that the average hearing aid user wears the instrument twelve to sixteen hours a day, and the resulting fatigue, with a poor design, can easily be beyond human endurance.

Motion picture theater owners have known for years that the public will go to some little trouble to avoid poor sound, as reference to the *SMPE Journal* will indicate. Few engineers realize that the theater owner's success depends on giving an *impression* of spending money with reckless abandon, whilst

A view from a Columbia Broadcasting System studio into the associated control room. The clean-cut design of the double glass observation window (which is typical of this network's studio design) provides excellent unobstructed vision between the two areas. The glass is set at carefully determined angles in order to avoid all possibility of any reflections that would hinder vision to the slightest degree. It is interesting to note that because of this slope, the curved, by-window-like end sections require especially fabricated glass. Rather than being simply sections of cylinders of proper radius, they must be particular sections of accurately dimensioned cones.



* *Audio Devices, Inc., 444 Madison Ave., N. Y. C.*

actually purchasing with great care. Hence it is significant that you will find no theater with equipment as bad as the average home radio set.

In the disc field, the influence of sound quality is mingled with the effect of an artist's fame and the vogue for a particular selection. It is the duty of the musical director and the artist to avoid a head-on comparison of technical quality undiluted by artistic considerations, but nevertheless it is recognized that good sound quality is one of the factors which affect record sales.

The inexperienced observer's reaction to sound quality of reduced fatigue factor shows his appreciation of better presence: "It sounds very natural", "the announcer sounds right here in the room", "this is very easy to listen to", "you can listen all day without getting tired", "it sounds so clear", are commonly used phrases.

Fatigue Testing Methods

There has been no prior discussion of fatigue testing methods before a learned society, and so a brief statement of these methods may be of interest. They are subjective methods — you may as well be warned in advance. Subjective does not mean qualitative, for the results can be reduced to rough quantitative relations.

The most accurate results can be secured in the hearing aid field. A series of identical-units, matched as to all performance elements except the one being tested, are given to a skilled but non-technical observer. One or two standard production units are always included as controls. Each instrument is worn several days, then the best two are worn for two days each. Instruments may be listed in order of performance under any given condition. If very poor units are being compared, or very sensitive observers used, the measure of performance is the consecutive number of hours the unit can be worn before fatigue compels shutting it off. Better units are compared on the basis of the amount of fatigue after a normal period of use. It is wise to diversify the type and degree of hearing loss when picking such a jury, for some are abnormally sensitive. For instance, certain nerve type cases were unable to wear a once very popular low-priced unit for more than an hour before fatigue became excessive.

Results from such tests correlate well, and it is possible to predict accurately the public's reaction to a new model on the basis of three sets of curves and a listening test. The listening test is intended to evaluate the effect of mechanical noises — particularly case, clothing, and cord noise.

Radio sets are no easier to study. Good designs may require a week of listening to evaluate, while mistakes may be rejected in one evening. A chance for direct comparison is also desirable if the models to be compared are unusually bad or unusually good. This places the evaluating burden upon a trained listener, usually an engineer. That the method can be successful is shown by the number of companies which have been able to turn out superior product by this means.

The motion picture field has relied on the trained listener also, for many years.

While comparing some 16 mm sound projectors intended for the educational market, it was found that student attention was direct function of reduction in fatigue. Even the most interesting films cannot hold the student if the reproduction is bad. Without verbalizing the matter very much, educators have been conscious of the need for good quality sound in both projectors and educational recorders.

Fatigue Factors

We now examine the fatigue factors in detail, realizing that noise is the worst. Different fields find different noise sources their main problem. In hearing aids, case vibration, clothing rubbing across the microphone opening, and cords rubbing on the clothing and body are of major importance. Tube microphonics can be objectionable if not guarded against.

In home radios, natural and man-made static are made more serious by the vogue for built-in antennas. The use of narrow band width is largely prompted by the discomfort caused by wide band reproduction in sets as ordinarily installed.

In phonograph work, record scratch has been a great deal more serious than it need be. Quieter pressing stocks have been available for years, but their universal use waits upon more intense competitive conditions than presently obtain.

Harmonic and intermodulation distortion have a strong fatigue effect which can be somewhat minimized by reduction in frequency range. Articulation tests do not provide a good index because they are generally carried on for too short a time. Most radios marketed today distort a great deal more than they need to for a given manufacturing cost. American records often have much higher intermodulation distortion than desirable, due to polishing and over-use of stampers. In this connection it is interesting to note that Roys[1] has shown that intermodulation is greatly increased by polishing, but that harmonic distortion is not significantly affected.

The mere raising of the upper cutoff frequency to 8,000 cycles is not in itself a guarantee of improved quality. The lower cutoff must be reasonably placed, and there should be no gaps in the response. The following efforts were labeled "high fidelity" some years ago, but they were still very unsatisfactory to listen to: One had response limits of 300 and 8,000 cycles; another had limits of 100 and 8,000 cycles with a gap between 2,000 and 4,000 cycles, and a third had limits of 100 and 7,000 cycles with strong peaks at 200 and 1,000 cycles. Peaks, valleys, unbalanced limits are all annoying to the ear.

The dynamic characteristics of a transducer are not always the same as the static characteristic. As Shorter[2] has shown, a loudspeaker may behave quite differently when handling interrupted tone than when handling steady tone. Many of us have had the experience of having a loudspeaker sound differently from what its response curve would indicate, and this work of Shorter's provides the explanation. Peaks so produced are ruinous to "presence", and have a strong fatigue effect. Loudspeaker manufacturers have often put peaks in the response, feeling that it was an inexpensive way of adding lows, or highs, or loudness efficiency to satisfy the set manufacturer. This has done more to impair radio receiver quality than almost any other defect.

While discussing dynamic characteristics, it might be well to note that microphones of nominally identical response but different construction may produce rather different sound — a matter which seems to have been given no engineering study. Other transducers may have a similar fault, though for better understood reasons. For example, a phonograph pickup which is not tracking properly will sound very unpleasant even though the nominal characteristics seem normal.

In a system with low distortion the result may be very fatiguing if the frequency range is not great enough for good presence. Good speech through a 2,000 cycle cutoff system can be understood for a while, but becomes painful in time.

Some Remarks on Testing Procedures

There is great need for getting down to fundamentals in the matter of testing procedures. All that have been reported have involved a series of multiple choice tests, with the procedure entirely beyond the control of the subject, and with no satisfactory explanation of what was going on. We may term this the "confused reaction". Unfortunately, radio set tone controls are not set by a series of multiple choice tests. The user handles the control him-

self, and he hears the effect as he changes it. Cause and effect are always obvious, and he has a much clearer understanding of the whole matter. We may term this the "considered reaction". The dealer may even make some educational comment on the effect of the control in question. It is perfectly true that the confused reaction is much easier to elicit, and that group testing is much cheaper. On the other hand, the greater significance of the individual test would seem to compel its use, if only for validating purposes.

In setting up a test, be sure that the program source is suited to the work. Only Olson[3] has paid attention to the acoustical conditions of the studio and the listening space. In Chinn and Eisenberg's work[4], for example, we are informed only that studios were used for each. But the standard broadcast studio is, intentionally, very bright. To originate and reproduce under the same abnormal conditions will produce an abnormal result. Increased band width makes exacting demands on every part of the system. Any study in which the acoustical conditions are not well defined is likely to lead to an inaccurate conclusion.

Be sure that program balance is correct for the band width employed. A large number of programs are acoustically balanced to sound well through a $19.95 mantel set, and when exposed to a wide range system the results are very painful. A similar problem is that of choosing the proper equalization in reproducing phonograph records. It is impossible to accept the results of a record test unless data on the equalization, pressing stock, amount of previous use of the records, etc., are supplied. The only published study of this nature is notably deficient in its facts, but the many unpublished studies seem to be equally faulty.

It is particularly necessary to be sure that the program material is as free from distortion as possible. Because of faulty processing many phonograph records have a great deal higher intermodulation distortion than would be indicated by measurements on the recording equipment alone. For example, one manufacturer's lacquer originals show three to four per cent intermodulation, while his production pressings run ten to twenty per cent. This is not necessarily characteristic of all American recording today, but it is typical of a certain class of organization. Pickups are a prolific source of distortion when abused, or when the stylus tip becomes worn. Receiving set demodulators are another potential source of trouble. Set designers have paid too little attention to this problem, so rigorous tests must be made before a tuner is used for listener preference testing.

Some programs do not require wide range to give excellent presence. The accordion is an excellent example of such a source. Other programs that do require wide range for presence, may irritate the listener so much that he prefers narrow range reproduction to ease the pain of listening. Certain ultra-modern forms of music unquestionably fall in this class. Therefore, it is a wise precaution not to attempt to test the popular music devotee with the three B's or vice versa. This is a legitimate restriction because each will normally listen only to his own preferred type of music.

One type of test, which should be more often considered, is that in which the response curve is shaped. We listen to a set at a rather lower acoustic level than that which exists in the auditorium or dance hall, and it is well known that this forces the ear to operate over a different part of the Fletcher-Munson curves[6], so that it loses low frequency sensitivity. A certain amount of bass boost is therefore reasonable. Likewise, it is not necessary to eliminate the higher frequencies to reduce the offensiveness of a small amount of distortion. A small amount of high frequency attenuation is quite effective. There is a certain tendency, amongst those with just enough knowledge to handicap them, to criticize high frequency attenuation as a fatal defect. The contrary is true. A few db of attenuation do not shift a tone below the threshold of hearing; they merely reduce its importance to the whole. We regard this expedient as an excellent interim way to render slightly distorted wide band reproduction more palatable.

Difference Limens

Gannett and Kerney made a study[5], of the discernibility of changes in program band width, which should be read by everyone interested in listener preference testing. Their tests were made with experienced listeners, and were designed to measure solely the ability to *discern* a change in band width with no question of preference involved. In testing with ordinary unskilled listeners, with both discernibility and preference involved, we would expect a less certain reaction in all cases. If we find a more certain reaction by unskilled listeners, then we may fairly conclude that there is something wrong with the program material in the tone band which is being thus forcibly rejected.

In order to compare the Gannett and Kerney results with those of Chinn and Eisenberg, it is necessary to simplify the latter's figures. Chinn and Eisenberg permitted a listener vote for either band width, and also a vote of no preference, whereas Gannett and Kerney permitted only an opinion on which of two band widths was being presented.

A vote of "no preference" is rather indefinite, it can mean either that (1), the listener could detect no difference or (2), the listener could detect the difference, but found both presentations equally acceptable. It is fair to assume that if the "no preference" choice were eliminated, those who had previously so voted would have to choose one of the remaining groups by guessing—which would put half of their votes in each group.

If we apply this technique to Chinn and Eisenberg's 5000 versus 8000 cycle test pair, we get the following orthodox

[*Continued on page 32*]

This group-accommodation studio "C" 26' x 40', is located on the ground floor of KOMO'S new broadcasting station, directly adjacent to the central operations corridor, which is entered through the door in the right background.

If you tried the recent Pictures at an Exhibition album by Horowitz, you'll find this one as good piano recording, as remarkable playing—though the music couldn't be more different. These are harpsichord pieces, for double keyboard, with the wide range of tone colors the harpsichord can give with its registration (like an organ). Most pianists just make them sound silly and precious; Horowitz makes them big, massive, as well as brilliant, as they sound in the original. (The sonatas are one-movement, varying from three or four minutes down to thirty seconds or so. Remarkably concentrated music, with plenty to it.)

QUALITY REPRODUCTION

[from page 11]

results, with all data in per cent:

Classical Music

Band	Average Listener	Musician	Gannett & Kerney
60-8,000	40	46	9
120-5,000	60	54	91

Male Speech

Band	Average Listener	Musician	Gannett & Kerney
60-8,000	64	76	80
120-5,000	36	24	20

Note that in the classical music group the choices are much more evenly balanced than Gannett & Kerney's discernability data. In the male speech group we have the first foretaste of special significance: the musicians are choosing the wider band almost as certainly as the probability of detectability would permit.

Applying the same method to the 8,000 and 10,000 cycle pair, we get a very interesting result:

Classical Music

Band	Average Listener	Musician	Gannett & Kerney
20-10,000	22	12	30
60- 8,000	78	88	70

Male Speech

Band	Average Listener	Musician	Gannett & Kerney
20-10,000	33	38	38
60- 8,000	67	62	62

In the music group it should be noted that the two bands are separated by less than one difference limen (one limen is defined as a 75:25 choice), but both average listener and musician show a vote separation of greater than one limen. They are detecting a difference more certainly than the experts were able to do (with good program material). We are forced to conclude that the music spectrum between 8,000 and 10,000 cycles was tagged by distortion products which marked it so strongly that average listeners were able to detect it with great certainty.

In the male speech group we see the same effect, to a lesser degree, in the average listener results.

This is a confirmation of our statement that the public will reject greater bandwidth with vigor if the added range is flavored with distortion products.

Objectives

Quality improvement can have only two commercial objectives: to increase the sale and use of radio sets, amplifiers, and components. It is obvious that there is a point of diminishing returns, and that going beyond it will be exceedingly unprofitable.

It is clear that a set with greater band width than the customer wants can still be used by turning down the tone control, whereas a set with inadequate band width built in cannot be so provided with band spread. On this basis we can estimate from the Chinn and Eisenberg data what would be the relation between upper cutoff frequency and the proportion of satisfied listeners.

Upper Cutoff Frequency	% of Satisfied Listeners
5 kc	50
8 kc	90
10 kc	99
12 kc	99.9

This assumes a system which is relatively free from fatigue producing factors.

We conclude from this table that

about one-half the radio sets and phonographs made should have satisfactory audio response up to 8 or 10 kc. If they are to use this bandwidth satisfactorily, a number of changes in component, broadcasting and phonograph record practice will be necessary.

Broadcast network lines will have to be improved. It is absurd to have studio input systems whose response is uniform to 15 kc, and transmitters whose response is in the same class, coupled by long lines which cut off sharply at 5 kc. Available are 8 kc lines, at an average increase of only about $300 per month per station. In many parts of the country today, the night time listener has available only 5-kc network programs.

The broadcaster who uses discs extensively will have to pay a little attention to the quality of his reproducers and records. Damaged reproducers or worn reproducer styli can cause intermodulation distortion of the order of 30%. Badly worn records, and records cut many years ago with extremely poor quality, are still in use. One has only to listen to discover that record reproduction is on a very low quality level in too many stations.

A few network broadcasters have taken the extraordinary step of using cheap, home-type wire recorders for their delayed broadcasting. We can see no possible excuse for such a degradation of quality.

Phonograph recording quality for 10 years has been adequate to meet the challenge, but many phonograph records cannot. We must make the production pressing nearer to the test pressing in quality.

Some manufacturers of inexpensive loudspeakers must mend their ways. Little has been done and nothing has been published on loudspeaker art can be traced directly to the secrecy tradition and lack of exchange of information by those in loudspeaker engineering. It is noteworthy that the BBC has published several times as much information on loudspeaker engineering as that released by all the loudspeaker manufacturers in America. The tricks of the trade will have to be regarded for what they are; "tricks" in the worst sense.

Finally, the set manufacturer will have to pay more attention to the audio part of his design. The present practice of having one of the technicians casually throw an audio system onto the tail end of an elaborately designed r-f circuit will have to stop. This policy seems to be at its worst in popular priced television sets.

Conclusion

This paper can best be summarized by saying that the typical set engineer

is very wrong in thinking that the auditory system is easy to deceive, and that perpetrating an acoustic fraud upon it will have no repercussions. Whereas the eye rebels very fast at unsatisfactory conditions, the ear is slow to anger. Even when very angry, it does not directly reveal the cause of its rage. Yet, in the end, it enforces its desires surprisingly well. Every time a listener yawns and turns off his set, his ears have won a victory.

Reference

[1] H. E. Roys, Intermodulation Distortion Analysis as Applied to Disc Recording and Reproducing; *Proc. I.R.E.*, vol. 35, no. 10, pp. 1149-1152, Oct., 1947.

[2] D. E. L. Shorter, Loudspeaker Transient Response; *BBC Quarterly*, vol. 1 no. 3, pp. 121-129, Oct. 1946.

[3] Harry F. Olson, Frequency Range Preference for Speech and Music, Talk before *Acoustical Society of America*, May 9, 1947.

[4] H. A. Chinn and Philip Eisenberg, Tonal Range and Sound Intensity Preferences of Broadcast Listeners; *Proc. I.R.E.*, vol. 33, no. 9, pp. 571-581, Sept. 1945.

[5] D. K. Gannett and Iden Kerney, The Discernability of Changes in Program Band Width; *BSTJ*, vol. 23, no. 1, pp. 1-10, Jan. 1944.

[6] Harvey Fletcher and W. A. Munson, Loudness; Its Definition, Measurement and Calculation; *Journal Acoustical Society of America*, vol. 5, no. 2, p. 91, Oct. 1933.

— Letters —

[from page 6]

inverters and the application (and effect of, etc.) of negative feedback.

In this country we were using amplifiers after the style of your 6AS7 one before the last war. The amplifier on page 19 Sept. '48 issue was originally designed here in England and is a good example of our standards. Very few designers now use tetrodes in HiFi amps; triodes are almost invariably used. Most amps have an output of 10-15 watts, less than 0.1% harmonic distortion and are flat within 0.1 db over at least 20-20,000 cs. A damping factor of at least 10 is considered essential, some (the one mentioned above for instance) have an output resistance as low as 0.5 ohm on 15 ohm winding, a damping factor of 30! Most HiFi systems have a "main" amp loading with about 1-5 volts input which is preceded by suitable "pre" amps.

The "main" amp is left "wide open" as far as frequency response goes, recording correction is done in the "pre" amp. Some people always want to "help" the composer but the majority leave the response alone, once set. The tone control controversy also rages in this country and we have many queer ways of making the response as nonlinear as possible.

Corner horns and acoustical labyrinths are now in common use, also bass reflex and a few infinite baffle types, also we