**Source: Qualcomm Incorporated, Dolby Laboratories, Inc.**

**Title: Instruction Set for P.SUPPL800 Tests**

**Document for: Discussion and Agreement**

**Agenda Item: 7.5**

# **1. Introduction**

The completion of IVAS-8a Selection Test Plan is scheduled in Audio SWG for this meeting.

In this contribution, we address an open item – text in brackets included in IVAS-8a Section 4.4 on Opinion Scales and in Annex A on the instruction set for P.SUPPL800 DCR experiments. We propose an alternative solution that, in our view, could be considered as the instruction set to be used in these experiments.

While the original application of P.800 DCR test in NB mono speech aims at testing coding artefacts hence any deviation from the original signal can be classified as a quality degradation, DCR testing of spatial audio coding focuses on signal quality and spatial image preservation at the same time. The coding may result in an impairment compared to the case when the coding is not present. An impairment may be a quality degradation or a difference (alteration) which does not necessarily have the character of a degradation. While signal quality degradations resulting from coding artefacts may be described as annoying, naïve listeners may not sense spatial localization inaccuracies or a reduction of spatial image as annoying degradation. In this case, forcing them to classify the degree of annoyance seems inappropriate. From this point of view, we may find that the classical instructions do not address the entire purpose of the tests properly.

# **2. Proposal**

Below instruction set is related to IVAS-8a, Annex A. A first basic instruction set is provided that is suitable e.g. for a paper printout for the case for labs when the listener seats are not equipped with individual computer monitors and several subjects listen simultaneously in the same room.

Also provided is an extended implementation example that illustrates all steps including listener familiarization as discussed in ITU-T P.SUPPL800. This implementation is applicable when labs offer the possibility to guide the subjects through the test individually using test seats equipped with computer monitors.

Basic instruction set:

|  |
| --- |
| **INSTRUCTIONS TO NAÏVE LISTENERS FOR P.SUPPL800 DCR TEST**  In this experiment you will be evaluating systems that might be used for future immersive telecommunication services using spatial audio. Spatial audio means that you can locate various sound sources around yourself. For example, a first talker may appear to talk from the left-hand side and a second talker from the right-hand side, a talker can be moving, etc.  In each trial, you will hear a *reference* audio sample followed by a *test* sample. The *test* sample has the same content as the *reference* sample, but it is possibly impaired after it has passed through a telecommunication system.  Your task is to evaluate the overall impairment of the second sample compared to the first sample, comprising both degradations in the sound quality (e.g., due to additional noise, roughness, clicks or other distortions), and/or differences in the spatial representation (e.g., sound source location, distance, spatial width, movement, etc.).  You should listen carefully to both samples within a trial. When they have finished, select the category that best describes your overall impression about the amount of any impairment you can perceive in the second sample relative to the first sample:  5 - No impairment  4 - Small impairment  3 - Moderate impairment  2 - Large impairment  1 - Very large impairment  Note that the level of impairments present in different *test* samples is expected to span the complete range of the rating scale during the experiment.  Please do not discuss your opinions with other listeners participating in the experiment. If you have any questions, please ask the test administrator. |

Extended implementation example:

Below is an implementation example of how the screens shown to a P.SUPPL800 listening test subject may look like. It comprises three parts:

* General introduction to the task of audio quality evaluation
* Familiarization with the test procedure
* Voting screen (after playback of sample pair)

Even if monitor-based guidance of the test subjects is assumed in this example, it is also advisable to use paper printouts of the screens during the initial instructions of the listeners by the test administrator.

Note that these screens are intended to give guidance to the listening labs. Adaptations corresponding to the properties of the test administration tool are envisioned while keeping the general concept.

Screen 1: General introduction to the task of audio quality evaluation

|  |
| --- |
| **Evaluation of the quality of future spatial audio telephony and conferencing systems**  In this experiment you will be evaluating systems that might be used for future immersive telecommunication services using spatial audio. Spatial audio means that you can locate various sound sources around yourself. For example, a first talker may appear to talk from the left-hand side and a second talker from the right-hand side, a talker can be moving, etc.  In each trial, you will hear a REFERENCE audio sample followed by a TEST sample. The TEST sample has the same content as the REFERENCE sample, but it is possibly impaired after it has passed through a telecommunication system.  Your task is to evaluate the OVERALL impairment of the second sample compared to the first sample, comprising both degradations in the sound quality (e.g., due to additional noise, roughness, clicks or other distortions), and/or differences in the spatial representation (e.g., sound source location, distance, spatial width, movement, etc.).  To familiarize yourself with the kind of spatial audio samples and the OVERALL impairment range, you will now hear a few sample pairs where the first sample is the REFERENCE sample and the second is the possibly impaired TEST sample. The TEST samples are expected to represent the complete range of OVERALL impairments caused by the different systems. |

Screen 2: Familiarization with the test procedure

|  |
| --- |
| **Familiarization with the test procedure**  The next part is to get you familiar with the test procedure in a short practice session. You will hear the sample pairs as before. However, now you will also be asked to register a score of the perceived OVERALL impairment. Again, you will hear the REFERENCE sample, then a short silence, then the impaired TEST sample.  Your task is to evaluate the OVERALL impairment of the second sample compared to the first sample, comprising both degradations in the sound quality (e.g., due to additional noise, roughness, clicks or other distortions), and/or differences in the spatial representation (e.g., sound source location, distance, spatial width, movement, etc.).  You should listen carefully to both samples within a trial. When they have finished, select the category that best describes your overall impression about the amount of any impairment you can perceive in the second sample relative to the first sample:  5 - No impairment  4 - Small impairment  3 - Moderate impairment  2 - Large impairment  1 - Very large impairment  Note that the level of impairments present in different *test* samples is expected to span the complete range of the rating scale during the experiment.  You will have 5 seconds to record your answer by pushing the button corresponding to your choice. There will be a short pause before the presentation of the next sample pair.  We will now begin with the short practice session. The actual tests will then take place in multiple test sessions, with short breaks in between each session. |

Screen 3: Voting screen (after playback of sample pair)

|  |
| --- |
| Please rate the OVERALL IMPAIRMENT of the second sample compared to the first sample:  5 - No impairment  4 - Small impairment  3 - Moderate impairment  2 - Large impairment  1 - Very large impairment |

**References**

[1] ITU-T Recommendation P.800, “Methods for subjective determination of transmission quality.”, 1996

[2] ITU-T Recommendation P.SUPPL800, “ITU-T Rec. P.800 use case examples”, 01/2023

[3] ITU-T Recommendation P.811, “Subjective test methodology for evaluating Speech oriented stereo communication systems over headphones”, 01/2019

[4] 3GPP SA4 IVAS Permanent Document IVAS-8a: Test Plan for Selection Phase, v.0.8.3