Albert Shuldiner
Senior VP & General Counsel

July 27, 2004

Ms. Marlene H. Dortch
Secretary
Federal Communications Commission
445 Twelfth Street, S.W.
Washington, D.C. 20554

Re: MM Docket No. 99-325

Dear Ms. Dortch:

iBiquity Digital Corporation ("iBiquity") hereby submits for filing in this docket recent test reports about the audio quality of iBiquity's HD Radio™ system. These test reports were analyzed by the National Radio Systems Committee ("NRSC") and were the subject of a report recently submitted to the Commission on behalf of the NRSC.

In 2003, iBiquity modified the HD Radio system to include a new audio compression technology known as HDC. Because the performance of the audio compression technology determines the audio quality of the digital system, the NRSC asked iBiquity to conduct a new round of testing designed to demonstrate that the HD Radio system using HDC could provide better audio quality than that obtained using existing analog AM and FM technology.

The tests were designed to compare the HD Radio system against unimpaired analog audio using a variety of analog receivers and typical radio formats. The test program also included regression tests of the HD Radio system with impairments. These regression tests were designed to confirm that the modified system including HDC continued to perform in a manner consistent with the system that had been previously tested by the NRSC.

As the attached test results demonstrate, both the AM and FM HD Radio systems provide audio quality that exceeds the quality of typical analog broadcasts. The tests confirm that the FM HD Radio system operating at either 96 or 64 kbps provides audio that is equivalent to a CD. The AM HD Radio system operating at 36 or 20 kbps provides audio quality approaching the quality of unimpaired analog FM. General population listeners subjectively evaluating the audio collected during the tests consistently scored digital above analog confirming that the HD Radio system will provide listeners with an improved listening experience.

If you have any questions about this submission or these test results, please do not hesitate to contact the undersigned.

Sincerely,

Albert Shuldiner

8865 Stanford Boulevard • Suite 202 • Columbia, Maryland 21045 • Phone 410 872 1536 • Fax 410 872 1560

Subjective test results for FM IBOC DAB Generation 3 Hardware

Unimpaired and impaired test conditions

Report to the National Radio Systems Committee
and
iBiquity Digital Corporation

February 25, 2004

Ellyn G. Sheffield
Sheffield Audio Consulting
Princeton, New Jersey 08540
(609) 737-9451
egsheffield@comcast.net

1. Introduction

This report describes results from unimpaired and impaired FM audio quality testing conducted at iBiquity Digital Radio Corporation between December 1 and January 13th 2004 under the supervision of Dr. Ellyn Sheffield. The experimental program was designed to subjectively test FM Generation 3 hardware, including unimpaired transmissions (Test I), transmissions with AWGN (Test B.1) and with multipath impairments (Test B.2). Unless otherwise noted, all procedures were followed as identified in the test plan "Proposal for Subjective Evaluation of Generation 3 HD Radio Hardware" dated December 1, 2003.

In order to maintain consistency between this study and past subjective tests of the FM IBOC system, methodologies replicated previous testing procedures to the greatest possible extent. As in the past, (a) audio samples were delivered over Sennheiser HD-600 headphones; (b) Tremetrics sound booths and presentation software used previously for subjective testing at Dynastat, Inc. were used at iBiquity; and (c) samples originally picked for Generation 2 testing were again presented to listeners for rating.

2. Experimental Methodology

2.1 Sound sample preparation

Audio samples were recorded directly onto wave files at iBiquity. This procedure was witnessed by a designated NRSC observer. Sound samples were recorded in groups, based on their post-processor settings (For details, see iBiquity's report entitled "Digital Performance Regression Tests of the iBiquity Generation 3 Digital HD Radio System in the AM & FM Bands: Subjective Audio Evaluation Sample Preparation Procedure", dated February 26, 2004). The recordings were sent to Dr. Sheffield, who parsed them into individual samples. She then edited, leveled and named them. All of these efforts were done in accordance with procedures identified in the FM IBOC Dab Laboratory and Field Testing Report, August 2001, Exhibit 4: Procedure for Editing and Leveling Sound Samples. Sound samples were loaded onto test computers and experiments were created using iBiquity's subjective testing software.

2.2 Listener Sample

Data from 40 qualified listeners (19 males and 21 females) are included in this report. In order to qualify as a listener, participants needed to pass a screening test (see Section 2.2: Screening Procedures). Additionally, a post-hoc analysis designed to eliminate obvious outliers was conducted on each listener's data. Five listeners were disqualified from the sample for failing to pass the screening test, one was eliminated as a result of post-hoc analysis. Listeners were recruited from the Columbia, Md. area by word-of-mouth, sending flyers to local business establishments, colleges and universities, and from referrals by listeners participating in the test. Table 2.2 shows the distribution of participants by age and gender.

	Male	Female
18-24	6	5
25-32	5	5
33-42	5	4
43-50	4*	6

Table 2.2: Distribution of participants
*Includes 1 NRSC member, above the age of 50.

2.3 Screening Procedure

Listeners completed a questionnaire and received instructions on the specific tasks they were asked to perform (see Attachment 1: Experimenter Script – Gen 3 FM testing). Prior to testing, participants completed a screening test. Screening was conducted to ensure that listeners were reliably able to distinguish between samples that differed substantially in quality. In order to pass the screening test, participants needed to answer 5 out of 6 screening questions correctly. In a triple-stimulus, double-blind screening procedure, participants were asked to listen to 3 samples: a "Reference sample", and 2 additional samples. One of the samples was identical to the reference, the other was different. At the beginning of the screening test, listeners were shown how to register their answers, and were given one example of the task, which they were asked to complete. After this training period, they were told that they should proceed with the screening test. They were encouraged to listen to the samples as many times as necessary to make their judgment. Figure 2.3 is a schematic which depicts the screen presentation layout. Table 2.3 lists the samples used for comparisons:

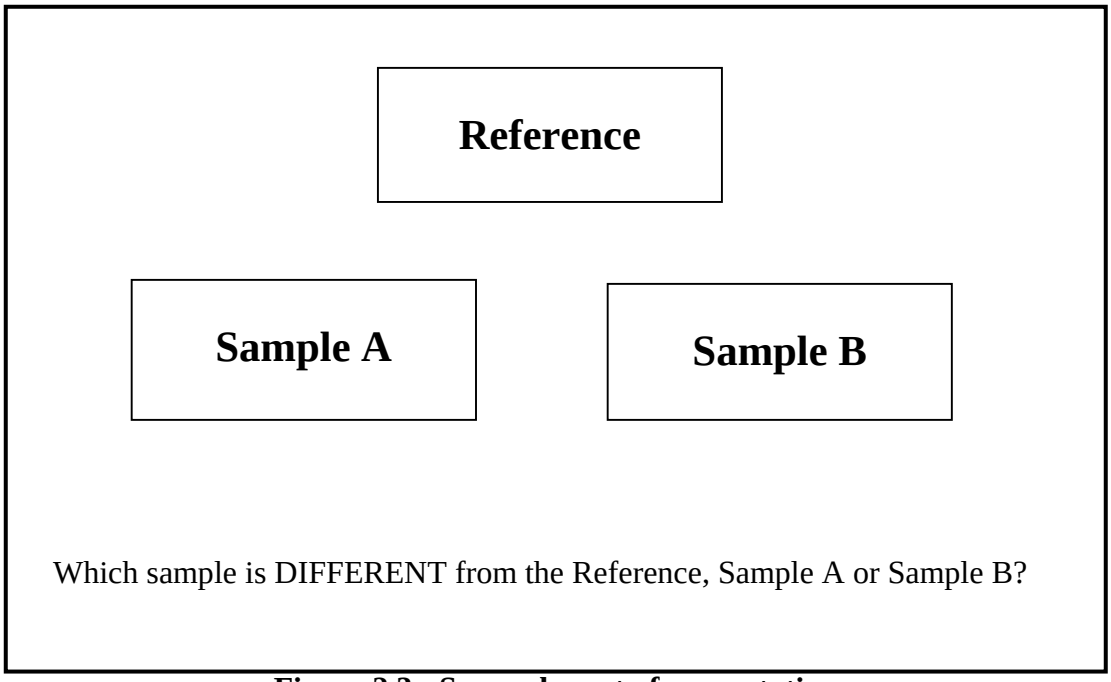


Figure 2.3: Screen layout of presentation

Sample	Reference	Sample A	Sample B
Speech – Woman	CD Source	CD Source	FM Rural Fast (Dephi)
Classical – Bach	CD Source	CD Source	FM AWGN (Delphi)
Speech – Brokaw	CD Source	Delphi Terrain Obstructed	CD Source
Rock – Clapton	CD Source	CD Source	AM Clean (Sony)
Rock – Travis	CD Source	HDC20kps	CD Source
Classical – 1812	CD Source	CD Source	FM Clean (Sony)

Table 2.3 Sound samples used in screening

2.4 Main Test

Listeners next participated in two ACR-MOS tests. The first test, comprised of 144 samples, exclusively included unimpaired audio samples. Participants listened to the samples, one-by-one, and rated each sample on its own merit on a 5-point scale (Excellent; Good; Fair; Poor; and Bad). They heard 48 samples, were given a 5 minute break, listened to another 48 samples, were given another break and then finished the test. After this period, participants were given a 10-minute break and were encouraged to walk around, relax, have a drink, etc. They then completed the second test, comprised of 54 impaired samples. (See Tables 2.4.1 and 2.4.2 for samples used during these tests). In order to minimize the effect of presentation order, the order of sample presentation was randomized uniquely for each participant, but each participant received all samples in each test.

Sample	Gen 3 - HDC96	Gen 3 - HDC64	Delphi	Pioneer	Tech	Sony	CD Source	AM
Amy Grant	1	1	1	1	1	1	1	1
Bach	1	1	1	1	1	1	1	1
Brokaw	1	1	1	1	1	1	1	1
Bizet's Carmen	1	1	1	1	1	1	1	1
Earth, Wind & Fire	1	1	1	1	1	1	1	1
Enya	1	1	1	1	1	1	1	1
Eric Clapton	1	1	1	1	1	1	1	1
Glockenspiel	1	1	1	1	1	1	1	1
Man	1	1	1	1	1	1	1	1
Medewski, Martin & Wood	1	1	1	1	1	1	1	1
Messiah	1	1	1	1	1	1	1	1
Paul Simon	1	1	1	1	1	1	1	1
Persian Music	1	1	1	1	1	1	1	1
Randy Travis	1	1	1	1	1	1	1	1
Saito	1	1	1	1	1	1	1	1
Tchaikovsky's 1812 Overture	1	1	1	1	1	1	1	1
Trumpet	1	1	1	1	1	1	1	1
Woman	1	1	1	1	1	1	1	1

Table 2.4.1: Unimpaired Test Samples

Sample	Impairment	Gen 1 – 96	Gen 3 – 96	Delphi	Pioneer	Technics	Sony
Brokaw	AWGN	1	1	1	1	1	1
Bach	AWGN	1	1	1	1	1	1
Prince	AWGN	1	1	1	1	1	1
Messiah	RF	1	1	1	1		
Fagen	RF	1	1	1	1		
Woman	RF	1	1	1	1		
Man	UF	1	1	1	1		
Cole	UF	1	1	1	1		
1812	UF	1	1	1	1		
Brokaw	TO	1	1	1	1		
Crowded House	TO	1	1	1	1		
Persian	TO	1	1	1	1		

Table 2.4.2: Impaired Test Samples

3. Results – Unimpaired Test Samples

3.1 Preliminary analyses

Preliminary analyses of variance (ANOVAs) were conducted to uncover any differences in participants' responses, based on their age and gender. There was no main effect of gender, indicating that females and males responded similarly to all transmissions (female MOS was 3.73; male MOS was 3.67). There was an effect of age, with younger listeners rating samples more critically than older listeners. This finding is not surprising and replicates results from past listening tests conducted for iBiquity.

3.2 Test results

The 5-point scale (Excellent through Bad) was translated into the following numerical values for analysis:

Rating	Numerical Value
Excellent	5.0
Good	4.0
Fair	3.0
Poor	2.0
Bad	1.0

Figure 3.2 shows MOS scores aggregated by genre. *Classical* includes Tchaikovsky's 1812 Overture, Bach, Bizet's Carmen, Enya, Handel's Messiah, and Saito. *Critical* includes the Trumpet, Glockenspiel, and Persian music. *Rock* includes Paul Simon, Randy Travis, Eric Clapton, Earth Wind and Fire, Medewski, Martin & Wood and Amy Grant. *Speech* includes the

Man, Woman and Tom Brokaw. In the classical and speech genres, participants rated HDC96 and HDC64 kbps significantly better than all FM analog transmissions. In the critical category, HDC96 and HDC64 were rated significantly better than the Sony, Technics and Delphi. Surprisingly, in all categories, the difference between the CD source, HDC96 and HDC64 was not statistically significant. Therefore, although participants heard a difference between the CD source material and FM analog transmissions, they never heard a difference between the CD source and the HDC transmissions. There were no differences between HDC96, HDC64 and FM analog transmissions in the Rock category. See Appendix 1 for individual ANOVA results and Newman Keuls post-hoc comparisons. See Appendix 3 for ratings of individual samples and error terms.

Figure 3.2 – MOS by Genre

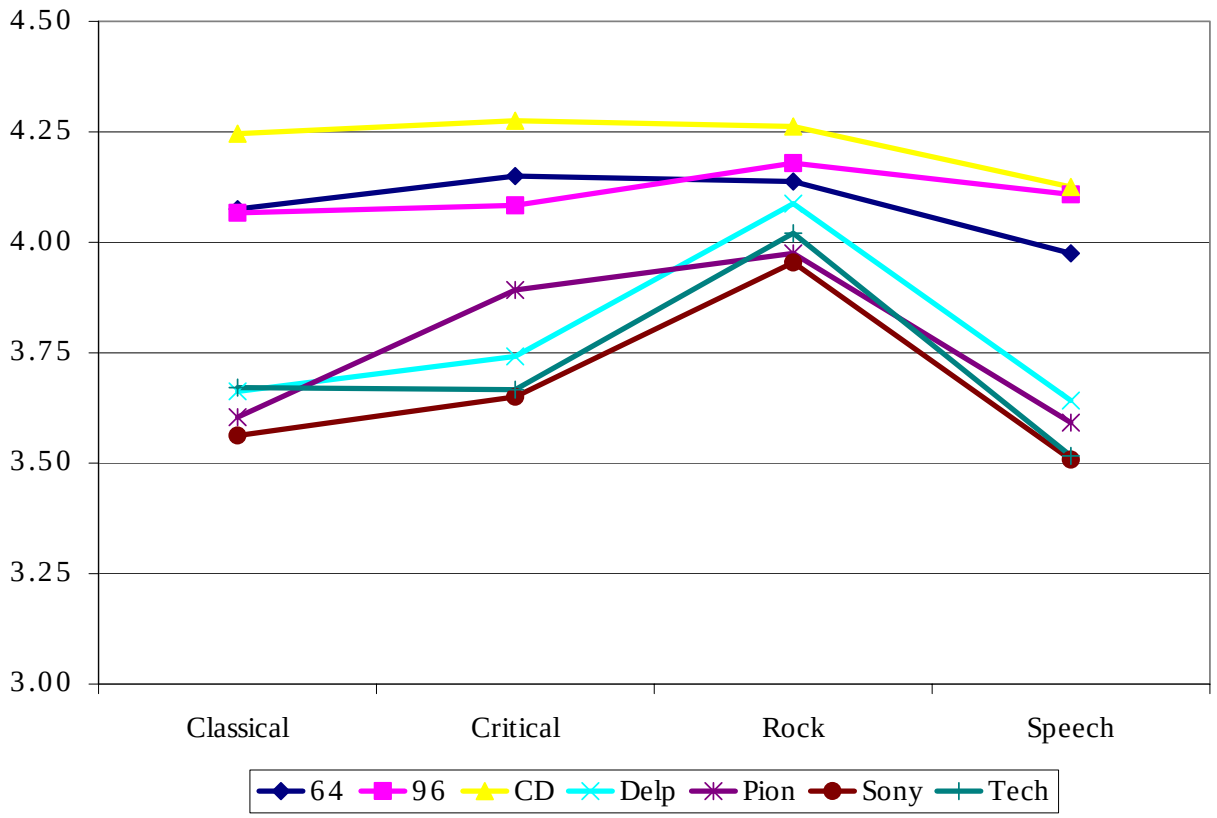


Table 3.2 lists all rating scores including low anchor samples (clean AM transmission). Results show that participants used the 5-point scale reasonably, scoring low anchor AM samples consistently in the "poor" range, CD source material in the "good" range, and all other transmissions between these points.

Receiver	Classical	Critical	Rock	Speech
CD	4.2	4.3	4.3	4.1
HDC96	4.1	4.0	4.2	4.1
HDC64	4.1	4.1	4.1	4.0
Delphi	3.6	3.8	4.1	3.6
Pioneer	3.6	3.8	4.0	3.6
Sony	3.6	3.6	4.0	3.5
Technics	3.7	3.7	4.0	3.5
AM	1.9	2.3	2.0	2.6

Table 3.2: Unimpaired MOS ratings by genre

4. Results – Impaired Test Samples

Figures 4.1 – 4.4 show results from impaired test conditions, divided by impairment type. These figures show participants' ratings of individual samples. In all cases participants rated HDC96 superior to FM analog transmissions. Additionally, in all impairment conditions except "urban fast", HDC96 and the Generation1 (96 kbps) performed identically. In "urban fast" Generation1 was rated slightly higher than HDC96. See Appendix 2 for individual ANOVA results, and Appendix 4 for ratings of individual samples and error terms.

Figure 4.1: Mean opinion scores (AWGN)

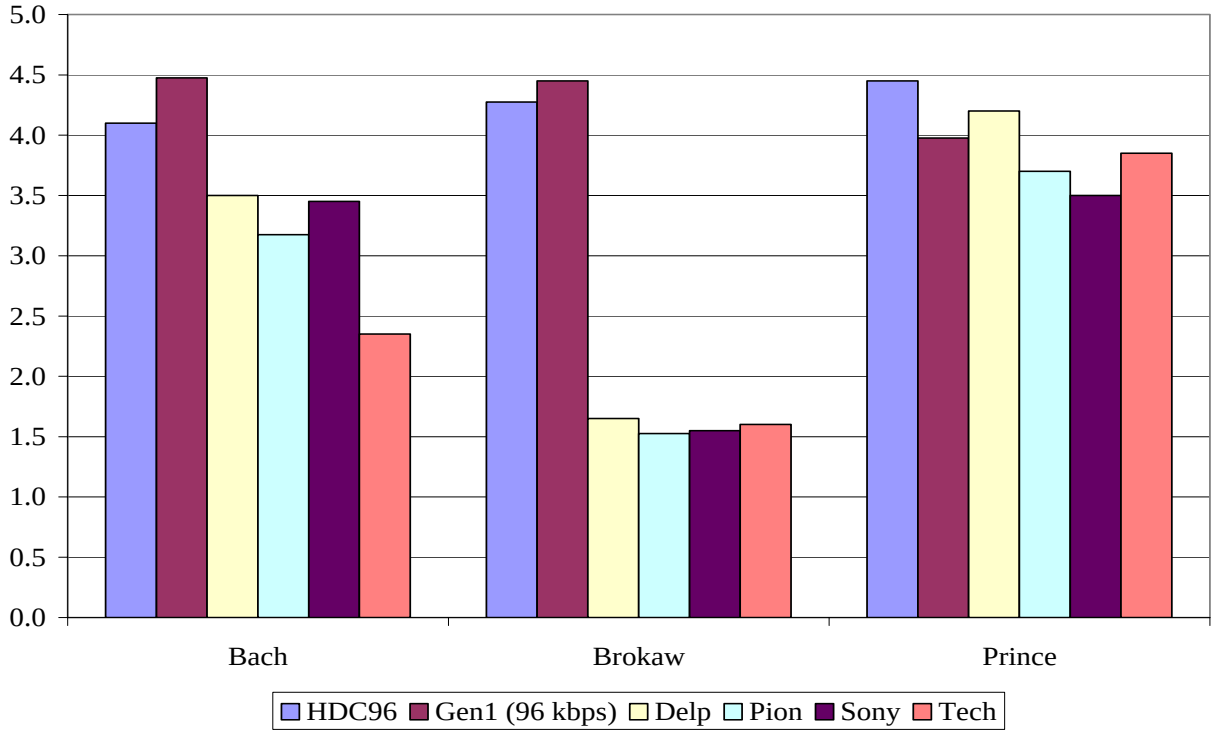


Figure 4.2: Mean opinion scores (Rural Fast)

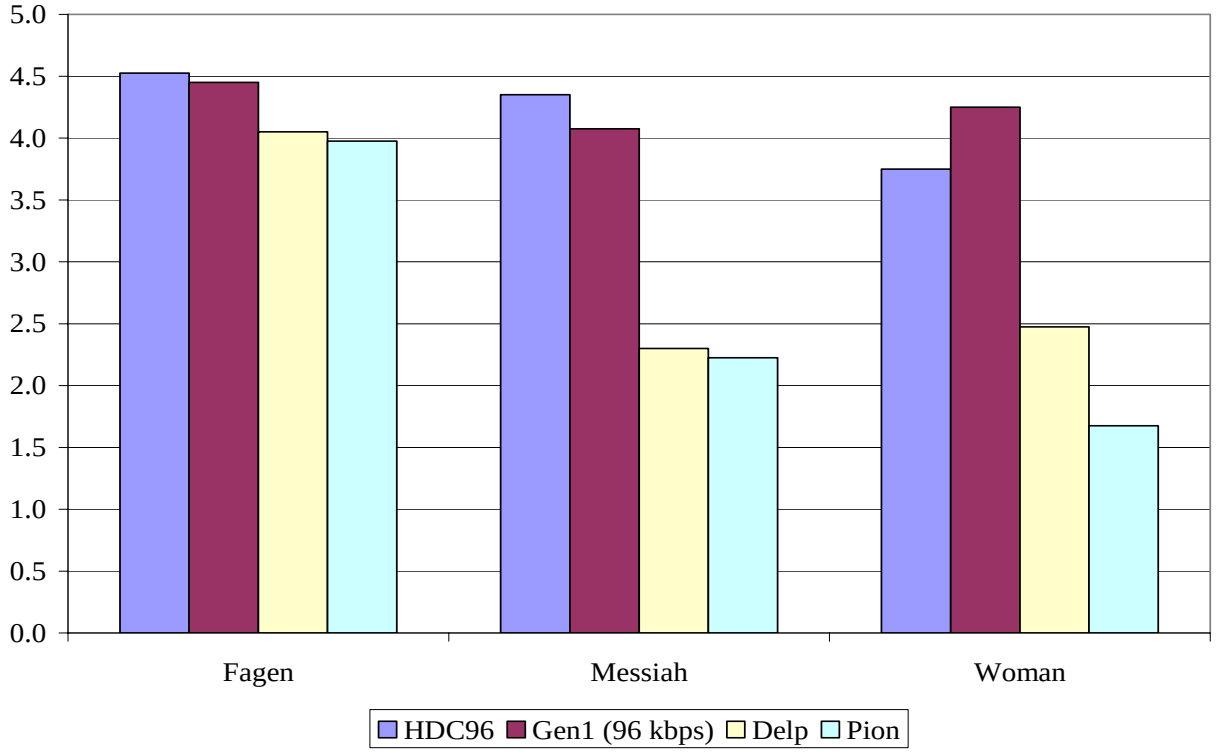


Figure 4.3: Mean Opinion Scores (Urban Fast)

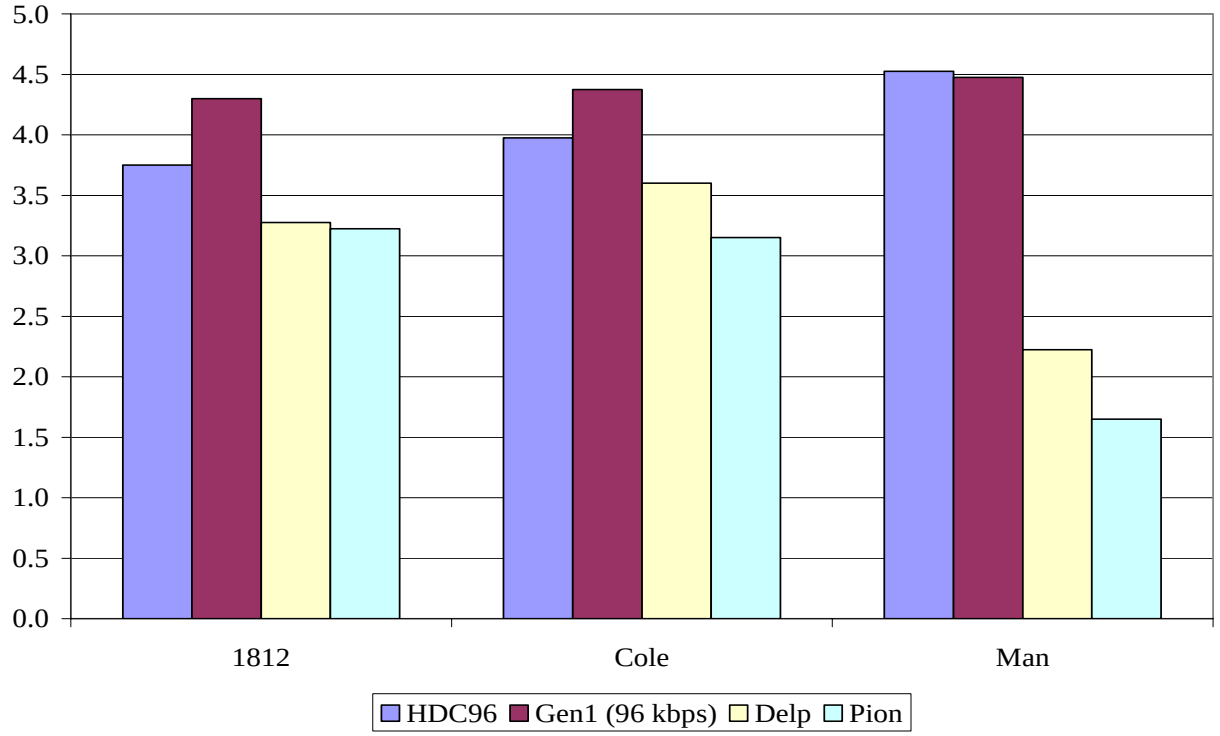
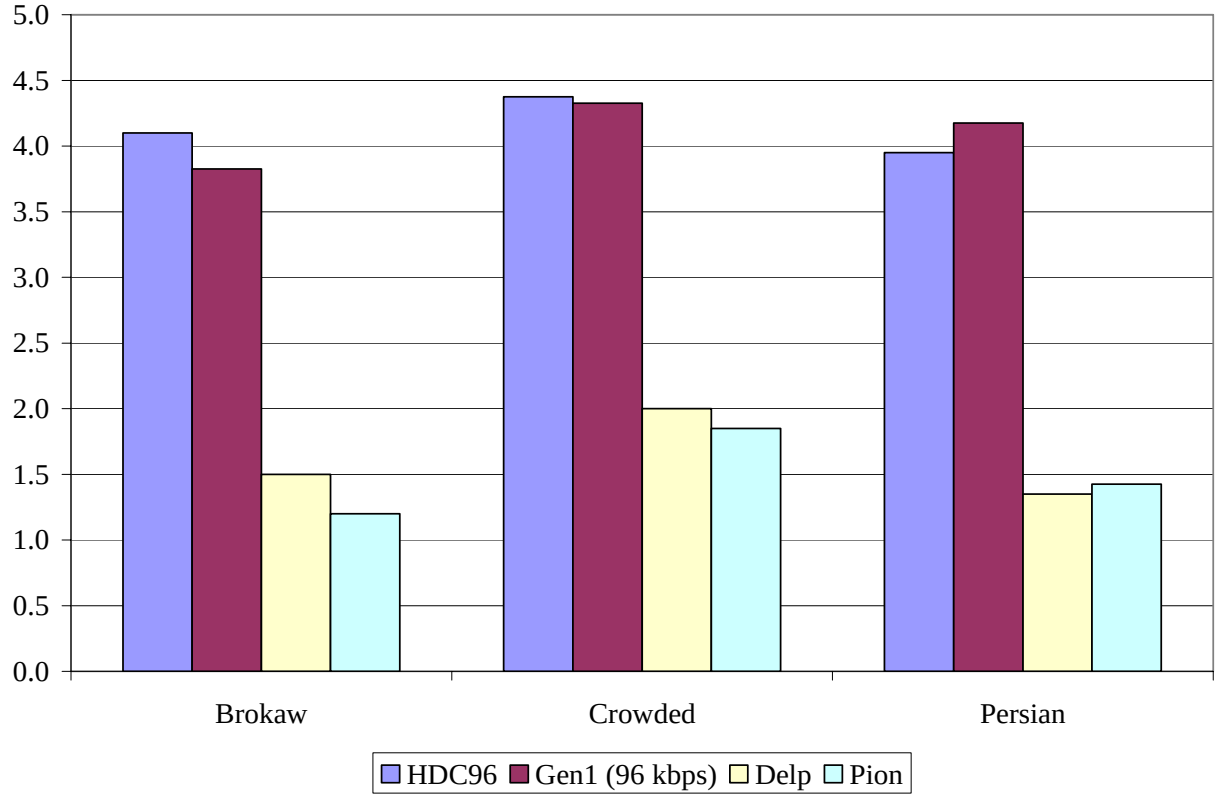


Figure 4.4: Mean Opinion Scores (Terrain Obstructed)



Attachment 1: Experimenter Script – Gen 3 FM testing

Welcome to our session. Today you will be participating in an audio test which should last approximately 2 hours. There are two parts to this test. In the first part of the test, you will hear a series of samples and compare them to a "reference". In this part you will play a Reference, an A and a B sample, and judge which of the samples is **different** from the reference. There will be six triplets to judge. In the second part you will first hear 144 short samples, in groups of 48. Please listen to the sample from start to finish. At the end of the clip you will be asked one question about it. After each session, the computer will tell you to take a 5 minute break. After listening to the 144 samples you will have a 10-minute break. That's your turn to go to the bathroom, have a drink, or just relax. In the final session you will be presented with another 56 samples. Once you start a session, you should continue until the program tells you to take a break, but you are also encouraged to take the test at your own pace. This may mean stopping between samples if you feel you need to "clear your head" for a few seconds.

For each session, you will be asked to rate the audio on a 5-point scale. In all cases, we want to remind you that we are not asking you to judge the material, or whether you like a particular cut. We know that you will have various feelings about the samples you are going to listen to. For this test, we are asking you to try to keep focused on only one thing: the quality of the transmission you are listening to.

Now we are going to begin. Any questions so far?

Appendix 1: Analysis of Variance Report for Unimpaired Samples

The following notes are intended to help the reader interpret the tables listed in this Appendix and Appendix 2.

(a) The Overall ANOVA allows us to see whether there is a statistical difference between any and all groups under test (i.e., HDC96, HDC64, Delphi, Pioneer, etc.)
(b) If the Overall ANOVA p-value < .05, then it is justifiable to use post-hoc Newman-Keuls Multiple comparison statistical tests to discern which specific groups differ from each other.
(c) The Newman-Keuls post-hoc tests are run at a p-value of 0.05, which is a standard p-value for multiple comparison statistical tests.
(d) The column to the far right, "Different from Groups" identifies which groups are statistically different from a group in question (the column to the far left). For example, in the first table, the CD source is different from AM, Sony, Pioneer, Delphi and Technics, but not different from HDC96 and HDC64.

Classical
Overall ANOVA: DF= 7,1920; F-Value=137.08; P-Value (alpha) =0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1912 MSE=0.878162

Group	Count	Mean	Different From Groups
AM	240	1.975	Sony, Pion, Delp, Tech, HDC96, HDC64, CD
Sony	240	3.5625	AM, HDC96, HDC64, CD
Pion	240	3.604167	AM, HDC96, HDC64, CD
Delp	240	3.6625	AM, HDC96, HDC64, CD
Tech	240	3.670833	AM, HDC96, HDC64, CD
HDC96	240	4.066667	AM, Sony, Pion, Delp, Tech
HDC64	240	4.075	AM, Sony, Pion, Delp, Tech
CD	240	4.245833	AM, Sony, Pion, Delp, Tech

Critical
Overall ANOVA: DF= 7,960; F-Value=45.20; P-Value (alpha) =0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=952 MSE=1.070842

Group	Count	Mean	Different From Groups
AM	120	2.25	Sony, Tech, Delp, Pion, HDC96, HDC64, CD
Sony	120	3.65	AM, HDC96, HDC64, CD
Tech	120	3.666667	AM, HDC96, HDC64, CD
Delp	120	3.741667	AM, HDC96, HDC64, CD
Pion	120	3.891667	AM, CD
HDC96	120	4.083333	AM, Sony, Tech, Delp
HDC64	120	4.15	AM, Sony, Tech, Delp
CD	120	4.275	AM, Sony, Tech, Delp, Pion

Rock
Overall ANOVA, DF=7,1920; F-Value =175.39; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1912 MSE=0.7847062

Group	Count	Mean	Different From Groups
AM	240	1.966667	Sony, Pion, Tech, Delp, HDC64, HDC96, CD
Sony	240	3.954167	AM, CD
Pion	240	3.975	AM, CD
Tech	240	4.020833	AM, CD
Delp	240	4.0875	AM
HDC96	240	4.179167	AM
HDC64	240	4.1375	AM
CD	240	4.2625	AM, Sony, Pion, Tech

Speech
Overall ANOVA, DF=7,960; F-Value =30.15; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=952 MSE=0.9181197

Group	Count	Mean	Different From Groups
AM	120	2.633333	Sony, Tech, Pion, Delp, HDC64, HDC96, CD
Sony	120	3.508333	AM, HDC64, HDC96, CD
Tech	120	3.516667	AM, HDC64, HDC96, CD
Pion	120	3.591667	AM, HDC64, HDC96, CD
Delp	120	3. 641667	AM, HDC64, HDC96, CD
HDC96	120	4.108333	AM, Sony, Tech, Pion, Delp
HDC64	120	3.975	AM, Sony, Tech, Pion, Delp
CD	120	4.125	AM, Sony, Tech, Pion, Delp

Appendix 2: Analysis of Variance Report for Impaired Samples

AWGN
Overall ANOVA, DF=5,720; F-Value =51.08; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=714 MSE=1.381174

Group	Count	Mean	Different From Groups
Tech	120	2.6	Delp, HDC96, Gen1
Pion	120	2.8	HDC96, Gen1
Sony	120	2.833333	HDC96, Gen1
Delp	120	3.116667	Tech, HDC96, Gen1
HDC96	120	4.275	Tech, Pion, Sony, Delp
Gen1	120	4.3	Tech, Pion, Sony, Delp

Rural Fast
Overall ANOVA, DF=3,480; F-Value =70.52; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=476 MSE=1.221639

Group	Count	Mean	Different From Groups
Pion	120	2.625	Delp, HDC96, Gen1
Delp	120	2.941667	Pion, HDC96, Gen1
HDC96	120	4.208333	Pion, Delp
Gen1	120	4.258333	Pion, Delp

Urban Fast
Overall ANOVA, DF=3,480; F-Value =82.23; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=476 MSE=0.9784139

Group	Count	Mean	Different From Groups
Pion	120	2.675	Delp, HDC96, Gen1
Delp	120	3.033333	Pion, HDC96, Gen1
HDC96	120	4.083333	Pion, Delp, Gen1
Gen1	120	4.383333	Pion, Delp, HDC96

Terrain Obstructed
Overall ANOVA, DF=3,480; F-Value =430.70; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=476 MSE=1.2236572

Group	Count	Mean	Different From Groups
Pion	120	1.491667	Gen1, HDC96
Delp	120	1.616667	Gen1, HDC96
Gen1	120	4.108333	Pion, Delp
HDC96	120	4.141667	Pion, Delp

Appendix 3: Ratings of individual samples in unimpaired condition

		CD	HDC96	HDC64	Delphi	Pioneer	Sony	Technics	AM
1812	MOS	4.2	3.5	3.4	3.2	2.9	3.0	3.3	1.6
	CI +/-	0.25	0.34	0.37	0.28	0.23	0.32	0.17	0.20
Bach	MOS	4.1	4.3	4.2	3.9	4.0	4.1	4.1	1.6
	CI +/-	0.25	0.25	0.24	0.26	0.28	0.31	0.11	0.20
Brokaw	MOS	3.7	4.0	4.1	3.4	3.3	3.3	3.4	2.6
	CI +/-	0.33	0.26	0.29	0.28	0.31	0.35	0.13	0.32
Carmen	MOS	4.4	4.2	4.4	4.0	4.0	4.2	4.0	2.1
	CI +/-	0.23	0.22	0.24	0.24	0.25	0.28	0.12	0.32
Clapton	MOS	4.7	4.3	4.5	4.5	4.4	4.4	4.2	1.8
	CI +/-	0.15	0.23	0.20	0.19	0.21	0.23	0.13	0.21
Enya	MOS	4.4	3.7	3.8	3.9	3.7	3.5	3.7	2.3
	CI +/-	0.21	0.35	0.27	0.26	0.30	0.29	0.13	0.35
Earth, Wind, Fire	MOS	4.1	4.1	3.9	3.6	3.8	3.8	3.7	2.0
	CI +/-	0.32	0.28	0.33	0.31	0.31	0.36	0.16	0.29
Glockenspiel	MOS	4.7	4.6	4.5	4.0	4.2	3.8	4.1	2.0
	CI +/-	0.20	0.18	0.25	0.29	0.23	0.33	0.15	0.37
Grant	MOS	4.3	4.2	4.0	4.2	4.2	4.1	4.2	1.8
	CI +/-	0.22	0.22	0.25	0.26	0.29	0.31	0.12	0.22
Man	MOS	4.5	4.6	4.1	4.0	3.8	3.7	3.7	2.9
	CI +/-	0.19	0.20	0.26	0.23	0.26	0.23	0.16	0.34
Messiah	MOS	4.3	4.4	4.5	3.6	3.6	3.4	3.6	2.2
	CI +/-	0.23	0.20	0.21	0.26	0.28	0.29	0.15	0.31
Medewski, Medin, Wood	MOS	4.1	4.0	4.0	4.1	3.9	3.6	4.1	2.3
	CI +/-	0.26	0.25	0.27	0.25	0.27	0.24	0.13	0.28
Persian	MOS	4.3	4.2	4.1	3.8	4.0	3.7	3.6	2.1
	CI +/-	0.28	0.31	0.36	0.29	0.31	0.27	0.18	0.27
Saito	MOS	4.0	4.3	4.3	3.4	3.4	3.3	3.4	2.1
	CI +/-	0.26	0.24	0.21	0.32	0.36	0.33	0.17	0.29
Simon	MOS	4.3	4.3	4.3	4.4	4.2	4.1	4.2	2.3
	CI +/-	0.27	0.27	0.20	0.27	0.29	0.32	0.15	0.34
Travis	MOS	4.2	4.3	4.1	3.7	3.6	3.8	3.9	1.7
	CI +/-	0.24	0.24	0.30	0.27	0.30	0.29	0.13	0.24
Trumpet	MOS	3.8	3.5	4.0	3.4	3.5	3.5	3.3	2.6
	CI +/-	0.34	0.31	0.28	0.32	0.41	0.40	0.16	0.34
Woman	MOS	4.2	3.7	3.8	3.5	3.7	3.5	3.5	2.4
	CI +/-	0.30	0.36	0.32	0.30	0.30	0.31	0.15	0.30

Appendix 4: Ratings of individual samples in impaired conditions

		AWGN			Rural Fast			Terrain Obstructed			Urban Fast		
		Bach	Brokaw	Prince	Fagen	Messiah	Woman	Brokaw	Crowded	Persian	1812	Cole	Man
HDC96	MOS	4.1	4.3	4.5	4.5	4.4	3.8	4.1	4.4	4.0	3.8	4.0	4.5
	CI +/-	0.25	0.24	0.22	0.22	0.21	0.32	0.24	0.21	0.31	0.27	0.27	0.17
Generation 1	MOS	4.5	4.5	4.0	4.5	4.1	4.3	3.8	4.3	4.2	4.3	4.4	4.5
	CI +/-	0.21	0.23	0.30	0.22	0.25	0.26	0.27	0.18	0.24	0.25	0.24	0.21
Delphi	MOS	3.5	1.7	4.2	4.1	2.3	2.5	1.5	2.0	1.4	3.3	3.6	2.2
	CI +/-	0.35	0.26	0.21	0.24	0.28	0.35	0.20	0.32	0.17	0.32	0.31	0.28
Pioneer	MOS	3.2	1.5	3.7	4.0	2.2	1.7	1.2	1.9	1.4	3.2	3.2	1.7
	CI +/-	0.25	0.24	0.30	0.29	0.32	0.32	0.13	0.25	0.21	0.30	0.32	0.21
Sony	MOS	3.5	1.6	3.5									
	CI +/-	0.31	0.25	0.30									
Technics	MOS	2.4	1.6	3.9									
	CI +/-	0.33	0.27	0.28									

Subjective test results for AM IBOC DAB Generation 3 Hardware

Unimpaired and impaired test conditions

Report to the National Radio Systems Committee
and
iBiquity Digital Corporation

March 29, 2004
(revised 4/12/04)

Ellyn G. Sheffield
Sheffield Audio Consulting
Princeton, New Jersey 08540
(609) 737-9451
egsheffield@comcast.net

1. Introduction

This report describes results from unimpaired and impaired AM audio quality testing conducted at iBiquity Digital Radio Corporation between December 8th and March 8th, 2004 under the supervision of Dr. Ellyn Sheffield. The experimental program was designed to subjectively test AM Generation 3 hardware, including unimpaired transmissions (Test G), and transmissions with AWGN (Test B.1). Unless otherwise noted, all procedures were followed as identified in the test plan "Proposal for Subjective Evaluation of Generation 3 HD Radio Hardware" dated December 1, 2003.

In order to maintain consistency between this study and past subjective tests of the AM IBOC system, methodologies replicated previous testing procedures to the greatest possible extent. As in the past, audio samples were delivered over loudspeakers, and samples originally picked for Generation 2 testing were again presented to listeners for rating. All analog AM samples were taken from recordings made previously for the NRSC AM IBOC Gen2 test program. For digital audio, samples were transmitted over unimpaired laboratory test channels through both the AM core and enhanced IBOC systems. Two reference recordings were included to act as high and low anchors (see 2.1 Sound sample preparation for details).

2. Experimental Methodology

2.1 Sound sample preparation

With the exception of AM low anchors, audio samples were recorded directly onto wave files at iBiquity. This procedure was witnessed by a designated NRSC observer. Sound samples were recorded in groups, based on their post-processor settings (for details, see iBiquity's report entitled "Digital Performance Regression Tests of the iBiquity Generation 3 Digital HD Radio System in the AM & FM Bands: Subjective Audio Evaluation Sample Preparation Procedure", dated February 26, 2004). In order to create the AM low anchors which did not include additional transmission impairments (e.g., additive White Gaussian noise), iBiquity generated poor quality low AM anchors in the following manner: A source wave file was loaded into Cool Edit 2000. The file was intentionally bass distorted using a preset utility in Cool Edit. The file was then saved in MPEG Layer 3 Format at a sample rate of 8kbps (mono). The file was then saved back to a wave format.

All recording groups were sent to Dr. Sheffield, who parsed them into individual samples. She then edited, leveled and named them. All of these efforts were done in accordance with procedures identified in the FM IBOC Dab Laboratory and Field Testing Report, August 2001, Exhibit 4: Procedure for Editing and Leveling Sound Samples. Sound samples were loaded onto test computers and experiments were created using iBiquity's subjective testing software.

2.2 Listener Sample

Data from 40 qualified listeners (19 males and 21 females) are included in this report. In order to qualify as a listener, participants needed to pass a screening test (see Section 2.2: Screening Procedures). Additionally, a post-hoc analysis designed to eliminate obvious outliers was conducted on each listener's data. Six listeners were disqualified from the sample for failing to pass the screening test. Listeners were recruited from the Columbia, Md. area by word-of-mouth, sending flyers to local business establishments, colleges and universities, and from referrals by listeners participating in the test. Table 2.2 shows the distribution of participants by age and gender.

	Male	**Female**
18-24	5	6
25-32	6	4
33-42	5	5
43-50	4*	5

Table 2.2: Distribution of participants
*Includes 1 NRSC member, above the age of 50.

2.3 Screening Procedure

Listeners completed a questionnaire and received instructions on the specific tasks they were asked to perform (see Appendix 1: Experimenter Script – Gen 3 AM testing). Prior to testing, participants completed a screening test. Screening was conducted to ensure that listeners were reliably able to distinguish between samples that differed substantially in quality. In order to pass the screening test, participants needed to answer 5 out of 6 screening questions correctly. In a triple-stimulus, double-blind screening procedure, participants were asked to listen to 3 samples: a "Reference sample", and 2 additional samples. One of the samples was identical to the reference, the other was different. At the beginning of the screening test, listeners were shown how to register their answers, and were given one example of the task, which they were asked to complete. After this training period, they were told that they should proceed with the screening test. They were encouraged to listen to the samples as many times as necessary to make their judgment. Figure 2.3 is a schematic which depicts the screen presentation layout. Table 2.3 lists the samples used for comparisons:

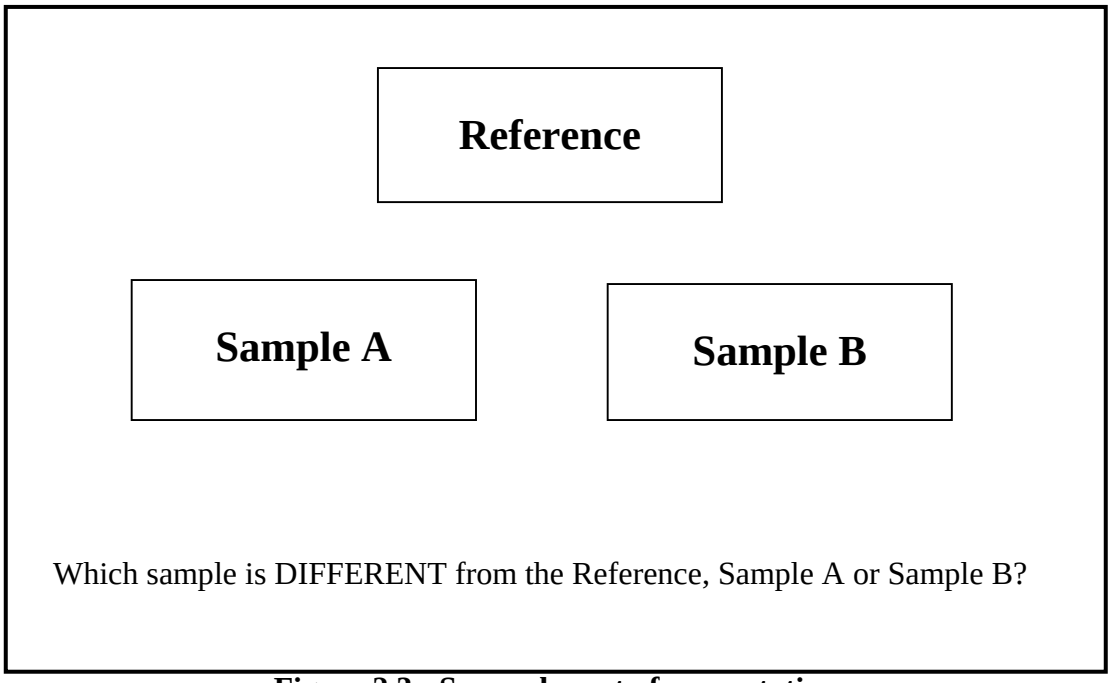


Figure 2.3: Screen layout of presentation

Sample	Reference	Sample A	Sample B
Speech – Woman	CD Source	CD Source	FM Rural Fast (Dephi)
Classical – Bach	CD Source	CD Source	FM AWGN (Delphi)
Speech – Brokaw	CD Source	Delphi Terrain Obstructed	CD Source
Rock – Clapton	CD Source	CD Source	AM Clean (Sony)
Rock – Travis	CD Source	HDC20kps	CD Source
Classical – 1812	CD Source	CD Source	FM Clean (Sony)

Table 2.3 Sound samples used in screening

2.4 Main Test

Listeners next participated in two ACR-MOS tests. The first test, comprised of 144 samples, exclusively included unimpaired audio samples. Participants listened to the samples, one-by-one, and rated each sample on its own merit on a 5-point scale (Excellent; Good; Fair; Poor; and Bad). They heard 48 samples, were given a 5 minute break, listened to another 48 samples, were given another break and then finished the test. After this period, participants were given a 10-minute break and were encouraged to walk around, relax, have a drink, etc. They then completed the second test, comprised of 60 impaired samples. (See Tables 2.4.1 and 2.4.2 for samples used during these tests).

Sample	Core Gen 3 – HDC20	Enhanced Gen 3 – HDC36	Pioneer	Technics	Delphi	Sony	FM High Anchor	AM Low Anchor
Classical								
Carmen	1	1	1	1	1	1	1	1
Handel Messiah	1	1	1	1	1	1	1	1
Ibert	1	1	1	1	1	1	1	1
Saito	1	1	1	1	1	1	1	1
Commercial								
Ballet Woman	1	1	1	1	1	1	1	1
Camera	1	1	1	1	1	1	1	1
From Richmond	1	1	1	1	1	1	1	1
Riverdance	1	1	1	1	1	1	1	1
Rock								
Clapton	1	1	1	1	1	1	1	1
CSNY	1	1	1	1	1	1	1	1
EWF	1	1	1	1	1	1	1	1
REO Speedwagon	1	1	1	1	1	1	1	1
Randy Travis	1	1	1	1	1	1	1	1
Suzanne Vega	1	1	1	1	1	1	1	1
Speech								
Female A1	1	1	1	1	1	1	1	1
Female C10	1	1	1	1	1	1	1	1
Male A1	1	1	1	1	1	1	1	1
Male B4	1	1	1	1	1	1	1	1

Table 2.4.1: Unimpaired Test Samples

Sample	Core Gen 3 (20 kbps)	Enhanced Gen 3 (36 kbps)	FM High Anchor	Pioneer	Technics	Delphi	Sony
Classical							
Ibert		1	1	1	1	1	1
Debussy	1		1	1	1	1	1
Commercial							
Imagine	1		1	1	1	1	1
Riverdance		1	1	1	1	1	1
Rock							
Santana		1	1	1	1	1	1
Fleetwood Mac	1		1	1	1	1	1
Speech							
Male A1	1		1	1	1	1	1
Female B2	1		1	1	1	1	1
Male B4			1	1	1	1	1
Female A1		1	1	1	1	1	1

Table 2.4.2: Impaired Test Samples (AWGN)[1]

[1] The AM Test Plan called for including 4 samples from the impaired Gen 1 test so that direct comparisons could be made between impaired Gen 3 and Gen 1 audio. Due to an error in the test setup, Gen 1 samples were inadvertently omitted. Because of this omission, no direct comparisons between Gen 1 and Gen 3 can be made.

3. Results – Unimpaired

3.1 Preliminary analyses

Preliminary analyses of variance (ANOVAs) were conducted to uncover any differences in participants' responses, based on their age and gender. There was no main effect of gender, indicating that females and males responded similarly to all transmissions (female MOS was 3.36; male MOS was 3.37), and no main effect of age, indicating that all age groups responded similarly (MOS varied from 3.31 in the youngest group to 3.42 in the oldest group).

3.2 Test results

The 5-point scale (Excellent through Bad) was translated into the following numerical values for analysis:

Rating	**Numerical Value**
Excellent	5.0
Good	4.0
Fair	3.0
Poor	2.0
Bad	1.0

Figure 3.2 shows MOS scores aggregated by genre. *Classical* includes Ibert, Carmen, Handel's Messiah, and Saito. *Commercial* includes Ballet Woman, Camera, From Richmond, and Riverdance. *Rock* includes Eric Clapton, Crosby, Stills, Nash and Young, Earth Wind and Fire, REO, Randy Travis and Suzanne Vega. *Speech* includes 2 female and 2 male samples (Female A1, FemaleC10, Male A1 and Male B4). In the classical genre, FM, HDC36 and HDC20 were rated equivalently and significantly better than all AM analog transmissions. In all other genres HDC20 and HDC36 were rated significantly lower than FM analog, but significantly higher than AM analog transmissions. Interestingly, in the Commercial and Speech genres, HDC20 and HDC36 were rated equivalently. However, in Rock and Classical, HDC36 was rated significantly higher than HDC20. See Appendix 2 for individual ANOVA results and Newman Keuls post-hoc comparisons. See Appendix 3 for ratings of individual samples and error terms.

Figure 3.2 – MOS by Genre

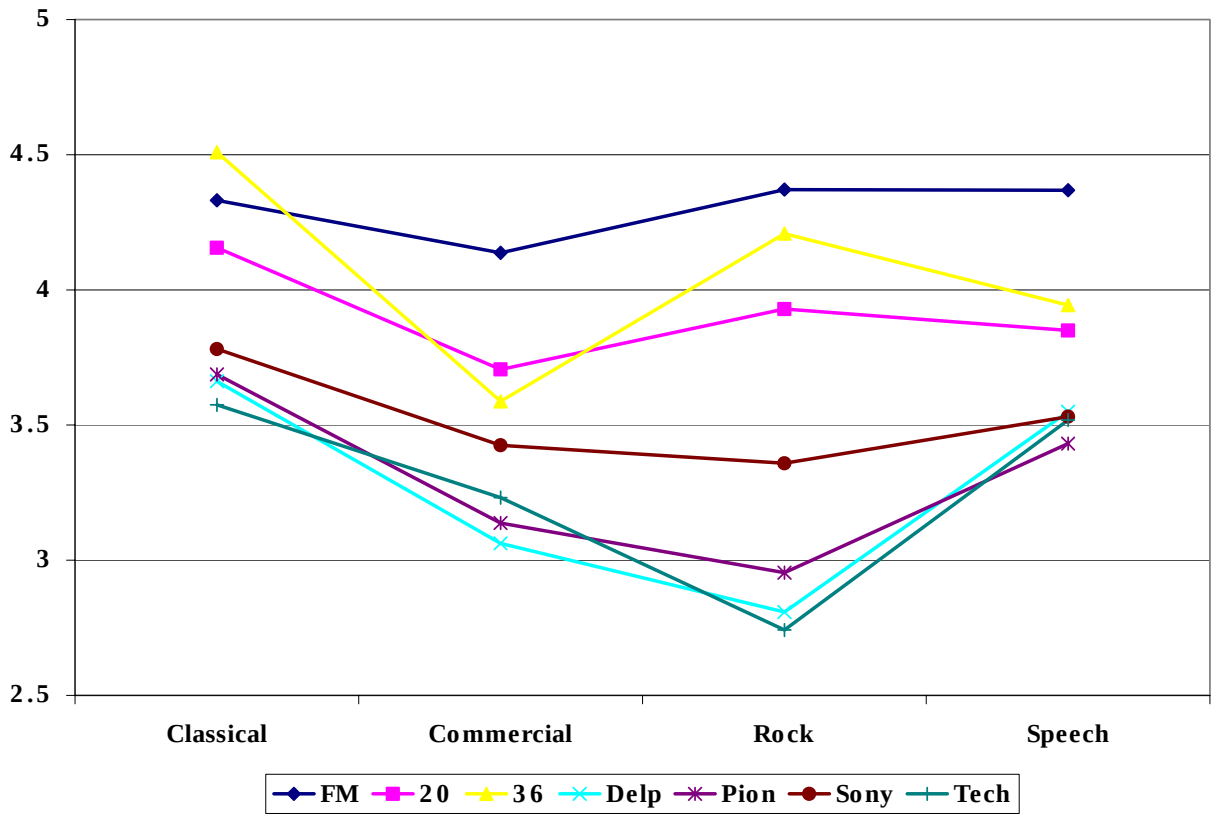


Table 3.2 lists all rating scores including AM low anchor samples. Results show that participants used the 5-point scale broadly, scoring low anchor AM samples consistently in the "bad" (1.0-1.5) range, FM source material in the "good to excellent" (4.0 – 4.5) range, and the majority of other transmissions between these points.

Receiver	Classical	Commercial	Rock	Speech
FM	4.3	4.1	4.4	4.4
HDC36	4.5	3.6	4.2	3.9
HDC20	4.2	3.7	3.9	3.9
Delphi	3.7	3.1	2.8	3.6
Pioneer	3.7	3.1	3.0	3.4
Sony	3.8	3.4	3.4	3.5
Technics	3.6	3.2	2.7	3.5
AM Low Anchor	1.6	1.2	1.3	1.4

Table 3.2: Unimpaired MOS ratings by genre

4. Results – Impaired

Figures 4.1 shows results from impaired test conditions in enhanced mode. Participants rated HDC36 superior to AM analog transmissions in the classical, rock and speech categories. In the commercial genre, HDC36 and Sony were rated the same, and significantly better than the Pioneer, Technics and Delphi. Additionally, in the classical and rock genres, FM analog and HDC36 were rated the same, while HDC36 was rated lower than FM analog in commercial and speech. See Appendix 4 for individual samples and error terms.

Figure 4.1: Mean opinion scores (Enhanced Mode)

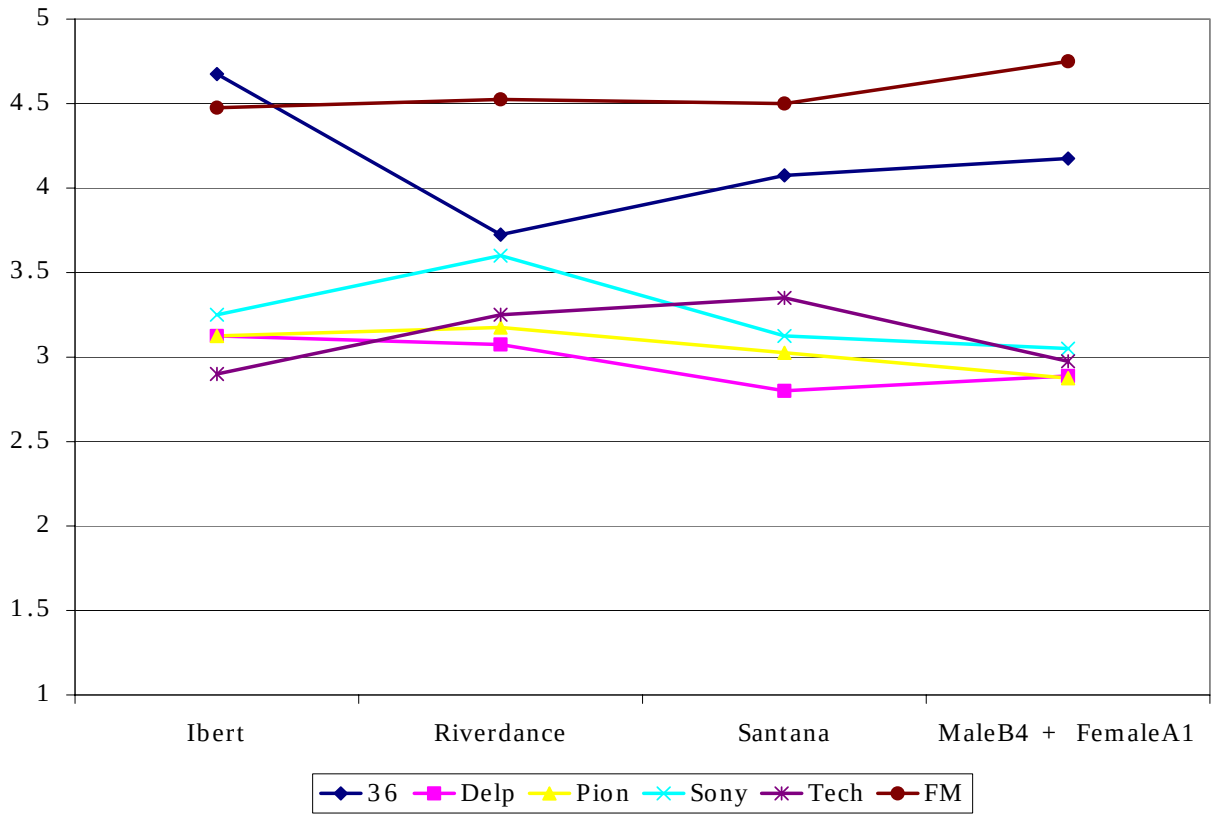
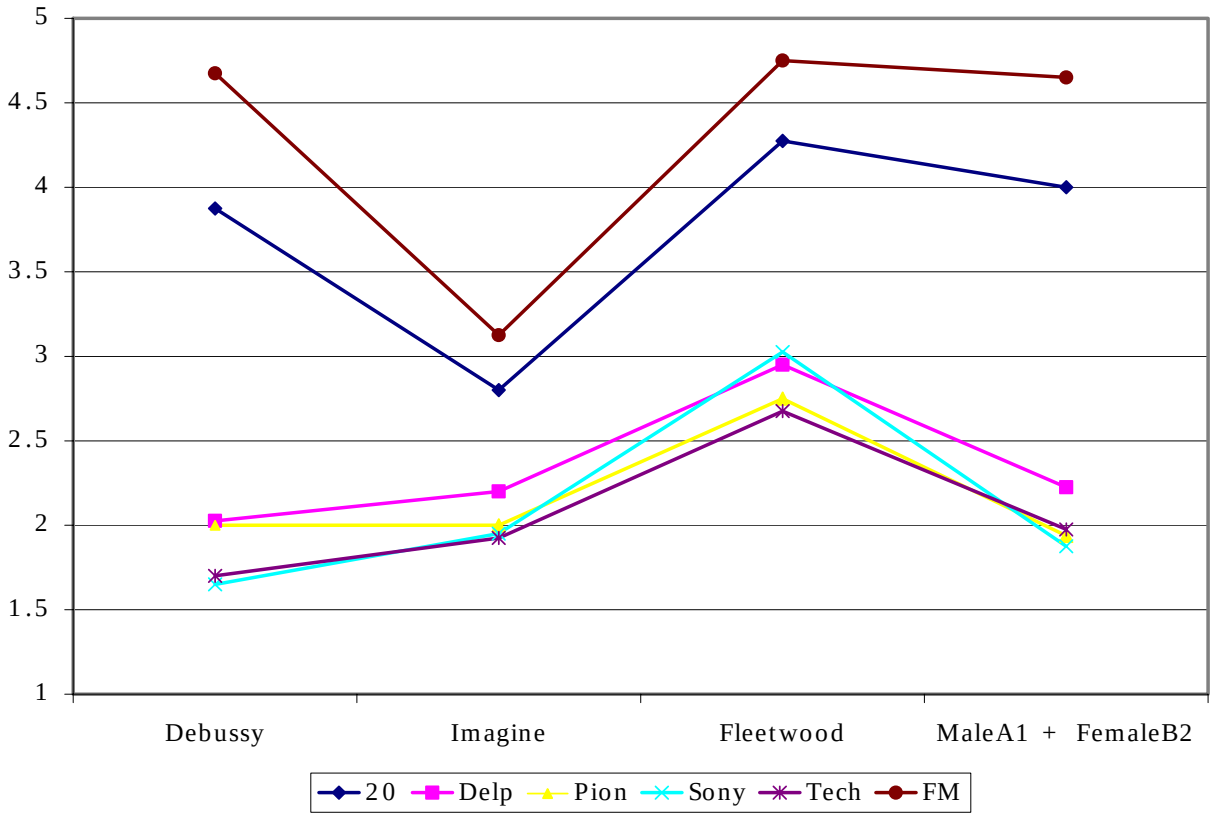


Figure 4.2 shows results from impaired test conditions in core mode. This figure shows that in all genres HDC20 was rated superior to AM analog transmissions. However, HDC20 was rated lower than FM in all genres except the commercial genre, where they were rated equivalently. There is a noticeable dip in participants' ratings for HDC20 in the commercial genre. However, it is worthy to note that the FM analog transmission follows the same pattern. Thus, it is likely that participants were not commenting on the coder's treatment of the sample, but rather on the programmatic material or the original quality of recording.

Figure 4.2: Mean Opinion Scores (Core Mode)



Appendix 1: Experimenter Script – Gen 3 AM testing

Welcome to our session. Today you will be participating in an audio test which should last approximately 2 hours. There are two parts to this test. In the first part of the test, you will hear a series of samples and compare them to a "reference". In this part you will play a Reference, an A and a B sample, and judge which of the samples is **different** from the reference. There will be six triplets to judge. In the second part you will first hear 144 short samples, in groups of 48. Please listen to the sample from start to finish. At the end of the clip you will be asked one question about it. After each session, the computer will tell you to take a 5 minute break. After listening to the 144 samples you will have a 10-minute break. That's your turn to go to the bathroom, have a drink, or just relax. In the final session you will be presented with another 60 samples. Once you start a session, you should continue until the program tells you to take a break, but you are also encouraged to take the test at your own pace. This may mean stopping between samples if you feel you need to "clear your head" for a few seconds.

For each session, you will be asked to rate the audio on a 5-point scale. In all cases, we want to remind you that we are not asking you to judge the material, or whether you like a particular cut. We know that you will have various feelings about the samples you are going to listen to. For this test, we are asking you to try to keep focused on only one thing: the quality of the transmission you are listening to.

Now we are going to begin. Any questions so far?

Appendix 2: Analysis of Variance Report for Unimpaired Samples

The following notes are intended to help the reader interpret the tables listed in this Appendix and Appendix 2.

(a) The Overall ANOVA allows us to see whether there is a statistical difference between any and all groups under test (i.e., HDC36, HDC20, Delphi, Pioneer, etc.)
(b) If the Overall ANOVA p-value < .05, then it is justifiable to use post-hoc Newman-Keuls Multiple comparison statistical tests to discern which specific groups differ from each other.
(c) The Newman-Keuls post-hoc tests are run at a p-value of 0.05, which is a standard p-value for multiple comparison statistical tests.
(d) The column to the far right, "Different from Groups" identifies which groups are statistically different from a group in question (the column to the far left). For example, in the first table, FM is different from AM Low, Sony, Pioneer, Delphi and Technics, but not different from HDC36 and HDC20.

Classical
Overall ANOVA: DF= 7,1280; F-Value=183.77; P-Value (alpha) =0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1272 MSE=0.6923202

Group	Count	Mean	Different From Groups
AM Low Anchor	160	1.625	Tech, Delp, Pion, Sony, HDC20, FM, HDC36
Sony	160	3.78125	AM Low, HDC20, HDC36, FM
Pion	160	3.6875	AM Low, HDC20, HDC36, FM
Delp	160	3.6625	AM Low, HDC20, HDC36, FM
Tech	160	3.575	AM Low, HDC20, HDC36, FM
HDC36	160	4.5	AM Low, Sony, Pion, Delp, Tech, HDC20
HDC20	160	4.15625	AM Low, Sony, Pion, Delp, Tech, HDC36
FM High Anchor	160	4.33125	AM Low, Sony, Pion, Delp, Tech

Commercial
Overall ANOVA: DF= 7, 1240; F-Value=118.68; P-Value (alpha) =0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1232 MSE=.8018094

Group	Count	Mean	Different From Groups
AM Low Anchor	160	1.25833	Sony, Tech, Delp, Pion, HDC36, HDC20, FM
Sony	160	3.425	AM Low, Delp, Pion, HDC20, FM
Tech	160	3.23125	AM Low, HDC36, HDC20, FM
Delp	160	3.0625	AM Low, Sony, HDC36, HDC20, FM
Pion	160	3.1375	AM Low, Sony, HDC36, HDC20, FM
HDC36	160	3.5875	AM Low, Delp, Pion, Tech, FM
HDC20	160	3.70625	AM Low, Delp, Pion, Tech, Sony, FM
FM High Anchor	160	4.1375	AM Low, Delp, Pion, Tech, Sony, HDC20, HDC36

Rock
Overall ANOVA, DF=7,1920; F-Value =293.29; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1912 MSE=0.7991022

Group	Count	Mean	Different From Groups
AM Low Anchor	240	1.3375	Tech, Delp, Pion, Sony, HDC20, HDC36, FM
Sony	240	3.35833	AM Low, Tech, Delp, Pion, HDC20, HDC36, FM
Tech	240	2.74166	AM Low, Pion, Sony, HDC20, HDC36, FM
Delp	240	2.08033	AM Low, Sony, HDC20, HDC36, FM
Pion	240	2.94541	AM Low, Tech, Sony, HDC20, HDC36, FM
HDC36	240	4.20833	AM Low, Tech, Delp, Pion, Sony, HDC20, FM
HDC20	240	3.92916	AM Low, Tech, Delp, Pion, Sony, HDC36, FM
FM High Anchor	240	4.37083	AM Low, Tech, Delp, Pion, Sony, HDC20, HDC36

Speech
Overall ANOVA, DF=7,1320; F-Value =30.15; P-Value (alpha) = 0.000000
Newman-Keuls Multiple-Comparison Test
Alpha=0.050 Error Term=S DF=1312 MSE=0.7756431

Group	Count	Mean	Different From Groups
AM Low Anchor	160	1.425	Pion, Tech, Sony, Delp, HDC20, HDC36, FM
Sony	160	3.5312	AM Low, HDC20, HDC36, FM
Tech	160	3.51875	AM Low, HDC20, HDC36, FM
Delp	160	3.55	AM Low, HDC20, HDC36, FM
Pion	160	3.43125	AM Low, HDC20, HDC36, FM
HDC36	160	3.94375	AM Low, Pion, Tech, Sony, Delp, FM
HDC20	160	3.85	AM Low, Pion, Tech, Sony, Delp, FM
FM High Anchor	160	4.36875	AM Low, Pion, Tech, Sony, Delp, HDC20, HDC36

Appendix 3: Individual Samples and Error Terms (Unimpaired)

UNIMPAIRED									
		FM	HDC36	HDC20	Delphi	Pioneer	Sony	Technics	AM Low
BalletWoman	MOS	4.4	3.7	4.1	3.1	3.2	3.5	3.2	1.3
	CI (+/-)	0.25	0.33	0.25	0.26	0.24	0.29	0.27	0.22
Camera	MOS	3.9	3.4	2.6	3.3	3.4	3.5	3.3	1.4
	CI (+/-)	0.31	0.26	0.30	0.26	0.24	0.25	0.25	0.17
Carmen	MOS	4.9	4.6	4.3	3.9	3.8	4.1	3.5	1.8
	CI (+/-)	0.10	0.16	0.24	0.32	0.26	0.23	0.26	0.27
Clapton	MOS	4.9	4.7	4.5	2.7	3.0	3.4	2.8	1.2
	CI (+/-)	0.10	0.16	0.20	0.30	0.28	0.24	0.29	0.13
CSNY	MOS	4.7	4.4	4.3	3.1	3.1	3.6	2.8	1.3
	CI (+/-)	0.16	0.22	0.20	0.28	0.28	0.17	0.26	0.14
EWF	MOS	3.8	3.7	3.6	2.9	3.0	3.1	2.6	1.3
	CI (+/-)	0.33	0.33	0.33	0.29	0.23	0.24	0.28	0.16
FemaleA1	MOS	4.4	3.7	3.8	3.0	2.9	3.4	3.2	1.2
	CI (+/-)	0.25	0.32	0.34	0.31	0.34	0.29	0.27	0.13
FemaleC10	MOS	4.3	4.3	4.1	3.6	3.4	3.7	3.5	1.6
	CI (+/-)	0.25	0.22	0.29	0.29	0.21	0.26	0.26	0.18
Ibert	MOS	4.4	4.6	4.2	4.0	4.0	3.9	3.9	1.6
	CI (+/-)	0.23	0.17	0.23	0.26	0.25	0.22	0.23	0.23
MaleA1	MOS	4.2	3.8	3.5	3.8	3.8	3.3	3.8	1.6
	CI (+/-)	0.27	0.29	0.31	0.25	0.26	0.23	0.28	0.21
MaleB4	MOS	4.6	4.0	4.0	3.8	3.6	3.8	3.6	1.5
	CI (+/-)	0.21	0.29	0.31	0.28	0.33	0.26	0.28	0.22
Messiah	MOS	4.4	4.3	3.8	3.4	3.5	3.7	3.5	1.7
	CI (+/-)	0.20	0.21	0.30	0.29	0.23	0.24	0.29	0.28
REO	MOS	4.0	3.4	3.2	2.2	2.1	2.9	2.5	1.2
	CI (+/-)	0.30	0.34	0.32	0.23	0.25	0.27	0.26	0.22
Richmond	MOS	3.9	3.5	4.2	2.9	3.2	3.5	3.2	1.2
	CI (+/-)	0.33	0.31	0.27	0.29	0.28	0.27	0.28	0.13
Riverdance	MOS	4.4	3.8	4.0	3.0	2.9	3.3	3.2	1.1
	CI (+/-)	0.19	0.23	0.27	0.32	0.29	0.27	0.29	0.08
Saito	MOS	3.7	4.6	4.3	3.5	3.5	3.4	3.4	1.5
	CI (+/-)	0.33	0.20	0.22	0.35	0.24	0.27	0.25	0.21
Travis	MOS	4.2	4.5	3.8	2.8	3.0	3.5	2.7	1.5
	CI (+/-)	0.29	0.28	0.22	0.27	0.29	0.26	0.29	0.24
Vega	MOS	4.7	4.6	4.3	3.2	3.6	3.7	3.1	1.5
	CI (+/-)	0.17	0.17	0.27	0.31	0.31	0.28	0.32	0.23

Appendix 4: Individual Samples and Error Terms (Impaired)

ENHANCED						
		FemaleA1	**MaleB4**	**Ibert**	**Riverdance**	**Santana**
FM	MOS	4.8	4.7	4.5	4.5	4.5
	CI (+/-)	0.22	0.19	0.20	0.17	0.19
HDC36	MOS	4.3	4.1	4.7	3.7	4.1
	CI (+/-)	0.28	0.36	0.16	0.23	0.25
Delphi	MOS	2.6	3.2	3.1	3.1	2.8
	CI (+/-)	0.28	0.27	0.27	0.30	0.31
Pioneer	MOS	3.0	2.8	3.1	3.2	3.0
	CI (+/-)	0.28	0.28	0.26	0.30	0.31
Sony	MOS	3.0	3.1	3.3	3.6	3.1
	CI (+/-)	0.28	0.28	0.30	0.26	0.25
Technics	MOS	3.0	3.0	2.9	3.3	3.4
	CI (+/-)	0.28	0.31	0.25	0.24	0.27

CORE						
		FemaleB2	**MaleA1**	**Debussy**	**Imagine**	**Fleetwood**
FM	MOS	4.6	4.7	4.7	3.1	4.8
	CI (+/-)	0.24	0.16	0.16	0.33	0.17
HDC20	MOS	4.0	4.0	3.9	2.8	4.3
	CI (+/-)	0.30	0.33	0.21	0.28	0.28
Delphi	MOS	2.1	2.4	2.0	2.2	3.0
	CI (+/-)	0.24	0.31	0.25	0.25	0.29
Pioneer	MOS	1.9	2.0	2.0	2.0	2.8
	CI (+/-)	0.31	0.23	0.26	0.24	0.29
Sony	MOS	1.8	2.0	1.7	2.0	3.0
	CI (+/-)	0.26	0.23	0.21	0.27	0.28
Technics	MOS	1.8	2.1	1.7	1.9	2.7
	CI (+/-)	0.30	0.26	0.25	0.28	0.28

Albert Shuldiner
Senior VP & General Counsel

April 14, 2004

Mr. David Wilson
Consumer Electronics Association
2500 Wilson Boulevard
Arlington, Virginia 22201

Re: NRSC Gen 3 Audio Tests

Dear Dave:

The NRSC already has received reports on the test procedures and results of the Gen 3 audio tests conducted for the HD Radio™ system. The material previously submitted to the NRSC did not contain the BLER curves or "objective results" required by the NRSC test procedures. Attached to this letter you will find the BLER curves for both the AM and FM tests that were completed in November of last year. These curves can be compared with the results obtained by the ATTC in previous HD Radio tests.

The NRSC procedures also require that iBiquity identify any differences between the current Gen 3 system and the Gen 1 system previously tested by the NRSC. Other than the substitution of HDC audio compression technology for the previous system codec, the Gen 3 operations are consistent with the operations of Gen 1, as detailed in the respective Appendices A to iBiquity's FM[1] and AM[2] reports.

Please let me know if you have any questions about this material.

Sincerely,

Albert Shuldiner

cc: David Layer (NAB)

---

[1] FM IBOC DAB Laboratory and Field Testing, Report to the National Radio Systems Committee dated August 2001.

[2] AM IBOC DAB Laboratory and Field Testing, Third Report to the National Radio Systems Committee dated January 4, 2002.

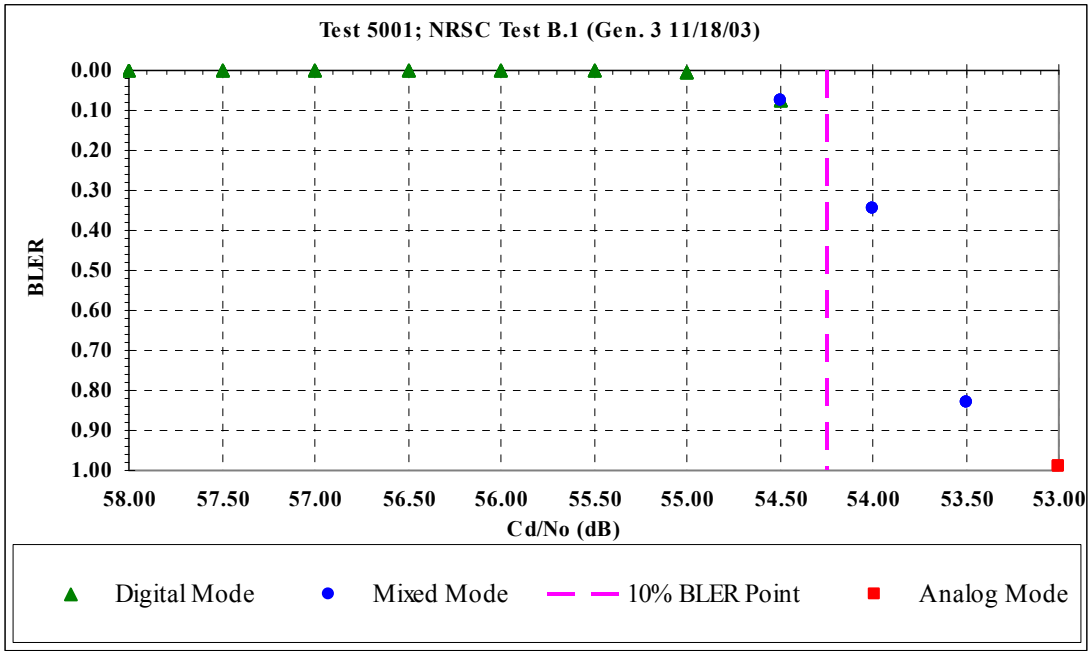


Figure 1 - iBiquity Generation 3 FM IBOC System Performance (AWGN Present)

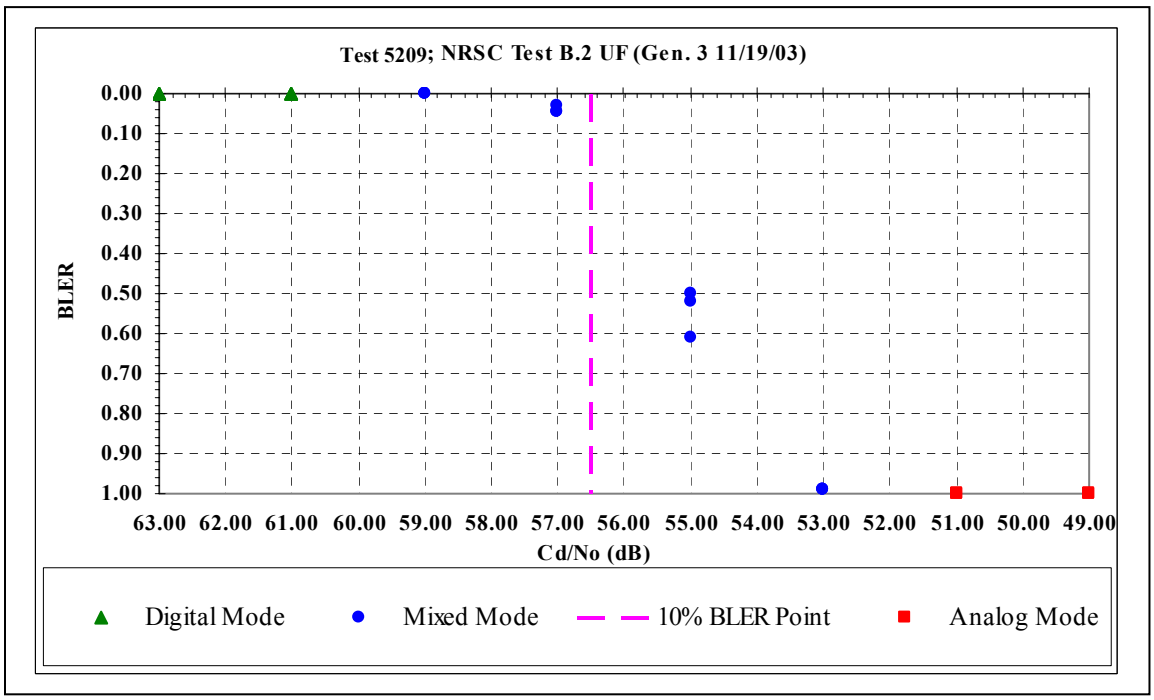


Figure 2 - iBiquity Generation 3 FM IBOC System Performance (Urban Fast Multipath & AWGN Present)

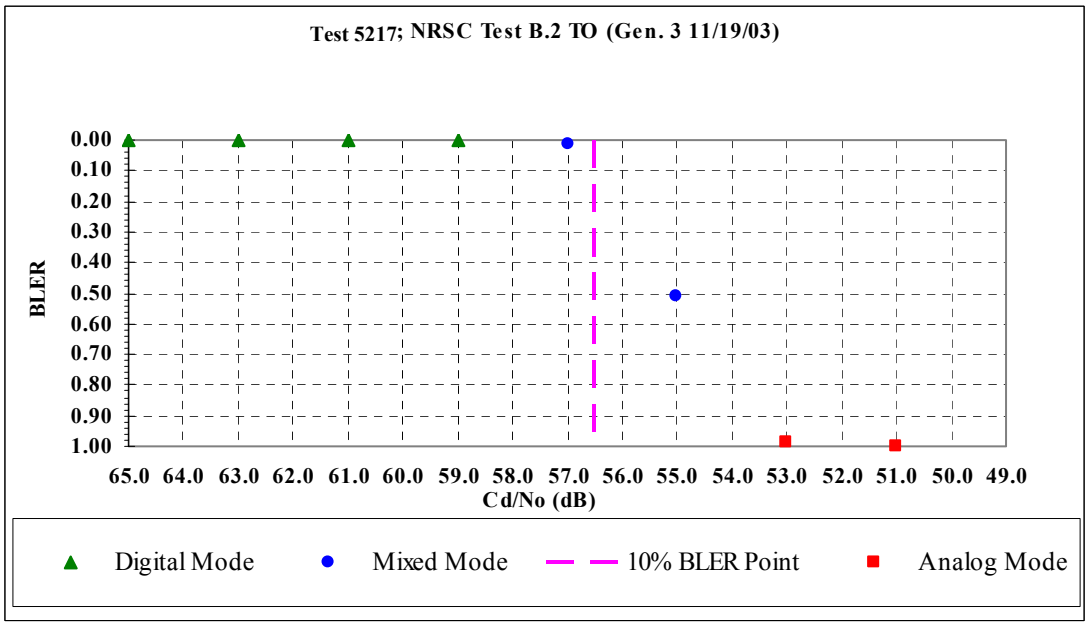


Figure 3 - iBiquity Generation 3 FM IBOC System Performance (Terrain Obstructed Multipath & AWGN Present)

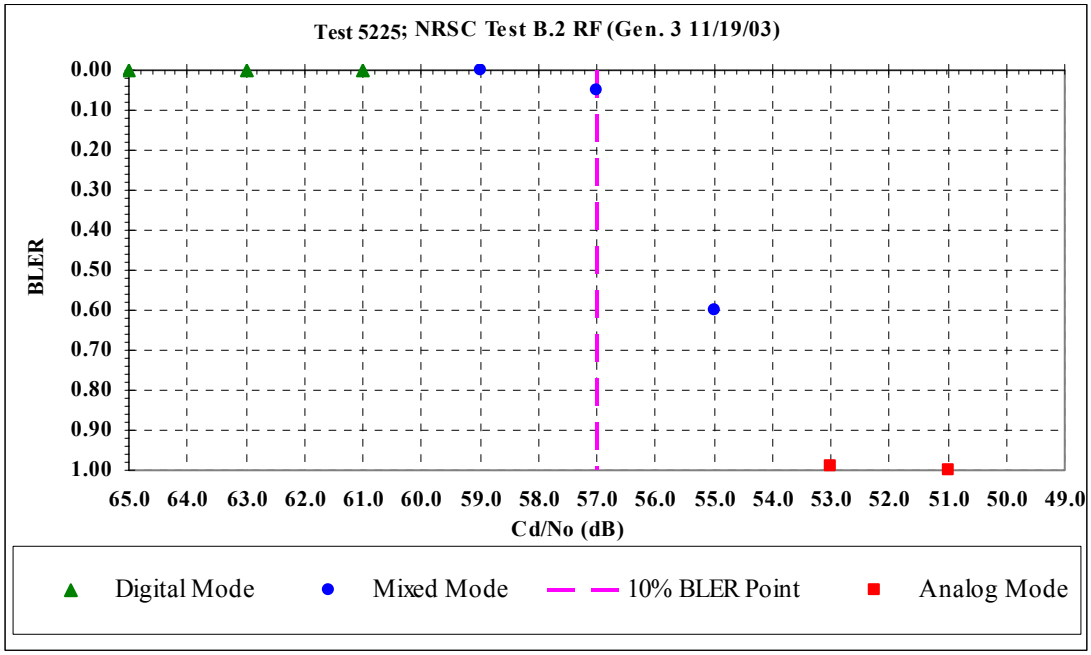


Figure 4 - iBiquity Generation 3 FM IBOC System Performance (Rural Fast Multipath & AWGN Present)

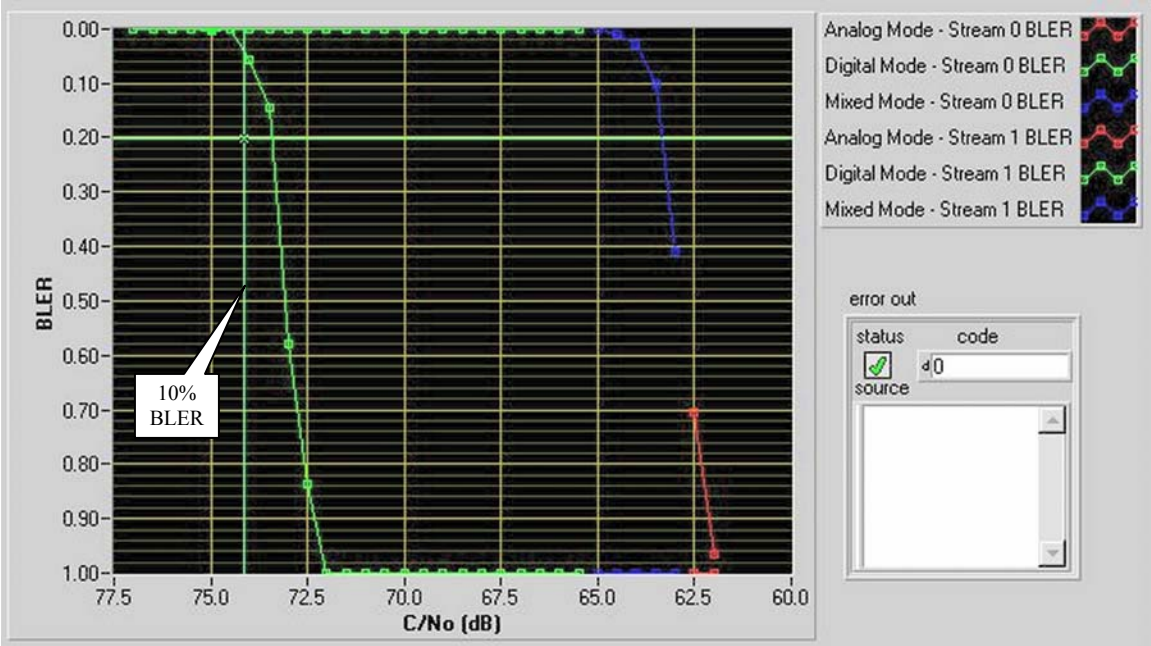


Figure 5 - iBiquity Generation 3 AM IBOC System Performance (AWGN Present)