A

APPENDIX A

LDR TEST PRINCIPLES AND METHODOLOGY

Lucent Digital Radio, Inc.

20 Independence Blvd

Warren, NJ 07059 USA

LDR recognizes that the test methodology used to evaluate disparate systems is of great importance to perform a fair comparison and to adequately reflect consumer acceptance of these systems.

LDR's test methodology is designed to compare the LDR IBOC system with other digital broadcast systems and with the present analog systems. It includes four major components that are of great importance for obtaining accurate results: (1) proper sampling and characterization of the analog receiver population, (2) audio clip selection and processing prior to transmission, (3) using subjective audio testing which is the accepted "golden" method for audio evaluation, (4) judicious selection of the signal and impairment conditions for laboratory testing so that they adequately represent real life scenarios. Field tests are used for final verification of the lab results.

Following is the summary of the LDR approach that we believe results in the most accurate and realistic tests and that correlate well with the consumer preference.

1. Sampling of the population of the present receivers

In comparing digital systems to analog systems, special consideration must be given to the embedded base of the analog receivers.

LDR recognizes that while digital system performance can be predicted quite accurately based on the system parameters such as error correction capability, noise figure, *etc.*, an analog system

needs to be considered relative to the many types of the existing receivers. This requires that the analog system performance be evaluated based not on a simulated performance of an idealized (or "typical") receiver, but rather on the measured performance of a representative sample of the population of existing receivers.

The receiver selection for subjective tests should be based on an objective evaluation of a large representative sample of the population, computation of the receiver statistics, and selection of a smaller number of units that best represent the computed statistics. Such sampling process is necessary to reduce the number of receivers that have to be used for very tedious subjective tests, and still adequately represent the receiver population.

In evaluating the receiver statistics, LDR observed there are large differences in performance between various receiver categories (home, auto, portable) and considerable variance within these groups. For example, according to a CEMA report [1], the first adjacent channel rejection can vary by 30dB among various receiver categories and 10dB within the automotive receiver category alone (in a relatively small group of receivers consisting of 16 receivers in total including five automotive receivers).

Consequently, LDR expects that significantly different test results may be reported by different laboratories depending on the receiver selection process. Care must be taken in extrapolating the results to predict the system performance as it relates to the listening experience of a broad population of consumers.

2. Uniform subjective quality measure – need for subjective testing - ACR

[1] Consumer Electronics Manufacturing Association, 1999, *"FM Receiver Interference Test – Laboratory Test Report"*

LDR recognizes that comparison between different systems under a variety of conditions is best performed as a series of subjective tests with uniform test criteria that *directly compare the systems on the same scale*. Subjective tests have two clear advantages: (1) they allow us to directly relate the overall signal quality to consumer preference, (2) they allow us to obtain *quantitative* results in certain impairment conditions where the objective tests cannot be performed or are very unreliable. Thus, subjective tests consolidate a great number of disparate performance characteristics in an objective form that can be more efficiently used in a decision making process.

For example, the quantitative subjective tests may be used to develop a *quality distribution* measure that relates the quality of the received audio to the distance from the transmitter. Such distribution provides information on the *aggregate listening experience* within the entire area of coverage.

The figure below illustrates an example of quality distribution for two systems.

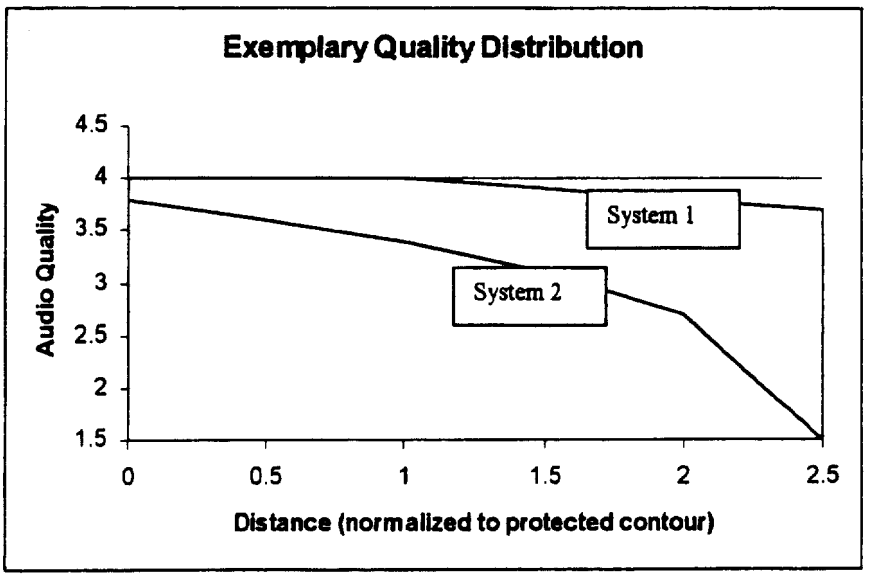

Figure A-1: Exemplary quality distribution chart

The figure shows two quality distribution curves, one for each system. The initial values represent clean (virtually unimpaired) signal quality, thus they relate to fidelity measure. As receivers are removed away from the transmitter, the signal degradation is reflected in diminishing quality. The shaded area represents the aggregated difference between the listening experience in the entire area.

There exists broad consensus that the subjective tests are the most compelling and accurate methodology for the purpose of comparing disparate systems under a variety of conditions (for example, the voice quality of digital cellular and PCS systems are tested subjectively). However, they are also the most time-consuming, expensive and, historically, have not been applied in the broadcast industry. This is one of the reasons why the quantitative performance data for analog receivers in fading environment is not widely available and is largely anecdotal and, consequently, not accurate. LDR has made the necessary investment to perform subjective tests that result in quantitative quality measure for signals with various impairments. LDR has used the subjective test methodology called Absolute Category Rating (ACR) Mean Opinion Score (MOS) for all tests, except for a few, well justified cases.

More details on the ACR subjective testing methodology used by LDR is provided with the subjective test results (Appendix F).

3. Selection of the sound samples

It is important to perform the evaluation in a way that reflects both the consumer preferences and likely acceptance of the system. In particular, the audio samples for the evaluation should

represent a variety of the broadcast material. LDR recognizes that certain sound samples can be designed to expose the soft spots of digital systems and are excellent tools for laboratory diagnostics. However, the consumer acceptance is more accurately predicted with audio material that is both strenuous and representative of commonly played material. In addition, when comparing an analog system with a digital system, it is important to appropriately preprocess the analog signals prior to transmission.

LDR has selected the sound samples that represent actual broadcast material and at the same time are demanding in the performance of both analog and the digital systems. In selecting the sound samples, appropriate weight has been given to the proportions of the broadcast content (*e.g.* classical, rock, jazz, talk, *etc.*). All sound samples were processed before the transmission by the appropriate sound processing (using professional high quality studio equipment) as it would normally be done according to the particular sound/station profile.

More details on the selection of sound samples and processing performed by LDR are provided with the subjective test results (Appendix F).

4. Evaluation methodology – choices of the test cases

The evaluation performed by LDR is based on a combination of field tests, lab measurements, emulation, simulation, and data analysis, with strong emphasis on the laboratory measurements using fading simulators.

In the scientific and engineering community, it is generally accepted that field-testing of a mobile radio system is time-consuming and often inconclusive, due to uncertainty in the

statistical signal variations. Therefore, the field tests are a final checkpoint that is used after extensive lab tests. In contrast, laboratory tests can be precisely controlled by using fading simulators, interference generators, etc., which can be programmed to produce a variety of channel conditions. The laboratory test setup not only can produce the signals that well represent the actual impairments encountered in a variety of terrain and man made environments, but also permit a precise and quantitative characterization of the mobile receiver that can be used for absolute and relative comparisons. Furthermore, the precise data collected in lab tests can be interpolated and extrapolated to determine the performance in cases that are not explicitly covered by the tests. LDR believes that systems should be compared on a common test setup under judiciously selected set of conditions representing the critical signal conditions and impairments. The system that performs better under these conditions will also perform better in the field. This applies to both analog and digital systems.

For these reasons, LDR's tests focus mainly on laboratory tests using fading simulators, noise and interference generators. Furthermore, certain extrapolation of the results was done through careful analysis. Checkpoints were provided to verify the analytical results by comparing with measured results in selected points.

The critical test cases include impairments due to fading, co-channel and first and second adjacent channel interference. The tests were performed under various signal strength (signal to noise ratio) conditions, both within and beyond the protected contour, up to the receiver sensitivity limit. Unimpaired (clean channel) performance was also tested for the reference. In addition, special attention has been given to compatibility issues. In particular, the host compatibility and the first adjacent channel compatibility were extensively tested and analyzed.

At present, LDR is progressing with more measurements and field tests and further refinement and more complete results will follow.

5. Evaluation steps - highlights

1. Collect many receivers representing, in relative proportion, different receiver groups and manufacturers.
2. Measure the performance of each receiver using objective criteria (*e.g.* sensitivity, first adjacent channel interference rejection, *etc.*)
3. Compute the statistics: average the best 25% of the receivers separately for non-mobile and mobile receivers and separately for all the key performance measures (*e.g.* first adjacent rejection, second adjacent rejection, sensitivity, *etc.*)
4. Select a small number of receivers from each group (mobile, portable, home) that corresponds to the average in the group or the average of the best 25% receivers for this group.
5. Using subjective tests, evaluate the performance of the selected receivers with the impairments related to the specific performance measure used for the selection (*i.e.* if the receiver was used as representing an average performance for the adjacent channel interference rejection, use this receiver to characterize the average receiver performance in the presence of the adjacent channel interference). The performance evaluation is done separately for each impairment condition, using variety of sound samples and a variety of receivers. The data is tabulated for each impairment condition, for each receiver and each

audio sample. Finally, for better illustration, the results are processed and an average performance is computed.

6. Performance versus distance from the transmitter will be plotted based on signal propagation, according to the FCC propagation curves. In the illustrative example, the FCC propagation curves were used that reflect typical signal propagation loss assuming protected contour radius of 80km and antenna height of 200m. Accordingly, the approximate power of the desired signal is assumed +54dBu at the protected contour distance, +36dBu at twice the distance, +22dBu at three times the distance, and +10dBu at four times the distance.

© Lucent Digital Radio, Inc.

B

APPENDIX B

SELECTION OF AUDIO SAMPLES & FM PROCESSING

Lucent Digital Radio, Inc.

20 Independence Blvd

Warren, NJ 07059, USA

© Lucent Digital Radio, Inc.

Introduction

LDR uses general population subjective listening tests as one tool to evaluate its PAC and IBOC technologies for broadcast radio applications. The test audio cuts – known variously as the *program material*, *subjective audio*, and *audio samples* – are the primary stimulus in these listening experiments. As such, every aspect of the test cuts, including content, length, editing, volume and processing, may have a first order effect on the precision and accuracy of subjective audio testing. This is evidenced in the treatment the Radiocommunication Assembly (ITU-R) gives to program material selection in its recommendations on subjective audio testing. Presented here is the rationale for the selection and processing of audio samples used in LDR's recent FM and PAC audio subjective testing.

Selection of Audio samples

As the ITU-R suggests in *Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems*[1], critical audio material is necessary for effective comparative testing of audio transmission systems. Used in this context, the term *critical* connotes the ability of the audio material to reveal the limitations of a system under test. Additionally, it is important that the selection of critical audio material include component samples that specifically challenge each system under test – though not necessarily at the same time.

[1] *Recommendation ITU-R BS.1116-1 (1994-1997)*

Additionally, the ITU-R recommends that the test material [2]:

1. May include any material that is potential broadcast content.
2. Should not distract a subject from the task of evaluation.
3. Be normalized for loudness.
4. Should not include specially contrived to "break" a particular system.

When considering general population comparisons of perceptually processed audio, LDR also believes the following:

1. Test material should substantially represent a significant range of broadcast material that is ecologically.
2. The critical nature of the test material should be typically found in broadcast material.
3. The test material source should be high fidelity, preferably digitally recorded, mastered and replicated to CD quality[3] or better.

Audio experts from Moulton Laboratories and Digital Media Services (Moulton Labs), Groton, MA and Lucent Technologies' Multimedia Perceptual Assessment Center (MPAC), Holmdel, NJ have helped LDR select critical program materials for general population testing. In accordance with these recommendations, the selected test audio represents the categories of broadcast

---

[2] *See Section 6, Recommendation ITU-R BS.1116-1*

[3] *Stereo format, 16 bits @ 44.1 kS/s per channel*

programming defined by combinations of male vocal, female vocal and instrumental performances within each of two program genres – classical and rock/pop. Two speech samples, male and female are also included. While these genres do not represent the total breadth of broadcast programming, each one is complex enough to overlap the nature and content of other types of music such as jazz, folk, R&B and urban music.

The selections made for testing are listed in Table B-1. Of the eight sample cuts, five are edited into two lengths. The length of a test cut used for a particular subjective test is based upon the design of the listening experiment. Differing lengths do not significantly alter the nature or content of the sample, but in some circumstances enable the subject to more quickly judge and report audio quality. The shorter length sample is a sub-sample of the larger.

© Lucent Digital Radio, Inc.

Code	Description	Source	Long Version (Seconds)	Short Version (Seconds)
FC	Classical Female Vocal	Moulton Labs, Critical Listening Excerpts CD, cut 3 (Kyoko Saito)	24	18
RI	Rock Instrumental	Jeff Beck, Who Else CD, cut 1 (*What Mama Said*)	22	15
CI	Classical Instrumental	Moulton Labs, Critical Listening Excerpts CD, from cut 11 (Bang & Olufsen Test Sequence – Britten: *Young Person's Guide to the Orchestra*)	27	20.5
MC	Blues/Jazz Male Vocal	Moulton Labs, Critical Listening Excerpts CD, cut 11 (Bang & Olufsen Test Sequence -- Robert Cray Band)	24	13
SF	Female Speech	From EIA SQAM CD, cut 48	6.5	N/A
FR	Rock Female Vocal	Susan Vega, Nine Objects of Desire CD, cut 3 (*Caramel*)	16	N/A
MR	Rock Male Vocal	Crowded House, Woodface CD, cut 5 (*Weather with You*)	21	21
SM	Male Voice-over	Gravity Music Library, cut 1 (*Introduction*)	12	N/A

Table B-1: LDR Subjective Testing Audio Selections

In Table B-2 the audio samples are related to their representative radio formats.

Code	Description	**Primary Formats**	**Sonically Similar Formats**
FC	Classical Female Vocal	classical	
RI	Rock Instrumental	rock, metal, alternative	fusion, urban
CI	Classical Instrumental	classical	
MC	Blues/Jazz Male Vocal	rock, jazz, R&B	country
SF	Female Speech	news, talk radio, educational	
FR	Rock Female Vocal	rock, pop, alternative	jazz, folk
MR	Rock Male Vocal	rock, pop, alternative	R&B, easy listening, rock oldies
SM	Male Voiceover	commercial advertisement, promotional	pop

Table B-2: Test Audio Selections by Format

The matrix of Table B-3 shows the pertinent audio characteristics of each audio cut that make it suitable as critical listening material.

Code	Description	**Primary Critical Characteristics**	**Secondary Critical Characteristics**
FC	Classical Female Vocal	Transparency, spatial ambience, dynamic range	Reproduction of piano
RI	Rock Instrumental	Complexity, instrumental transients, high bandwidth, stereo imaging	Clarity, critical low frequency content
CI	Classical Instrumental	Instrumental complexity, reproduction of traditional, orchestral instruments	Orchestral dynamics
MC	Blues/Jazz Male Vocal	Clarity, vocal transients	Vocal dynamic range, bandwidth
SF	Speech Female	Sibilance, intermittent silence, high peak to average ratio	Vocal transients
FR	Rock Female Vocal	Vocal complexity and detail, discrete stereo imaging, tightly balanced vocal-instrumental arrangement	
MR	Rock Male Vocal	Moderate studio processing, electronic instrument content	Modern vocal processing
SM	Male Voiceover	Voiceover intelligibility , heavy vocal processing	Heavy pop instrumental processing

Table B-3: Critical audio properties of audio selections

Processing of audio samples

General Editing

For use ease of use in editing, reproduction and storage, LDR engineers converted all sample audio cuts from CD audio format to 16-bit @ 44.1 kS/s stereo wave (*.WAV) files. In this format, the editing engineer has used Syntrillium's Cool Edit Pro, a software editing tooling, to precisely cut each sample sound file. The editor has removed any abrupt truncations at beginnings and ends of each file using envelope shaping. This shaping has been confined to the initial one-half and final one seconds of the sample.

Analog FM Processing

LDR's IBOC laboratory test bed replicates the performance of high performance AM and FM stations and provides careful control of RF signal levels delivered to the analog and digital radios under test. To meet the latter condition, LDR's test engineers used low-noise, synthesized RF signal generators in lieu of commercial AM and FM exciter units. These generators not only facilitate consistent control of desired and interferer signal levels across the entire AM and FM broadcast frequency ranges, but also ensure low noise and low distortion host modulation. To closely simulate state-of-the-art FM audio transmission in the laboratory, LDR uses an ORBAN model 8200 FM processor for all desired-signal FM baseband processing. This unit performs multiband audio program processing as well as generation of the composite stereo baseband signal. In a typical test configuration, the analog composite output of the ORBAN 8200 directly drives the FM modulation input of the desired-channel RF signal generator.

The ORBAN 8200 provides pre-programmed audio processing settings designed to address the wide range of possible FM program formats. With the help of an MPAC audio expert, LDR

matched one ORBAN 8200 processing profile to each of the test audio samples for use in generating analog FM test signals. The MPAC expert evaluated each audio sample over a range of processor pre-set possibilities considered suitable for that audio cut's genre. For each sample the processing profile providing the best FM system transmission sound was chosen as the FM audio processing for that cut. For this process, the expert listener monitored the FM signal using a high-end home tuner[4], integrated audio amplifier system and Sennheiser HD-600 headphones.

Code	Description	**FM Audio Processing (Orban 8200 Setting)**	**Additional Treatment**
FC	Classical Female Vocal	2 dB Pur Classical	none
RI	Rock Instrumental	Rock Dense	intro tone burst
CI	Classical Instrumental	Classical	none
MC	Blues/Jazz Male Vocal	Jazz	intro tone burst
SF	Speech Female	News/Talk	none
FR	Rock Female Vocal	Jazz	intro tone burst
MR	Rock Male Vocal	Rock Open	intro tone burst
SM	Male Voiceover	Pop Medium	intro tone burst

Table B-4: FM Processor settings

A crucial function of audio source processors like the ORBAN 8200 is control of the FM signal's RF spectral occupancy. This is accomplished through transient and level control in the source

[4] *Denon Model TU-1500RD*

audio. The processor softens attack envelopes, may compress the audio dynamic range and also may perform composite FM signal processing. In a sophisticated device such as the 8200, this processing is often very complicated and characterized by many parametric operations in multiple frequency bands. Most of the test samples are cut from the middle of the parent audio passage and, as a result, rapidly transition from lead-in silence to an appreciable faction of full-scale volume. Thus, the very act of sampling the audio material creates a boundary condition in the test material that may cause a FM audio processor to modulate the sample's loudness. This effect could be heard clearly in a number of LDR's sample cuts when processed directly by the 8200. As this effect is an artifact, it should be removed, or at least mitigated. Although there are several solutions, the most straightforward is dynamically to "prepare" the processor. To accomplish this, each audio sample that exhibited an immediate, moderate to high volume level as it begins was prefaced with a two second, multi-tone burst at –4 dB total power relative to full scale. The length and frequency content of the burst was sufficient to set up a processing history in the 8200. This confined the 8200's response to a sample's beginning amplitude burst to the region of the multi-tone burst. The burst prefixes were removed from the sound samples before use in subjective testing. The final column of Table 4 indicates those subjective cuts that are processed with a lead-in multi-tone burst.

IBOC Processing

Other than that intrinsic to perceptual coding, no source program audio processing was used for transmission of FM IBOC signals.

Loudness Normalization

Experimenters normalized the volume levels of all test samples before use in listening tests. After an initial leveling, a second experimenter confirmed results by ear and adjustments were made as necessary. Within groups of audio samples generated from the same source test cuts, volume was matched to within less than 1 dB. Between audio samples originating from different source test cuts, relative volume was set less accurately, but within the same range of listening comfort.

Conclusion

With the recommendations of industry audio experts, Lucent Digital Radio has carefully selected and processed source audio samples for general population listening tests. These test cuts:

1. Represent a broad range of FM programming formats within eight different samples of realistic, broadcast-like content.
2. Contain critical content capable of stressing the capabilities of both IBOC and FM audio transmission systems without being contrived "system breakers."
3. Represent today's consumer standard for the highest audio quality, since most are from digitally mastered CD sources

For FM transmission, each test audio source has been matched with appropriate FM broadcast audio processing to ensure realistic laboratory simulation of high quality FM transmission. In addition, FM processor artifacts resulting from short sample duration have been minimized by adding multi-tone prefacing bursts to some cuts as necessary.

For FM IBOC transmission, the test audio samples require no processing beyond standard editing.

As selected and treated, LDR's test audio samples will support good methodology for general population testing of analog and digital broadcast audio.

C

US OFFICE PRODUCTS

APPENDIX C

ITS AUDIT REPORT ON LDR TESTING PROCEDURES

Lucent Digital Radio, Inc.

20 Independence Blvd

Warren, NJ 07059, USA

© Lucent Digital Radio, Inc.

U.S. DEPARTMENT OF COMMERCE
National Telecommunications and Information Administration
INSTITUTE FOR TELECOMMUNICATION SCIENCES
325 Broadway
Boulder, Colorado 80303-3328

Reply to the attention of: ITS.D4

January 20, 2000

Independent Observer's Report

The Institute for Telecommunication Sciences (ITS) of the U.S. Department of Commerce's National Telecommunications and Information Administration was funded by Lucent Digital Radio (LDR) under a Cooperative Research and Development Agreement to act as the independent observer of the testing of the LDR proposed technology for In-Band/On-Channel (IBOC) Digital Audio Broadcasting (DAB).

Laboratory Test

ITS inspected the test laboratory at LDR and found the equipment to be of good quality. It appeared that good engineering practice was followed in the assembly of the laboratory testbed. The calibration of the instruments by the manufacturers were kept up-to-date and documented. Careful preparation for the tests was evident. The test procedure began with a thorough, documented calibration check (+ or - 0.1 dB) of the instruments before each test. Test scripts to help the test personnel correctly make all necessary adjustments and equipment settings were used. Logs were kept of instrument settings and results obtained (audio recordings).

Reference audio clips were recorded during tests for subjective rating by expert listening panels. ITS inspected the audio listening laboratory at Lucent, Murray Hill, NJ and reviewed the rating procedures. An ITS staff person went through the training given to the expert listeners. The subjective audio listening facilities, equipment and procedures appeared to be scientificly sound.

ITS personnel were not present during the host compatibility laboratory tests but were present during a subsequent test. ITS inspected the calibration check document and the test log of the host compatibility laboratory test. The trained ITS staff person completed one audio listening test session in which he rated the relative quality of paired audio recordings from the host compatibility laboratory test.

Field Tests

ITS inspected the mobile receiving vehicle (van) and FM radio transmitter site used in the field tests. ITS observed LDR calibration of the receiving system including antennas and reviewed LDR calibration data for the transmitter. It appeared that good engineering practice was followed in the assembly and calibration of the measurement van.

Test Report

ITS reviewed the host compatibility test report. The report provides an accurate description of the test equipment, test setup, and test procedures.

KCAllen

Kenneth C. Allen, ITS
Principal Observer