|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | |  | | |
| **STUDY GROUP 12** | | |
| **Original: English** | | |
| **Question(s):** | | 13/12, 14/12 | | Krakow, 27-29 November 2017 | |
| **CONTRIBUTION** | | | | | |
| **Source:** | | Huawei Technologies Co. Ltd. | | | |
| **Title:** | | Subjective test methodology for assessing impact of initial loading buffering delay on user experience (P.QUITS) | | | |
| **Purpose:** | | Proposal | | | |
| **Contact:** | | Bin Wang Huawei China | | | Tel: +86 2138906807 Fax: +xx E-mail: i.wangbin@huawei.com |
| **Contact:** | | Lily Ji Huawei China | | | Tel: +86 2138902979 Fax: +xx E-mail: lily.jili@huawei.com |
| **Contact:** | | Paul Coverdale Huawei Canada | | | Tel: +1 613 820 6643 Fax:  E-mail: coverdale@sympatico.ca |

|  |  |
| --- | --- |
| **Keywords:** | Initial buffering, subjective test, P.QUITS |
| **Abstract:** | This contribution provides input material for P.QUITS |

This contribution proposes initial baseline text for P.QUITS “Subjective test methodolgy for assessing impact of initial loading delay on user experience”. It is based on contribution SG12-C123, submitted to the September 2017 SG12 meeting.

**Subjective Test Plan for assessing user experience of initial loading of streaming video (P.QUIT)**

**Table of Contents**

1. List of Definitions 4

2. Introduction 5

3. Subjective Rating Tests for Lab 5

3.1. Test platform 5

3.1.1. Audiovisual Content 5

3.1.2. Audiovisual Material 6

3.1.3. Material Lengths 6

3.2. Test Method: ACR 6

3.2.1. Quality rating questions and scale 7

3.3. Test Design 7

3.3.1. Training phase 8

3.3.2. Length of Sessions 8

3.3.3. Test Procedure 8

3.4. Subjective Test Environment and Set-up 9

3.4.1. Common Properties 9

3.4.2. PC Tests 10

3.4.2.1. Test Environment 10

3.4.2.2. Display Specification and Set-up 10

3.4.2.3. Input requirements 11

3.4.2.4. Viewing Distance 11

3.4.2.5. Viewing Conditions 12

3.4.2.6. Listening Conditions 12

3.4.3. Mobile Equipment Tests 12

3.4.3.1. Test Environment 12

3.4.3.2. Device Specification and Set-up 13

3.4.3.3. Viewing Distance 13

3.4.3.4. Viewing Conditions 13

3.4.3.5. Listening Conditions 14

3.5. Subjects and Subjective Test Control 14

3.5.1. Instructions for Subjects 15

3.6. Subjective Test Data 15

4. References 16

List of Acronyms

ACR Absolute Category Rating

FHD Full HD (1920x1080)

GOP Group of Pictures

HD High Definition (television)

HRC Hypothetical Reference Circuit

ITU International Telecommunications Union

ITU-R ITU Radiocommunications Standardization Sector

ITU-T ITU Telecommunications Standardization Sector

MOS Mean Opinion Score

PVS Processed Video Sequence

QHD Quad HD (2560x1440)

SRC Source Reference Channel or Circuit

VQEG Video Quality Experts

# List of Definitions

Intended frame rate is defined as the number of video frames per second physically stored for some representation of a video sequence. This may be constant or may change with time.

Source frame rate (SFR) is the intended frame rate of the original source video sequences. The source frame rate is constant.

Frame rate is the number of (progressive) frames displayed per second (fps).

Refresh rate is defined as the rate at which the computer monitor is updated.

Initial buffering is defined as the period of time between the initiation of the video playback and the start of the video playing; during this time a ‘stalling’ indicator is shown.

Stalling is defined as any event where the video pauses for some period of time and then restarts without losing any video information.

# Introduction

This document defines the procedure for running subjective tests to be used for developing a model to estimate the percentage of users giving up watching a video and the perceived quality, due to the initial loading time of streaming video.

There are three main methods of collecting the users’ experience data: laboratory test, crowd souring test in field, and commercial video website/application statistics.

This document focuses on lab test first, and then expands to crowd sourcing test in the field which includes casual environments (such as office, café, living room, etc.).

# Subjective Rating Tests for Lab

## Test platform

The video user’s behavior and perceived quality on video initial loading is strongly related to user's expectation. The environment, mood, preference of the video content will affect user’s experience, even with the same initial loading time. So the test platform should perform similarly as with online OTT video service, to simulate a real video view experience.

The following figure shows the snapshots of YouTube desktop website and mobile application. The test platform includes desktop and mobile platform. The UI of both platforms should have two key components, navigator bar and video main page. Through the navigator bar, subjects can browse different video categories, and the related videos will be shown in the video main page in a list. If subjects select a specific video, the video will start to play with a randomized initial loading time.

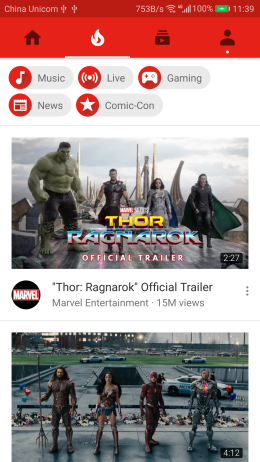
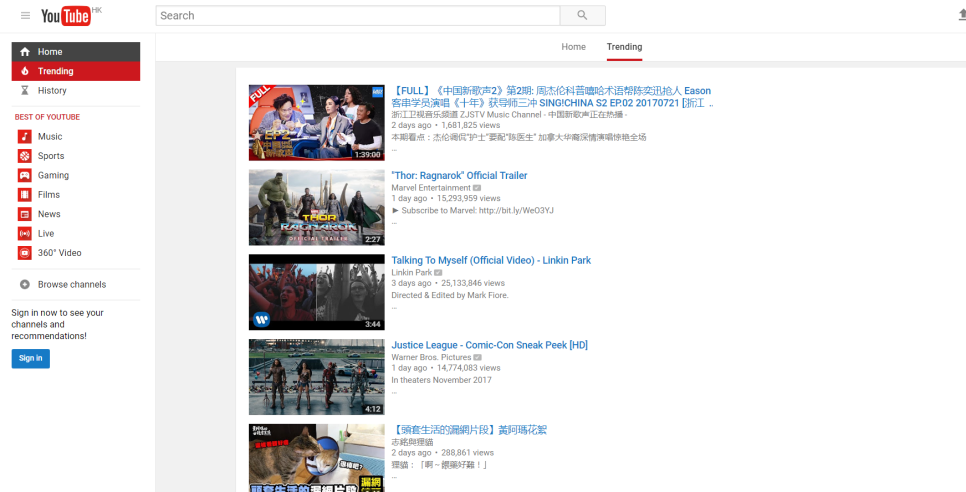


Figure 1 - snapshots of YouTube desktop website and mobile application

### Audiovisual Content

The audiovisual materials will be representative of a range of video clips on major VoD websites (such as YouTube, Netflix, etc.). The list below identifies the types of video material that form the basis for test platform navigator bar.

1. Movies, movie trailers
2. TV Shows
3. Sports
4. Gaming
5. Music video
6. Animation
7. News
8. Documentaries

### Audiovisual Material

Audiovisual material shall provide well-synchronized audio and video content, i.e. lip-synchronicity must be guaranteed for non-dubbed video material and must be good enough for dubbed video material. Audiovisual source material should also provide a natural combination of the presented audio and video signal content. For example, scenarios where we see musicians playing classical instruments while we listen however to the song Thriller from Michael Jackson should be avoided, same as subtitles. Dubbed video material is allowed in countries where users are typically watching dubbed video.

Audiovisual material can be downloaded from the major OTT VoD site. Although the encoding scheme of different OTTs is different, for example the bitrate of typical 1080p video for YouTube is about 3Mbps, but 7Mbps for Netflix. So it is recommended that the test platform use the audiovisual materials downloaded from the same OTT. The recommended bitrate for different video resolution is showed in the following table.

Table 1 - Recommended bitrates for different resolution

|  |  |
| --- | --- |
| Video Resolution | Bitrate Range |
| 360p | 220-450kbps |
| 480p | 375-750kbps |
| 720p | 1050-2100kbps |
| 1080p | 1875 kbps – 10 Mbps |
| 1440p (2K) | 1875 kbps – 15 Mbps |
| 2160p (4K) | 10 Mbps – 30 Mbps |

Audiovisual material will also cover the most popular video resolutions: 360p, 480p, 720p, 1080p, 2K, and 4K. The highest video resolution for desktop test platform is 4k, and 2k for mobile test platform.

### Source Material Lengths

The source material can be divided into short video and long video.

The duration of short videos commonly is less than 5 minutes. For example, movie trailers, highlight of sports & gaming, music video are short videos.

The duration of long videos commonly is more than 20 minutes. For example, movies, TV show, replay of sports, animation, documentaries are long videos.

Both short videos and long videos shall be provided in the test platform.

## Test Method: ACR

The subjective tests will be performed using the Absolute Category Rating method [ITU-T Rec. P.910 (2008)].

In the ACR test method, each video clip is presented once at a time and rated individually according to the time pattern shown in Figure 1. Subjects cannot re-play a sequence. After each presentation, subjects are asked to judge:

* Overall experience
* Audiovisual integral quality
* Loading experience
* Content preference

using the ACR scale.

The test presentation order is selected by subjects (see Section 3.3.3).

Grey

Grey

PVS.i

PVS.j

PVS.k

voting

voting

voting

Figure 2 - Presentation pattern for the ACR method.

### Quality rating questions and scale

The subjects will be asked a series of rating questions:

* Loading: What is your opinion of the loading experience? Here Loading consists of initial buffering delay before begin of play-out and stalling/rebuffering during play-out.
* Content: What is your opinion of the video content? Are you interested in it?

All subjective ratings are reported on the 5-point discrete scale as shown in Figure 2.

No question about the semantic content will be asked in addition to the overall quality question.

|  |  |
| --- | --- |
| 5 | Excellent |
| 4 | Good |
| 3 | Fair |
| 2 | Poor |
| 1 | Bad |

Figure 3 - Rating scale (numerical values will be used in addition to the labels)

Labels “Excellent” to “Bad” are translated to the main language of the country in which the test is conducted.

## Test Design

For the purposes of these subjective tests the following definitions are used to describe the constituent parts of a subjective test. These are shown in Table 2.

* **Session** – the time when the subject is browsing VoD simulation platform, viewing PVSs and voting (with no breaks)
* **Test** – the entire subjective test process, which may consist of multiple sessions.

Table 2 Breakdown of a Subjective Test

|  |
| --- |
|  |
|  | Test | | | |
|  | Instru- ction | Session 1 | Break | Session 2 |
| Notes |  | Includes training PVSs (see Section 3.3.1) |  |  |
| Time (mins) | 15 | <30 | 5-10 | <30 |

### Training phase

The first session of each visit should include a training phase. This phase consists of showing the subject(s) how to use the test platform, which includes getting familiar with the UI of VoD simulation platform, browsing and playing N video contents interested in, rating tool and to adapt to the types of video material, initial loading time, and to the quality range of the tests.

Typically, N= 2 to 4, yielding 10 min training. However, N may depend on the duration of the PVSs used in the tests. In practice, PVSs with duration shorter (e.g. 3 min) than the duration of test PVSs may be used.

### Length of Sessions

For a single test, the time of actively viewing videos or hearing audio and voting will be limited to a maximum of 60 minutes. The total visit time, including instructions, warm-up, and play, will be limited to 1 hour 45 minutes. Each test should be split into 2–3 sessions between which subjects have a short break (5-10 minutes). A session will last no longer than 30 minutes. For tests with 5-min SRCs, a session may be extended to 30-35min.

### Pretest Procedure

A pre-test is needed to verify that the subject behaves normally. It would be possible to, for example test when a subject will actually abort a session with extremely long initial loading delay.

Step 1. Browsing

The subject is allowed to browse the VoD simulation platform; find the video he/she would like to view by filtering different video categories.

Step 2. Selecting

Once the subject finds the video, he/she can click for preparing playback, the resolution used will be selected according to the simulated network connection (see Annex III. Q1 & Q2).

Step 3. Playback

The video will start to play with 3 minutes initial loading time. If the subject will not abort this session, this meaning behaves abnormally. So the subject will not pass the pretest procedure.

### Test Procedure

The test procedure will try to simulate the subjects watching a video-on-demand service as much as possible, as to guide the subjects "we are trialing an online video service, not a boring lab test", so the subjects will generate a similar expectation with OTT video.

Step 1. Browsing

The subject can browse the VoD simulation platform; find the video he/she would like to view by filtering different video categories.

Step 2. Selecting

Once the subject finds the video, he/she can click for preparing playback, the resolution used will be selected according to the simulated network connection (see Annex III. Q1 & Q2). .

Step 3. Generating initial loading time



For each playback, the initial loading time is randomized controlled by application/website, from 0 ~ 20 seconds, and in normal distribution.

Step 4. Playback

The video will start to play with generated initial loading time. During the loading period, if the subject feels the loading time is too long, he/she can abort waiting.

If the subject selects a long video with duration large than 5 minutes, the maximum view duration is limited to 3 minutes with notification "trial time exceed'.

If the subject selects a short video with duration less than 5 minutes, there is no view duration limitation.

Step 5. Voting

Once the playback ends, or the subject aborts the playback, the questionnaire [see Section 3.2.2] pops up for the subject rating the quality.

## Subjective Test Environment and Set-up

[Copied from P.NATS Phase 2 Subjective Test Procedure Section 3.3, with minor modification]

Two types of subjective tests will be carried out, namely using PC (e.g, computer playback) or mobile equipment (e.g. mobile phone).

### Common Properties

The following conditions must be met for both static and mobile tests.

It should be ensured that:

1. playback mechanism is guaranteed to play at frame rate without dropping frames,
2. playback mechanism does not add visible artifacts,
3. playback mechanism does not add visible startup delay

The tests must be conducted indoors – see the following sections for detailed requirements.

### PC Tests

PC tests will be conducted using a computer where test sequences are loaded from a hard disk and presented on a computer monitor.

#### Test Environment

The test room will conform to the following requirements:

* Audiovisual: ITU-T P.911
* Video-only: ITU-T P.910

| Table 4 – Test Environment Parameters | |
| --- | --- |
| Parameter | Setting |
| Peak luminance of the screen | 100-200 cd/m2 |
| Ratio of luminance of inactive screen to peak luminance | ≤0.05 |
| Ratio of the luminance of the screen, when displaying only black level in a completely dark room, to that corresponding to peak white | ≤0.1 |
| Ratio of luminance of background behind picture monitor to peak luminance of picture | ≤0.2 |
| Chromaticity of background | D65 |
| Background room illumination | ≤20 lux |
| Environmental background noise level | ≤35 dBA |
| Maximum observation angle relative to normal (in case several viewers are used with a single display) | 30 |
|  | |

#### Display Specification and Set-up

The subjective experiments will use a full UHD-1 PC monitor; provided that the monitor meets the required specifications (below) and is color calibrated for video.

The video sequences must be displayed on a window of the screen (1:1 pixel ratio without up-sampling) at native resolution. If the screen resolution is slightly higher, the color of the background in the screen should be black.

If upscaling results in pillar-/letterboxing, the total width/height of the black borders should not exceed 20% of the total display width/height. (For example, a 3840x2160 sequence may be displayed on a 4096x2160 monitor.)

Given that the subjective tests will use different display technologies, it is necessary to ensure that each test laboratory selects appropriate display specification and common set-up techniques are employed. Due to the fact that most consumer grade displays employ some kind of display processing that will be difficult to account for in the models, all subjective facilities shall use a full resolution display.

Table 5 – Monitor Specifications

| **Monitor Feature** | **Specification** |
| --- | --- |
| Diagonal Size | 14 - 27 inches |
| Dot pitch | < 0.30 |
| Resolution | Native resolution (no scaling allowed) |
| Gray to Gray Response Time (if specified by manufacturer, otherwise assume response time reported is white-black) | < 30 ms  (<10 ms if based on white-black) |
| Color Temperature | 6500K |
| Calibration | Yes (colors and brightness) |
| Bit Depth | >=8 bits/color |
| Refresh Rate | >= 60 Hz |

The display shall be set-up using the following procedure:

* Use the auto setting to set the default values for luminance, contrast and color shade of white.
* Adjust the brightness according to Rec. ITU-T P.910, but do not adjust the contrast (it might change balance of the color temperature).
* Set the gamma to 2.2.
* Set the color temperature to 6500 K.

Any post-processing done by the monitor (e.g., frame interpolation, motion smoothing, …) must be deactivated.

It must be ensured by each proponent that the playout device in combination with the display can show the processed sequences without further introduction of artifacts (e.g., frame interpolation, frame dropping, stuttering).

#### Input requirements

The display may be connected to a playout device via HDMI, SDI, or any other link that does not change the signal or introduce lossy compression.

#### Viewing Distance

The instructions given to subjects will request subjects to maintain a specified viewing distance from the display device by asking them to maintain their back in contact with the chair at all times during the test. This will minimize the variation of viewing distance during the test.

The viewing distance should be 1.5H or 3H, where H = Picture Height (picture is defined as the display size), as determined by the test matrix.

When the resolution changes as part of the HRC (“adaptive streaming”), then any upscaling will be performed in the processing of the sequences, so that all PVS use the same resolution for the test.

#### Viewing Conditions

Only one subject should be seated in front of the viewing device.

The test room will conform to the requirements specified in Section 3.3.2.1.

It is recommended that subjects be seated facing the center of the video display at the specified viewing distance. Subject's eyes are positioned opposite to the video display's center (i.e. if possible, centered both vertically and horizontally).

#### Listening Conditions

For long-duration tests, audio will be presented using headphones to ensure similar listening conditions between different labs, which is difficult to achieve with loudspeakers.

When listening is carried out with headphones, audio will be played using a diotic presentation (both ears receive the same mono signal) or binaural presentation (each ear receives one channel of a stereo signal). Headphones should be diffuse-field equalized headphones.

Artificial background noise (e.g. Hoth noise) will not be used. The environmental noise of the room should be lower than the value specified in Section 3.3.2.1.

Presentation (listening) level should be 73dB (SPL) at both ears when using headphones.

### Mobile Equipment Tests

Mobile equipment audiovisual tests will be conducted using a mobile phone where test sequences are loaded from either of the following sources:

1. the phone’s internal memory
2. a network resource

The question pertaining to the quality of the stimulus (see 3.1.1) may be presented by one of the following means:

1. on the phone display
2. on a paper-based questionnaire

#### Test Environment

The test room will conform to the following requirements.

| Table 4 – Mobile Test Environment | |
| --- | --- |
| Parameter | Setting |
| Peak luminance of the screen | 100-200 cd/m2 |
| Ratio of luminance of inactive screen to peak luminance | ≤0.05 |
| Ratio of the luminance of the screen, when displaying only black level in a completely dark room, to that corresponding to peak white | ≤0.1 |
| Ratio of luminance of background behind picture monitor to peak luminance of picture | ≤0.2 |
| Chromaticity of background | D65 |
| Background room illumination | ≤20 lux |
| Environmental background noise level | ≤35 dBA |

#### Device Specification and Set-up

The test will be carried out on mobile phones. The devices will have the specifications given in the following table.

Table 5 – Mobile Device Specifications

| **Device Feature** | **Device 1** | **Device 2** |
| --- | --- | --- |
| Model | Samsung Galaxy S7 (no edge) | Google Nexus 6p |
| Diagonal display size | 5.1” | 5.7” |
| Display resolution | 2560 x 1440 | 2560 x 1440 |
| Display type | SuperAMOLED | AMOLED |

The sequences may be loaded from a network source during the test rather than from the device’s internal memory. It must be ensured (e.g. by using a WiFi connection) that this loading process does not introduce a noticeable startup delay or unwanted stalling effects for any given sequence.

The phone should be charged during the session to avoid depleting the battery.

#### Viewing Distance

The viewing distance should be at most 8H where H = display height (in landscape mode), according to subjects’ preference.

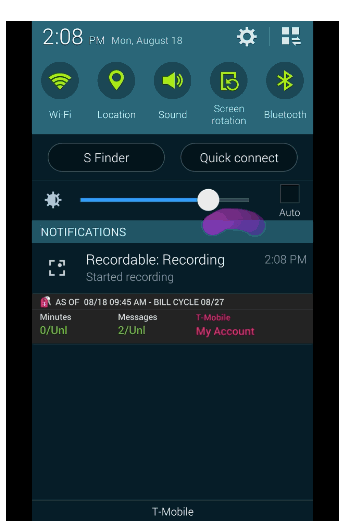
#### Viewing Conditions

Only one subject should be seated in front of the viewing device.

To ensure that the viewing distance is met, the device should be mounted on a desk or wall. Alternatively, the phone may be handheld by subjects, but this is discouraged. In the latter case, the phone should be held above a desk on which the region for the allowed viewing distance is marked.

Subjects will be seated facing the device at a horizontal angle between 0 and 45°. If the angle is higher, the viewing distance should be lower (5H) to accommodate for the smaller resting point of vergence.

The device display must be set to a brightness level that fulfills the requirements from Table 4. Automatic brightness correction must be disabled by pulling down the notifications drawer from the top of the screen, and unchecking the “Auto” box, as shown in Figure 3.



**Figure 3 – Disabling auto brightness mode**

#### Listening Conditions

See 3.3.2.3.

## Subjects and Subjective Test Control

Each (audio or video or audiovisual) test will require at least 50 “valid” subjects (see section 3.4.2 for details). It is recommended to have a 50-50 split or near 50-50 split between female and male subjects. If the parity between male and female participants cannot be achieved, then a maximum of 60-40 split is permitted.

Only non-expert subjects will participate. The term non-expert is used in the sense that the subjects’ work does not involve audio or picture quality and they are not experienced assessors. They must not have participated in a subjective quality test over a period of six months.

Prior to participation in a video test, subjects must be screened for the following:

* normal (20/30) visual acuity with or without corrective glasses (per Snellen test or equivalent), both for mobile and PC/TV tests using near or far tests, respectively
* normal color vision (per Ishihara test or equivalent).
* familiarity with the language sufficient to comprehend instruction and to provide valid responses using the semantic judgment terms expressed in that language.

For audiovisual tests, it is expected that some content will include partly voice-over parts or voice (e.g. commentary). Therefore, subjects who understand the language used in the audiovisual material should be considered for the audiovisual tests.

Subjects should report normal hearing in order to participate in audio and audiovisual tests.

### Instructions for Subjects

For many labs, obtaining a reasonably representative sample of subjects is difficult. Therefore, obtaining and retaining a valid data set from each subject is important. The following procedures are highly recommended to ensure valid subjective data:

Write out a set of instructions that the experimenter will read to each test subject. The instructions should clearly explain why the test is being run, what the subject will see/hear, and what the subject should do. Pre-test the instructions with non-experts to make sure they are clear; revise as necessary.

Explain that it is important for subjects to pay attention to the sequences shown on each trial.

There are no “correct” ratings. The instructions should not suggest that there is a correct rating or provide any feedback as to the “correctness” of any response. The instructions should emphasize that the test is being conducted to learn subjects’ judgments of the quality of the samples, and that it is the subject’s opinion that determines the appropriate rating.

Paying subjects helps keep them motivated.

Subjects should be instructed to watch/listen carefully to the entire sequence before voting. The screen should say when to vote (e.g., “vote now”).

Detailed example instructions to subjects are provided in Annex I.

# References

[[AVHD-PNATS2 ToR](https://www.itu.int/ifa/c/irg/avqa/pnats2avhd/documents/ToR/)] AVHD-PNATS2 Term of Reference, TD 105 (GEN/12), January 2017

[ITU-T Rec. P.910 (2008)] “Subjective video quality assessment methods for multimedia applications”

Annex I  
EXAMPLE INSTRUCTIONS TO THE SUBJECTS

[Copied from P.NATS Phase 2 Subjective Test Procedure Annex II, with minor modification]

This is a template for subject instructions. The instructions must be handed to the subjects in written form*.* They must be translated to the local language of the lab in which the tests are being conducted.

**Introduction:** Thanks for coming in today to participate in our study. The study is about the quality of *video-on-demand service*; it is being sponsored and conducted by companies that are developing testing new technologies to enhance consumers' online video experience, including video quality and smooth viewing. These companies are interested in what looks good to you, the potential user of next-generation devices.

**Vision Tests:** Before we get started, we would like to check your vision in two tests, one for acuity and one for color vision. (*These tests will probably differ for the different labs, so one common set of instructions is not possible. Insert instructions here.*)

**Overall Goal:** We are going to ask you to browse the video-on-demand platform, select the videos you are interested in, and enjoy watching. Then judge each of them for “quality” – we will explain more below about what we mean by “quality.” It takes time to load the video, sometimes there are loading problems. So you can abort/go back/reload the video. We will ask you to rate the video quality and smooth playback of each one after you have seen it.

**Setup:**

* When we get started with the study, please sit at *(location)*. The video-on-demand service will be displayed on the *(screen/phone)*.
* Please keep your head reasonably close to the position indicated by the mark. This is because the videos might look a little different from different positions, and we would like everyone to judge the videos from about the same position.
* (*Optional for mobile tests:*) Please do not press any of the physical buttons on the phone. This may turn off the phone display or exit the test, making your ratings invalid.

**Process:** Each video will be (*insert number, maybe between X and Y if needed*) seconds *(minutes)* long. You will need to select the video first, and press the *(insert details)* button to start the audio-video playback and once the audio-video has finished you should then vote by (*insert method details, e.g. using screen*). This process will be repeated until you have seen and voted on (*insert number*) sequences, then we’ll have a break.  Then there will be another similar session.

**Task:** Your task is to judge the *overall* *quality* of each sequence – not the content.Parts of the playback where the video is not playing and a waiting indicator is shown are part of the test.  Any video quality loss is also part of the test. You should consider both of these things as part of your overall quality judgment.

Due to limited duration of the video, some videos may end abruptly in a middle of scene. Please do not consider this abrupt ending in your judgment.

There is no wrong answer in this task; just rely on your own judgment.

**Rating:** After judgin*g* the *(overall/video)* quality of a sequence, please rate the quality of the sequence. Here is the rating scale we would like you to use: *(insert picture of scale)* Please indicate your rating by adjusting the cursor on the scale accordingly.

**Practice:** At the start of the test the first sequences you see will be practice sequences so you can get a feel for the setup and how to make your ratings. After several of these practice sequences you will then be told when the test starts properly.

**Questions:** If you have questions, please ask the experiment leader.

**Subject Consent:** *(example, may be different for each lab)*The *(name of experiment)* Experiment is being conducted at the (*name of your lab*) lab. The purpose, procedure, and risks of participating in the *(name of experiment)* Experiment have been explained to me. I voluntarily agree to participate in this experiment. I understand that I may ask questions, and that I have the right to withdraw from the experiment at any time. I also understand that (*name of lab*) lab may exclude me from the experiment at any time. I understand that any data I contribute to this experiment will not be identified with me personally, but will only be reported as a statistical average.

Signature of participant Signature of experimenter

Name of participant Date Name of experimenter

Annex II  
MOBILE SUBJECTIVE TEST LAB SETUP

[Copied from P.NATS Phase 2 Subjective Test Procedure Annex III]

The following points should be considered when running a subjective test using mobile phones.

1. To minimize dirt and greasy marks on the screen the screen should be wiped before each subject starts their test and during each break (mobile tests should have the same breaks as PC based tests).
2. The mobile phone should be mounted to help maintain the distance to the viewer and the angle of the screen. This will also help avoid fatigue. Cycle mounts are useful for this and can be mounted to an adjustable boom mic stand (see photos in Figure 4). Car mounts can be fixed on the table as well (see photos in Figure 5).NOTE: An audio extension lead should be used to avoid a heavy headphone connection directly into the phone..

**Figure 4 – Netscout Mobile Test Setup**

****

**Figure 5 – T-Labs Mobile Test Setup**

** **

**Figure 6 – OPTICOM Mobile Test Setup**

1. Settings that are recommended:
   1. Sound:
      1. Music effects control: SoundAlive (default)  
          [other options: MusicFX?]
   2. Display
      1. Brightness: Max (disable auto)
      2. Smart stay : disabled
      3. Screen timeout: 10 mins or maximum
      4. Screen mode: Standard  
         [other options: Adapt Display (default), Dynamic, Standard, Professional photo, Cinema]
      5. Auto adjust screen tone : Disabled  
         [is enabled by default]
   3. Power Saving
      1. Power saving mode: Off
      2. Ultra power saving mode: Off

Annex III  
demographics for Laboratory and Crowd sourcing Tests

In order to understand the subjects' background on OTT VoD experience, the following question will be asked after the test.

Q1. How fast is your internet connection at home? (Single select)

|  |  |
| --- | --- |
| Less than 2Mbps | 1 |
| 2~10Mbps | 2 |
| 10 ~ 50Mbps | 3 |
| 50 ~ 100Mbps | 4 |
| More than 100Mbps | 5 |
| Not sure | 88 |
| Do not have internet connection | 99 |

Q2. What generation of mobile network are you currently using on your smartphone? (Single select)

|  |  |
| --- | --- |
| 2G | 1 |
| 3G | 2 |
| 4G-LTE | 3 |
| Not sure | Close |
| Do not use smartphone |  |

Q3. What devices do you typically used for streaming video at home / on the go (Multi select)

|  |  |  |
| --- | --- | --- |
|  | Television | 1 |
|  | Tablet | 2 |
|  | Smartphone | 3 |
|  | Laptop | 4 |
|  | Personal Computer | 5 |
|  | Other, please specify\_\_\_\_\_ | 6 |

Q4a. In the past 3 months, how often did you watch online videos via your phone? (Single select)

|  |  |  |
| --- | --- | --- |
|  | Less than once per day | 1 |
|  | Once per day | 2 |
|  | 2-3 times per day | 3 |
|  | 4-5 times per day | 4 |
|  | More than 5 times per day | 5 |

Q4b. In the past 3 months, on average how much time did you spend each session on online videos on your phone? (Single select)

|  |  |  |
| --- | --- | --- |
|  | Less than 5 minutes | 1 |
|  | 5-9 minutes | 2 |
|  | 10-19 minutes | 3 |
|  | 20-29 minutes | 4 |
|  | 30-59 minutes | 5 |
|  | 1-2 hours | 6 |
|  | More than 2 hours | 7 |

Q5. What is your favourite video? (Multi select)

|  |  |
| --- | --- |
| Movies, Movie trailers | 1 |
| TV Shows | 2 |
| Sports | 3 |
| Gaming | 4 |
| Music video | 5 |
| Animation | 6 |
| News | 7 |
| Documentaries | 8 |
| Live | 9 |
| Other, please specify\_\_\_\_\_ | 99 |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_