

TABLE OF CONTENTS

	Page
TABLE OF CONTENTS	III-1
EXECUTIVE SUMMARY	III-3
ACKNOWLEDGMENTS	III-4
1.0 INTRODUCTION	III-5
2.0 SUBJECTIVE TESTS OF BASIC RECEIVED QUALITY	III-7
2.1 Objectives	III-7
2.2 Methodology	III-7
2.3 Results - Basic Received Quality Assessments	III-15
2.4 Comparative - Basic Received Quality	III-24
3.0 SUBJECTIVE TESTS OF IMPAIRMENT/INTERFERENCE	III-31
3.1 Objectives	III-31
3.2 Methodology	III-31
3.3 Results - Impairment/Interference Assessments	III-35
3.4 Comparative - Impairment/Interference	III-40
4.0 SUBJECTIVE TESTS OF RECEIVER SCAN CONVERSION	III-43
4.1 Objectives	III-43
4.2 Methodology	III-43
4.3 Results - Receiver Scan Conversion Assessments	III-45
4.4 Comparative - Receiver Scan Conversion	III-55
5.0 AUXILIARY DATA TRADEOFF	III-57
6.0 COMMENTS BY THE GRAND ALLIANCE	III-59
6.1 Introduction	III-59
6.2 Basic Received Quality	III-59
6.3 Interference to NTSC	III-60
6.4 Receiver Scan Conversion	III-60
6.5 Video Quality/Auxiliary Data Tradeoff	III-61
6.6 Conclusions	III-61
6.7 Acknowledgments	III-62

APPENDICES

			III-6
Appendix	A	Review of Basic Received Quality Test Material	III-6
Appendix	B	Review of Impairment Test Material (ATV-to-NTSC)	III-8
Appendix	C	Review of Receiver Scan Conversion Material	III-9
Appendix	D	Correspondence	III-11
Appendix	E	Summary of Responses to Quality Questionnaires	III-12
Appendix	F	Summary of Responses to Impairment Questionnaires (ATV-to-NTSC)	III-13
Appendix	G	Summary of Responses to Receiver Scan Conversion Questionnaires	III-14
Appendix	H	1080I - Basic Received Quality Analyses	III-15
Appendix	I	720P - Basic Received Quality Analyses	III-20
Appendix	J	Core Analysis - Interlaced	III-24
Appendix	K	Core Analysis - Progressive	III-24
Appendix	L	Lower-Adjacent Channel Interference (ATV-to-NTSC) Analyses	III-27
Appendix	M	Co-Channel Interference (ATV-to-NTSC) Analyses	III-31
Appendix	N	Receiver Scan Conversion Analyses	III-33

EXECUTIVE SUMMARY

The Advanced Television Evaluation Laboratory (ATEL) is a facility of the Communications Research Centre (CRC) in the Department of Industry, Government of Canada. The ATEL was selected by the U.S. Federal Communications Commission's Advisory Committee on Advanced Television Service to participate in the establishment of an advanced television broadcast standard for North America.

The *digital* HDTV Grand Alliance System represents the joint effort of American Telephone and Telegraph Company, David Sarnoff Research Center, General Instrument, Massachusetts Institute of Technology, Philips Electronics North America, Thomson Consumer Electronics, and Zenith Electronics Corporation.

Subjective tests described in this report evaluated the performance of the *digital* HDTV Grand Alliance System in terms of basic received quality, impairment/interference into NTSC, and scan conversion between 1080 interlaced and 720 progressive formats.

The basic received quality of the *digital* HDTV Grand Alliance System operating in both the 1080 interlaced and 720 progressive modes was examined. The overall performance of the System exceeded the Advisory Committee's target specifications.

Interference into NTSC tests were performed for both the Lower-Adjacent and Co-Channels. In both cases the System exceeded the Advisory Committee's target specifications. Upper-Adjacent Channel tests were deemed invalid, as the receiver used to record the test material was not representative of the median receiver selected by the experts. A receiver with baseband outputs matching the median receiver was not available.

At the discretion of the Task Force on Digital Specific Testing, Receiver Scan Conversion tests were done for both 1080 interlaced to 720 progressive conversion and for 720 progressive to 1080 interlaced conversion. In both cases, the overall performance of the System exceeded the Advisory Committee's target specifications.

At the discretion of the Task Force on Digital Specific Testing, Auxiliary Data Tradeoff testing at the ATEL was waived.

Further details are contained in the main body and appendices of this report.

ACKNOWLEDGMENTS

The Advanced Television Evaluation Laboratory (ATEL) and its staff are grateful to the representatives of the Grand Alliance member organizations, for their participation and support in preparation, testing, and reporting.

The ATEL acknowledges contributions by Advisory Committee members in the development of test procedures and test materials, and members of the Digital Specific Task Force for their work in support of the video subjective testing. The ATEL expresses its appreciation to the Advisory Committee and, in particular, Jim Gaspar, Chairman of PS/WP-6 for effective management of test material selection, and to Mark Richer, Chairman of SSWP2 for effective management of the test program.

The ATEL further acknowledges contributions of the Advanced Television Test Centre, and in particular Peter Fannon and Tom Gurley for their cooperation, support, and timely preparation of the random tapes for subjective testing.

This report was produced at the ATEL by Metin Akgun, Annu Chopra, Carolyn Cooke, Philip Corriveau, Andre Kennedy, Robert Leafloor, Tanya McCreith, Ivano Pagliarello, Ron Renaud, Lew Stelmach and Susan Van Dusen.

1.0 INTRODUCTION

Advanced Television Evaluation Laboratory (ATEL). The Advanced Television Evaluation Laboratory is a facility of the Communications Research Centre in the Department of Industry, Government of Canada. The ATEL conducts subjective assessments of new video technologies to support the development of standards, and to promote the successful introduction of new video services.

The ATEL is equipped with world class facilities for assessments of conventional (NTSC) and high-definition (HDTV) video systems. These assessments are carried out in compliance with rigorous international standards. The staff at the ATEL are recognized experts in the scientific, technical and operational aspects of video assessment.

The ATEL was selected by the U.S. Federal Communications Commission's (FCC) Advisory Committee on Advanced Television Service to participate in the establishment of an advanced television broadcast standard for North America. A first round of subjective tests were completed at the ATEL in 1992, in collaboration with the Advanced Television Test Center (ATTC) and CableLabs, to evaluate the performance of six candidate systems. Of these, two systems used analog transmission (Advanced Compatible Television and Narrow-MUSE), and four used digital transmission (Advanced Digital HDTV, Channel Compatible DigiCipher, HD-DigiCipher and Digital Spectrum Compatible HDTV). Based on the work of the ATTC, CableLabs and the ATEL, it was decided that none of the systems were ready for standardization and that improvements should be implemented. A further decision was made to proceed only with digital systems. The four digital proponents were encouraged to collaborate and combine the best parts of their respective systems.

In the second round, the four digital systems merged into a single *digital* HDTV Grand Alliance System. The *digital* HDTV Grand Alliance System represents the joint effort of American Telephone and Telegraph Company, David Sarnoff Research Center, General Instrument, Philips Electronics North America, Thomson Consumer Electronics, and Zenith Electronics Corporation. The ATEL was engaged to conduct subjective assessments of this new *digital* HDTV Grand Alliance System. The results of these assessments are summarized in this report.

Objectives. The purpose of the subjective tests described in this report was to evaluate the performance of the *digital* HDTV Grand Alliance System in terms of basic received quality, scan conversion between 1080 interlaced and 720 progressive formats, and impairment/interference into NTSC. The System is currently under review as the sole candidate for a terrestrial HDTV transmission standard.

Chronology of Subjective Tests. Basic Received Quality tests of the *digital* HDTV Grand Alliance System were conducted at the ATEL during the period from June 23 to August 3, 1995. Impairment/Interference tests were conducted during the period from May 29 to June 5, 1995. Receiver Scan Conversion tests were conducted during the period from August 14 to August 16, 1995.

Report Contents. This report contains the results of tests conducted by the ATEL on the *digital* HDTV Grand Alliance System.

It also contains comments by the Grand Alliance on these test results. The results of additional tests on the Grand Alliance System are provided elsewhere by:

- Task Force on Digital Specific Tests (SSWP2)
- Cable Television Laboratories (CableLabs)
- Advanced Television Test Center (ATTC)
- Task Force on Audio (SSWP2)
- Hitachi America Ltd.
- IBM
- Task Force on Field Testing (SSWP2)

Test Plans. The ATEL conducted its tests in accordance with the requirements and test procedures developed and approved by the FCC Advisory Committee. These include the relevant portions of the "Grand Alliance System Test Procedures" (SSWP2-1306, "Test Plans"), and the "Test Management Plan" (SSWP2-0124). Since these documents are sizable and are in the FCC Advisory Committee's public record, they are not reproduced in this report.

The Test Plans used by the ATEL were those approved by the FCC Advisory Committee as of March 24, 1995. Since that time, including during the course of the Grand Alliance test period, the following interpretations or other actions on the Test Plans were approved by the Chairman of SSWP2:

- The Upper-Adjacent Channel Interference (ATV-to-NTSC) results were deemed invalid and suppressed (see Section 3.1 and Appendix D document SSWP2-1462).
- The Digital Specific Test on Scan Conversion was performed, see Section 4.0.
- The Digital Specific Test on Auxiliary Data Tradeoff was waived, see Section 5.0.

Grand Alliance. The comments of the Grand Alliance about the tests contained in this report are included in their entirety (see Section 6.0).

The Grand Alliance was provided with an office with an audio and video feed for the purpose of viewing the tests. They were also provided with the opportunity to review all test material under the same conditions as the non-expert viewers, and to witness the facility verification and certification. The Grand Alliance was fully satisfied with the conduct of the tests, and the procedures followed for testing, analysis and reporting.

Information about the *digital* Grand Alliance HDTV System is contained in the documentation provided by the Grand Alliance to the FCC Advisory Committee. It is sizable and available in public record; therefore, it is not reproduced in this report (see "Grand Alliance HDTV System Specification," December 7, 1994, as amended).

2.0 SUBJECTIVE TESTS OF BASIC RECEIVED QUALITY

2.1 Objectives

The Basic Received Quality tests were conducted to assess the subjective quality of image sequences that were encoded, modulated, transmitted, demodulated and decoded by the *digital* HDTV Grand Alliance System. Two modes of operation of the *digital* HDTV Grand Alliance System were tested: 1080 interlaced and 720 progressive (abbreviated 1080I and 720P, respectively).

2.2 Methodology

Viewers. Viewers were recruited through a local university and were screened for visual acuity (normal or corrected-to-normal), contrast sensitivity (normal), color vision (normal), and English comprehension. Those who met the screening criteria were permitted to participate in the tests. All viewers, but for two in the 720P and two in the 1080I tests, were non-experts, defined as individuals with no previous exposure to video subjective evaluation experiments at the ATEL, or professional experience in television technology. Post hoc analyses indicated that the image quality ratings for these viewers were consonant with the mean ratings for the other viewers.

Separate groups of viewers were used in the 1080I and 720P Basic Received Quality tests. Details of the composition of the two groups are presented in Table 1. A maximum of five viewers participated in any given session.

TABLE 1
VIEWER CHARACTERISTICS

BASIC RECEIVED QUALITY TEST	TAPE ORDER 1			TAPE ORDER 2			COMBINED		
	N	MALE/ FEMALE	MEAN AGE	N	MALE/ FEMALE	MEAN AGE	N	MALE/ FEMALE	MEAN AGE
1080I	13	8/5	23.6	14	7/7	22.4	27	15/12	23.0
720P	14	10/4	32.1	13	5/8	26.5	27	15/12	29.4

Test Material. A list of the 26 sequences used in the subjective tests of Basic Received Quality is presented in Table 2. Each sequence consisted of the central 10 seconds of a 15 second video clip. These sequences were selected by a panel of experts (FCC Advisory Committee on Advanced Television Service, Planning Subcommittee-Working Party 6) to ensure that a broad range of image attributes were represented. Of these 26 sequences, ten were retained from the first round of ATV testing, and are referred to as "core" sequences. They are marked with an asterisk in Table 2.

The columns labelled "1035I" and "720P" in Table 2 indicate the origin of the video material as follows: "Pixar" was scanned using a high resolution scanner; "Camera" was captured using an HDTV camera in the SMPTE 240M standard; "Computer Graphics" was rendered using animation software; "Film" was obtained from an HDTV telecine; "Converted" was converted from 1035I to 720P using digital signal processing techniques. Conversion resulted in a slight vertical stretch and loss of content, and minor horizontal misalignment of the 720P images, which was corrected by creating a distinct projector setup used exclusively for the display of these converted images. A small vertical and horizontal misalignment was also present for the 720P sequences S14A and M16A. This was the result of small spatial differences in the direct rendering of the 1035I and 720P source images, also used during Round 1 of testing. These inherent spatial differences were not entirely correctable.

Reference sequences for the both 1080I and 720P tests of the *digital* HDTV Grand Alliance System were always displayed in the 1035I source format. Test sequences were generated by processing (encoding, modulating, transmitting, demodulating and decoding) source sequences through the *digital* HDTV Grand Alliance System. For the 1080I mode of testing, the 1035I source sequences were used as input to the *digital* HDTV Grand Alliance System. The extra 45 lines were obtained by duplicating the top 40 lines of the 1035I image frame and adding 5 black lines at the bottom (i.e. $1035 + 40 + 5 = 1080$). Due to limitations of the Sony HDD-1000 VTR, it was not possible to record Test sequences with 1080I lines. Therefore, for storage, display and subjective evaluation, the extra 45 lines were stripped away, such that the 1080I mode of the *digital* HDTV Grand Alliance System was represented with 1035I lines. For the 720P mode of testing, the 720P source sequences were used as input to the *digital* HDTV Grand Alliance System. Test sequences were displayed in 720P.¹ For a detailed description of the material see Appendix A.

During testing, color discrepancies were identified in Pixar originated sequences (S1, S5, S6, S8, S9) between 1080I and 720P modes. The color of the Pixar originated sequences (for both Reference and Test) was shifted slightly towards red for 1080I as compared to 720P. The latter sequences matched those of Round 1 core sequences. For further details see Appendix D, document SSWP2-1465. In the opinion of staff at the ATEL and the ATTC, the magnitude of the color shift was small enough as to have no important effects on the performance of the *digital* HDTV Grand Alliance System, nor to preclude comparison of results from the current round of

¹ During testing of the 720P System, the Hitachi projector was switched between 1035I (Reference sequence) and 720P (Test sequence) seamlessly without perceptible jitter.

ATV testing with the first round of testing. For this reason, no further consideration is given to this issue in the present report.

TABLE 2
SOURCE SEQUENCES FOR BASIC RECEIVED QUALITY TESTS

NO	ID	PICTURE	IMAGE ATTRIBUTES	1035I	720P
1	S1*	Metal Table & Chairs	Luma Resolution	Pixar	Pixar
2	S5*	Tulips	Chroma Resolution	Pixar	Pixar
3	S6	Sculptures	Chroma Artifacts	Pixar	Pixar
4	S8*	Toys	Chroma Dyn Range & Compression Noise	Pixar	Pixar
5	S9*	Girls With Toys	Peripheral Performance And Texture	Pixar	Pixar
6	S14A	Cheshire Cat ²	Spatial Resolution	Computer Graphics	Computer Graphics
7	M1*	Window	Dynamic Luma Resolution	Video Camera	Converted ³
8	M2*	Fax Machine	Dynamic Luma Resolution	Video Camera	Converted ³
9	M4*	Mannequins	Chroma Artifacts	Video Camera	Converted ³
10	M5*	Living Room	Motion Rendition	Video Camera	Converted ³
11	M6*	Den	Motion Rendition In-Scene-Movement	Video Camera	Converted ³
12	M10*	Woman And Room	Motion Rendition & In-Scene-Movement	Video Camera	Converted ³
13	M16A	Rotating Pyramids ²	Motion Graphics	Computer Graphics	Computer Graphics
14	M35	Crosswalk	Motion, Multiple Object	Video Camera	Converted ³
15	M36	Georgetown Loop	Motion With Texture	Video Camera	Converted ³
16	M37	Buckingham Palace ⁴	Chroma Saturation	Video Camera	Converted ³
17	M38	Snow Tires / Trees	Luma Resolution	Video Camera	Converted ³
18	M39	End Zone	Crowd Scene, Slow Zoom	Video Camera	Converted ³
19	M40	Dream Team ⁴	Fast Motion, Scene Cuts	Video Camera	Converted ³
20	M41	Golf ⁵	Dissolve, Chroma Saturation, Zoom	Video Camera	Converted ³
21	M44	Mirror	Film Transfer 24fps	Film (24fps)	Converted ³
22	M45	Christa	Film Transfer 24fps	Film (24fps)	Converted ³
23	M46	Fountain	Film Transfer 30fps	Film (30fps)	Converted ³
24	M47	Clock #1	Computer Graphics, Pan & Zoom	Computer Graphics	Converted ³
25	M48	Connections	Motion Graphics	Computer Graphics	Converted ³
26	M49	Picnic With Ants ⁶	Noise Impairment	Pixar	Pixar

* Denotes Core Sequences

S in the ID column indicates a "still" sequence and M indicates a "motion" sequence

² S14A and M16A are corrected versions of S14 and M16.

³ Material converted from 1035I to 720P.

⁴ Sequence contains a scene cut.

⁵ Sequence contains a cross-fade.

⁶ Special image with noise encroaching from all sides.

Design and Procedures. Tests were conducted in two phases, 1080I was tested in the first phase, and 720P was tested in the second phase, using separate groups of viewers (see Table 1).

The design of each phase of testing was identical. There were three factors: Picture, Replicate and Tape Order. Picture and Replicate were varied within subjects and Tape Order was varied between subjects. Picture refers to the 26 sequences listed in Table 2. Replicate refers to the number of times a condition occurred during a session; each condition was rated twice per session for both Reference and Test. This yielded 52 trials per session (26 Pictures x 2 Replicates). Tape Order refers to the random order of the 52 trials; two Tape Orders were used. During each session, viewers completed all 52 trials plus 8 practice trials. Five practice trials were completed at the start of testing, and 3 were completed after a 30 minute rest-break midway through the session.

The layout of a Basic Received Quality assessment trial is shown schematically in Figure 1, and is based on the double-stimulus continuous quality scale method described in ITU-R Recommendation 500⁷. Each trial consisted of a pair of Reference and Test sequences presented twice in succession. When sequence A was a Reference, sequence B was a Test, and vice versa. Viewers were not informed whether A or B was the Reference or the Test sequence.

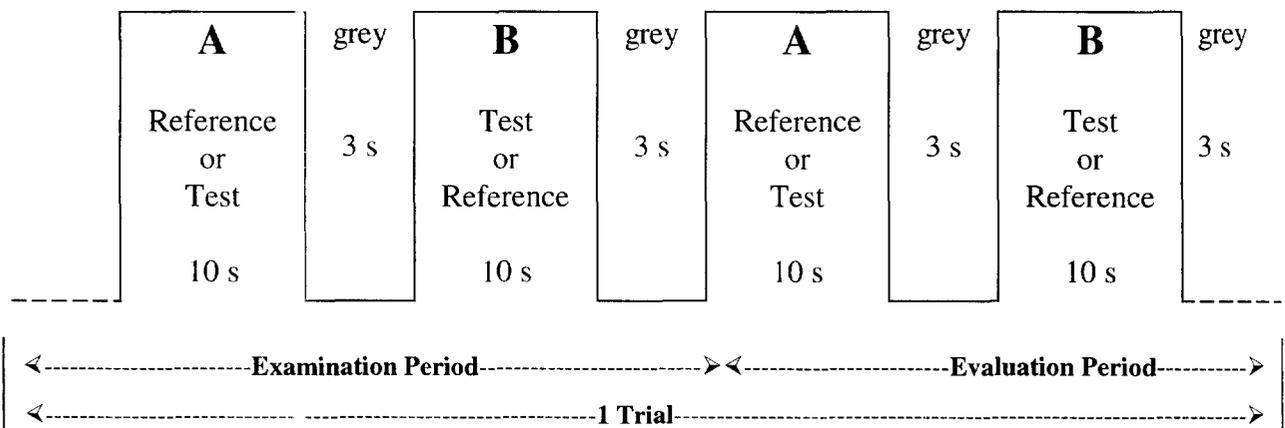


FIGURE 1. Layout of a Basic Received Quality assessment trial.

Viewers were instructed to rate the perceived image quality of the “A” and the “B” sequences using scales shown in Figure 2. These judgement scales were 100 mm in length. The labels “Excellent”, “Good”, “Fair”, “Poor” and “Bad” were printed at the locations shown in Figure 2. Numerical values in brackets are presented for the readers convenience only, and were not provided to the viewers.

⁷ ITU-R International Telecommunications Union Radiocommunications Sector is an organization of the International Telecommunications Union, a United Nations agency. Formerly it was called the International Radio Consultive Committee (CCIR).

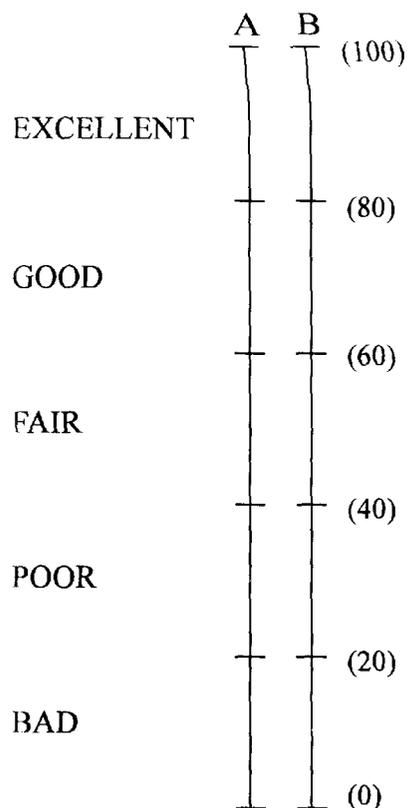


FIGURE 2. Rating scales used in the Basic Received Quality tests. Numerical values in brackets are presented for the readers convenience only and were not included on the forms used in the tests.

All viewers participated in a post session questionnaire. The results are summarized and presented in Appendix E.

Test Setup. The display was a Hitachi C65-4510R Multi-Scan rear projection HDTV monitor. Playback was from a Sony HDD-1000 VTR through a custom ATTC format converter. Audio tracks on the HDD-1000 VTR were used to record control signals. These provided advanced warning of changes in video format (i.e. from 1080I to 720P) to enable smooth blanking of the Hitachi display, and to enable correction of the slight vertical and horizontal misalignment in the converted 720P images. Audio announcements of trial numbers were presented using a Panasonic SV-3700 DAT Tape Deck.

Viewing Conditions. The conditions in the ATEL viewing room matched the specifications shown in Table 3. The layout of the viewing room is shown in Figure 3. The lightwall provided uniform illumination of the background surrounding the video display screen, and is shown schematically in Figure 4.

TABLE 3
VIEWING CONDITIONS FOR HDTV SUBJECTIVE ASSESSMENTS

CONDITION	SPECIFICATION ⁸	MAINTAINED VALUE
Peak monitor luminance (PML)	150 - 250 ⁹ cd/m ²	68 - 72 cd/m ²
Monitor luminance (maximum, at beam cut-off under ambient lighting conditions)	2 % of PML	1.5 - 2 % PML
Monitor luminance (maximum, at black-level in dark room)	1 % of PML	.5 - 1 % of PML
Monitor white color temperature	6500 ⁰ K	6400 - 6600 ⁰ K
Luminance of controlled monitor surround	15 % of PML	14 - 16 % of PML
Color temperature of controlled monitor surround	6500 ⁰ K	6400 - 6600 ⁰ K
Size of controlled monitor surround	not specified	54 ⁰ H x 82 ⁰ W
Room illumination	low	9.5 - 10.5 lux
Color temperature of room lighting	6500 ⁰ K	6400 - 6600 ⁰ K
Wall colors	minimum color	white/grey
Ratio of viewing distance to picture height	3 H	3 H
Monitor size	not specified	65 in
Maximum off-center angle of view for individual viewer	30 ⁰	26 - 28 ⁰

⁸ Source for specification is Document SSWP2-1306.

⁹ It has been determined that these levels cannot be achieved with available HDTV display equipment.

VIEWING THEATER

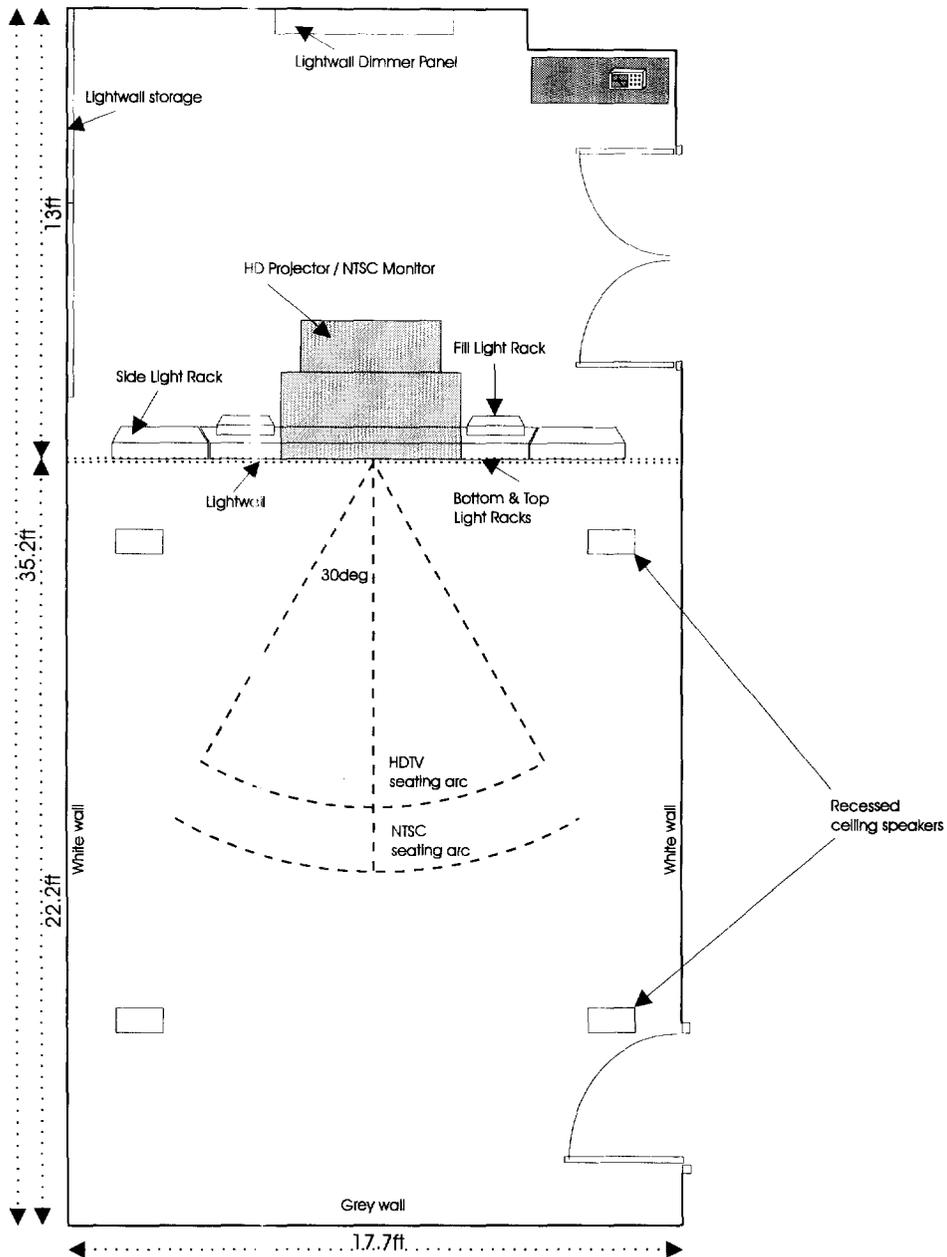


FIGURE 3. Controlled viewing environment.

LIGHTWALL

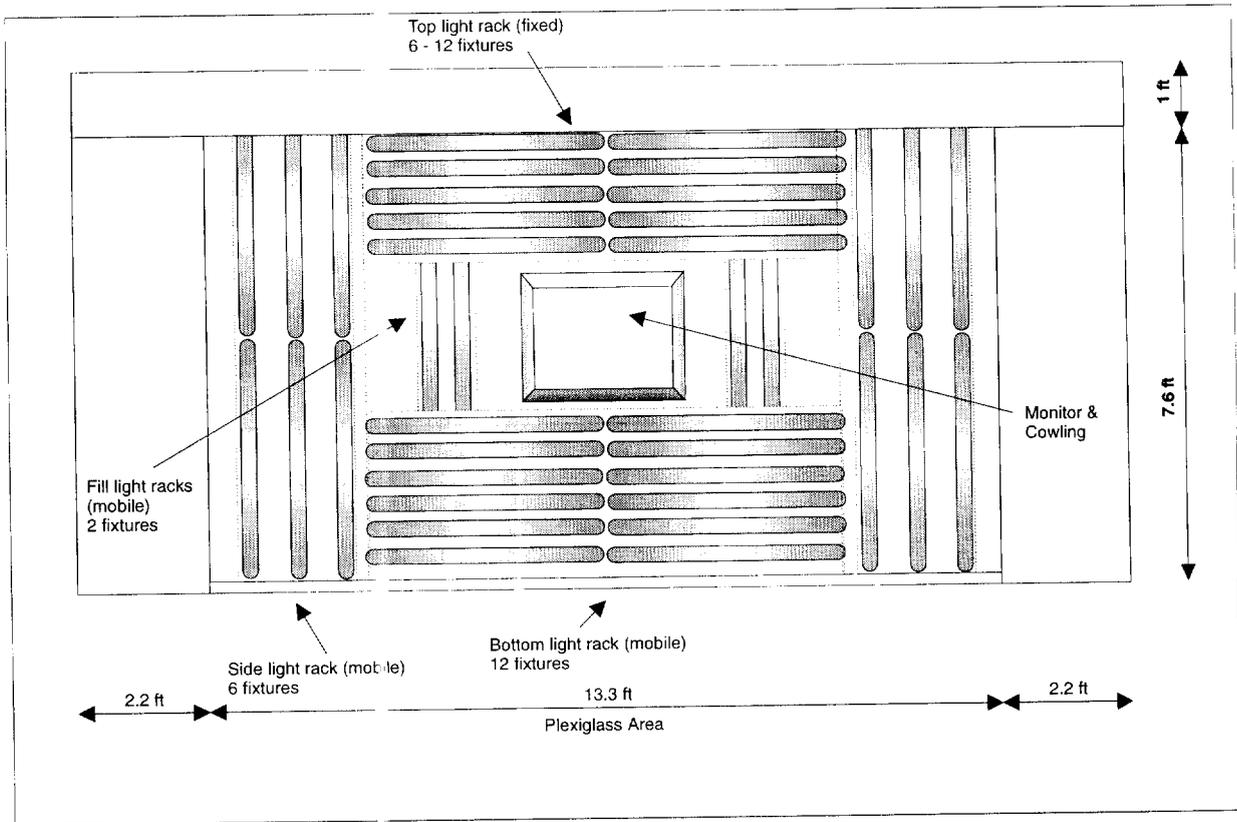


FIGURE 4. Diagram of the lighting arrangements for the lightwall.

2.3 Results - Basic Received Quality Assessments

Overview. The analyses examined the quality ratings of Reference and Test sequences for the 1080I and 720P modes of the *digital* HDTV Grand Alliance System.

The quality ratings, ranging from 0 to 100, were digitized for computer analysis from the continuous quality scales completed by the viewers. Judgements of "BAD" corresponded to values between 0 and 20; "POOR" corresponded to values between 20 and 40; "FAIR" corresponded to values between 40 and 60; "GOOD" corresponded to values between 60 and 80; and "EXCELLENT" corresponded to values between 80 and 100.

The results are organized in graphical and tabular form. The mean quality ratings for 1080I and 720P are shown in Figures 5 and 6, respectively. Difference scores (Test minus Reference) for 1080I and 720P are shown in Figure 7. The error bars in each figure represent the 95% confidence interval around the mean. Numerical values used to create these figures are shown on the facing pages in Tables 4, 5 and 6, respectively. In Figure 7, a score of 0 indicates that the mean rating of the Reference and Test sequence was equivalent. A rating greater than 0 indicates that the Test sequence was rated higher than the Reference sequence. A rating less than 0 indicates that the Test sequence was rated lower than the Reference sequence.

TABLE 4
BASIC RECEIVED QUALITY
1080I

ID	PICTURE	REFERENCE			TEST		
		MEAN RATING	STND. DEV.	CONF. INT.	MEAN RATING	STND. DEV.	CONF. INT.
S1*	Metal Table & Chairs	81.14	11.76	±3.21	82.14	10.86	±2.96
S5*	Tulips	80.47	11.72	±3.20	80.43	11.28	±3.08
S6	Sculptures	78.00	10.50	±2.87	79.04	10.39	±2.84
S8*	Toys	79.19	11.05	±3.02	78.97	12.38	±3.38
S9*	Girls With Toys	73.59	12.78	±3.49	73.47	12.32	±3.36
S14A	Cheshire Cat	80.82	10.79	±2.95	80.03	12.04	±3.29
M1*	Window	79.92	10.67	±2.91	78.30	10.61	±2.90
M2*	Fax Machine	78.60	11.71	±3.20	74.42	14.47	±3.95
M4*	Mannequins	73.76	14.69	±4.01	67.86	16.39	±4.47
M5*	Living Room	74.05	16.93	±4.62	66.54	19.70	±5.38
M6*	Den	75.73	11.05	±3.02	74.44	11.27	±3.08
M10*	Woman And Room	80.99	10.42	±2.84	77.51	12.59	±3.44
M16A	Rotating Pyramids	80.93	12.45	±3.40	78.91	11.44	±3.12
M35	Crosswalk	80.05	12.40	±3.39	76.38	12.56	±3.43
M36	Georgetown Loop	78.47	9.47	±2.59	74.58	12.25	±3.34
M37	Buckingham Palace	81.69	9.73	±2.66	73.71	17.94	±4.90
M38	Snow Tires / Trees	84.10	8.99	±2.45	77.35	12.66	±3.46
M39	End Zone	74.63	12.41	±3.39	72.09	13.98	±3.82
M40	Dream Team	77.53	14.68	±4.01	68.23	18.71	±5.11
M41	Golf	76.92	12.75	±3.48	77.00	11.23	±3.07
M44	Mirror	69.46	17.97	±4.91	68.67	17.26	±4.71
M45	Christa	65.35	15.11	±4.12	64.65	17.10	±4.67
M46	Fountain	78.93	12.99	±3.54	79.37	9.28	±2.53
M47	Clock #1	80.86	11.25	±3.07	78.86	13.53	±3.69
M48	Connections	85.18	10.06	±2.75	85.60	8.88	±2.42
M49	Picnic With Ants	77.28	14.91	±4.07	62.49	21.05	±5.75

*Denotes Core Sequences

**BASIC RECEIVED QUALITY
1080I**

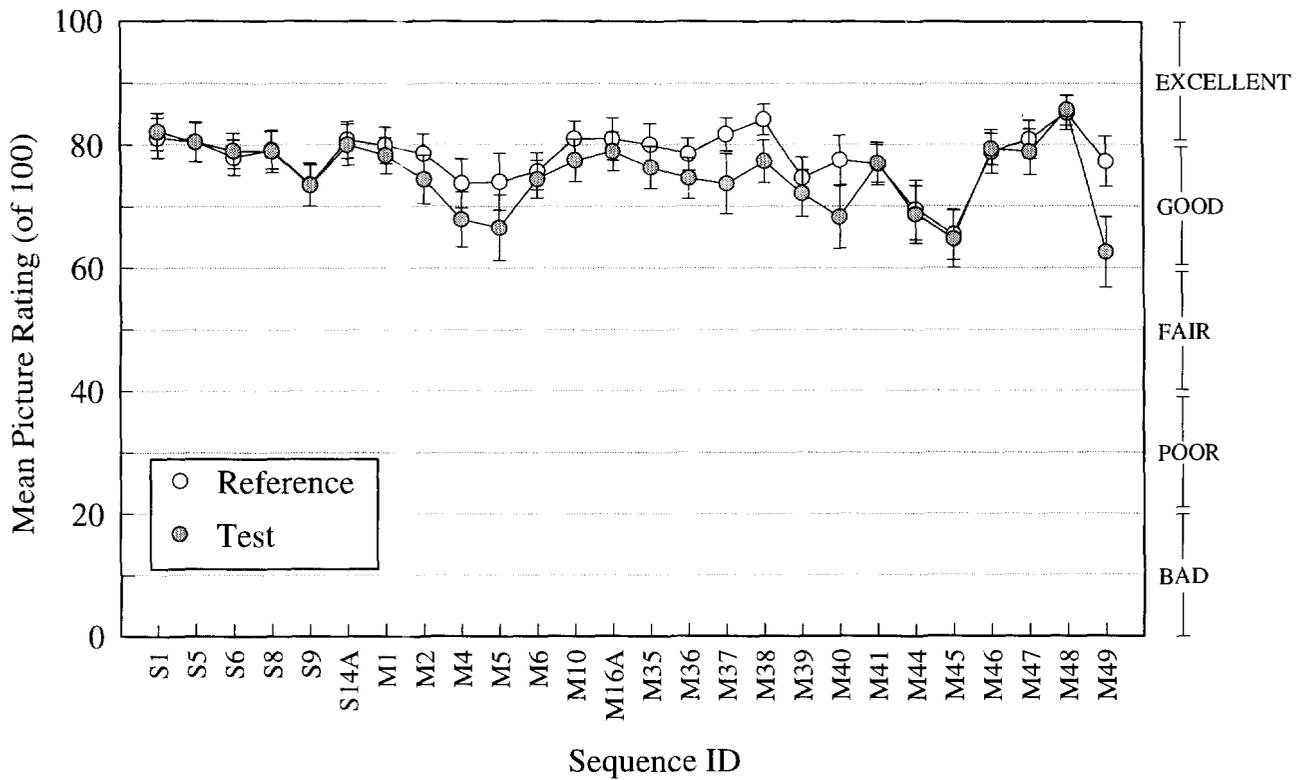


FIGURE 5. Mean quality ratings of Reference and Test sequences for the 1080I mode of the digital Grand Alliance HDTV System.

TABLE 5
BASIC RECEIVED QUALITY
720P

ID	PICTURE	REFERENCE			TEST		
		MEAN RATING	STND. DEV.	CONF. INT.	MEAN RATING	STND. DEV.	CONF. INT.
S1*	Metal Table & Chairs	79.22	12.45	±3.40	77.23	13.40	±3.66
S5*	Tulips	79.17	13.28	±3.63	76.10	15.05	±4.11
S6	Sculptures	75.31	12.24	±3.34	74.13	13.81	±3.77
S8*	Toys	75.86	12.89	±3.52	76.39	12.64	±3.45
S9*	Girls With Toys	68.83	16.43	±4.49	67.50	15.85	±4.33
S14A	Cheshire Cat	80.04	11.28	±3.08	77.26	12.67	±3.46
M1*	Window	79.36	9.81	±2.68	78.85	11.98	±3.27
M2*	Fax Machine	77.34	13.52	±3.69	76.44	13.57	±3.70
M4*	Mannequins	75.06	17.59	±4.80	70.64	19.42	±5.30
M5*	Living Room	72.66	14.30	±3.90	70.91	15.06	±4.11
M6*	Den	75.35	14.94	±4.08	75.05	12.90	±3.52
M10*	Woman And Room	78.40	12.89	±3.52	76.62	13.18	±3.60
M16A	Rotating Pyramids	82.19	13.31	±3.63	83.05	11.50	±3.14
M35	Crosswalk	77.54	15.86	±4.33	71.90	16.04	±4.38
M36	Georgetown Loop	80.50	13.29	±3.63	76.58	14.34	±3.91
M37	Buckingham Palace	82.31	12.43	±3.39	77.93	13.83	±3.78
M38	Snow Tires / Trees	85.11	9.85	±2.69	80.59	11.90	±3.25
M39	End Zone	72.57	15.82	±4.32	66.61	18.52	±5.06
M40	Dream Team	70.29	12.29	±3.35	73.05	16.02	±4.37
M41	Golf	79.89	10.39	±2.84	77.47	12.84	±3.51
M44	Mirror	68.26	16.84	±4.60	68.59	15.41	±4.21
M45	Christa	63.02	17.89	±4.88	62.42	17.85	±4.87
M46	Fountain	69.76	17.84	±4.87	69.32	18.85	±5.15
M47	Clock #1	81.00	12.65	±3.45	80.07	13.69	±3.74
M48	Connections	83.78	11.60	±3.17	83.49	9.95	±2.72
M49	Picnic With Ants	78.45	12.46	±3.40	51.49	22.33	±6.10

*Denotes Core Sequences

BASIC RECEIVED QUALITY 720P

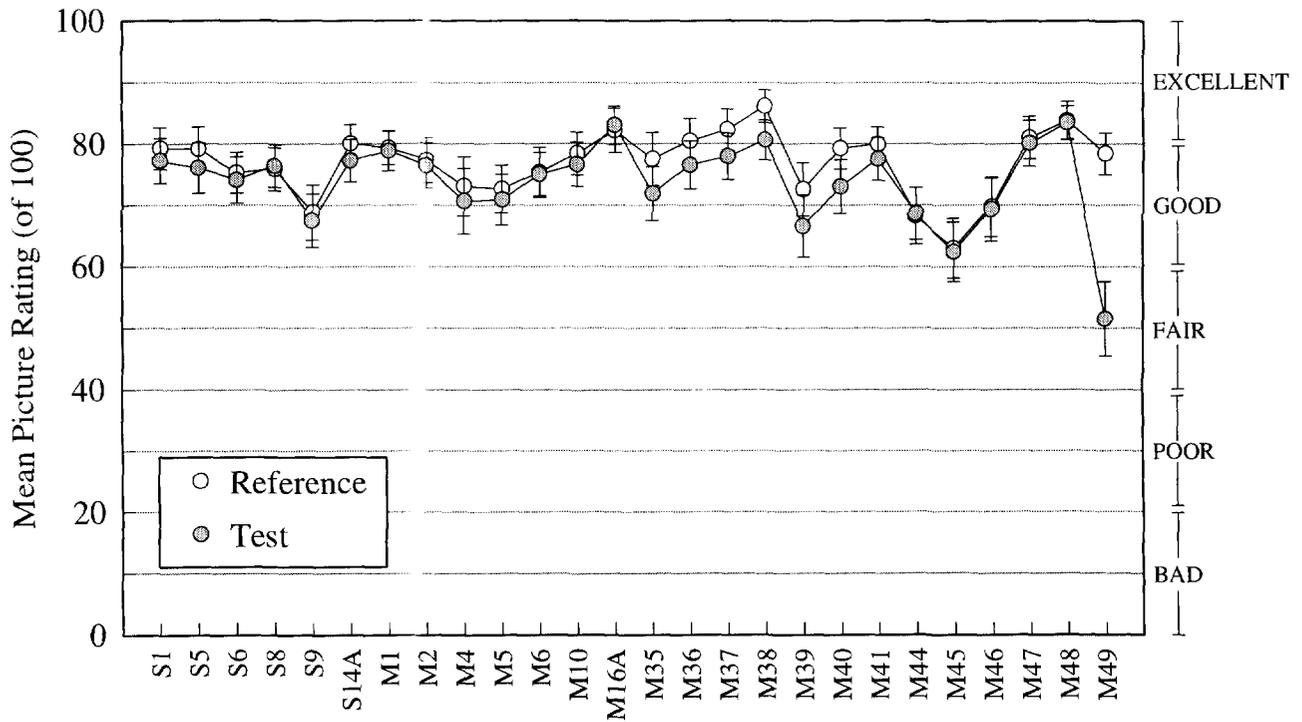


FIGURE 6. Mean quality ratings of Reference and Test sequences for the 720P mode of the *digital* Grand Alliance HDTV System.

TABLE 6
BASIC RECEIVED QUALITY
DIFFERENCE SCORES (TEST minus REFERENCE)

ID	PICTURE	1080I			720P		
		MEAN RATING	STND. DEV.	CONF. INT.	MEAN RATING	STND. DEV.	CONF. INT.
S1*	Metal Table & Chairs	1.00	5.70	±1.56	-1.99	9.33	±2.55
S5*	Tulips	-0.04	6.73	±1.84	-3.07	8.26	±2.25
S6	Sculptures	1.04	5.57	±1.52	-1.19	9.73	±2.66
S8*	Toys	-0.22	6.83	±1.86	0.53	8.27	±2.26
S9*	Girls With Toys	-0.12	5.79	±1.58	-1.33	8.50	±2.32
S14A	Cheshire Cat	-0.79	6.99	±1.91	-2.78	6.48	±1.77
M1*	Window	-1.61	7.28	±1.99	-0.50	7.19	±1.96
M2*	Fax Machine	-4.18	7.36	±2.01	-0.89	7.70	±2.10
M4*	Mannequins	-5.90	9.69	±2.65	-2.42	10.46	±2.85
M5*	Living Room	-7.51	10.57	±2.89	-1.75	9.52	±2.60
M6*	Den	-1.29	7.55	±2.06	-0.30	11.04	±3.01
M10*	Woman And Room	-3.49	7.21	±1.97	-1.78	6.65	±1.81
M16A	Rotating Pyramids	-2.03	8.47	±2.31	0.86	10.63	±2.90
M35	Crosswalk	-3.67	6.33	±1.73	-5.64	13.28	±3.63
M36	Georgetown Loop	-3.89	8.49	±2.32	-3.92	6.65	±1.82
M37	Buckingham Palace	-7.98	12.96	±3.54	-4.38	11.99	±3.27
M38	Snow Tires / Trees	-6.75	10.56	±2.88	-5.53	10.53	±2.87
M39	End Zone	-2.54	6.51	±1.78	-5.96	11.73	±3.20
M40	Dream Team	-9.30	14.39	±3.93	-6.24	12.17	±3.32
M41	Golf	0.08	7.51	±2.05	-2.42	10.24	±2.80
M44	Mirror	-0.79	7.36	±2.01	0.33	7.82	±2.14
M45	Christa	-0.70	7.70	±2.10	-0.60	9.03	±2.47
M46	Fountain	0.44	8.35	±2.28	-0.44	6.59	±1.80
M47	Clock #1	-2.00	8.14	±2.22	-0.92	10.24	±2.80
M48	Connections	0.43	6.80	±1.86	-0.29	9.25	±2.52
M49	Picnic With Ants	-14.79	16.80	±4.59	-26.95	22.93	±6.26

* Denotes Core Sequences

**BASIC RECEIVED QUALITY
DIFFERENCE SCORES (TEST minus REFERENCE)**

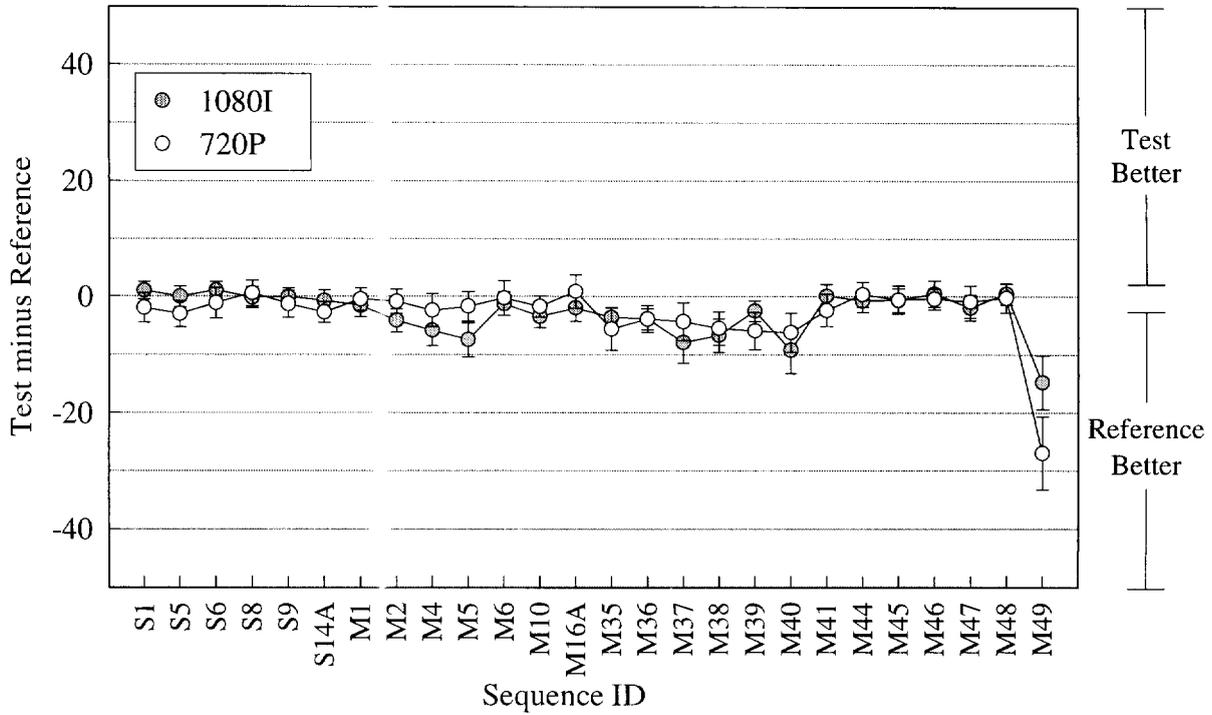


FIGURE 7. Differences in quality ratings (Test minus Reference) for the 1080I and 720P modes of the *digital* Grand Alliance HDTV System.

Statistical Analyses. The primary purpose of the statistical analyses was to determine whether differences between Reference and Test were statistically reliable; that is, whether the observed differences could be attributed to chance variation or to actual differences in rated quality. ANOVAs assessed overall effects of Tape Order, Reference *vs.* Test, Replicate and Picture. Newman-Keuls tests assessed effects on a sequence-by-sequence basis.

Statistical analyses were conducted in three stages: 1) Preliminary ANOVAs were conducted separately on Reference sequences and on Test sequences; 2) Comprehensive ANOVAs were conducted comparing Reference and Test sequences; 3) Pairwise Newman-Keuls comparisons were conducted on Reference and Test sequences. All analyses were performed separately for the 1080I and 720P modes of the *digital* HDTV Grand Alliance System. Details of the analyses are shown in Appendices H and I for the 1080I and 720P modes, respectively.

Preliminary ANOVAs (Tape Order x Replicate x Picture) and comprehensive ANOVAs (Tape Order x Reference *vs.* Test x Replicate x Picture) for both 1080I and 720P showed a large effect of Picture indicating that, overall, different sequences were rated at different levels of quality. For example, Christa (M45) was rated lower than Metal Table and Chairs (S1).

Comprehensive ANOVAs showed a significant main effect of Reference *vs.* Test, indicating that the Reference sequences were rated higher than the Test sequences. A significant 2-way interaction between Reference *vs.* Test and Picture indicated that the magnitude of the difference in rated picture quality between Reference and Test varied with Picture. For example, the difference between Reference and Test for Metal Table and Chairs (S1) was negligible, whereas the difference was larger for Dream Team (M40). Other interactions were examined and deemed to have no important implications for the evaluation of the *digital* HDTV Grand Alliance System.

Post hoc analyses were performed using the Newman-Keuls test. Significance levels are shown in Table 7. For 1080I, statistically significant effects at $p < .05$ were observed for sequences M2, M4, M5, M37, M38, M40 and M49. For 720P, statistically significant effects at $p < .05$ were observed for sequences M35, M38, M39, M40 and M49. For these sequences, the Test sequence was rated significantly lower in perceived quality than the Reference sequence.

The largest difference between Test and Reference occurred for sequence M49. M49 consisted of a central still image with noise encroaching from the sides. This type of sequence is known to be particularly stressful for image compression algorithms.

TABLE 7
SUMMARY OF POST HOC TESTS
REFERENCE vs. TEST

NEWMAN-KEULS			
ID	PICTURE	1080I	720P
S1*	Metal Table & Chairs	.70	.97
S5*	Tulips	.99	.79
S6	Sculptures	.99	.87
S8*	Toys	.87	.99
S9*	Girls With Toys	.93	.80
S14A	Cheshire Cat	.92	.88
M1*	Window	.97	.98
M2*	Fax Machine	.05	.99
M4*	Mannequins	.00	.61
M5*	Living Room	.00	.64
M6*	Den	.74	.89
M10*	Woman And Room	.36	.95
M16A	Rotating Pyramids	.93	.86
M35	Crosswalk	.27	.04
M36	Georgetown Loop	.08	.46
M37	Buckingham Palace	.00	.27
M38	Snow Tires / Trees	.00	.01
M39	End Zone	.54	.00
M40	Dream Team	.00	.01
M41	Golf	.93	.88
M44	Mirror	.55	.84
M45	Christa	.55	.69
M46	Fountain	.99	.78
M47	Clock #1	.92	.97
M48	Connections	.70	.89
M49	Picnic With Ants	.00	.00

* Denotes Core Sequences

Significance levels $p < .05$ are shaded.

2.4 Comparative - Basic Received Quality

To facilitate evaluation of the *digital* HDTV Grand Alliance System, the observed performance of the System was compared against target specifications obtained from the "Grand Alliance HDTV System Specification" manual (Chapter 9, version Dec. 7, 1994) and against test results from the first round of A TV testing.

Target Specifications. Target specifications for the *digital* HDTV Grand Alliance System are shown in Tables 8 and 9 for the 1080I and 720P modes, respectively. These indicate for each category the maximum value of the mean difference score, that is, the maximum amount by which the average rating of the Test sequences can fall below the average rating of the Reference sequences. For certain categories, no target specifications were indicated in the System Specification Manual. Observed performance levels of the System are shown alongside the target specifications. Note, that in all cases the *digital* HDTV Grand Alliance System exceeded the target specifications by a wide margin.

TABLE 8
TARGET SPECIFICATIONS
1080I

TEST SEQUENCE 1080I	TARGET SPECIFICATION ¹⁰	OBSERVED PERFORMANCE	
		MEAN (TEST minus REF.)	CONFIDENCE INTERVAL
Basic Material ¹¹	≥ -6.0 ¹²	-2.42	± 2.06
Noise & cuts (M37, M40, M41, M49)	≥ -20.0 ¹³	-8.00	± 3.53
Graphics (S14A, M16A, M47, M48)	≥ -20.0 ¹³	-1.10	± 2.08
Film 24fps (M44, M45)	≥ -5.0 ¹⁴	-0.74	± 2.05
All sequences	None	-2.95	± 2.29
All stills (S)	None	0.14	± 1.71
All motion (M)	None	-3.87	± 2.46

TABLE 9
TARGET SPECIFICATIONS
720P

TEST SEQUENCE 720P	TARGET SPECIFICATION ¹⁰	OBSERVED PERFORMANCE	
		MEAN (TEST minus REF.)	CONFIDENCE INTERVAL
Basic Material ¹¹	≥ -6.0 ¹²	-2.26	± 2.48
Noise & cuts (M37, M40, M41, M49)	≥ -20.0 ¹³	-10.00	± 3.91
Graphics (S14A, M16A, M47, M48)	≥ -20.0 ¹³	-0.78	± 2.50
Film 24fps (M44, M45)	≥ -5.0 ¹⁴	-0.13	± 2.30
All sequences	None	-3.06	± 2.69
All stills (S)	None	-1.64	± 2.30
All motion (M)	None	-3.49	± 2.81

¹⁰ Target specifications are reported in units, where 20 units = 1 grade.

¹¹ Basic Material consisted of the following 16 sequences; S1, S5, S6, S8, S9, M1, M2, M4, M5, M6, M10, M35, M36, M38, M39 and M46.

¹² Corresponds to ≤ 0.3 grade.

¹³ Corresponds to ≤ 1.0 grade.

¹⁴ Corresponds to ≤ 0.25 grade.