1.2.1 Single Variable Ranging Test

In single variable ranging the value of the single impairment being tested is slowly varied. The expert observers then decide, as described elsewhere [See Sections 1.4.1 and 1.4.2], on the threshold of visibility and the point of unusability.

The threshold of visibility and the point of unusability make up a range of impairment which the observers divide into steps which are used in possible subsequent rating tests.

1.2.2 Multiple Variable Ranging Test

In multiple variable ranging it is recognized that many impairments are characterized by more than one variable, e.g., multipath (Desired/Undesired and delay), airplane flutter (D/U, delay, and Doppler shift). For these impairments it is necessary to hold all variables but one constant for any given ranging test. This means that there will be a large number of ranging tests for multiple variable ranging tests.

1.2.3 NTSC Comparative Ranging

In these tests a ranging test on an NTSC signal will first be done up to the point of establishing the threshold of visibility for the impairment and then recording the result. The NTSC signal will then be replaced by the system under test [with the impairment and the system under test at the same power levels as were used for the NTSC test]. The observers will then establish the threshold of visibility for the system under test. This level of impairment for the system under test will be recorded relative to the level for NTSC, e.g., "the system under test reaches visibility at a level 3 db less than NTSC."

The above procedure will then be used to find the unusable point.

1.2.4 Main/Augmentation Ranging Test

In these tests, several ranges of main channel impairment will be investigated so that a curve can be drawn showing the range of augmentation

impairment for a given level of main channel impairment.

In augmentation systems, the complete television signal is carried by two separate carriers. It is necessary, therefore, to test the effects that differing amounts of an impairment may have on the two carriers. In these tests, a fixed amount of an impairment is added to the main channel carrier, and a varying amount is added to the augmentation carrier.

First the main channel is tested alone with a standard ranging test as described above. Then the augmentation channel is turned on. The impairment added to the augmentation channel is then varied over a wide range for each main channel impairment step. The two points at which demarcation of the parts of the picture which are carried by the two carriers can just be seen are recorded; this is done for all main channel impairment steps.

In these tests there can be two thresholds of visibility and two unusable points. The first set of visibility and unusable points occur when the augmentation impairment is increased. The process is then reversed, and augmentation impairment is gradually reduced. The point at which picture parts cease to match is now sought again. Depending on the main channel impairment, this level may or may not be found; the same is true of the unusable point.

1.3 Rating Tests-General Description

The purpose of rating tests is to ascertain subjective assessments by lay observers of picture quality and impairments. These judgments will be made from video tapes played back for groups of lay observers.

There are two types of rating tests: quality and transmission impairment. Quality tests assess the intrinsic picture reproduction of a system under test relative to a reference television picture. Transmission impairment tests allow observers to rate the degree of degradation of pictures subjected to various levels of impairment (such ratings being made relative to the same picture unimpaired).

1.3.1 Quality Rating Test

The purpose of the quality tests is to ascertain the impression of overall picture quality. In these tests, lay observers will compare the unimpaired picture quality of the system under test with the quality of an unimpaired 1125/59.94 [or 60] reference picture.

The actual method of testing is a variant of the double-stimulus continuous quality method described in CCIR Rec 500-3. In this method the reference picture and the same picture via the system under test are shown sequentially to the observers. This is then repeated. Each observer then individually rates the reference and the system under test.

Twenty lay observers shall view and rate the trials for each ATV system. Additional lay observers may be required if it is determined that the distribution of test results is too broad and that additional lay observer judgements might reasonably be expected to improve the distribution of results.

1.3.2 Transmission Impairment Rating Test

The purpose of the transmission impairment tests is to ascertain observer reaction to transmission impairment. In these tests the observers will compare an unimpaired picture with an impaired one (both pictures being in the format of the system under test).

The actual method of testing is the double-stimulus impairment method described in CCIR Rec 500-3. In this method the unimpaired picture is first shown to the observers followed by the same picture with the impairment added. Each observer then rates the system under test with respect to the unimpaired first picture.

Twenty lay observers shall view and rate the trials for a given impairment and ATV system. Additional lay observers may be required if it is determined that the distribution of test results is too broad and that additional lay observer judgements might reasonably be expected to improve the distribution of results.

1.4 Ranging Tests-Detailed Information

1.4.1 Determination of the Threshold of Visibility

This is the method by which the threshold of visibility is ascertained. The method for one trial is as follows:

The impairment level is set at a level well below the threshold of visibility.

The level is then increased in 1 db steps every three seconds. (Depending on the impairment, the step size and timing may have to be modified.) As the impairment become visible to each observer, he presses his voting button. When four of the five observers have so voted, the level is recorded, and the impairment begins to decrement in 1 db steps. Two short beeps will be sounded to alert the observers that decrementing has begun. When four of the five observers vote that the impairment is no longer visible, the level is recorded, two short beeps are sounded and the impairment begins to increment in 1 db steps. This cycle is repeated four times for a total of five cycles.

If the maximum and minimum levels of the last cycle are within 3 db of each other, then the threshold of visibility is computed by averaging the last two maximum and minimum levels.

If the maximum and minimum levels of the last cycle are not within 3 db then convergence must be verified as follows:

> The sum of the differences between the maximums and minimums for cycles two and three are compared to the sum of the differences of the maximums and minimums for cycles four and five. If the cycle four and five sum is smaller, then the threshold of visibility is the average of the maximum and minimums of cycles four and five. If the cycle two and three sum is smaller, then proper convergence has not been achieved, and the entire process must be repeated.

If the threshold of visibility persistently will not converge, an average of cycles four and five

shall be taken and the fact of poor convergence noted.

1.4.2 Determination of the Point of Unusability

The point of unusability found by taking the average of three iterations of the following:

> The impairment is set randomly at a point well below the point of unusability. It is then incremented until four of the five observers vote that the point of unusability has been reached.

On the second and third iterations, the impairment starting points are chosen to be random and different for each iteration.

1.4.3 Choice of Picture Material

The picture used for the Desired signal shall be chosen to be sensitive to the impairment being investigated. Generally a 30% grey flat field will be used.

The picture used for the Undesired shall be chosen to maximally interfere with the desired signal.

1.4.4 Determination of Rating Test Step Sizes

After the observers determine the threshold of visibility and the point of unusability, the test administrator shall inform the observers of the range between the two values. The observers shall then decide by consensus what test step sizes are to be used in subsequent rating tests. This information will be recorded in written form by the test administrator as part of the test results.

The step sizes shall be determined by the Method of Bisection. The observers determine a point that they feel is perceptually "halfway" between the threshold of visibility and the point of unusability. Each half of this range is then "halved" perceptually. The resulting quarters are again "halved" perceptually. In this way nine points defining eight steps of equal perceptual size are derived.

From these nine points, the observers shall select 5 to 7 points. These points shall include a step below threshold (impairment off), the step just below the point of unusability, and three to five intermediate steps which favor the low impairment end of the range which is the range of greatest interest.

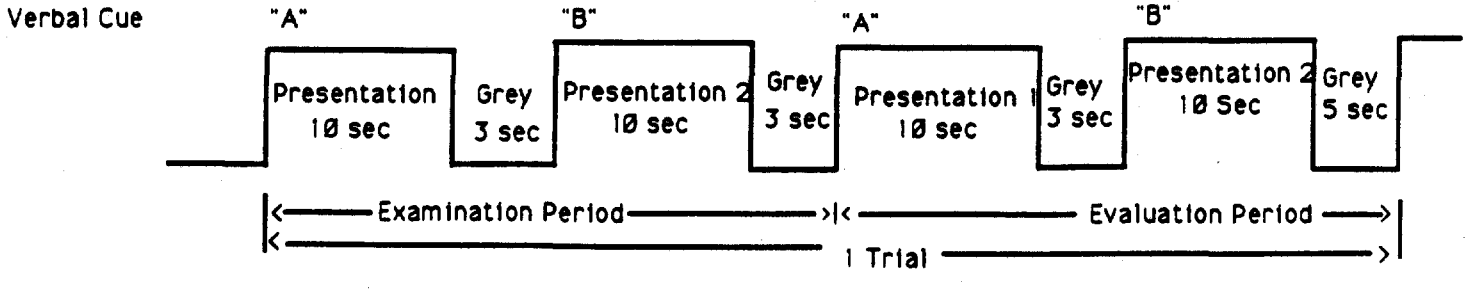
1.5 Quality Rating Tests-Detailed Information

1.5.1 Test Material

The test material, selected by PS/WP-6, shall consist of still and moving pictures in a pseudo-random order. The same two pseudo-random orderings shall be used for all proponents. Each of the two pseudo-random sequence shall be such that the same picture material is never used in two consecutive trials. All test trials shall be presented twice.

1.5.2 Description of a Single Trial

A single trial consists of two presentations of a given piece of test material. One presentation shows the test material in 1125/59.94 [or 60]/2:1 (SMPTE 240 M). This is the reference, although the observers are not told this. The other presentation is the system under test showing the same test material. The two presentations are shown twice separated by a period of mid-level grey. The following diagram shows the whole sequence:



Each presentation is identified to the observers by a spoken cue saying either "A" or "B" recorded on the video tape. On half the trials, "A" will be used to identify the reference and "B" will be used to identify the system under test. During the other trials "A" will designate the system under test and "B" will designate the reference. The "A" and "B" designations will change in a pseudo-random fashion during the session. The first and third presentation of a trial will always be called "A" and the second and fourth "B".

The "PRESENTATION 1" and "PRESENTATION 2" designations are arbitrary names for the reference and system under test. During the session either designation can be applied to either presentation as described below.

During a session both the "A" and "B" verbal cues and the order of presentation of the reference and system under test during a given trial (e.g., which system is "PRESENTATION 1" and which is "PRESENTATION 2") will change in a pseudo-random fashion. The same pseudo-random ordering will be used for all proponents.

Each trial will be identified by a verbal cue, "Trial xx", recorded on the video tape.

1.5.3 Test Session Description

Prior to the beginning of a test session, formal instructions are given to the observers. These instructions are both written and verbal. At this point the trials begin.

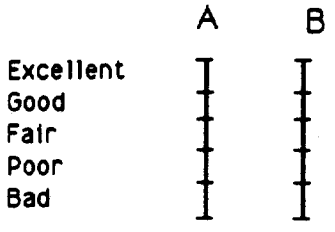
For a given ATV system a test block shall consist of four demonstration trials followed by five practice trials (The results from these five trials are not used.) followed by the actual trials.

Each observer shall participate in only a single session.

Normally a session shall be conducted with five observers.

1.5.4 Judgement Scale

The figure below shows the printed scale that the observers will use to rate each trial.



1.5.5 Instructions to Observers

The following representative instructions are given both in writing and presented verbally (via recording) to each group of observers:

"This experiment is one of a series being carried out to evaluate new kinds of television for reception in the home. In this experiment, we ask you to evaluate the overall quality of the pictures you see."

"The experiment will consist of a series of evaluation trials. Each trial will involve...."

"We will now show you four sample trials. The first two of these will show still pictures, and the second two will show moving pictures. These will illustrate the kinds of pictures you will be asked to evaluate."

[PRESENT DEMONSTRATION TRIALS HERE]

"In the experiment, we will ask you to judge the overall quality of the pictures you see. To do this, you.... (Describe rating scale and judgement task)"

"Have you any questions?"

[AFTER QUESTIONS, BEGIN PRACTICE AND ACTUAL TRIALS]

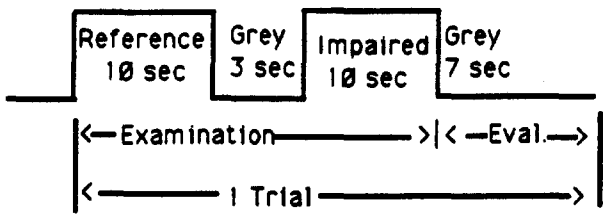
1.6 Transmission Impairment Rating Tests-Detailed Information

1.6.1 Test Material

The test material, selected by PS/WP-6, shall consist of still and moving pictures in a pseudo-random order. For a given number of steps the same two pseudo-random orderings shall be used for all proponents. Each of the two pseudo-random sequence shall be such that the same picture material is never used in two consecutive trials. All test trials shall be presented twice. The first actual trial on the video tape shall have readily discernable impairment.

1.6.2 Description of a Single Trial

A single trial consists of two presentations of a given piece of test material. The first presentation shows the test material on the system under test without impairment. This is the reference. The second presentation is the system under test showing the same test material with impairment. The two presentations are shown once separated by a period of mid-level grey. The following diagram shows a complete trial:



The first member of the pair is identified to the observers as the unimpaired reference; the second member of the pair is identified as the test picture.

Each trial is identified by a verbal cue recorded on the video tape as "Trial xx".

For a given number of steps the pseudo-random ordering of the chosen levels of impairments [see 1.4.4] and related pictures shall be the same for

all proponents for a given impairment. The first presentation for any test shall be the median impairment level identified in ranging tests.

1.6.3 Test Session Description

Prior to the beginning of a test session, formal instructions are given to the observers. These instructions are both written and verbal. At this point the trials begin.

For a given impairment a test block shall consist of four demonstration trials followed by five practice trials followed by the actual trials each of which shall be repeated. [This allows the repeatability of the observers to be checked.] The number of actual trials will vary depending on the number of steps selected during the ranging process.

Each observer shall participate in only a single session.

Normally a session shall be conducted with five observers.

On the average a test block for a given impairment will last about 30 to 35 minutes. To best utilize the lay observers it is desirable to have sessions which last from one to one and a half hours. Therefore two to three impairment block would normally be possible in a given session. A natural break lasting ten minutes is placed between blocks.

1.6.4 Judgement Scale

The figure below shows the printed scale that the observers will use to rate each trial.

circle one

5 Imperceptible
4 Perceptible, but not Annoying
3 Slightly Annoying
2 Annoying
1 Very Annoying

1.6.5 Instructions to Observers

The following representative instructions are both given in writing and presented verbally (via recording) to each group of observers:

"This experiment is one of a series being carried out to evaluate new kinds of television for reception in the home. In this experiment, we ask you to evaluate the effects of certain kinds of impairments that can occur in television reception"

"The experiment will consist of a series of evaluation trials. Each trial will involve...."

"We will now show you four sample trials. The first two of these will show still pictures, and the second two will show moving pictures. These will demonstrate the kinds of impairments you may see in the experiment to follow."

[PRESENT DEMONSTRATION TRIALS HERE]

"In the experiment, we will ask you to judge the degree of impairment to the pictures you will see. To do this, you.... (Describe rating scale and judgement task)"

"Have you any questions?"

[AFTER QUESTIONS, BEGIN PRACTICE AND ACTUAL TRIALS]

2.0 DEFINITIONS

2.1 Expert Observer

An expert observer is an individual who has had "recent extensive experience in observing picture quality or impairments, particularly of the type being studied in the subjective test." (CCIR Rec 500-3).

Expert observers shall have normal color vision and visual acuity. Prior to being used as an expert observer, an individual shall have passed a color blindness test (Ishihara or equivalent), a visual acuity test (Snellen Chart or equivalent) that shows the observer has at least 20/20 vision (corrected) in both eyes, and a contrast sensitivity test (Pelli-Robson or equivalent).

Normally expert observers will be engineers or technicians who work in the television industry. However, for certain impairments an "expert" may be someone who works in the entertainment industry in a creative capacity who would be accustomed to critical inspection of television pictures, e.g., a tape or film editor might be a very good expert observer in looking for transition artifacts in the "Video Cut" testing.

2.2 Lay Observer

A lay observer is an individual with no special technical training or experience in television picture quality or impairments.

Lay observers shall have normal color vision and visual acuity. Prior to be using used as a lay observer an individual shall have passed a color blindness test (Ishihara or equivalent), a visual acuity test (Snellen Chart or equivalent) that shows the observer has at least 20/20 vision in both eyes, and a contrast sensitivity test (Pelli-Robson or equivalent).

Normally lay observers are regular television viewers.

2.3 D/U Ratio

The D/U ratio is the ratio normally expressed in db between the Desired and Undesired signal.

2.4 Unusable

This word is used to describe a television picture that an observer would not deem "watchable" by the public.

2.5 Advanced Television System (ATV)

For the purposes of this document, advanced television systems (ATV) refers to any television system whose proponent represents it as having performance superior to that of NTSC.

2.6 Presentation

A presentation is a single picture or motion sequence be it a reference or test picture.

2.7 Test Trial

A test trial is a sequence of reference and test presentation(s) along with an evaluation period.

2.8 Demonstration Trial

A test trial designed for observer orientation. No test results are inferred from demonstration trials.

2.9 Block

A block is a series of contiguous test trials that all deal with a single impairment or with the quality of a single ATV system. There shall be no intermingling of impairments or ATV systems within blocks. Each block shall begin with demonstration trials to familiarize the observers with the test.

2.10 Session

A session is one or more blocks of tests conducted with a single group of observers. No more than thirty to forty minutes of actual testing shall occur within a session without at least a ten minute break.

Within a given session, each block of tests shall involve independent randomizations of the test pictures and the relative orderings of the impairment levels.

3.0 VIEWING CONDITIONS

3.1 General

All subjective tests shall be conducted in controlled viewing environments. The monitors/display devices shall be calibrated prior to any testing, and the layout and room conditions shall be adhered to throughout the testing.

3.2 Test Viewing Conditions

3.2.1 Viewing Room

The room shall be dimly lit. The exact room illumination shall be such that the requirements below for a dark screen shall be met.

The color temperature of the room lighting shall be D_{6500}.

The room decor shall be neutral and low key. No bright colors shall be used.

Comfortable chairs shall be provided for each observer.

The room shall be quiet and free from aural distractions.

3.2.2 Multiple NTSC Receivers for Ranging Tests

There shall be 24 representative NTSC television receivers.

NTSC images shall be evaluated at a distance of 5 to 6 picture heights. To the extent possible the observers shall be ± 30° from a perpendicular at the center of the kinescope face.

The display(s) shall be adjusted for a peak screen luminance of 70 ± 10 candelas per square meter. At beam cut-off the screen black shall be ≤ 1.2 candelas per square meter. With the receiver power turned off the screen shall be ≤ 0.6 candela per square meter. It is recognized that not all receivers will be capable of meeting these requirements.

The brightness and contrast of all receivers

shall be set using a standard PLUGE signal. The chroma phase and gain of all receivers shall be set using SMPTE 75% color bars.

3.2.3 ATV Display Device

The ATV display device shall be a large screen multi-standard video projector type display.

The 5 observers shall be located at a distance of 3 picture heights from the display along a radius extending from the center of the display screen. They shall be located within a cone ±30° from a perpendicular extending from the center of the viewing screen.

The wall behind the display device shall be painted a neutral gray and shall be dimly illuminated by D_{6500} light sources. The luminance of these walls shall be ≈12 cd/m².

The brightness and contrast of the display shall be adjusted with a standard PLUGE signal. The red, blue, and green gains shall be checked with 75% color bars. The overall color temperature of the display shall be adjusted for a peak white color temperature of 6500° K.

The aspect ratio of the display shall be masked for 16:9. The ATV images from all proponents systems shall be displayed at the same height.

At beam cutoff the screen luminance shall not exceed 2% of the peak luminance. With the display power off the screen luminance shall not exceed 1.75 cd/m².

The following shall be a goal for the ATV display:

> The peak luminance of the ATV display shall be adjusted for 150-200 cd/m².

3.3 Use of Receivers and ATV Display

All viewing where the NTSC picture is the Desired signal shall be done using NTSC display(s). All viewing where the ATV signal is the Desired signal shall be done using the ATV Display Device.

It should be noted that during testing it will be necessary for the observers to move back and forth between the NTSC display(s) and the ATV Display especially during NTSC Comparative Ranging Tests.

4.0 VIDEO TAPING OF RANGING TESTS

4.1 General

Video tapes shall be made showing the impairment at each of the nine levels set forth in 1.4.4. The video material recorded shall be the same as that viewed by the expert observers during the ranging process. These tapes may not be made at the exact time the ranging tests are conducted.

4.2 Recording of Specific Levels of Impairment

The test administrator may, at his discretion, record still and moving pictures of a system under test at various levels of impairment for the purpose of documenting performance at impairment levels which may be of interest to regulatory bodies.

4.3 Rating Master Tapes

Master tapes shall be recorded for any future rating tests. These tapes shall contain the selected test pictures impaired at the levels determined during the ranging process. These master tapes shall contain the five to seven levels of impairment (as derived in Section 1.4.4) which will actually be used to make the rating test tapes.

5.0 TEST MATERIAL

5.1 ATV Material

ATV systems are such that their characteristics change greatly when there is motion in the transmitted picture. For this reason each test shall be conducted with both still and moving pictures. Still and moving pictures shall be transmitted for every test by the ATV system regardless of whether it is the desired or undesired signal.

For the rating tests, the following visual test materials will be used:

ATV Quality:	9 still pictures and 14 motion sequences
ATV Impairment:	2 still pictures and 1 motion sequence

5.2 NTSC Material

Normally the NTSC carrier shall transmit 75% color bars or Matrix pattern (described in the Objective Test Plan) when it is the Undesired interferer. When NTSC is the Desired signal, a 30% flat field shall be used. When, however, there exists the possibility that the NTSC Desired may be subjected to non-linear interference such as the UHF Taboo testing, 75% color bars may also be used for the NTSC Desired.

6.0 PREPARATION OF RATING VIDEO TAPES

6.1 General Requirements

Each video tape shall have at its beginning the following:

* 1 minute of color bars and 400 hz 0 level tone on the designated audio program channel

* 30 seconds of a slate containing the following information:

Copyright Notice
Serial Number of the Tape
Name of Proponent System Under Test and Scanning Standard
Type of Test A. Quality or B. Transmission Impairment
If Transmission Impairment, then the impairment(s) being tested
Preparation Date of the Tape

* 10 seconds of black

NOTE: If the tape is quality rating, then the above shall be in 1125/59.94 [or 60]. If the tape is transmission impairment rating, the tape shall be in the scanning standard of the system under test.

6.2 Randomization of Test Sequence

The two rating tapes for a given type of test shall use the same test materials, one tape in each of two pseudo-random orders. The two pseudo-randomizations shall be designed to counteract, when the results of the tests are analyzed, any effects arising from adjacency of picture material or impairment level or from learning curve. Also as described in 1.6.2 above, the same relative ordering of impairment levels shall be used in Transmission Impairment Rating Tapes for a given impairment.

7.0 DATA ANALYSIS

7.1 General

[The exact methods to be used by the ATTC to reduce, analyze, and present the data is still under study. Comments are sought on this matter.] All analysis shall employ recognized statistical methods.

7.2 Number of Observers in Rating Tests

After the required number of observers have viewed a particular rating test, the results shall be analyzed and the standard deviation calculated. If the standard deviation is large, the test administrator may determine that additional testing is desirable or may decide that there is a flaw in the test design. If so, he may modify the test method and/or material and suggest retesting.

8.0 TEST METHODS FOR INDIVIDUAL IMPAIRMENTS

8.1 General

This section details the particular test method to be used for a given impairment. In some cases an impairment will be first tested by ranging and then by rating so its name will appear more than once in the following lists. [The parentheses refer to the Objective Test Plan and the CableLabs Test Plan.]

8.2 Single Variable Ranging Tests

* Carrier-to-Random Noise Ratio (19.1.1)
* Discrete Frequency Interference (19.3.3.2.4)
* Effects of Cable Second Order Intermodulation (3.4)
* Effects of Cable Third Order Intermodulation (3.4)
* Effects of Cable High Level Sweep (5.4)
* Effects of Cable Hum and Low Frequency Noise (2.9)
* Cable ICPM (7.4)

8.3 Multiple Variable Ranging Tests

* Multipath (19.2.1)
* Airplane Flutter (19.2.1)
* Effects of Cable Microreflections (2.7)

8.4 NTSC Comparative Ranging

* Upper Adjacent Channel Interference (19.3.1)
* Co-channel Interference (19.3.1)
* Lower Adjacent Channel Interference (19.3.1)
* UHF Taboo Interference (19.3.3.2.3)
* Carrier-to-Impulsive Noise Ratio (19.1.2)

NOTE: Interference Tests shall be conducted with NTSC as Desired and ATV as Undesired, ATV as Desired and NTSC as Undesired, and with ATV as both Desired and Undesired, if appropriate.

8.5 Main/Augmentation Ranging Tests

* Side Panel/Center Area Differential Random Noise Susceptibility (19.1.3)

8.6 Quality Rating Tests

* All Proponent Systems

8.7 Transmission Impairment Rating Tests

* Carrier-to-Random Noise Ratio
* Upper Adjacent Channel Interference
* Lower Adjacent Channel Interference
* Co-channel Interference
* UHF Taboos [NTSC as Desired only]
* Multipath
* Cable Microreflections

NOTE: Interference Tests shall be conducted with NTSC as Desired and ATV as Undesired, ATV as Desired and NTSC as Undesired, and with ATV as both Desired and Undesired, if appropriate.

Advisory Committee on
Advanced Television (ATV) Service

Doc. No. _____

Date _____

PS/WP-3-0117

THIRD INTERIM REPORT OF THE

SPECTRUM UTILIZATION AND ALTERNATIVES WORKING PARTY

of the

PLANNING SUBCOMMITTEE

of the

ADVISORY COMMITTEE ON ADVANCED TELEVISION SERVICE

January 10, 1990

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY

As previously reported, the Spectrum Utilization and Alternatives Working Party of the Planning Subcommittee of the Advisory Committee on Advanced Television Service has divided its studies and analyses into three broad areas: (1) issues associated with accommodating an Advanced Television (ATV) system within existing VHF and/or UHF television allocations, (2) issues associated with accommodating an ATV system in spectrum above 1 GHz, and (3) possible impacts of an Advanced Television Service (ATS) on the spectrum utilization of various broadcast support services. Also, as previously reported, Working Party 3 has organized itself into various Specialist Groups to carry out its assigned work within these three broad areas.

During this reporting period, Specialist Groups 6, 7, and 9 carried out activities relating to the first area noted above. Earlier spectrum studies conducted by WP-3 only considered cochannel and adjacent channel interference; i.e., taboos were ignored. Those studies demonstrated that providing each existing station with additional spectrum for an augmentation or simulcast channel for transmitting ATV would require a system that is both benign and robust relative to existing NTSC signals. That is, the system must have a low potential for causing interference to existing NTSC reception and should be relatively immune to interference from NTSC stations.

Although these previously reported studies showed that sufficient spectrum might be made available within the present VHF/UHF television broadcast allocations to accommodate all, or