ETSI TR 126 997 V19.0.0 (2026-02)

TECHNICAL REPORT

LTE;
5G;
Codec for Immersive Voice and Audio Services (IVAS);
Performance characterization
(3GPP TR 26.997 version 19.0.0 Release 19)

Reference
RTR/TSGS-0426997vj00

Keywords
5G,LTE

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI IPR online database.

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found at 3GPP to ETSI numbering cross-referencing.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, modal verbs have the following meanings:

shall indicates a mandatory requirement to do something

shall not indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

should indicates a recommendation to do something

should not indicates a recommendation not to do something

may indicates permission to do something

need not indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

can indicates that something is possible

cannot indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

will indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

will not indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

might indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

might not indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

is (or any other verb in the indicative mood) indicates a statement of fact

is not (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

1 Scope

The present document provides information on the Immersive Voice and Audio Services (IVAS) codec Selection Phase which was run using the floating point (3GPP TS 26.258). Experimental test results from the subjective quality testing are reported to illustrate the behaviour of the IVAS codec. Additional information is provided on implementation complexity of the IVAS codec and objective test results.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document.*

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 26.250: "Codec for Immersive Voice and Audio Services (IVAS); General overview".

[3] 3GPP TS 26.251: "Codec for Immersive Voice and Audio Services (IVAS); C code (fixed-point)".

[4] 3GPP TS 26.252: "Codec for Immersive Voice and Audio Services (IVAS); Test Sequences".

[5] 3GPP TS 26.253: "Codec for Immersive Voice and Audio Services (IVAS); Detailed Algorithmic Description including RTP payload format and SDP parameter definitions".

[6] 3GPP TS 26 254: "Codec for Immersive Voice and Audio Services (IVAS); Rendering".

[7] 3GPP TS 26.255: "Codec for Immersive Voice and Audio Services (IVAS); Error concealment of lost packets".

[8] 3GPP TS 26.256: "Codec for Immersive Voice and Audio Services (IVAS); Jitter buffer management".

[9] 3GPP TS 26.258: "Codec for Immersive Voice and Audio Services (IVAS); C code (floating-point)".

[10] Recommendation ITU-T P.800: "Methods for subjective determination of transmission quality".

[11] Supplement ITU-T P.Suppl29: "ITU-T P.800 – Use Cases".

[12] Recommendation ITU-R BS.1534-3: "Method of the subjective assessment of intermediate quality level of audio systems".

[13] ITU-T Handbook of subjective testing practical procedures, 2011.

[14] Audio File Format Specifications: WAVE, https://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/WAVE.html.

[15] AFsp Package https://www-mmsp.ece.mcgill.ca/Documents/Downloads/AFsp/.

[16] IEEE Recommended Practice for Speech Quality Measurements, in IEEE Transactions on Audio and Electroacoustics, vol. 17, no. 3, pp. 225-246, September 1969, doi: 10.1109/TAU.1969.1162058.a.

[17] Recommendation ITU-R BS.2051-3 (05/2022): Advanced sound system for programme production.[18] 3GPP TR 26.952: "Codec for Enhanced Voice Services (EVS); Performance characterization".

[19] 3GPP TS 26.445: "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description".

[20] 3GPP TR 26.996: "Immersive Audio for Split Rendering Scenarios; Performance characterization".

[21] Recommendation ITU-T P.191 (03/2023): Software tools for speech and audio coding standardization.

[22] Recommendation ITU-R BS.1770-4 (10/2015): Algorithms to measure audio programme loudness and true-peak audio level.

[23] Recommendation ITU-T P.1401, "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models," January 2020.

3 Definitions of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the terms given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

CL	Cross-check Laboratory
CuT	Codec under Test
DCR	Degradation Category Rating
DTX	Discontinuous transmission
ESDRU	Energy-based Spatial Distortion Reference Unit
EVS	Enhanced Voice Services
FB	Fullband
FEC	Frame Erasure Concealment
FOA	First-Order Ambisonics
GAL	Global Analysis Laboratory
HL	Host Laboratory
HOA2	Higher-Order Ambisonics, 2nd order
HOA3	Higher-Order Ambisonics, 3rd order
HRTF	Head Related Transfer Function
ISM	Independent Stream with Metadata
IVAS	Immersive Voice and Audio Services
ISAR	Immersive Audio for Split Rendering Scenarios
JBM	Jitter Buffer Management
LFE	Low Frequency Enhancement
LL	Listening Laboratory
MASA	Metadata-Assisted Spatial Audio

MC	Multi-channel
MNRU	Modulated Noise Reference Unit
NB	Narrowband
OBA	Object Based Audio
PLC	Packet Loss Concealment
SBA	Scene Based Audio
SDRU	Spatial Distortion Reference Unit
SNR	Signal-to-Noise Ratio
SWB	Super Wideband
TC	Transport Channels
ToR	Terms of Reference
WB	Wideband
WMOPS	Weighted Millions of Operations Per Second

4 General

4.1 Overview of the IVAS Codec Work Item

This clause provides an overview of the objectives before the actual work started, as a historical background. The standardized IVAS codec fulfilled all project objectives.

NOTE: The subsequent text in italics cites justification and objective from the IVAS Work Item Description. Since it is a direct citation, the tense of this text might generally refer to the future.

The introduction of 4G/5G high-speed wireless access to telecommunications networks, combined with the availability of increasingly powerful hardware platforms, will enable advanced communications and multimedia services to be deployed more quickly and easily than ever before. Immersive services, applications and devices, as envisioned in TR 22.891 are expected to provide an immersive user experience which, when compared to existing media services, will deliver a quantum leap in the quality of experience. Immersiveness will be a core service attribute of eXtended Reality (XR) ranging from VR over MR to AR, as described in TRs 26.918, 26.928 and 26.998. An immersive audio-visual experience implies, for the audio component, that a spatial sound impression is convincingly consistent with the presented visual scene and/or the real-world scene in the AR/MR case. In addition, the user should be able to move, within certain limits defined by the application, throughout the scene, and the audio component will adjust to reflect the user's spatial orientation/position.

This WID will develop solutions specifically for conversational and non-conversational use cases where immersive content originates and is consumed in end-user devices.

The 3GPP Enhanced Voice Services (EVS) codec has delivered a highly significant improvement in user experience with the introduction of super-wideband (SWB) and full-band (FB) speech and audio coding, together with improved packet loss resiliency. For a truly immersive experience though, extended audio bandwidth is just one of the dimensions required and support beyond the mono and multi-mono currently offered by EVS is ideally required to immerse the user in a convincing virtual world in a resource-efficient manner. In addition, the currently specified stereo codecs in 3GPP, e.g., Enhanced aacPlus (eAAC+) and AMR-WB+ provide suitable quality and compression for stereo content in an adequate bit rate range, but lack the conversational features (e.g. sufficiently low latency) needed for conversational voice and teleconferencing. These coders also lack immersive audio support that is necessary for immersive services, including e.g., live streaming, XR and immersive teleconferencing.

The purpose of this work item is therefore to fill this technology gap and to address the increasing demand for rich and immersive multimedia services. In addition, teleconferencing applications over 4G/5G will benefit from this next generation codec used as an improved conversational coder supporting multi-stream coding (e.g., channel, object and scene-based audio). Use cases for this next generation codec include, but are not limited to, conversational voice, multi-stream teleconferencing, VR conversational and user generated live and non-live content streaming, AR/MR. The approach proposed is to build upon the EVS codec with the goal of developing a single codec with attractive features and performance (e.g. excellent audio quality, low delay, spatial audio coding support, appropriate range of bit rates, high-quality error resiliency, practical implementation complexity). In the scope of 3GPP the predominant audio rendering instrument is envisaged to be headphones but configurations with e.g. tablet speaker playback may also be of relevance.

The overall objective of this work item is to develop a single general-purpose audio codec for immersive 4G and 5G services and applications including the XR use cases envisioned in TRs 26.918 and 26.928 and possibly relying on devices described in 26.998. The following objectives should be achieved with the work item:

- *The solution is expected to meet the terms of reference (design constraints, performance requirements) developed as part of this WI.*
- *The solution is expected to handle encoding/decoding/rendering of speech, music and generic sound.*
  - *It is expected to support encoding of channel-based audio (e.g. mono, stereo or 5.1) and scene-based audio (e.g., higher-order ambisonics) inputs including spatial information about the sound field and sound sources. The solution is expected to provide support for diegetic and non-diegetic input.*
  - *It is expected to provide a decoder for the encoded format and a renderer with sufficiently low motion to sound latency.*
- *The solution is expected to operate with low latency to enable conversational services over 4G/5G.*
- *The solution is expected to support high error robustness under various transmission conditions from clean channels to channels with packet loss and delay jitter and to be optimized for 4G/5G.*
- *The solution is expected to provide support for a range of service capabilities, e.g., from mono to stereo to fully immersive audio encoding/decoding/rendering.*
- *The solution is expected to be implementable on a wide range of UEs and other end-user devices to address various needs in terms of balancing user experience and implementation complexity / cost.*
- *The solution is expected to provide support for immersive Real-time Communication services including MTSI and potentially streaming services offered by the 5G system through the definition of a new immersive audio media component. Support for MTSI services is also accomplished by the provision of bit-exact EVS operation as part of the solution.*

The developments under this work item should lead to a set of new specifications defining among others textual description of the IVAS codec. Following 3GPP practice, fixed-point and floating-point C code and associated test vectors should also be part of this set of specifications. RTP payload format, SDP parameter definitions, jitter buffer management, rendering and packet loss concealment should be specified as part of the set of the codec specifications.

The codec for Immersive Voice and Audio Services is part of a framework comprising besides encoder and decoder, renderer and a number of auxiliary functions associated with the support of stereo and immersive audio formats. The IVAS codec is an extension of the 3GPP Enhanced Voice Services (EVS) codec; it provides full and bit exact EVS codec functionality for mono speech/audio signal input.

On top of that the IVAS codec is optimized for encoding and decoding of stereo and immersive audio formats, using tools such as Single Channel Element (SCE) coding, Channel Pair Element (CPE) coding and multi-channel coding by means of the Multi-channel Coding Tool (MCT). The stereo modes comprise a hybrid time-domain/DFT-domain/MDCT-domain coding scheme including inter channel alignment (ICA). Immersive audio formats comprise multi-channel audio (5.1, 5.1.2, 5.1.4, 7.1, 7.1.4 setups), scene-based audio (Ambisonics up to order 3), metadata-assisted spatial audio (MASA), and object-based audio (Independent Stream with Metadata (ISM) up to 4 ISMs). In addition, the following combined immersive audio formats are supported: object-based audio with scene-based audio (OSBA, up to 4 ISMs with Ambisonics) and object-based audio with metadata-assisted spatial audio (OMASA, up to 4 ISMs with MASA).

The codec features VAD/DTX/CNG for rate efficient stereo and immersive conversational voice transmissions, an error concealment mechanism to combat the effects of transmission errors and lost packets. Jitter buffer management is also provided.

The IVAS codec operates on 20 ms audio frames. It is capable of switching its bit rate upon command instantly at (active) frame boundaries.

In addition, split rendering functionality was added by the 3GPP work item on “Immersive Audio for Split Rendering Scenarios” (ISAR) [33].

NOTE: The subsequent text in italics cites justification and objective from the ISAR Work Item Description [33]. Since it is a direct citation, the tense of this text might generally refer to the future.

Work currently carried out under the MeCAR and related work items assumes a common XR Baseline Client architecture. An essential characteristic is that a functional split is envisioned between a Presentation Engine comprising a set of composite renderers that are controlled by a Scene Manager and an XR Runtime performing a set of functions that interface with a platform to perform commonly required operations, e.g. post-rendering, prior to final output. The relevant interface between Presentation engine and end device may be a 5G physical interface between, e.g., between a smartphone or 5G EDGE and a lightweight device (AR glasses) like those considered in 5G EDGe-Dependent AR (EDGAR) and 5G Wireless Tethered AR UEs as described in TR 26.998 or those considered under FS_SmarTAR.

The functional split assumed in split renderer architectures is a result of stringent implementation and operational requirements applicable for rendering of XR media on XR devices. For head-tracked immersive audio, the need to rely on a split renderer architecture, may depend on various factors among which the round-trip latency between the renderer in the presentation engine and the lightweight device is a decisive parameter. There are scenarios where this latency may be substantial which may prefer a split rendering approach with pose correction in the end device for binaural audio in a similar way as for video unless decoding and head-tracked binaural audio rendering on the lightweight device does not exceed its strict complexity constraints. In other scenarios, that latency may be sufficiently low, in which case the head-tracked binaural rendering can exclusively be done in the presentation engine. It is notable that the transmission over the interface may generally be bit rate constrained and dependent on the specific physical interface.

Binaural audio rendering comprises of signal processing functionalities that may include:

- *Binauralization of audio input based on head rotation (3DoF),*
- *Binauralization of audio input based on listener position and head rotation (6DoF),*
- *Room acoustics synthesis.*

Audio input to be rendered may be a combination of diegetic immersive (3D audio) and non-diegetic sounds. The diegetic immersive sounds need to be binauralized using the up-to-date head rotation data. The head rotation data is typically originating from the head-tracker available from the lightweight end device. The room acoustic synthesis can be performed using room impulse response data or parametric representation thereof, typically supplied to the Presentation Engine.

Depending on constraints and design preferences of the lightweight device (AR glasses, earbuds, etc.) and the properties of the interface between Presentation Engine and end device, solutions are needed that among more are compliant with TRs 26.928 and 26.998. The solutions shall address given interface characteristics and not impose any new requirements for them.

Another aspect is the currently ongoing standardization of the EVS Codec Extension for Immersive Voice and Audio Services (IVAS) codec. While low complex rendering for lightweight devices is not a specific design objective, the IVAS codec work item should ideally provide solutions that would enable using IVAS services over head-tracked lightweight clients meeting relevant requirements.

Bearing in mind the evolution of the AR/XR technologies, it would be desirable to design low complex solutions for head-tracked binaural audio rendering on lightweight devices that under certain limitations are agnostic in a sense that the pre-renderer component in the presentation engine could be connected with any immersive binaural audio framework through suitable APIs.

The solutions to be specified are intended to add to the number of rendering options to enable immersive audio services on a broad range of devices, including light-weight AR glasses or earbuds. The pre-rendering part of the solutions is expected to become non-mandatory but shall fulfill the relevant requirements set out under this work item. It should interface through a fully specified intermediate bitstream with a fully specified split rendering decoder. For end device implementations claiming support of a specific solution, a fully compliant implementation of at least the split rendering decoder shall be required. Other end device implementations not claiming support of a specific solution remain at the discretion of the implementor.

The overall objective of this work item is to develop solutions for immersive binaural audio on head-tracked devices that are compatible with the envisaged split architectures (MeCAR, 26.998). The solutions should consider low-complex and lightweight devices and demonstrate operational benefits over solutions with full decoding and rendering in the end device. The following objectives should be achieved with the work item:

- *Provide format specification for intermediate representation(s).*

- *Provide functional requirements for (pre-)renderer operations to be carried out by Presentation Engine.*
- *Define suitable APIs.*

- *Provide encoder, bitstream and decoder specification for intermediate representations including audio with and without post-rendering control metadata.*
- *Provide a specification for decoded intermediate representations to provide binaural audio output with and without head-tracker input and post-rendering control metadata.*
- *Consider potential solutions offered by the IVAS work item, and specify the necessary interfaces.*

The work item shall in a first phase identify and agree relevant requirements to be documented in a TR. This shall cover:

- *Design constraints related to complexity and memory as well as constraints related to relevant interfaces between presentation engine and end device such as bit rate, latency, down- and upstream traffic characteristics.*
- *Design constraints related to functional capability requirements such as rendering of non-diegetic sounds, 3DoF rendering of diegetic immersive sounds, 6DoF rendering of diegetic immersive sounds, including simultaneous rendering of different sound categories, and room acoustics synthesis.*
- *Performance requirements.*

The solution(s) are characterized for the range of relevant interface characteristics between presentation engine and lightweight device. The case where the immersive audio is decoded and rendered within the end device should be considered as a reference.

The requirements will be documented in a first technical report. The developments under this work item shall lead to a new specification defining among others textual descriptions of the involved renderers and codec (incl. frame loss concealment) of the intermediate representation(s). The performance of the developed solutions in relation to the requirements will be documented in a second technical report. Solutions meeting the ISAR split rendering requirement may be added to the set of IVAS codec specifications (by means of CRs) if they are found suitable for IVAS. The developed solutions should also be referenced in the MeCAR specification.

Specific split rendering solutions for IVAS should comprise a non-mandatory default split rendering encoder for the specified internal and stand-alone IVAS renderers. In addition, for a given specific solution there should be specified interfaces offering the possibility either to connect a given (proprietary) renderer for IVAS to the intermediate representation encoder or to use proprietary pre-renderers/intermediate encoders to produce compliant intermediate bitstreams. Such proprietary solutions shall be compliant with the relevant requirements documented in the first technical report. ISAR end device implementations for IVAS claiming support of a specific solution shall be required to have at least a fully compliant split rendering decoder. Other decoder/post-renderer implementations not claiming support of a specific ISAR solution for IVAS remain at the discretion of the implementor.

A special feature of the IVAS renderer is that it supports split operation with pre-rendering and transcoding to a head-trackable intermediate representation that can be transmitted to a post-rendering end-device. This enables moving a large part of the processing load and memory requirements for IVAS decoding and rendering to a (more) capable node/UE while offloading the final rendering end-device. For the split rendering bitstreams, bit-rates ranging from 256 kbps to 768 kbps are supported.

4.2 Presentation of the Following clauses

Clause 5 outlines the Terms of reference for the IVAS project. In clause 6, the selection process in 3GPP is presented. An overview of selection [and characterization tests] can be found in clause 7. The subjective tests provide statistical data which are subject to variations; important notes about interpretation of results are described in clause 8.

The actual subjective test results are presented in clause 9, and objective evaluations are presented in clause 10.

5 Terms of Reference

3GPP sets the codec Terms of Reference as Design Constraints and Performance Requirements.

The design constraints specified in the IVAS-4 Permanent Document (Annex E.3) set the framework for the IVAS codec in terms of capability and resource usage. As such they list functionalities that are divided into mandatory, recommended and optional features to be provided by IVAS codec candidates. In the final standard, all modes have an equal status and they together form the IVAS codec.

Codec features were defined as follows:

The IVAS codec is an extension of the 3GPP Enhanced Voice Services (EVS) codec [2]. It provides full and bit exact EVS codec functionality for mono speech/audio signal input. It further provides:

- Encoding and decoding of stereo and immersive audio formats such as multi-channel audio, scene-based audio (Ambisonics), metadata-assisted spatial audio (MASA), object-based audio (ISM), and their combination.
- VAD/DTX/CNG for rate efficient stereo and immersive conversational voice transmissions
- Error concealment mechanisms to combat the effects of transmission errors and lost packets. Jitter buffer management is also provided.
- The IVAS codec operates on 20 ms audio frames. In addition, rendering is possible with 5 ms granularity.
- Support for bit rate switching upon command.
- Stereo and immersive audio coding at the following discrete bit rates [kbps]: 13.2, 16.4, 24.4, 32, 48, 64, 80, 128, 160, 192, 256, 384, and 512, with supported bit rate ranges listed in Table 4.2-1.
- Support for WB, SWB and FB audio, with the supported bitrate range listed in Table 4.2-2.

Table 4.2-1: Ranges of supported bitrates for stereo and immersive coding of the IVAS codec

Input audio format	Range of supported bitrates [kbps]
Stereo	13.2 – 256
Scene-based audio (SBA)	13.2 – 512
Metadata assisted spatial audio (MASA)	13.2 – 512
Object-based audio (ISM)	13.2 – 512
Multi-channel audio (MC)	13.2 – 512
Combined ISM and MASA (OMASA)	13.2 – 512
Combined ISM and SBA (OSBA)	13.2 – 512

Note, that for the object-based audio format (ISM) the range of supported bitrates varies based on the number of objects as follows: 13.2 kbps – 128 kbps for 1 ISM, 16.4 kbps – 256 kbps for 2 ISMs, 24.4 kbps – 384 kbps for 3 ISMs, 24.4 kbps – 512 kbps for 4 ISMs.

Table 4.2-2: Supported audio bandwidth per input audio format and bitrate

Input audio format	Bitrates supporting WB [kbps]	Bitrates supporting SWB [kbps]	Bitrates supporting FB [kbps]
Stereo	13.2 – 256	13.2 – 256	32 – 256
Scene-based audio (SBA)	13.2 – 512	13.2 – 512	32 – 512
Metadata assisted spatial audio (MASA)	13.2 – 512	13.2 – 512	32 – 512
Object-based audio (ISM), 1 object	13.2 – 128	13.2 – 128	16.4 – 128
Object-based audio (ISM), 2 objects	13.2 – 256	24.4 – 256	32 – 256
Object-based audio (ISM), 3 objects	24.4 – 384	24.4 – 384	48 – 384
Object-based audio (ISM), 4 objects	24.4 – 512	24.4 – 512	64 – 512
Multi-channel audio (MC)	13.2 – 512	13.2 – 512	32 – 512
Combined ISM and MASA (OMASA)	13.2 – 512	13.2 – 512	32 – 512
Combined ISM and SBA (OSBA)	13.2 – 512	13.2 – 512	32 – 512

The IVAS-4 Permanent Document (Annex E.3) also sets constraints on maximum algorithmic delay (40 ms); frame length (20 ms); maximum computational complexity (10*EVS); memory limits (10*EVS); and limit of the output gain.

A full description of the performance requirements can be found in IVAS-3 Permanent Document: Performance Requirements (see Annex E.3).

6 Selection Process

3GPP runs codec selection as a rigorous process, outlined below.

Codec selection in 3GPP follows pre-defined procedures. Proponents are obliged to provide certain information about their candidate to facilitate selection, and strict rules are set prior to selection to provide guidance on selecting the candidate to be standardized. Verification serves the purpose of cross-check and provision of additional (technical) information.

Selection Deliverables are specified in IVAS-6 Permanent Document (Annex E.5).

Proponents were required to provide the following information about their candidate for selection (named selection deliverables):

- Candidate Codec Executable (floating-point)
- Candidate Codec Source Code (floating-point)
- High level technical description of the candidate algorithm
- Report covering the compliance to Design Constraints
- Draft overview specification
- IPR declaration
- Funding payment (proponents paid for selection testing)
- Legal framework to cover use model parameters, unprocessed audio test material, processed audio test material, and test results
- Optional additional information incl. demo material

Selection rules are specified in IVAS-5 Permanent Document (Annex E.4).

The strict 3GPP selection process involved the following rules (which were agreed before selection) to determine the candidate to be standardized:

- Provision of full set of selection phase deliverables
- Compliance with design constraints
- Codec performance analysed based on test conditions, see IVAS-8a (Annex E.8).

In the 3GPP SA4#125 meeting the selection deliverables and selection test results were reviewed and based on this information, 3GPP SA4 selected the IVAS codec candidate jointly developed by the IVAS Public Collaboration as the 3GPP IVAS standard. Subsequently the SA#101 plenary meeting approved the IVAS codec selection and the set of IVAS specifications [4], [5], [6], [7], [8], and [9].

7 Introduction to the Testing of the IVAS codec

7.1 IVAS Selection Phase Testing

7.1.1 General methodology

7.1.1.1 Overview

The following test methodologies were used in the IVAS Selection test: P.800 [10] under consideration of P.Suppl29 [11] was used in experiments designed to evaluate the Immersive conversation use case scenarios, and BS.1534 [12] was used in experiments designed to evaluate the Generic immersive audio use case scenarios. The high-level configuration of experiments for both methodologies is outlined below.

7.1.1.2 P.800

- Test duration should not exceed 2 hours per listening panel. Typical value of voting period was used for estimation of test durations, but actual voting period is not specified.
- Randomizations constructed under randomized blocks experimental design described in [13].
- 6 categories for each test. Categories are defined for each experiment separately.
- 6 samples/category (1 for each listening panel) plus 1 sample/category for preliminaries.
- 30 naïve listeners, 6 listening panels (5 listeners per panel), each panel with an independent randomization
- 180 votes for each condition.
- Total number of conditions for each experiment: 36
- Number of trials: number of test conditions x 6 talkers/categories = 216 trials.

7.1.1.3 BS.1534

- Number of items per experiment: 12
- 14 experienced listeners
- Maximum total number of conditions: 8
- Number of anchor conditions: 2
  - Direct
  - 7 kHz low-pass anchor

NOTE: the exact number of anchors may vary depending on actual experiment.

7.1.2 Opinion Scales

7.1.2.1 ITU-T P.800 DCR

For ITU-T P.800 DCR tests, the opinion scale as defined in Table 7.1-1 is used.

Table 7.1-1: Opinion scale for ITU-T P.800 DCR test

Impairment	Scale
No impairment	5
Small impairment	4
Moderate impairment	3
Large impairment	2
Very large impairment	1

The following instructions (translated properly to the listening labs language and be given to the listeners) are used. The instructions given to the listeners shall be provided for information in the listening lab report.

INSTRUCTIONS TO NAÏVE LISTENERS FOR P.800 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.

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

7.1.2.2 ITU-R BS.1534

For BS.1534 tests, the continuous quality scale (QCS) according to [12] is used.

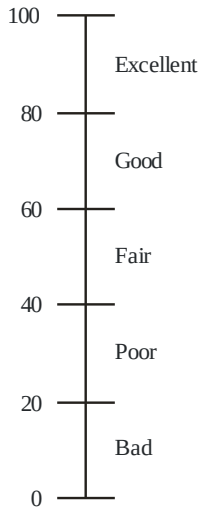


Figure 7.1-1: BS.1534 Continuous Quality Scale

7.1.3 Material

7.1.3.1 Overview

All audio material was sampled at 48 kHz with Full Band (FB) content. The audio material was delivered to the HL as 16-bit little endian WAVE format files [14] following the naming convention provided in the IVAS Processing Plan (IVAS-7a, see Annex E.6). For multi-track audio, the audio tracks are ordered according to Table 5 of IVAS Processing Plan (IVAS-7a). Additionally, it should be verified that the audio material can be processed with the AFsp package tools [15].

The following categories of audio content were used in IVAS Selection Test using P.800:

- Clean speech: Except for experiment P800-6 (1 object), each sample contains two (or more) different talkers in conversation scenario. The talkers transition from one to another as in natural conversation, possibly with partial overlap.
- Speech with background: the details about the environment are specified in IVAS-8a Annex E.
- Music and Mixed content – categories specified in IVAS-8a clause 4.5.3.

The following category of audio content was be used in IVAS Selection Test using BS.1534:

- Generic audio – critical generic audio items including speech with and/or without background, music, mixed.

7.1.3.2 Speech Material for P.800 testing

Except for the experiment P800-3 (Music and mixed content stereo experiment), P.800 test experiments have used artificially created immersive audio. LLs provided clean speech mono audio samples and music and mixed content stereo samples. SA4 would provide scene descriptions and scripts to create the immersive audio.

The recording SNR should be in accordance with P.800 at least 40 dB but preferably 50 dB or higher, the leading and trailing inactivity portions should be shorter than 20 ms. The reverberation time RT60 should be in accordance to P.800 less than 500 ms, preferably below 200 ms. The length of the sentences should typically correspond to the length of traditional Harvard sentences [16].

7.1.3.3 Background Material

- Immersive conversation use case scenario (P.800 testing): A mix-based approach using separate background recordings was used. The minimum lengths of noise files shall be 80 s. The following guideline is applied to the noise types used.
    - Car noise is intended to test the performance of the codec under steady state background noise and should be recorded in a moving car. A constant speed between 80 km/h (50 mph) and 110 km/h (70 mph) is recommended. The make and model of the car should be reasonably common in the country of the recording. Typically, the windows of the car should be closed, and the radio turned off.
    - Office noise is intended to represent a typical office environment. This noise type should also contain typical office sounds, such as keyboard noise, computer fans, telephones ringing, printers, air conditioner, etc.
    - Street noise is intended to represent a typical street environment. It should contain unsteady traffic noise for example recorded at traffic lights where cars stop, human noise such as steps. It should not contain speech, but baby cries are allowed.
- Generic immersive audio use case scenario (BS.1534 testing): Primarily, full recordings of complete immersive scenes including background will be used. A mix-based approach might be used in addition.

7.1.3.4 Music and Mixed Content Material for P.800 testing

Music and mixed content samples shall contain meaningful contents and the duration of each sample shall be approximately 8 s and at least 7 s. The following categories were used:

- Classical music
- Modern instrumental music
- Modern vocal music
- Radio Jingle
- Movie Trailer
- Advertisement

This means that LL shall provide 7 samples per category, 6 for evaluation and 1 for preliminaries.

7.1.3.5 Critical Generic Audio Items for BS.1534 testing

7.1.3.5.1 Steps of Critical Test Item Selection

The following steps were performed:

- Call for test material according to the generic audio signal categories described below.
- Material Collection entity collects candidate material submitted in response to the call and selects a number of critical items to be used in the Selection test.
- Material Collection entity selects a limited set of training items to be used in a training phase.

7.1.3.5.2 Test Material

First, a call was sent out for test material according to a number of generic audio signal categories as specified below. All 3GPP members were invited to submit test material to the Material Collection entity. The submitting organization assigned the items to the below-mentioned audio signal categories. Then, the Material Collection entity identified 12 critical items per experiment, plus four items for training, which are representative for assumed typical IVAS application scenarios.

Generic audio signal categories:

Stereo – generic stereo audio signals with a focus on music categories:

- Pop, with and/or without vocals
- Classic, with and/or without vocals
- Single instruments
- a capella vocals, solo and/or choir
- Mixed speech and music
- Speech with and/or without background noise

Multi-Channel (5.1 and 7.1.4) – generic channel-based audio signals from produced content:

- Music including concerts with live audience
- Film soundtracks with and/or without speech dialogue
- Effects (e,g, nature, city/transport sounds)

Scene-Based Audio / MASA – generic immersive audio signals in the form of complex scenes, captured and/or produced content which may or may not include speech:

- Nature sounds (e.g. forest, water, wind)
- City sounds (e.g. traffic, bus, train)
- Music including concerts with live audience
- Babble-like sound (e.g. market, restaurant, conference)
- Event/Sport-like sound
- Conferencing scene with and/or without background noise/music

Object-Based Audio - Realistic immersive audio signals, e.g.:

- Scenarios comprising voice, music, background objects.
- Conversational scenarios of several talkers with or without background, with or without partial overtalk of no more than two talkers. Talkers may be moving around the scene at natural pace. However, it is not expected that all talkers are active all the time, with unnaturally rapid displacements.

The length in time of the items is 10s at a maximum.

The Material Collection entity is reported to SA4 a list indicating the number of proposed items per submitting organization.

7.1.3.5.3 Training material

Limited material was used in the training phase in which the subjects familiarize with the testing methodology and environment.

The training was conducted with four sound items. These items were identified by the Material Collection entity and shall not be re-used in the blind grading phase. The training phase shall be executed as a separate short BS.1534 session.

7.1.4 Listening Systems and Listening Environments

The IVAS Selection Test has used the following listening systems:

- Stereo headphones for binaural listening, e.g.:
  - Beyerdynamic DT 770 Pro for P.800 experiments
  - Sennheiser HD 650 for BS1534 experiments
- Loudspeaker listening system – 7.1+4 loudspeaker setup [17].

7.1.5 Experimental Procedure for P.800 experiments

Initially the experimenter provided a written copy of the experiment instructions to the listeners. When the listeners have had acknowledged that they understand the instructions, they were presented with a practice session to rate the preliminary conditions. After the practice session was completed, the experimenter asked if there are any questions. Only questions about the rating procedures or the meaning of the instructions should be answered. Any technical questions on matters such as the experimental methodology or details of the types of distortions they are rating must not be answered.

7.1.6 Subjective Experiments

The purpose of the 23 experiments (Experiments P800-1 – P800-9, and BS1534-1a – BS1534-7b) is to evaluate the performances of the IVAS codec candidate algorithm with respect to the performance requirements and objectives defined in (IVAS-3).

The details provided in this clause and in corresponding Annexes of IVAS-8a are those that are specific to each particular experiment. Generic information can be found in IVAS-8a clause 4. LLs were asked to use the information in IVAS-8a clause 4 in conjunction with the information given in IVAS-8a clause 5 and its Annexes.

Table 7.1-2 shows a high-level overview of P.800 experiments. Table 7.1-3 shows a high-level overview of BS.1534 experiments. Finally, Table 7.1-4 shows allocation of experiments to LLs and languages proposed by LLs for each P.800 experiment.

Detail conditions for each subjective experiment are defined in IVAS-8a Annex E for P.800 experiments and in IVAS-8a Annex F for BS.1534 experiments.

Table 7.1-2: High-level overview of P.800 experiments

Exp	Input format	Source material	Listening environment	Bitrates kbps	FER/jitter	DTX
P800-1	Stereo	Clean speech	Headphones	≤ 48	≤ 5%	Y
P800-2	Stereo	Speech+Background	Headphones	≤ 64	≤ 5%	Y
P800-3	Stereo	Mixed & Music	Headphones	≤ 64	≤ 5%	Y
P800-4	FOA	Clean speech	Headphones	≤ 96	≤ 5%	N
P800-5	FOA	Speech+Background	Headphones	≤ 96	≤ 5%	Y
P800-6	1 Object	Clean speech	Headphones	≤ 64	≤ 5%	Y
P800-7	2 Objects	Clean speech	Headphones	≤ 64	≤ 5%	Y
P800-8	MASA	Clean speech	Headphones	≤ 80	≤ 5%	N
P800-9	MASA	Speech+Background	Headphones	≤ 80	≤ 5%	Y

Table 7.1-3: High-level overview of BS.1534 experiments

Exp	Input format	Source material	Listening environment	Bitrates kbps
BS1534-1a	Stereo	Generic Audio	Headphones	48, 64
BS1534-1b	Stereo	Generic Audio	Headphones	96, 128
BS1534-2a	5.1	Generic Audio	5.1	64, 96
BS1534-2b	5.1	Generic Audio	5.1	128, 160
BS1534-3a	7.1.4	Generic Audio	7.1 + 4	128, 160
BS1534-3b	7.1.4	Generic Audio	7.1 + 4	384, 512
BS1534-4a	FOA	Generic Audio	Headphones	96, 128, 160
BS1534-4b	HOA2	Generic Audio	Headphones	160, 192
BS1534-5a	HOA3	Generic Audio	Headphones	192, 256
BS1534-5b	HOA3	Generic Audio	7.1 + 4	384, 512
BS1534-6a	Objects	Generic Audio	Headphones	48, 64, 96
BS1534-6b	Objects	Generic Audio	Headphones	96, 128, 192
BS1534-7a	MASA	Generic Audio	Headphones	96, 128
BS1534-7b	MASA	Generic Audio	Headphones	192, 256

Notes:

- Stereo may include binauralized samples (without head tracking).
- For inputs 7.1+4, FOA, HOA2, HOA3, Objects & MASA vertical dimension is assumed in the samples.
- DTX on/off is assumed within the same experiment, where DTX on is used for relevant conditions.
- All experiments except for stereo P.800 experiments are assumed Full Band experiments, i.e., the direct reference condition is always FB. P.800 stereo experiments are SWB experiments.

Minimum requirements for P.800 experiments: 6 talkers (3 male + 3 female) per experiment, 14 single sentences per talker.

Table 7.1-4 shows allocation of LLs so that each experiment is conducted twice, each time by a different LL. For P.800 experiments, each experiment is run twice with different languages.

Table 7.1-4: Allocation of experiments to LLs and P.800 languages

Exp Source material Listening environment			Languages			
			Force Technology	Head Acoustics/ IKS	MQ University	Mesaqin
			Lab a	Lab b	Lab c	Lab d
P800-1	Clean speech	Headphones	JAP			FR
P800-2	Speech+Background	Headphones		GER		MAN
P800-3	Mixed & Music	Headphones	DAN			MAN
P800-4	Clean speech	Headphones	JAP		ENG	
P800-5	Speech+Background	Headphones	DAN	GER		
P800-6	Clean speech	Headphones	JAP		ENG	
P800-7	Clean speech	Headphones	DAN			MAN
P800-8	Clean speech	Headphones	DAN	GER		
P800-9	Speech+Background	Headphones	JAP			FR
BS1534-1a	Generic Audio	Headphones	x			x
BS1534-1b	Generic Audio	Headphones		x		x
BS1534-2a	Generic Audio	5.1	x	x		
BS1534-2b	Generic Audio	5.1	x	x		
BS1534-3a	Generic Audio	7.1 + 4	x	x		
BS1534-3b	Generic Audio	7.1 + 4	x	x		
BS1534-4a	Generic Audio	Headphones	x			x
BS1534-4b	Generic Audio	Headphones		x		x
BS1534-5a	Generic Audio	Headphones	x			x
BS1534-5b	Generic Audio	7.1+4	x	x		
BS1534-6a	Generic Audio	Headphones		x		x
BS1534-6b	Generic Audio	Headphones		x		x
BS1534-7a	Generic Audio	Headphones		x		x
BS1534-7b	Generic Audio	Headphones		x		x
Total						

Where

- JAP = Japanese
- FR = French
- GER = German
- MAN = Mandarin
- DAN = Danish
- ENG = English

7.2 IVAS Characterization Phase Testing

7.2.1 General methodology

7.2.1.1 Overview

The following test methodologies were used in the IVAS Characterization test:

- P.800 DCR [10] under consideration of P.Suppl29 [11];
- BS.1534 [12];
- P.800 ACR [10] under consideration of P.Suppl29 [11|;
- A preference testing methodology for testing room acoustics synthesis.

The high-level configuration of experiments for envisaged methodologies is outlined below.

7.2.1.2 P.800 DCR

- Considerations of P.Suppl29 [11] shall be taken into account;
- Test duration should not exceed 2 hours per listening panel. The typical value of voting period was used for estimation of test durations, but actual voting period is not specified;
- Randomizations constructed under randomized blocks experimental design described in [13];
- 6 categories for each test. Categories are defined for each experiment separately;
- 6 samples/category (1 for each listening panel) plus 1 sample/category for preliminaries;
- 30 naïve listeners, 6 listening panels (5 listeners per panel), each panel with an independent randomization;
- 180 votes for each condition;
- Total number of conditions for each experiment: 36;
- Number of trials: number of test conditions x 6 talkers/categories = 216 trials;
- P.800 experiments are performed with native listeners of the tested language.

7.2.1.3 BS.1534

- Number of items per experiment: 12;
- 14 experienced listeners;
- Maximum total number of conditions: 8;
- Number of anchor conditions: 2.
  - Direct;
  - 3.5 kHz low-pass anchor.

The exact number of conditions may vary depending on actual experiment. Each BS.1534 experiment comprises a training phase in which the subjects familiarize themselves with the testing methodology and environment [12].

7.2.1.4 P800 ACR

- Considerations of P.Suppl29 [11] shall be taken into account as far as they can be interpreted for ACR test methodology;
- Test duration should not exceed 2 hours per listening panel;

- Randomizations constructed under randomized blocks experimental design described in [13];
- 6 sample categories for ACR test;
- 6 samples/category (1 for each listening panel) plus 1 sample/category for preliminaries;
- 30 naïve listeners, 6 listening panels (5 listeners per panel), each panel with an independent randomization;
- 180 votes for each condition;
- Total number of conditions for the ACR experiment: 60;
- Number of trials: number of test conditions x 6 talkers/categories = 360 trials;
- P.800 ACR experiment is performed with native listeners of the tested language-

Initially the experimenter should provide a written copy of the experiment instructions to the listeners. When the listeners have acknowledged that they understand the instructions, they will be presented with a practice session to rate the preliminary conditions. After the practice session has been completed, the experimenter should ask if there are any questions. After practice session all samples are played back in full quality (Spatial Fullband) to give listeners the feeling of the best possible quality. After listening through the material, actual listening test starts. There are comfort pauses after listening every 120 samples.

7.2.1.5 Preference test methodology for testing room acoustics

For room acoustics testing preference test methodology is used. The listeners are requested to provide preference scoring based on quality rating.

- 4 different rooms;
- 4 audio items per room;
- 4 conditions per test: IVAS rendering with HRTF (reference), IVAS rendering with BRIR, IVAS rendering with synthetic reverb, rendering with BRIR (Python implementation);
- Total: 64 trials.

Each testing experiment is preceded with a training phase in which the listeners should familiarize themselves with the environment and methodology. Unipolar preference scale is used for rating.

7.2.2 Opinion Scales

7.2.2.1 ITU-T P.800 DCR

The same opinion scale, instructions, and voting screen as in 7.1.2.1 were used.

7.2.2.2 ITU-R BS.1534

For BS.1534 tests, the continuous quality scale (QCS) according to [12] is used. See 7.1.2.2 for details.

The following instructions for BS.1534 testing were provided in IVAS-8b:

Familiarization or training phase

The first step in the listening tests is to become familiar with the testing process. This phase is called a training phase, and it precedes the formal evaluation phase.

The purpose of the training phase is to allow you, as an evaluator, to achieve the following two objectives:

- **Part A**: to become familiar with all the sound excerpts under test and their quality level ranges; and
- **Part B**: to learn how to use the test equipment and the grading scale.

In Part A of the training phase, you will be able to listen to all sound excerpts that have been selected for the tests in order to illustrate the whole range of possible qualities. The sound items, which you will listen to, will be more or less critical depending on the bit rate and other "conditions" used. Figure 3 shows the user interface. You may click on different buttons to listen to different sound excerpts including the reference excerpts. In this way you can learn to appreciate a range of different levels of quality for different programme items. The excerpts are grouped on the basis of common conditions. Three such groups are identified in this case. Each group includes four processed signals.

In Part B of the training phase, you will learn to use the available playback and scoring equipment that will be used to evaluate the quality of the sound excerpts.

During the training phase you should be able to learn how you, as an individual, interpret the audible impairments in terms of the grading scale. You should not discuss your personal interpretation of the scale with the other assessors at any time during the training phase. However, you are encouraged to explain artefacts to other assessors.

No grades given during the training phase will be taken into account in the true tests.

Blind grading phase

The purpose of the blind grading phase is to invite you to assign your grades using the quality scale. Your grades should reflect your subjective judgement of the quality level for each of the sound excerpts presented to you. Each trial will contain 9 signals to be graded. Each of the items is approximately 10 s long. You should listen to the reference, anchor, and all the test conditions by clicking on the respective buttons. You may listen to the signals in any order, any number of times.

Use the slider for each signal to indicate your opinion of its quality. When you are satisfied with your grading of all signals you should click on the "register scores" button at the bottom of the screen.

You will use the following quality scale when assigning your grades. The grading scale is continuous from "excellent" to "bad". A grade of 0 corresponds to the bottom of the "bad" category, while a grade of 100 corresponds to the top of the "excellent" category.

In evaluating the sound excerpts, please note that you should not necessarily give a grade in the "bad" category to the sound excerpt with the lowest quality in the test. However, one or more excerpts must be given a grade of 100 because the unprocessed reference signal is included as one of the excerpts to be graded.

7.2.2.3 ITU-T P800-ACR

For ITU-T P.800 ACR tests, the opinion scale as defined in Table 7.2-1 is used.

Table 7.2-1: Opinion scale for ITU-T P.800 ACR test

Overall Quality	Scale
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

The following instructions (translated properly to the listening labs language and be given to the listeners) are used. The instructions given to the listeners shall be provided for information in the listening lab report.

INSTRUCTIONS TO NAÏVE LISTENERS FOR P.800 ACR 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 *test* sample. The *test* sample is possibly impaired after it has passed through a telecommunication system.

Your task is to evaluate the overall quality of the *test* sample, comprising both degradations in the sound quality (e.g., harshness / roughness, loss of bandwidth, unnatural speech, clicks or other distortions), and/or lack of quality in the spatial representation (e.g., unnatural sound source movement, spatial instability, lack of spatial width or other spatial problems).

You should listen carefully the sample within a trial. When sample playback has finished, select the category that best describes your overall impression about the overall quality of the sample you just heard:

5 - Excellent

4 - Good

3 - Fair

2 - Poor

1 - Bad

Note that the level of quality 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.

P.800 ACR Voting screen (after playback of sample pair)

Please rate the OVERALL QUALITY of the sample:

5 - Excellent

4 - Good

3 - Fair

2 - Poor

1 - Bad

7.2.2.4 Preference test for room acoustics

A unipolar preference scale is used for rating.

The following instructions are provided to the listeners:

INSTRUCTIONS TO THE LISTENERS FOR THE IVAS ROOM ACOUSTICS SYNTHESIS EVALUATION TEST

In this experiment you will be evaluating systems that might be used for future immersive telecommunication services using spatial audio. Spatial audio is an audio experience where you can perceive sound as if coming from the space around you, outside and around your head. Spatial audio can be enriched by synthetic reverberation and acoustic reflections that simulate the acoustics of a room, creating a more natural and immersive experience.

For this test you will be wearing a VR headset and headphones. The VR scene used in the test allows for three degrees of freedom (3DoF). That means that you can change your position only by rotation, and not by moving around. Please familiarize yourself with the equipment during the training session and adjust it appropriately. It is highly recommended to perform the tests while sitting on a swivel chair. It is advised not to take this test while standing.

Please take time to observe the scene around you and listen to all the stimuli representing different CuT (conditions under test). Note that it will take a couple of seconds for the playback to start. The CuT provided have been processed with different reverb pathways from the IVAS decoder. A stimulus rendered without room acoustics is

provided as a reference. Please rank your CuT preference regarding room acoustic synthesis quality and matching to the visual environment shown, as compared to the reference without room acoustics. Please consider factors such as timbral quality, spatial depth, and spatial envelopment and how do these relate to the image presented.

7.2.3 Material

7.2.3.1 Overview

All audio material is sampled at 48 kHz with Full Band (FB) content. The audio material was delivered to the HL as 16-bit little endian WAVE format files [14] following the naming convention provided in the IVAS Processing Plan for Characterization Phase (IVAS-7b). For multi-track audio, the audio tracks are ordered according to Table 5 of IVAS Processing Plan for Characterization Phase (IVAS-7b). Additionally, it should be verified that the audio material can be processed with the AFsp package tools [15].

The following categories of audio content was used in IVAS Characterization Test using P.800 DCR:

- Clean speech: Except for experiment evaluating 1 object input, each sample contains two (or more) different talkers in conversation scenario. The talkers transition from one to another as in natural conversation, possibly with partial overlap;
- Speech with background;
- Music and Mixed content;
- Generic audio.

The following category of audio content was used in IVAS Characterization Test using BS.1534:

- Generic audio – critical generic audio items including speech with and/or without background, music, mixed content.

Since it was expected that a significant amount of new material has to be checked and processed for the IVAS Codec Characterization test, test material proponents were encouraged to submit material identical to the Selection testing if possible (i.e. if format, test methodology and test lab match the setup of the Selection phase testing).

7.2.3.2 Material for P.800 DCR testing

7.2.3.2.1 Speech Material for P.800 DCR testing

Except for Music and mixed content categories, P.800 DCR test experiments have used artificially created immersive audio. LLs have provided clean speech mono audio samples. SA4 has provided scene descriptions and scripts to create the immersive audio.

The recording SNR was in accordance with P.800 at least 40 dB but preferably 50 dB or higher. The leading and trailing inactivity portions should be shorter than 20 ms. The reverberation time RT60 was in accordance with P.800 less than 500 ms, preferably below 200 ms. The length of the sentences typically corresponds to the length of traditional Harvard sentences [16].

7.2.3.2.2 Background Material for P.800 DCR testing

A mix-based approach using separate background recordings has been used. The minimum lengths of noise files were 80 s. The following guideline is applied to the noise types used:

- Car noise is intended to test the performance of the codec under steady state background noise and should be recorded in a moving car. A constant speed between 80 km/h (50 mph) and 110 km/h (70 mph) is recommended. The make and model of the car should be reasonably common in the country of the recording. Typically, the windows of the car should be closed, and the radio turned off;
- Office noise is intended to represent a typical office environment. This noise type should also contain typical office sounds, such as keyboard noise, computer fans, telephones ringing, printers, air conditioner, etc;

- Street noise is intended to represent a typical street environment. It should contain unsteady traffic noise for example recorded at traffic lights where cars stop, human noise such as steps. It should not contain speech, but baby cries are allowed.

7.2.3.2.3 Music and Mixed Content Material for P.800 DCR testing

Music and mixed content samples has contained meaningful contents, and the duration of each sample shall be approximately 8 and at least 7seconds. The following categories were used:

- Classical music;
- Modern instrumental music;
- Modern vocal music;
- Radio Jingle;
- Movie Trailer;
- Advertisement.

LLs have provided music and mixed content stereo samples for the stereo experiments. This means that LLs have provided 7 samples per category: 6 for evaluation and 1 for preliminaries. Music and mixed content audio samples for the other P.800 DCR experiments were collected and selected by MC, similarly, as done for the Generic Audio Items Selection for BS.1534 experiments.

7.2.3.2.4 Audio Material for 3- and 4-object categories in P.800 DCR testing

Audio material for 3- and 4-object categories have been collected and selected by MC, similarly, as done for the Generic Audio Items Selection for BS.1534 experiments. The collected audio material shall consist of complete audio scenes falling in the following categories:

- Speech + effects (scene with 3 objects);
- Speech + music (scene with 3 objects);
- Music or effects (scene with 3 objects);
- Speech + effects (scene with 4 objects);
- Speech + music (scene with 4 objects);
- Music or effects (scene with 4 objects).

7.2.3.3 Critical Generic Audio Items for BS.1534 testing

7.2.3.3.1 Steps of Critical Test Item Selection

The following steps have been conducted:

- Call for test material according to the generic audio signal categories described below;
- MC has collected candidate material submitted in response to the call and selects a number of critical items to be used in the Characterization test;
- MC has selected a limited set of training items to be used in a training phase.

7.2.3.3.2 Test Material

First, a call has been sent out for test material according to a number of generic audio signal categories as specified below. All 3GPP members were invited to submit test material to MC. The submitting organization has assigned the items to the below-mentioned audio signal categories. Then, MC has identified 12 critical items per experiment, plus four items for training, which are representative for assumed typical IVAS application scenarios.

Generic audio signal categories:

- Stereo – generic stereo audio signals with a focus on music categories:
  - Pop, with and/or without vocals
  - Classic, with and/or without vocals
  - Single instruments
  - A capella vocals, solo and/or choir
  - Mixed speech and music
  - Speech with and/or without background noise
- Multi-Channel (5.1, 5.1+2, 5.1+4, 7.1 and 7.1+4) – generic channel-based audio signals from produced content:
  - Music including concerts with live audience
  - Film soundtracks with and/or without speech dialogue
  - Effects (e,g, nature, city/transport sounds)
- Scene-Based Audio / MASA – generic immersive audio signals in the form of complex scenes, captured and/or produced content which may or may not include speech:
  - Nature sounds (e.g. forest, water, wind)
  - City sounds (e.g. traffic, bus, train)
  - Music including concerts with live audience
  - Babble-like sound (e.g. market, restaurant, conference)
  - Event/Sport-like sound
  - Conferencing scene with and/or without background noise/music
- Object-Based Audio - Realistic immersive audio signals, e.g.:
  - Scenarios comprising voice, music, background objects.
  - Conversational scenarios of several talkers with or without background, with or without partial overtalk of no more than two talkers. Talkers may be moving around the scene at natural pace.

The length in time of the items was 10 s at a maximum.

MC has further maintained and reported to SA4 a list indicating the number of proposed items per submitting organization.

In case the submitted material was insufficient/inadequate to conduct the tests, MC has added the missing test items.

In order to streamline the work of the Material Collection entity, the BS.1534 Selection test samples were alse used for the BS.1534 Characterization tests as far as possible, especially for the following formats:

- Stereo
- Ambisonics FOA, HOA2, HOA3
- Multi-channel 5.1, 7.1, 5.1+2, 5.1+4, 7.1+4 (7.1, 5.1+2, 5.1+4 to be derived from 7.1+4)
- Objects: 1-4 objects
- MASA 1-2 TCs ← MASA material expected to be derived from FOA/HOA2 material

In order to facilitate this, the material proponents are asked to re-submit the identical material as for the Selection phase testing.

7.2.3.3.3 Training material

Limited material has been used in the training phase in which the subjects familiarize themselves with the testing methodology and environment.

The training was conducted with four sound items. These items were identified by MC and may be re-used in the blind grading phase. The training phase was executed as a separate short BS.1534 session.

7.2.4 Listening Systems and Listening Environments

The IVAS Characterization Test has used the following listening systems:

- Stereo headphones for binaural listening, e.g.:
  - Beyerdynamic DT 770 Pro for P.800 experiments
  - Sennheiser HD 650 for BS.1534 experiments
- Loudspeaker listening system – 5.1, 5.1+2, 5.1+4, 7.1, 7.1+4 loudspeaker setup [17].
- All P.800 tests are carried out via headphones.
- VR setup for 3DoF binaural room acoustics synthesis listening comprising of:
  - A gaming-category PC running Unity game engine, capable to also run 3 IVAS decoder instances and a lightweight Python renderer in parallel,
  - A VR headset, e.g., Meta Quest or HTC Vive Pro,
  - Headphones requirements for binaural tests apply.

7.2.5 Subjective Experiments

7.3.5.1 General Consideration of Experiments

The main goal of the IVAS Characterization Test is to evaluate the aspects of the IVAS codec that were not tested in the Selection phase and validate the fixed-point implementation. Some of the aspects have been tested in formal subjective evaluation while other aspects were evaluated informally. The following aspects have been tested in formal subjective experiments:

- The IVAS fixed-point implementation and the interoperability between the floating-point implementation and the fixed-point implementation
- The integrated IVAS renderer is used, rendering to the playout configuration.
- Stereo; it may include binauralized samples (without head tracking).
- Multi-channel configurations 5.1, 5.1+2, 5.1+4, 7.1, and 7.1+4
- Objects (1-4 ISMs including metadata/rendering)
- FOA, HOA2, HOA3
- MASA 1 TC, 2 TC
- Combined input formats:
  - Objects + MASA (OMASA), 1-4 ISMs
  - Objects + SBA (OSBA), 1-4 ISMs
- JBM
- Packet loss conditions derived from delay and error profiles

- Binaural rendering configurations, e.g.
  - room effects,
  - head rotation,
  - 6 degrees-of-freedom (DoF) and directivity
- EVS-coded mono downmix of stereo input (13.2 and 24.4 kbps)

The subjective material in the P.800 DCR experiments comprising fixed-point validation accommodates all three input signal levels. Low and high signal levels are integrated in different audio samples within categories following Table 7.2-2.

Table 7.2-2: Assignment of input levels to audio samples within categories

Samples	Input Level (in LKFS)					
	cat 1	cat 2	cat 3	cat 4	cat 5	cat 6
s01	-16	-36	-36	-26	-26	-16
s02	-16	-16	-36	-36	-26	-26
s03	-26	-16	-16	-36	-36	-26
s04	-26	-26	-16	-16	-36	-36
s05	-36	-26	-26	-16	-16	-36
s06	-36	-36	-26	-26	-16	-16
s07	-16	-26	-36	-16	-26	-36

The subjective material in the BS.1534 experiments is level-adjusted to the nominal level of -26 dB LKFS.

Evaluation of different configurations of the IVAS codec require different testing methodologies, outlined in the following clauses and annexes.

7.2.5.2 P.800 DCR listening test layout

The following layout is a generic layout. The actual layouts of P.800 DCR experiments may be slightly different. They are described in Annex D.

Table 7.2-3: P.800 DCR testing under clean-channel conditions incl. a comparison of fixed-point and floating-point code

Label	Condition	Bitrate [kbps]	DTX
c01	Reference	-	-
c02	MNRU Q = xx dB	-	-
c03	MNRU Q = xx dB	-	-
c04	MNRU Q = xx dB	-	-
c05	MNRU Q = xx dB	-	-
c06	ESDRU $\alpha = xx$	-	-
c07	ESDRU $\alpha = xx$	-	-
c08	ESDRU $\alpha = xx$		
c09	ESDRU $\alpha = xx$	-	-
c10	IVAS FL enc / FX dec	13.2	
c11	IVAS FX enc / FL dec	16.4	
c12	IVAS FL enc / FX dec	24.4	
c13	IVAS FX enc / FL dec	32.0	
c14	IVAS FL enc / FX dec	48.0	
c15	IVAS FX enc / FL dec	64.0	
c16	IVAS FL enc / FX dec	80.0	
c17	IVAS FX enc / FL dec	96.0	
c18	IVAS FL enc / FX dec	128.0	
c19	IVAS FL	13.2	
c20	IVAS FL	16.4	
c21	IVAS FL	24.4	
c22	IVAS FL	32.0	
c23	IVAS FL	48.0	
c24	IVAS FL	64.0	
c25	IVAS FL	80.0	
c26	IVAS FL	96.0	
c27	IVAS FL	128.0	
c28	IVAS FX	13.2	
c29	IVAS FX	16.4	
c30	IVAS FX	24.4	
c31	IVAS FX	32.0	
c32	IVAS FX	48.0	
c33	IVAS FX	64.0	
c34	IVAS FX	80.0	
c35	IVAS FX	96.0	
c36	IVAS FX	128.0	

7.2.5.3 BS.1534 listening test layouts

The following layouts are generic layouts. The actual layouts of BS.1534 experiments may be slightly different. They are described in Annex D.

Table7.2-4: High Bitrate MUSHRA Tests

Label	Condition	Bitrate [kbps]	DTX
c01	Reference	-	-
c02	LP 3.5 kHz	-	-
c03	EVS	1x 64.0	Off
c04	EVS	1x128.0	Off
c05	IVAS	64.0	Off
c06	IVAS	96.0	Off
c07	IVAS	128.0	Off
c08	IVAS	256.0	Off

Table 7.2-5: Low Bitrate MUSHRA Tests

Label	Condition	Bitrate [kbps]	DTX
c01	Reference	-	-
c02	LP 3.5 kHz	-	-
c03	EVS	1x 16.4	Off
c04	EVS	1x 32.0	Off
c05	IVAS	16.4	Off
c06	IVAS	24.4	Off
c07	IVAS	32.0	Off
c08	IVAS	48.0	Off

7.2.5.4 Allocation of experiments

The following tables show high-level overview of the experiments with allocation of experiments to LLs. Table 7.2-6 shows overview of P.800 DCR experiments with languages used for each experiment. Table 7.2-7 shows overview of BS.1534 experiments. Finally, Table 7.2-8 shows overview of experiments run with different methodologies.

Detail conditions for each subjective experiment are defined in Annex D.

Table 7.2-6: Allocation of P.800 DCR experiments to LLs and corresponding languages

Exp	Objective	Input format	Source material	FX validation	DTX	FER	Bitrates [kb/s]	Language	LL
P800-1	FX, RD	Stereo	All	Yes	Off	0%	13.2-128	FR	VA
P800-2	DTX, FE	Stereo	Al	Yes	On	5%	13.2-128	JAP	NTT
P800-3	FX, RD	FOA	All	Yes	Off	0%	13.2-256	FR	Orange
P800-4	DTX, FE	HOA2	All	Yes	On	5%	16.4-384	FR	Orange
P800-5	FX, RD	HOA3	All	Yes	Off	0%	32-512	ENG	Dolby
P800-6	FX, RD	MC 5.1, 7.1	Clean speech, mixed/music	Yes	Off	0%	24.4-256	MAN	Mesaqin
P800-7	FX, RD	MC 5.1+4, 7.11+4	Clean speech, mixed/music	Yes	Off	0%	32-384	DAN	Force
P800-8	FE	MC (mixed CICP)	Clean speech, mixed/music	Yes	Off	5%	16.4-384	CZ	Mesaqin
P800-9	FX, RD	1-2 Objects	All	Yes	Off	0%	13.2-128	MAN	Mesaqin
P800-10	FX, RD	3-4 Objects	speech+ effects, speech + music, music	Yes	Off	0%	24.4-192	FR	VA
P800-11	DTX, FE	1-4 Objects	Clean speech, speech+ effects, speech + music, music	Yes	On	5%	24.4-192	DAN	Force
P800-12	FX, RD	MASA 1 TC	All	Yes	Off	0%	13.2-128	MAN	Mesaqin
P800-13	FX, RD	MASA 2 TC	All	Yes	Off	0%	13.2-128	DAN	Force
P800-14	DTX, FE	MASA 1,2 TC	All	Yes	On	4, 8%	13.2-128	FIN	Nokia
P800-15	FX, RD	OSBA, 1-2 Objects	All	Yes	Off	0%	32-512	ENG	Dolby
P800-16	FX, RD	OSBA, 3-4 Objects	All	Yes	Off	0%	32-512	SWE	Ericsson
P800-17	RD, FE	OSBA, 1-4 Objects	All	Yes	Off	5%	13.2-512	GER	FhG
P800-18	FX, RD	OMASA, 1-2 Objects	All	Yes	Off	0%	13.2-128	SWE	Ericsson
P800-19	FX, RD	OMASA, 3-4 Objects	All	Yes	Off	0%	13.2-128	GER	FhG
P800-20	RD, FE	OMASA, 1-4 Objects	All	Yes	Off	5%	13.2-512	FIN	Nokia
P800-21	JBM/FE	Stereo	All	Yes	On	I1.O1, I1.O2	24.4-96	CZ	Mesaqin
P800-22	JBM/FE	1-2 Objects	All	Yes	Off	I1.O1, I1.O2	24.4-96	DAN	Force
P800-23	JBM/FE	FOA	All	Yes	Off	I1.O1, I1.O2	24.4-96	DAN	Force

- JAP = Japanese
- FR = French

- GER = German
- MAN = Mandarin
- DAN = Danish
- ENG = English
- FIN = Finnish
- SWE = Swedish
- CZ = Czech

Table 7.2-7: Allocation of BS.1534 experiments to LLs

Exp	Objective	Input format	Source material	Bitrates [kb/s]	Listening environment	LL
BS1534-1	Compare to EVS	Stereo	Generic audio	16.4-48	Headphones	Force
BS1534-2	Compare to EVS	Stereo	Generic audio	64-256	Headphones	Huawei
BS1534-3	Compare to EVS	FOA	Generic audio	16.4-48	Headphones	Huawei
BS1534-4	Compare to EVS	FOA	Generic audio	64-256	Headphones	Panasonic/NTT
BS1534-5	Compare to EVS	HOA3	Generic audio	64-256	7.1 + 4	Dolby
BS1534-6	Compare to EVS	5.1	Generic audio	16.4-48	5.1	Ericsson
BS1534-7	Compare to EVS	5.1, 7.1	Generic audio	64-256	Headphones	Dolby
BS1534-8	Compare to EVS	5.1+2, 5.1+4	Generic audio	64-256	Headphones	Nokia
BS1534-9	Compare to EVS	7.1+4	Generic audio	64-256	7.1 + 4	FhG
BS1534-10	Compare to EVS	1-2 Objects	Generic audio	64-256	Headphones	Qualcomm
BS1534-11	Compare to EVS	3-4 Objects	Generic audio	24.4-64	Headphones	Dolby
BS1534-12	Compare to EVS	3-4 Objects	Generic audio	64-256	Headphones	Philips
BS1534-13	Compare to EVS	MASA 1 TC	Generic audio	16.4-48	Headphones	Qualcomm
BS1534-14	Compare to EVS	MASA 1 TC	Generic audio	64-256	Headphones	Dolby
BS1534-15	Compare to EVS	MASA 2 TC	Generic audio	64-256	Headphones	Nokia
BS1534-16	Compare to EVS	OSBA 1-4 Objects	Generic audio	16.4-48	Headphones	FhG
BS1534-17	Compare to EVS	OSBA 1-4 Objects	Generic audio	64-256	Headphones	Dolby
BS1534-18	Compare to EVS	OMASA 1-4 Objects	Generic audio	64-256	Headphones	Nokia
BS1534-19	Stereo downmix	Stereo	Generic audio	13.2, 24.2	Headphones	Orange
BS1534-20	6-DoF and directivity	4 Objects	Generic audio	64, 512	Headphones	Ericsson

Table 7.2-8: Allocation of experiments using different methodologies

Exp	Objective	Input format	Source material	Listening environment	Bitrates [kb/s]	LL
ACR-1	16, 32, 48 kHz	Stereo, FOA, HOA3, MASA, ISM2	All	Headphones	13.2-256	Nokia
ROOM-1	Room acoustics	FOA	Generic audio	Headphones	80	Philips
ROOM-2	Room acoustics	FOA	Generic audio	Headphones	96	Philips
ROOM-3	Room acoustics	MC – 5.1	Generic audio	Headphones	96	Philips
ROOM-4	Room acoustics	MC – 5.1.2	Generic audio	Headphones	128	Philips

Legend:

- Objective – main objectives of the experiment
- FE (in the “Objective” column) – performance in noisy channel
- DTX (in the “Objective” column) – performance in DTX on
- FX (in the “Objective” column) – fixed-point validation
- RD (in the “Objective” column) – rate-distortion curve
- JBM (in the “Objective” column) – including Jitter Buffer Management
- Tan (in the “Objective” column) – Tandeming
- EVS (in the “Objective” column) – comparison to EVS
- P800 – P.800 DCR test
- ACR – P.800 ACR test
- ROOM – room acoustics test
- All – means clean speech, speech with background, mixed content, and music

8 Important Notes about the Interpretation of Test Results

8.1 P.800 Testing

Mean Opinion Scores can only be representative of the test conditions in which they were recorded (speech/music material, processing, listening conditions, language, and cultural background of the listening subject). Listening tests performed with other conditions than those used in the testing could lead to a different set of MOS results. On the other hand, the relative performances of different codecs under test is considered more reliable and less impacted by cultural difference between listening subjects than absolute MOS values. When looking at the relative differences of the codecs in the same test, it should be noted that a difference of typically 0.15 - 0.2 MOS between two test results would not usually be found statistically significant; appropriate statistical significance tests such as Student's T-test should be used to get an accurate figure of statistically significant difference between conditions within an experiment.

The subjective testing is conducted using limited amount of source material in order to keep the size of the experiment within reasonable limits. Sometimes this can cause some irregularities to the test results. Also, the performance of the tested codecs is not always known when designing the test, thus balancing the test conditions may not always be perfect. This may result in imperfect utilisation of the ranking scale and difficulties to discriminate the codecs with quality very close to each other.

Furthermore, in a number of experiments both clean and erroneous channel conditions were presented in the same experiment. It can be expected that the separation of the different clean channel conditions is less in those experiments compared to experiments where only clean channel conditions are presented. During the setup of the listening experiments SA4 experts made every effort to minimize effects like scale saturation and alike. However, the large number of conditions to be tested and the limited number of experiments that could be conducted made certain compromises unavoidable.

The resolution of the testing is limited. The listeners only use a scale from 1 to 5 to rank the different codecs. However, during the tests presented in the present document, we are characterising a large number of different IVAS modes, most of which are very high quality codecs and this may cause sometimes a "saturation" effect in the test, i.e. the listeners cannot discriminate the different codecs because of the limited range and scale.

Taking into account the comments presented above, the reader is advised to exercise some precautions when looking and comparing the individual scores of the tests. Usually, looking at the whole picture and overall trends in the test in question may give better interpretation of the performance of the codecs. This precaution should be especially taken into account when looking at the experiments conducted using impaired channels which may present rather big variability of results over the limited amount of tested conditions.

Throughout the present document, test results are presented in the following graphical forms.

- **Line-graphs** - Summary results are presented in line-graphs which compare the Reference codecs and the IVAS codec for various test parameters, (e.g. Bit-rate, Frame Error Rate, DTX on/off, etc.). The line-graphs only include conditions from within a test, i.e., no comparisons are made across tests. The graphs include 95% confidence intervals.

8.2 BS.1534 Testing

In contrast to P.800, the BS.1534 (MUSHRA) test methodology uses the continuous quality score (CQS). The CQS consists of identical graphical scales which are divided into five equal intervals with the adjectives “Excellent”, “Good”, “Fair”, “Poor” and “Bad” as also shown in Figure 7.1-1 from top to bottom. The subject records his/her assessment of the quality in a suitable form, typically with the use of sliders on an electronic display. The assessor (subject) is asked to rate the quality of all stimuli, according to the five-interval CQS. The assessments for each test condition are converted linearly to normalized scores in the range 0 to 100, where 0 corresponds to the bottom of the scale (bad quality).

The MUSHRA method has the advantage of displaying many stimuli at the same time so that the assessor is able to carry out any comparison between them directly. In order to achieve reliable results, experienced assessors are chosen to carry out the listening tests. In combination with suitable test material (which was selected by an expert panel), both also contributes to a high resolution of the tests.

The size of a listening panel is not solely a consideration of the desired resolution but a large number of experienced subjects participating contributes to the generalization of the test results. Although 14 experienced subjects participated in the assessment per test for the IVAS selection tests, it should be understood that the results are in principle only valid for precisely that group of experienced listeners actually involved in the actual test. Taking this account, the reader is advised to exercise some precautions when looking and comparing the individual scores of the tests. Usually, looking at the whole picture and overall trends in the test in question may give better interpretation of the performance of the codecs. In order to illustrate the general trends, also graphs combining results from both (independent) participating labs are given.

Throughout the present document, test results are presented in the following graphical forms.

- **Point-graphs, Bar-graphs** - Summary results are presented in point-graphs/bar-graphs which display mean value and 95% confidence intervals. The graphs include the IVAS codec as well as the Reference codecs, anchors and hidden reference. The point-graphs only include conditions from within a test, i.e., no comparisons are made across tests.

As stated already for the P.800 tests, appropriate statistical significance tests such as Student's T-test should be used to get an accurate figure of statistically significant difference between conditions within an experiment.

8.3 Analysis of Selection Phase Results

The Selection phase GAL report (Annex F.1) provides the following analysis of test results:

For both test designs (P800 and BS1534), the ToR requirements in IVAS-8a (Annex E.8) are defined in the same way. All codec-under-test (CuT) conditions have to perform either better than (BT) or not worse than (NWT) certain reference conditions. Due to the uncertainty of the auditory data, it is necessary to consider not only the averaged per-condition results, but also the variance/standard deviation and the number of votes. For this reason, Student's t-test for independent groups is evaluated on each result of the test and the corresponding reference condition. The t-statistic (referred as "T-Stat" in the following) is calculated without the assumption that the variances of test and reference are equal (Welch's t-test).

T-Stat is compared to t(M), the single-sided 95% confidence level of the inverse t-distribution, which depends on the degree of freedom M, i.e., the number of votes per condition. The check result R is then determined according to equation 1. Note that R might also show that the CuT is worse than (WT) the reference.

$$R = \begin{cases} BT & \text{T-Stat} > t(M) \\ NWT & -t(M) \leq \text{T-Stat} \leq t(M) \\ WT & \text{T-Stat} < -t(M) \end{cases} \quad (1)$$

Based on clause 3.3.6.2 of IVAS-8a, it is indicated which and how many requirements have to be met for each condition in each experiment. The tabular data shown in the following clauses utilizes status flags for this purpose:

- FAIL: requirement was not met (cell in tabular data marked in red)
- PASS: requirement was met (cell in tabular data not marked)
- EXCEED: requirement was met and is significantly better (cell in tabular data marked in blue).

Table 8.3-1 illustrates how the status is determined for each combination of requirement and check result.

Table 8.3-1: Status based on check result

Requirement	Check result	Status
NWT	WT	FAIL
NWT	NWT	PASS
NWT	BT	EXCEED
BT	WT	FAIL
BT	NWT	FAIL
BT	BT	PASS

In some cases, a ToR requirement is defined as the logical disjunction (OR) of two separate checks, which lead to two status values. The resulting status with respect to the requirement is determined depending on the involved criterion types as defined in Table 8.3-2.

Table 8.3-2: Status disjunction per criteria

Status #1	Status #2		Result status
(NWT check)	(NWT check)	(BT check)	
EXCEED	EXCEED		EXCEED
EXCEED	PASS		PASS
EXCEED	FAIL		PASS
PASS	PASS		PASS
PASS	FAIL		PASS
FAIL	FAIL		FAIL
EXCEED		PASS	EXCEED
EXCEED		FAIL	PASS
PASS		PASS	PASS
PASS		FAIL	PASS
FAIL		PASS	PASS
FAIL		FAIL	FAIL

NOTE: Each ToR requirement includes at least one NWT check, but never a disjunction of two BT checks. Thus, the cases listed in Table 8.3-2 cover all possible cases of the ToR evaluation.

8.4 Analysis of Characterization Phase Results

The Characterization phase GAL report (Annex F.2) provides the following analysis of test results:

Statistical tests are used to compare the subjective scores of the codecs under test (CuTs) against the scores for specified reference conditions. Due to the uncertainty of the auditory data, it is necessary to consider not only the averaged per-condition results, but also the variance and the number of votes. The tests are conduced as specified in clause 3.3.6.2 of IVAS-8b [1]. Each subjective experiment contains a number of tests to be computed by the GAL.

For the P-800 and BS.1534 experiments, single-sided Student’t t-tests at 95% confidence level are conducted. For P.800 DCR experiments, *independent* groups t-tests are used. The t-statistic (referred as "T-Stat" in the following) is calculated without the assumption that the variances of test and reference are equal (Welch's t-test). For BS.1534 experiments, *dependent* group t-tests are used. T-Stat is compared to t(M), the single-sided 95% confidence level of the inverse t-distribution, which depends on the degree of freedom M (related to the number of votes per condition). The check result R is then determined as follows:

$$R = \begin{cases} BT & \text{T-Stat} > t(M) \\ NWT & -t(M) \leq \text{T-Stat} \leq t(M) \\ WT & \text{T-Stat} < -t(M) \end{cases} \quad (1)$$

Note that this approach, which considers both lower and higher means, is equivalent to a two-sided statistical test at 90% confidence interval.

Result R indicates the following:

- **NWT**: CuT "not worse than" Reference applies if the statistical test does not identify a significant difference between score means.
- **BT**: CuT "better than" Reference applies if the statistical test indicates a significant higher mean score in the CuT.
- **WT**: CuT "worse than" Reference applies if the statistical test indicates a significant lower mean score in the CuT.

IVAS-8b (Annex E.9) specifies the use of a t-test to establish statistical significance for each pair of conditions to be compared in a given listening test. While this simple approach of hypothesis testing is straightforward, it is known to increase the Type-I error in the precense of multiple comparisons. This limitation must be considered especially when interpreting the P.800 characterization results, where the number of comparisons per tests is typically around 20 but can

for some tests be above 50. With such large numbers of pairwise comparisons per test, at a 95% significance ($\alpha = 0.05$), it will be common to see WT/BT results where no actual difference exists. As stated in BS1534-3 [12], the probability Pr of at least one such error is $1 - (1 - 0.05)^N$, where N is the number of pairwise comparisons (note that for N=20, $Pr = 0.64$ and for N=50, $Pr = 0.92$). When interpreting the results of these tests, it is therefore important to consider the p-values of the BT/WT results in the context of N.

9 IVAS Performances

9.1 Mono

The IVAS codec is an extension of the 3GPP Enhanced Voice Services (EVS) codec; it provides full and bit exact EVS codec functionality for mono speech/audio signal input. The performance of the EVS codec is documented in TR 26.952 [17].

9.2 Stereo

9.2.1 Overview

In Selection phase, five experiments have been conducted to evaluate the performance of the IVAS codec with stereo content. While the experiments P800-1, P800-2 and P800-3 were conducted as P.800 DCR tests, the experiments BS1534-1a and BS1534-1b were conducted as BS.1534 tests. All experiments were conducted using headphone presentation.

- Selection Experiment P800-1: Stereo clean speech under clean and impaired channel conditions, headphone presentation
- Selection Experiment P800-2: Stereo speech + background under clean channel conditions, DTX off and on, headphone presentation
- Selection Experiment P800-3: Stereo mixed and music under clean and impaired channel conditions, headphone presentation
- Selection Experiment BS1534-1a: Stereo generic audio, 48 and 64 kbps, headphone presentation
- Selection Experiment BS1534-1b: Stereo generic audio, 96 and 128 kbps, headphone presentation

In Characterization phase, five additional experiments have been conducted. While the experiments P800-1, P800-2 and P800-21 were conducted as P.800 DCR tests, the experiments BS1534-1 and BS1534-2 were conducted as BS.1534 tests.

- Characterization Experiment P800-1: Stereo speech + background and mixed/music under clean channel conditions, headphone presentation
- Characterization Experiment P800-2: Stereo speech + background and mixed/music under impaired channel conditions, DTX on, headphone presentation
- Characterization Experiment P800-21: Stereo speech + background and mixed/music under impaired channel conditions using JBM, DTX on, headphone presentation
- Characterization Experiment BS1534-1: Stereo generic audio, 16.4 – 48 kbps, headphone presentation
- Characterization Experiment BS1534-2: Stereo generic audio, 64 – 256 kbps, headphone presentation

9.2.2 Selection Experiment P800-1 (Stereo/Binaural, Clean Speech, Headphone Presentation)

Selection Experiment P800-1 evaluates IVAS for Stereo/Binaural clean speech under clean and impaired channel conditions with headphone presentation. See Annex C.1 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-1 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.2-1: Statistical overview on the results of P800-1

		Type	CuT						EVS Reference				Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#													
a	c25	1	13.2	off		NWT	3.93	1.08	c10	2x8	3.45	1.06	4.27	BT	EXCEED
		2	13.2	off		BT	3.93	1.08	c09	2x7.2	3.31	1.15	5.33	BT	PASS
	c26	1	16.4	off		NWT	4.14	1.07	c11	2x9.6	3.83	1.02	2.87	BT	EXCEED
		2	16.4	off		BT	4.14	1.07	c10	2x8	3.45	1.06	6.18	BT	PASS
	c27	1	24.4	off		NWT	4.41	0.9	c12	2x13.2	4.38	0.76	0.32	NWT	PASS
		2	24.4	off		BT	4.41	0.9	c11	2x9.6	3.83	1.02	5.69	BT	PASS
	c28	1	32	off		NWT	4.61	0.66	c13	2x16.4	4.57	0.72	0.61	NWT	PASS
		2	32	off		BT	4.61	0.66	c12	2x13.2	4.38	0.76	3.13	BT	PASS
	c29	1	48	off		NWT	4.72	0.68	c15	2x32	4.67	0.62	0.73	NWT	PASS
		2	48	off		BT	4.72	0.68	c14	2x24.4	4.72	0.56	0.08	NWT	FAIL
	c30	1	13.2	off	5%	NWT	3.33	1.22	c17	2x8	2.88	1.14	3.58	BT	EXCEED
		2	13.2	off	5%	BT	3.33	1.22	c16	2x7.2	2.83	1.17	3.97	BT	PASS
	c31	1	16.4	off	5%	NWT	3.52	1.11	c18	2x9.6	3.11	1.21	3.36	BT	EXCEED
		2	16.4	off	5%	BT	3.52	1.11	c17	2x8	2.88	1.14	5.4	BT	PASS
	c32	1	24.4	off	5%	NWT	3.74	0.99	c19	2x13.2	3.62	1.12	1.09	NWT	PASS
		2	24.4	off	5%	BT	3.74	0.99	c18	2x9.6	3.11	1.21	5.38	BT	PASS
	c33	1	32	off	5%	NWT	3.9	1.09	c20	2x16.4	3.64	1.07	2.3	BT	EXCEED
		2	32	off	5%	BT	3.9	1.09	c19	2x13.2	3.62	1.12	2.43	BT	PASS
	c34	1	48	off	5%	NWT	4.15	0.86	c22	2x32	3.83	1.05	3.19	BT	EXCEED
		2	48	off	5%	BT	4.15	0.86	c21	2x24.4	4.09	1.04	0.56	NWT	FAIL
	c35	1	24.4	on		NWT	4.4	0.84	c23	2x13.2	4.23	0.85	1.87	BT	EXCEED
	c36	1	13.2	on	5%	NWT	3.27	1.19	c24	2x8	2.77	1.13	4.14	BT	EXCEED
d	c25	1	13.2	off		NWT	3.5	0.99	c10	2x8	2.07	0.84	14.75	BT	EXCEED
		2	13.2	off		BT	3.5	0.99	c09	2x7.2	2.1	0.92	13.9	BT	PASS
	c26	1	16.4	off		NWT	3.82	0.88	c11	2x9.6	3.22	1.07	5.78	BT	EXCEED
		2	16.4	off		BT	3.82	0.88	c10	2x8	2.07	0.84	19.32	BT	PASS
	c27	1	24.4	off		NWT	4.21	0.86	c12	2x13.2	3.82	0.97	4.09	BT	EXCEED
		2	24.4	off		BT	4.21	0.86	c11	2x9.6	3.22	1.07	9.67	BT	PASS
	c28	1	32	off		NWT	4.31	0.69	c13	2x16.4	4.11	0.8	2.54	BT	EXCEED
		2	32	off		BT	4.31	0.69	c12	2x13.2	3.82	0.97	5.52	BT	PASS
	c29	1	48	off		NWT	4.56	0.63	c15	2x32	4.32	0.77	3.24	BT	EXCEED
		2	48	off		BT	4.56	0.63	c14	2x24.4	4.38	0.68	2.51	BT	PASS
	c30	1	13.2	off	5%	NWT	2.92	1.11	c17	2x8	1.76	0.72	11.74	BT	EXCEED
		2	13.2	off	5%	BT	2.92	1.11	c16	2x7.2	1.69	0.69	12.55	BT	PASS
	c31	1	16.4	off	5%	NWT	3.23	1	c18	2x9.6	2.64	1.06	5.48	BT	EXCEED
		2	16.4	off	5%	BT	3.23	1	c17	2x8	1.76	0.72	16.05	BT	PASS
	c32	1	24.4	off	5%	NWT	3.39	1.01	c19	2x13.2	3.12	0.97	2.55	BT	EXCEED
		2	24.4	off	5%	BT	3.39	1.01	c18	2x9.6	2.64	1.06	6.89	BT	PASS
	c33	1	32	off	5%	NWT	3.6	0.96	c20	2x16.4	3.29	1	2.97	BT	EXCEED
		2	32	off	5%	BT	3.6	0.96	c19	2x13.2	3.12	0.97	4.69	BT	PASS
	c34	1	48	off	5%	NWT	3.76	1.03	c22	2x32	3.61	0.9	1.52	NWT	PASS
		2	48	off	5%	BT	3.76	1.03	c21	2x24.4	3.76	0.87	0.05	NWT	FAIL
	c35	1	24.4	on		NWT	4.14	0.81	c23	2x13.2	3.86	0.95	2.99	BT	EXCEED
	c36	1	13.2	on	5%	NWT	2.97	1.03	c24	2x8	1.82	0.74	12.22	BT	EXCEED

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.2-2: Summary of the results of P800-1

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c25	13.2	off		NWT c10 OR BT c09	EXCEED
	c26	16.4	off		NWT c11 OR BT c10	EXCEED
	c27	24.4	off		NWT c12 OR BT c11	PASS
	c28	32	off		NWT c13 OR BT c12	PASS
	c29	48	off		NWT c15 OR BT c14	PASS
	c30	13.2	off	5%	NWT c17 OR BT c16	EXCEED
	c31	16.4	off	5%	NWT c18 OR BT c17	EXCEED
	c32	24.4	off	5%	NWT c19 OR BT c18	PASS
	c33	32	off	5%	NWT c20 OR BT c19	EXCEED
	c34	48	off	5%	NWT c22 OR BT c21	PASS
	c35	24.4	on		NWT c23	EXCEED
	c36	13.2	on	5%	NWT c24	EXCEED
d	c25	13.2	off		NWT c10 OR BT c09	EXCEED
	c26	16.4	off		NWT c11 OR BT c10	EXCEED
	c27	24.4	off		NWT c12 OR BT c11	EXCEED
	c28	32	off		NWT c13 OR BT c12	EXCEED
	c29	48	off		NWT c15 OR BT c14	EXCEED
	c30	13.2	off	5%	NWT c17 OR BT c16	EXCEED

	c31	16.4	off	5%	NWT c18 OR BT c17	EXCEED
	c32	24.4	off	5%	NWT c19 OR BT c18	EXCEED
	c33	32	off	5%	NWT c20 OR BT c19	EXCEED
	c34	48	off	5%	NWT c22 OR BT c21	PASS
	c35	24.4	on		NWT c23	EXCEED
	c36	13.2	on	5%	NWT c24	EXCEED

The following diagrams show the results for a range of conditions from experiment P800-1 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c09 – c15 for EVS conditions and c25 – c29 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c16 – c22 for EVS conditions and c30 – c34 for IVAS conditions.

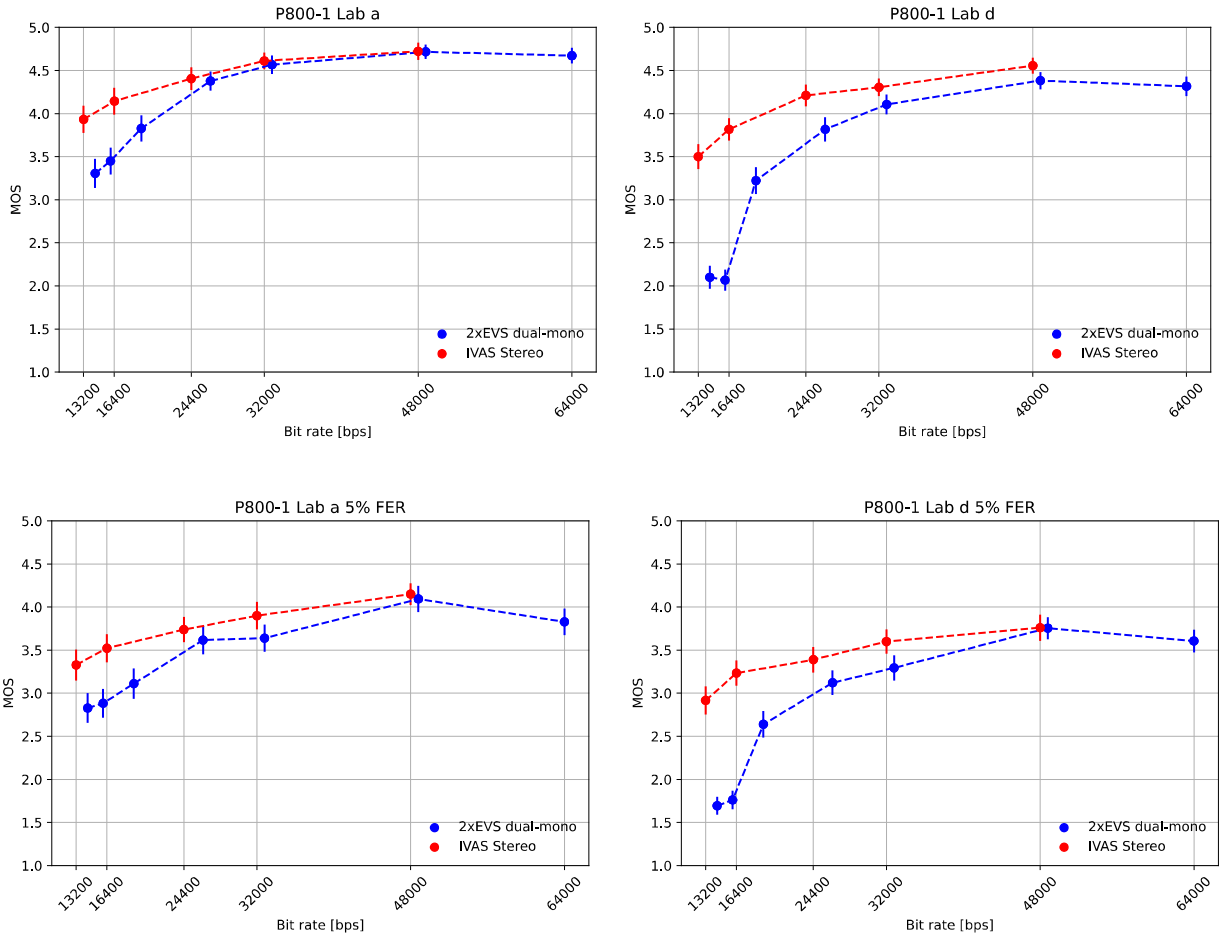


Figure 9.2-1: P800-1 (stereo/binaural, clean speech, headphone presentation) rate-distortion curves for clean and impaired channel conditions, DTX off

9.2.3 Selection Experiment P800-2 (Stereo, Speech+Background, Headphone Presentation)

Selection Experiment P800-2 evaluates IVAS for Stereo speech + background under clean channel conditions with DTX off and on using headphone presentation. See, Annex C.2 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-2 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.2-3: Statistical overview on the results of P800-2

		Type	CuT					EVS Reference				Evaluation		
		Value	Bitrate	DTX	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#												
b	c26	1	13.2	off	NWT	3.28	0.95	c12	2x8	2.45	1.12	7.62	BT	EXCEED
		2	13.2	off	BT	3.28	0.95	c11	2x7.2	2.57	1.11	6.52	BT	PASS
	c27	1	16.4	off	NWT	3.54	0.95	c13	2x9.6	3.43	0.98	1.15	NWT	PASS
		2	16.4	off	BT	3.54	0.95	c12	2x8	2.45	1.12	10.01	BT	PASS
	c28	1	24.4	off	NWT	4.04	0.85	c14	2x13.2	3.81	0.89	2.54	BT	EXCEED
		2	24.4	off	BT	4.04	0.85	c13	2x9.6	3.43	0.98	6.39	BT	PASS
	c29	1	32	off	NWT	4.21	0.75	c15	2x16.4	4.19	0.77	0.21	NWT	PASS
		2	32	off	BT	4.21	0.75	c14	2x13.2	3.81	0.89	4.62	BT	PASS
	c30	1	48	off	NWT	4.71	0.5	c17	2x32	4.41	0.73	4.54	BT	EXCEED
		2	48	off	BT	4.71	0.5	c16	2x24.4	4.54	0.6	2.95	BT	PASS
	c31	1	64	off	NWT	4.72	0.48	c18	2x48	4.82	0.47	-1.89	WT	FAIL
		2	64	off	BT	4.72	0.48	c17	2x32	4.41	0.73	4.76	BT	PASS
	c32	1	13.2	on	NWT	3.33	1.02	c20	2x8	2.48	1.06	7.78	BT	EXCEED
		2	13.2	on	BT	3.33	1.02	c19	2x7.2	2.39	1.02	8.72	BT	PASS
	c33	1	16.4	on	NWT	3.72	0.88	c21	2x9.6	3.34	0.99	3.84	BT	EXCEED
		2	16.4	on	BT	3.72	0.88	c20	2x8	2.48	1.06	12.07	BT	PASS
	c34	1	24.4	on	NWT	4	0.72	c22	2x13.2	3.66	1	3.75	BT	EXCEED
		2	24.4	on	BT	4	0.72	c21	2x9.6	3.34	0.99	7.27	BT	PASS
	c35	1	32	on	NWT	4.16	0.78	c23	2x16.4	4.1	0.83	0.66	NWT	PASS
		2	32	on	BT	4.16	0.78	c22	2x13.2	3.66	1	5.31	BT	PASS
	c36	1	48	on	NWT	4.52	0.59	c25	2x32	4.28	0.73	3.47	BT	EXCEED
		2	48	on	BT	4.52	0.59	c24	2x24.4	4.44	0.67	1.24	NWT	FAIL
d	c26	1	13.2	off	NWT	3.93	1.02	c12	2x8	2.55	1.1	12.38	BT	EXCEED
		2	13.2	off	BT	3.93	1.02	c11	2x7.2	2.71	1.16	10.63	BT	PASS
	c27	1	16.4	off	NWT	4.21	0.88	c13	2x9.6	3.67	1.14	5.06	BT	EXCEED
		2	16.4	off	BT	4.21	0.88	c12	2x8	2.55	1.1	15.84	BT	PASS
	c28	1	24.4	off	NWT	4.43	0.78	c14	2x13.2	4.12	0.97	3.41	BT	EXCEED
		2	24.4	off	BT	4.43	0.78	c13	2x9.6	3.67	1.14	7.46	BT	PASS
	c29	1	32	off	NWT	4.59	0.67	c15	2x16.4	4.41	0.78	2.4	BT	EXCEED
		2	32	off	BT	4.59	0.67	c14	2x13.2	4.12	0.97	5.44	BT	PASS
	c30	1	48	off	NWT	4.73	0.54	c17	2x32	4.56	0.7	2.69	BT	EXCEED
		2	48	off	BT	4.73	0.54	c16	2x24.4	4.63	0.62	1.65	NWT	FAIL
	c31	1	64	off	NWT	4.71	0.52	c18	2x48	4.77	0.52	-1.11	NWT	PASS
		2	64	off	BT	4.71	0.52	c17	2x32	4.56	0.7	2.37	BT	PASS
	c32	1	13.2	on	NWT	3.91	1.02	c20	2x8	2.67	1.13	10.94	BT	EXCEED
		2	13.2	on	BT	3.91	1.02	c19	2x7.2	2.62	1.11	11.51	BT	PASS
	c33	1	16.4	on	NWT	4.27	0.84	c21	2x9.6	3.68	1.18	5.45	BT	EXCEED
		2	16.4	on	BT	4.27	0.84	c20	2x8	2.67	1.13	15.23	BT	PASS
	c34	1	24.4	on	NWT	4.4	0.74	c22	2x13.2	4.24	0.88	1.82	BT	EXCEED
		2	24.4	on	BT	4.4	0.74	c21	2x9.6	3.68	1.18	6.92	BT	PASS
	c35	1	32	on	NWT	4.54	0.71	c23	2x16.4	4.37	0.8	2.16	BT	EXCEED
		2	32	on	BT	4.54	0.71	c22	2x13.2	4.24	0.88	3.49	BT	PASS
	c36	1	48	on	NWT	4.62	0.59	c25	2x32	4.51	0.77	1.54	NWT	PASS
		2	48	on	BT	4.62	0.59	c24	2x24.4	4.57	0.68	0.74	NWT	FAIL

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.2-4: Summary of the results of P800-2

Lab	Cond.	Bitrate	DTX	ToR	Status
b	c26	13.2	off	NWT c12 OR BT c11	EXCEED
	c27	16.4	off	NWT c13 OR BT c12	PASS
	c28	24.4	off	NWT c14 OR BT c13	EXCEED
	c29	32	off	NWT c15 OR BT c14	PASS
	c30	48	off	NWT c17 OR BT c16	EXCEED
	c31	64	off	NWT c18 OR BT c17	PASS
	c32	13.2	on	NWT c20 OR BT c19	EXCEED
	c33	16.4	on	NWT c21 OR BT c20	EXCEED
	c34	24.4	on	NWT c22 OR BT c21	EXCEED
	c35	32	on	NWT c23 OR BT c22	PASS
	c36	48	on	NWT c25 OR BT c24	PASS
d	c26	13.2	off	NWT c12 OR BT c11	EXCEED
	c27	16.4	off	NWT c13 OR BT c12	EXCEED
	c28	24.4	off	NWT c14 OR BT c13	EXCEED
	c29	32	off	NWT c15 OR BT c14	EXCEED
	c30	48	off	NWT c17 OR BT c16	PASS
	c31	64	off	NWT c18 OR BT c17	PASS
	c32	13.2	on	NWT c20 OR BT c19	EXCEED
	c33	16.4	on	NWT c21 OR BT c20	EXCEED
	c34	24.4	on	NWT c22 OR BT c21	EXCEED
	c35	32	on	NWT c23 OR BT c22	EXCEED

	c36	48	on	NWT c25 OR BT c24	PASS

The following diagrams show the results for a range of conditions from experiment P800-2 as rate-distortion curves. The first two diagrams only show results for active coding conditions (DTX off), i.e. conditions c11 – c18 for EVS conditions and c26 – c31 for IVAS conditions. The second two diagrams show results for DTX conditions (DTX on), i.e. conditions c19 – c25 for EVS conditions and c32 – c36 for IVAS conditions.

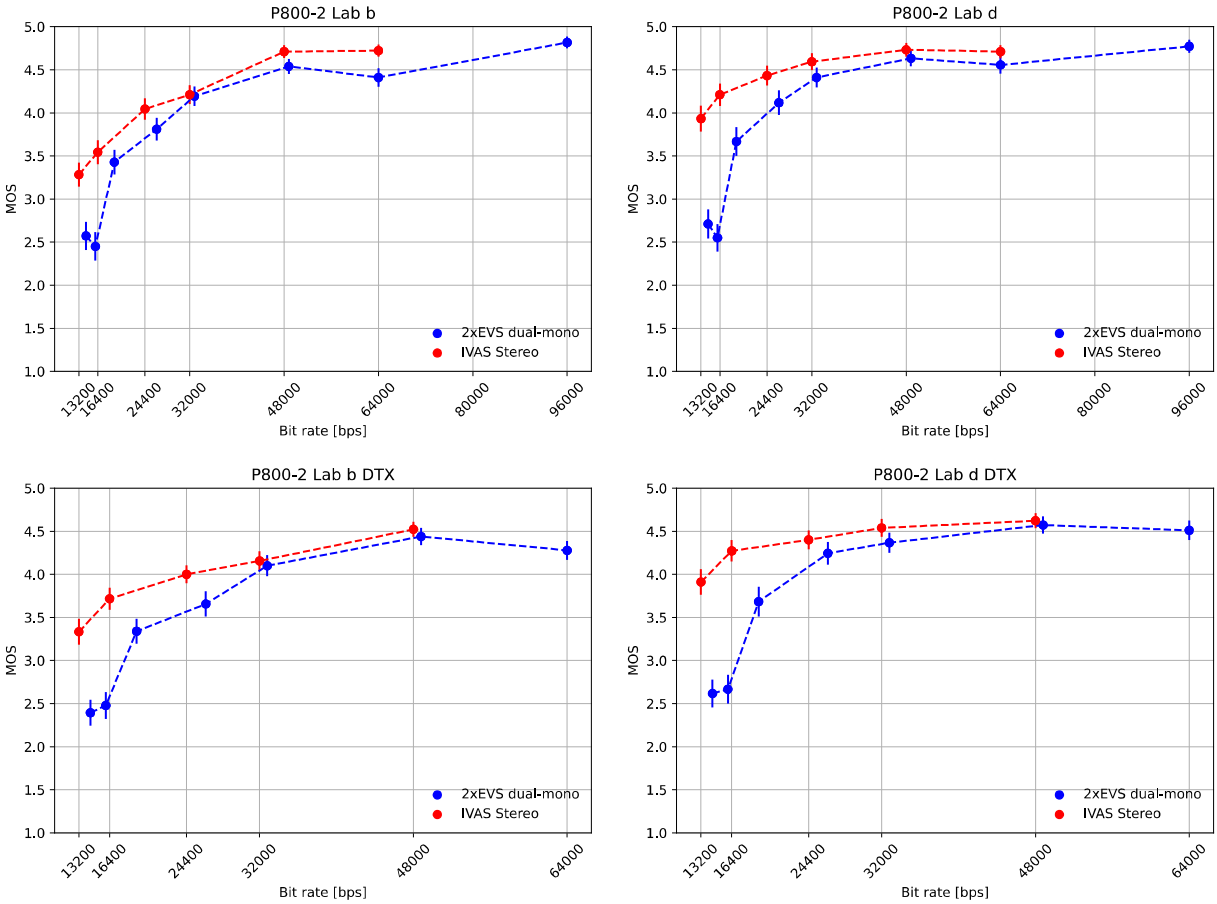


Figure 9.2-2: P800-2 (stereo, speech + background, headphone presentation) rate distortion curves for conditions with DTX off and on

9.2.4 Selection Experiment P800-3 (Stereo, Mixed and Music, Headphone Presentation)

Selection Experiment P800-3 evaluates IVAS for Stereo mixed and music under clean and impaired channel conditions using headphone presentation. See Annex C.3 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-3 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.2-5: Statistical overview on the results of P800-3

		Type	CuT						EVS Reference					Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	Bitrate	DTX	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#														
a	c24	1	13.2	off		NWT	3.26	1.16	c10	2x8		1.95	1	11.43	BT	EXCEED
		2	13.2	off		BT	3.26	1.16	c09	2x7.2		1.89	1.02	11.83	BT	PASS
	c25	1	16.4	off		NWT	3.79	0.93	c11	2x9.6		2.96	1.25	7.18	BT	EXCEED
		2	16.4	off		BT	3.79	0.93	c10	2x8		1.95	1	18	BT	PASS
	c26	1	24.4	off		NWT	4.13	0.88	c12	2x13.2		3.41	1.17	6.62	BT	EXCEED
		2	24.4	off		BT	4.13	0.88	c11	2x9.6		2.96	1.25	10.32	BT	PASS
	c27	1	32	off		NWT	4.24	0.76	c13	2x16.4		3.69	1.09	5.54	BT	EXCEED
		2	32	off		BT	4.24	0.76	c12	2x13.2		3.41	1.17	8.06	BT	PASS
	c28	1	48	off		NWT	4.44	0.71	c15	2x32		4.14	0.92	3.42	BT	EXCEED
		2	48	off		BT	4.44	0.71	c14	2x24.4		4.12	0.87	3.86	BT	PASS
	c29	1	64	off		NWT	4.47	0.65	c16	2x48		4.55	0.64	-1.22	NWT	PASS
		2	64	off		BT	4.47	0.65	c15	2x32		4.14	0.92	3.87	BT	PASS
	c30	1	13.2	off	5%	NWT	2.74	1.07	c18	2x8		1.66	0.94	10.28	BT	EXCEED
		2	13.2	off	5%	BT	2.74	1.07	c17	2x7.2		1.6	0.84	11.32	BT	PASS
	c31	1	16.4	off	5%	NWT	3.26	1.14	c19	2x9.6		2.71	1.18	4.5	BT	EXCEED
		2	16.4	off	5%	BT	3.26	1.14	c18	2x8		1.66	0.94	14.6	BT	PASS
	c32	1	24.4	off	5%	NWT	3.42	1.13	c20	2x13.2		2.78	1.16	5.34	BT	EXCEED
		2	24.4	off	5%	BT	3.42	1.13	c19	2x9.6		2.71	1.18	5.85	BT	PASS
	c33	1	32	off	5%	NWT	3.44	1.11	c21	2x16.4		2.98	1.1	3.96	BT	EXCEED
		2	32	off	5%	BT	3.44	1.11	c20	2x13.2		2.78	1.16	5.56	BT	PASS
	c34	1	48	off	5%	NWT	3.91	0.94	c23	2x32		3.41	1.12	4.6	BT	EXCEED
		2	48	off	5%	BT	3.91	0.94	c22	2x24.4		3.33	1.12	5.3	BT	PASS
	c35	1	24.4	on		NWT	4.17	0.87	c12	2x13.2	off	3.41	1.17	7	BT	EXCEED
		2	24.4	on		BT	4.17	0.87	c11	2x9.6	off	2.96	1.25	10.71	BT	PASS
	c36	1	13.2	on	5%	NWT	2.76	1.16	c18	2x8	off	1.66	0.94	9.98	BT	EXCEED
		2	13.2	on	5%	BT	2.76	1.16	c17	2x7.2	off	1.6	0.84	10.92	BT	PASS
d	c24	1	13.2	off		NWT	3.46	1.05	c10	2x8		2.18	0.98	11.97	BT	EXCEED
		2	13.2	off		BT	3.46	1.05	c09	2x7.2		2.04	0.95	13.44	BT	PASS
	c25	1	16.4	off		NWT	4.06	0.85	c11	2x9.6		3.29	1.2	6.97	BT	EXCEED
		2	16.4	off		BT	4.06	0.85	c10	2x8		2.18	0.98	19.43	BT	PASS
	c26	1	24.4	off		NWT	4.29	0.72	c12	2x13.2		3.64	1.04	6.89	BT	EXCEED
		2	24.4	off		BT	4.29	0.72	c11	2x9.6		3.29	1.2	9.51	BT	PASS
	c27	1	32	off		NWT	4.42	0.67	c13	2x16.4		3.97	0.89	5.41	BT	EXCEED
		2	32	off		BT	4.42	0.67	c12	2x13.2		3.64	1.04	8.5	BT	PASS
	c28	1	48	off		NWT	4.58	0.56	c15	2x32		4.33	0.79	3.46	BT	EXCEED
		2	48	off		BT	4.58	0.56	c14	2x24.4		4.43	0.67	2.39	BT	PASS
	c29	1	64	off		NWT	4.63	0.57	c16	2x48		4.61	0.53	0.46	NWT	PASS
		2	64	off		BT	4.63	0.57	c15	2x32		4.33	0.79	4.13	BT	PASS
	c30	1	13.2	off	5%	NWT	3.12	1.17	c18	2x8		1.88	0.99	10.85	BT	EXCEED
		2	13.2	off	5%	BT	3.12	1.17	c17	2x7.2		1.83	1	11.29	BT	PASS
	c31	1	16.4	off	5%	NWT	3.48	1.06	c19	2x9.6		2.79	1.25	5.64	BT	EXCEED
		2	16.4	off	5%	BT	3.48	1.06	c18	2x8		1.88	0.99	14.7	BT	PASS
	c32	1	24.4	off	5%	NWT	3.9	0.92	c20	2x13.2		3.21	1.16	6.3	BT	EXCEED
		2	24.4	off	5%	BT	3.9	0.92	c19	2x9.6		2.79	1.25	9.64	BT	PASS
	c33	1	32	off	5%	NWT	4.03	0.93	c21	2x16.4		3.52	1.21	4.55	BT	EXCEED
		2	32	off	5%	BT	4.03	0.93	c20	2x13.2		3.21	1.16	7.47	BT	PASS
	c34	1	48	off	5%	NWT	4.14	0.84	c23	2x32		3.76	1.09	3.74	BT	EXCEED
		2	48	off	5%	BT	4.14	0.84	c22	2x24.4		3.83	1.04	3.08	BT	PASS
	c35	1	24.4	on		NWT	4.38	0.79	c12	2x13.2	off	3.64	1.04	7.58	BT	EXCEED
		2	24.4	on		BT	4.38	0.79	c11	2x9.6	off	3.29	1.2	10.09	BT	PASS
	c36	1	13.2	on	5%	NWT	3.09	1.15	c18	2x8	off	1.88	0.99	10.65	BT	EXCEED
		2	13.2	on	5%	BT	3.09	1.15	c17	2x7.2	off	1.83	1	11.09	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.2-6: Summary of the results of P800-3

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c24	13.2	off		NWT c10 OR BT c09	EXCEED
	c25	16.4	off		NWT c11 OR BT c10	EXCEED
	c26	24.4	off		NWT c12 OR BT c11	EXCEED
	c27	32	off		NWT c13 OR BT c12	EXCEED
	c28	48	off		NWT c15 OR BT c14	EXCEED
	c29	64	off		NWT c16 OR BT c15	PASS
	c30	13.2	off	5%	NWT c18 OR BT c17	EXCEED
	c31	16.4	off	5%	NWT c19 OR BT c18	EXCEED
	c32	24.4	off	5%	NWT c20 OR BT c19	EXCEED
	c33	32	off	5%	NWT c21 OR BT c20	EXCEED
	c34	48	off	5%	NWT c23 OR BT c22	EXCEED
	c35	24.4	on		NWT c12 OR BT c11	EXCEED
	c36	13.2	on	5%	NWT c18 OR BT c17	EXCEED
d	c24	13.2	off		NWT c10 OR BT c09	EXCEED
	c25	16.4	off		NWT c11 OR BT c10	EXCEED

	c26	24.4	off		NWT c12 OR BT c11	EXCEED
	c27	32	off		NWT c13 OR BT c12	EXCEED
	c28	48	off		NWT c15 OR BT c14	EXCEED
	c29	64	off		NWT c16 OR BT c15	PASS
	c30	13.2	off	5%	NWT c18 OR BT c17	EXCEED
	c31	16.4	off	5%	NWT c19 OR BT c18	EXCEED
	c32	24.4	off	5%	NWT c20 OR BT c19	EXCEED
	c33	32	off	5%	NWT c21 OR BT c20	EXCEED
	c34	48	off	5%	NWT c23 OR BT c22	EXCEED
	c35	24.4	on		NWT c12 OR BT c11	EXCEED
	c36	13.2	on	5%	NWT c18 OR BT c17	EXCEED

The following diagrams show the results for a range of conditions from experiment P800-3 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c09 – c16 for EVS conditions and c24 – c29 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c17 – c24 for EVS conditions and c30 – c34 for IVAS conditions.

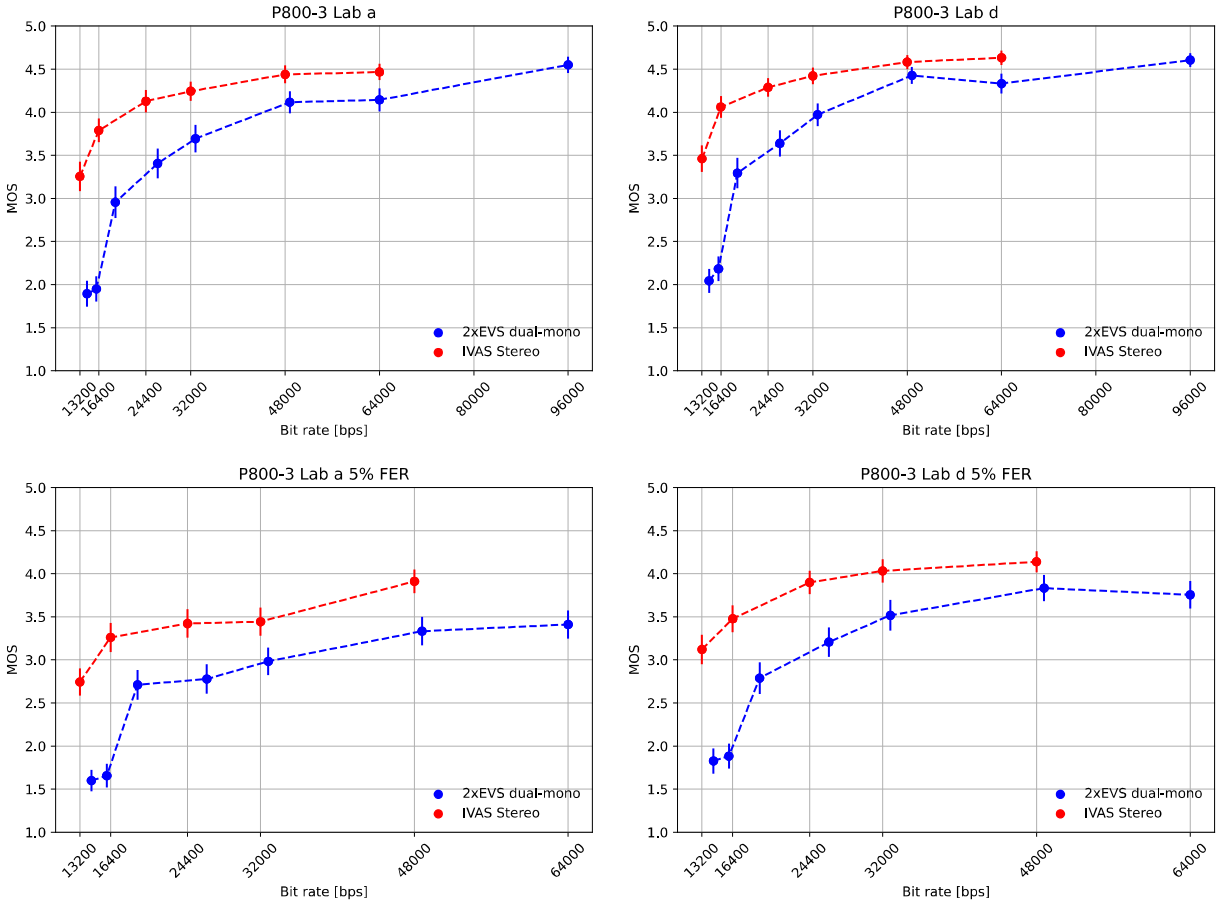


Figure 9.2-3: P800-3 (stereo, mixed & music) rate distortion curves for clean and impaired channel conditions

9.2.5 Selection Experiment BS1534-1a (Stereo, Generic Audio, 48 and 64 kbps, Headphone Presentation)

Selection Experiment BS1534-1a evaluates IVAS for Stereo generic audio at 48 and 64 kbps using headphone presentation. See Annex C.10 for details.

The averaged results per condition for experiment BS1534-1a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown

grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

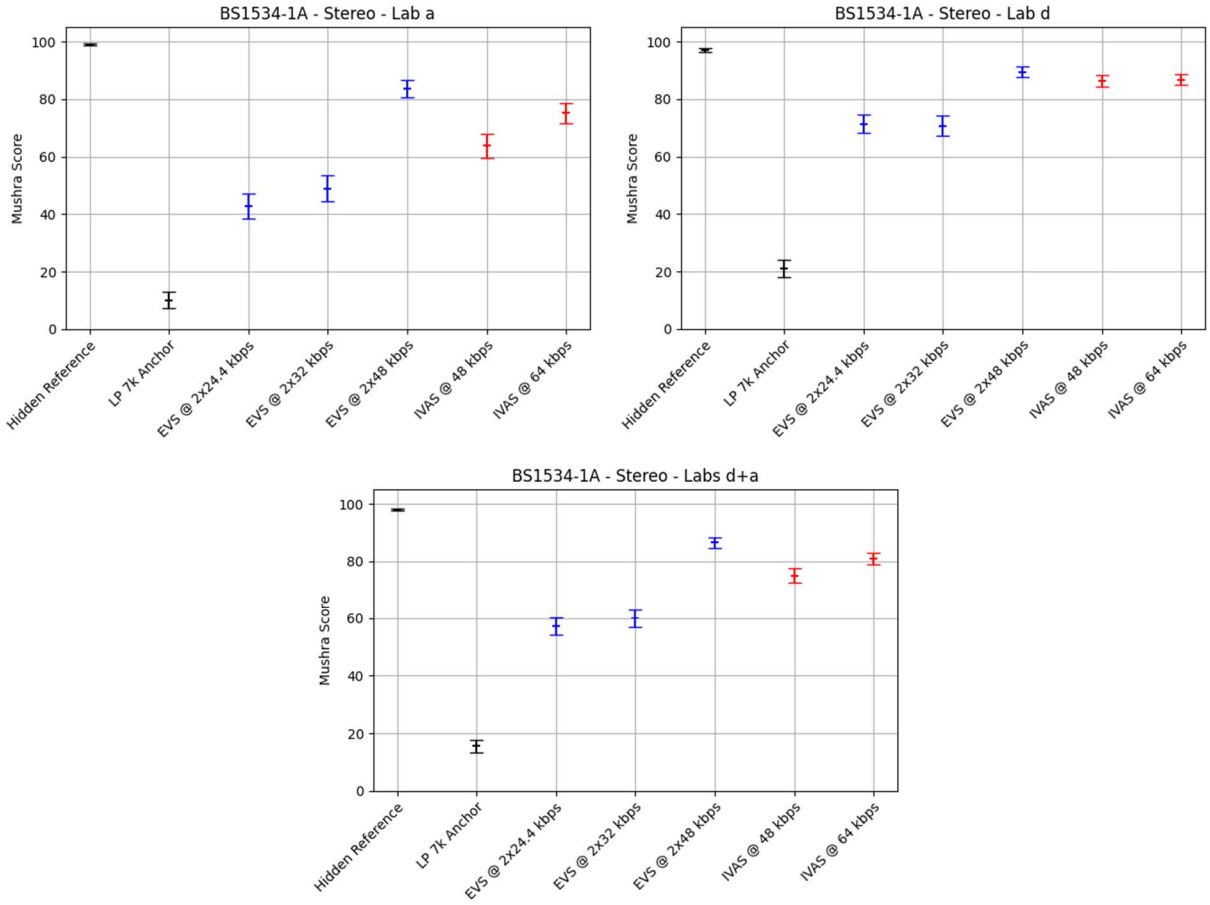


Figure 9.2-4: BS1534-1a (Stereo, generic audio, 48 and 64 kbps, headphone presentation) MUSHRA plots for labs a and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-1a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.2-7: Statistical overview on the results of BS1534-1a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	48	NWT	63.8	27.3	c04	2x32	48.9	29.2	4.81	BT	EXCEED
		2	48	BT	63.8	27.3	c03	2x24.4	42.9	28.1	6.92	BT	PASS
	c07	1	64	NWT	75.2	22.8	c05	2x48	83.6	19.4	-3.64	WT	FAIL
		2	64	BT	75.2	22.8	c04	2x32	48.9	29.2	9.18	BT	PASS
d	c06	1	48	NWT	86.5	13.3	c04	2x32	70.7	23.2	7.66	BT	EXCEED
		2	48	BT	86.5	13.3	c03	2x24.4	71.4	21.4	7.76	BT	PASS
	c07	1	64	NWT	86.8	11.6	c05	2x48	89.5	11.8	-2.1	WT	FAIL
		2	64	BT	86.8	11.6	c04	2x32	70.7	23.2	8.08	BT	PASS
d+a	c06	1	48	NWT	75.1	24.3	c04	2x32	59.8	28.5	7.5	BT	EXCEED
		2	48	BT	75.1	24.3	c03	2x24.4	57.1	28.7	8.77	BT	PASS
	c07	1	64	NWT	81	19	c05	2x48	86.6	16.3	-4.06	WT	FAIL
		2	64	BT	81	19	c04	2x32	59.8	28.5	11.36	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.2-8: Summary of the results of BS1534-1a

Lab	Cond.	Bitrate	ToR	Status
a	c06	48	NWT c04 OR BT c03	EXCEED
	c07	64	NWT c05 OR BT c04	PASS
d	c06	48	NWT c04 OR BT c03	EXCEED
	c07	64	NWT c05 OR BT c04	PASS
d+a	c06	48	NWT c04 OR BT c03	EXCEED
	c07	64	NWT c05 OR BT c04	PASS

9.2.6 Selection Experiment BS1534-1b (Stereo, Generic Audio, 96 and 128 kbps, Headphone Presentation)

Selection Experiment BS1534-1b evaluates IVAS for Stereo generic audio at 96 and 128 kbps using headphone presentation. Se IVAS-8a, Annex F.2 for details.

The averaged results per condition for experiment BS1534-1b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

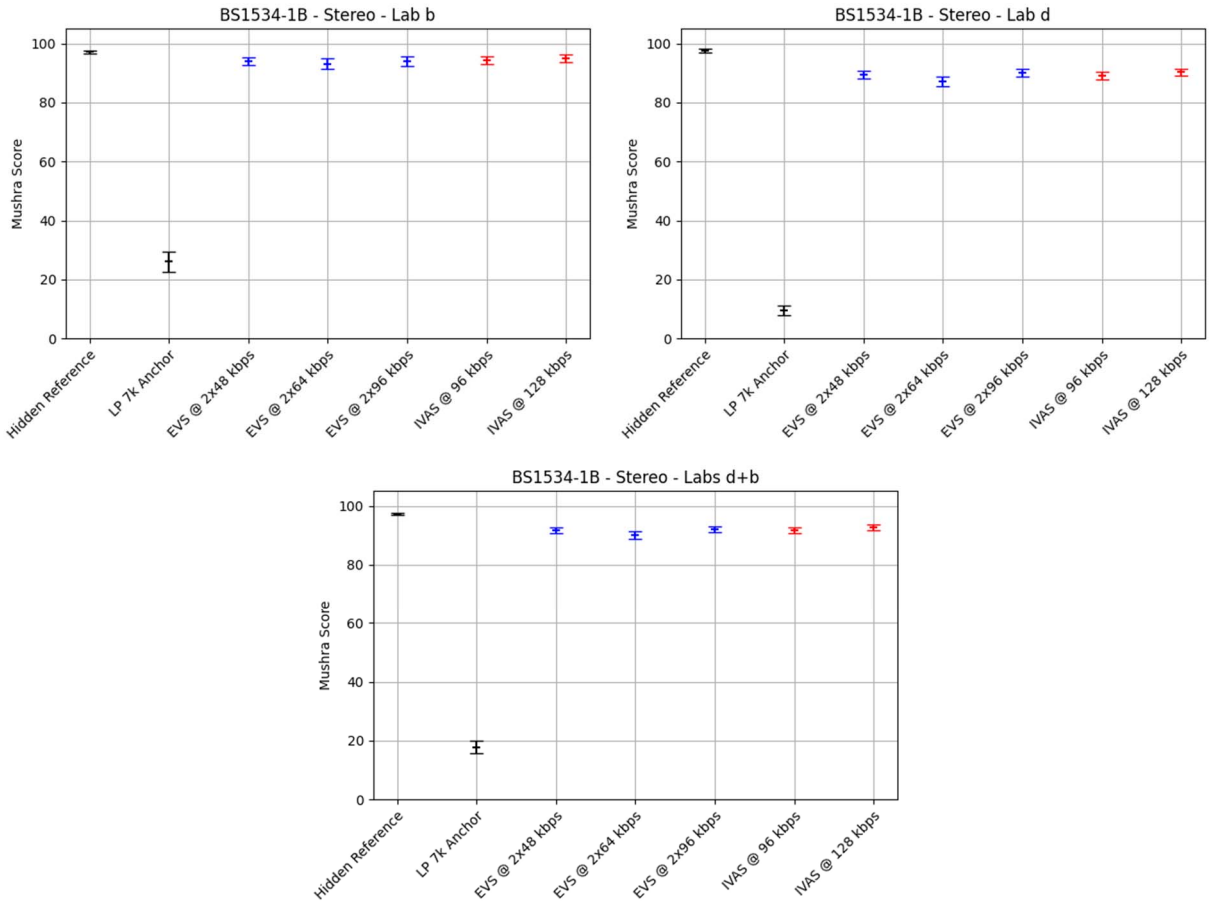


Figure 9.2-5: BS1534-1b (Stereo, generic audio, 96 and 128 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-1b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.2-9: Statistical overview on the results of BS1534-1b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c06	1	96	NWT	94.4	9	c04	2x64	93.1	11.9	1.13	NWT	PASS
		2	96	BT	94.4	9	c03	2x48	94	9.5	0.37	NWT	FAIL
	c07	1	128	NWT	95	8.5	c05	2x96	93.9	10.9	0.98	NWT	PASS
		2	128	BT	95	8.5	c04	2x64	93.1	11.9	1.68	BT	PASS
d	c06	1	96	NWT	89.2	8.9	c04	2x64	87.2	10.5	1.8	BT	EXCEED
		2	96	BT	89.2	8.9	c03	2x48	89.4	8	-0.3	NWT	FAIL
	c07	1	128	NWT	90.4	7.5	c05	2x96	90.2	8.7	0.26	NWT	PASS
		2	128	BT	90.4	7.5	c04	2x64	87.2	10.5	3.16	BT	PASS
d+b	c06	1	96	NWT	91.8	9.3	c04	2x64	90.2	11.6	1.98	BT	EXCEED
		2	96	BT	91.8	9.3	c03	2x48	91.7	9	0.06	NWT	FAIL
	c07	1	128	NWT	92.7	8.3	c05	2x96	92.1	10	0.9	NWT	PASS
		2	128	BT	92.7	8.3	c04	2x64	90.2	11.6	3.24	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.2-10: Summary of the results of BS1534-1b

Lab	Cond.	Bitrate	ToR	Status
b	c06	96	NWT c04 OR BT c03	PASS
	c07	128	NWT c05 OR BT c04	PASS
d	c06	96	NWT c04 OR BT c03	PASS
	c07	128	NWT c05 OR BT c04	PASS
d+b	c06	96	NWT c04 OR BT c03	PASS
	c07	128	NWT c05 OR BT c04	PASS

9.2.7 Characterization Experiment P800-1 (Stereo, Speech and Mixed/Music, Clean Channel, Headphone Presentation)

Characterization Experiment P800-1 evaluates the performance of the IVAS fixed-point and floating-point implementation for Stereo speech and mixed/music under clean channel conditions with headphone presentation. See Annex E.9 for details.

Characterization Experiment P800-1 was conducted by VoiceAge Corporation (language: FR).

The complete statistical evaluation of the requirement ToR tests for characterization experiment P800-1 is given in the following table.

The aggregated results per condition are listed in Table 9.2-11 and are shown in Figure 9.2-6. The results aggregated over FX/FL are shown in Figure 9.2-7. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.2-12. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.2-11: results per condition for experiment P800-1

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.63	0.60	0.09
c02	MNRU Q = 27dB	180	3.33	1.24	0.18
c03	MNRU Q = 22 dB	180	2.51	1.36	0.20
c04	MNRU Q = 17 dB	180	1.56	0.98	0.14
c05	MNRU Q = 12 dB	180	1.16	0.51	0.07
c06	ESDRU α=0.7	180	4.02	0.81	0.12
c07	ESDRU α=0.5	180	3.03	1.09	0.16
c08	ESDRU α=0.3	180	2.64	1.11	0.16
c09	ESDRU α=0.1	180	2.02	0.98	0.14
c10	IVAS FL enc / FX dec, 13.2 kbps, DTX off	180	2.73	1.18	0.17
c11	IVAS FX enc / FL dec, 16.4 kbps, DTX off	180	3.03	1.07	0.16
c12	IVAS FL enc / FX dec, 24.4 kbps, DTX off	180	3.61	0.95	0.14
c13	IVAS FX enc / FL dec, 32 kbps, DTX off	180	3.92	0.94	0.14
c14	IVAS FL enc / FX dec, 48 kbps, DTX off	180	4.35	0.80	0.12
c15	IVAS FX enc / FL dec, 64 kbps, DTX off	180	4.57	0.64	0.09

c16	IVAS FL enc / FX dec, 80 kbps, DTX off	180	4.61	0.60	0.09
c17	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.56	0.66	0.10
c18	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.62	0.60	0.09
c19	IVAS FL, 13.2 kbps, DTX off	180	2.66	1.09	0.16
c20	IVAS FL, 16.4 kbps, DTX off	180	3.23	1.11	0.16
c21	IVAS FL, 24.4 kbps, DTX off	180	3.73	0.94	0.14
c22	IVAS FL, 32 kbps, DTX off	180	3.82	0.97	0.14
c23	IVAS FL, 48 kbps, DTX off	180	4.28	0.83	0.12
c24	IVAS FL, 64 kbps, DTX off	180	4.48	0.66	0.10
c25	IVAS FL, 80 kbps, DTX off	180	4.54	0.68	0.10
c26	IVAS FL, 96 kbps, DTX off	180	4.61	0.57	0.08
c27	IVAS FL, 128 kbps, DTX off	180	4.62	0.61	0.09
c28	IVAS FX, 13.2 kbps, DTX off	180	2.60	1.09	0.16
c29	IVAS FX, 16.4 kbps, DTX off	180	3.19	1.08	0.16
c30	IVAS FX, 24.4 kbps, DTX off	180	3.65	0.99	0.14
c31	IVAS FX, 32 kbps, DTX off	180	3.78	1.02	0.15
c32	IVAS FX, 48 kbps, DTX off	180	4.33	0.72	0.11
c33	IVAS FX, 64 kbps, DTX off	180	4.45	0.63	0.09
c34	IVAS FX, 80 kbps, DTX off	180	4.59	0.60	0.09
c35	IVAS FX, 96 kbps, DTX off	180	4.54	0.68	0.10
c36	IVAS FX, 128 kbps, DTX off	180	4.59	0.58	0.09

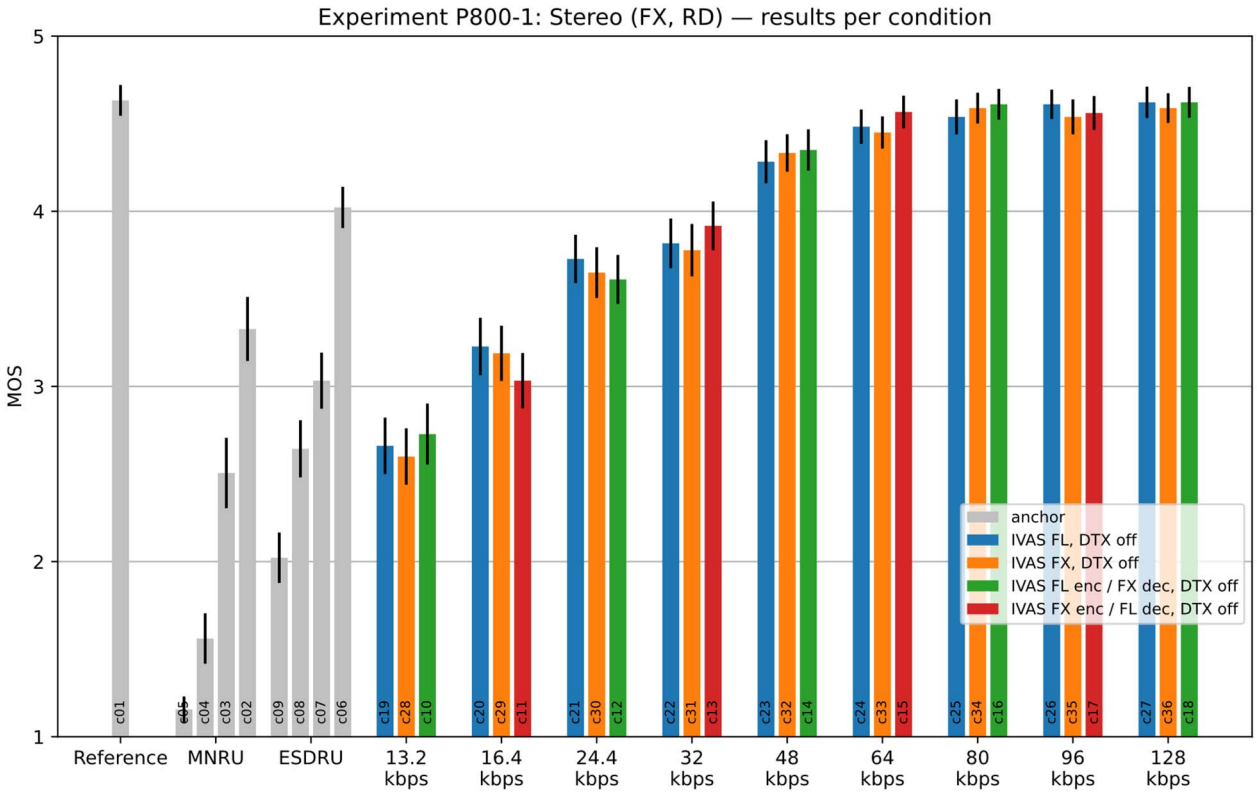


Figure 9.2-6: results per condition for experiment P800-1

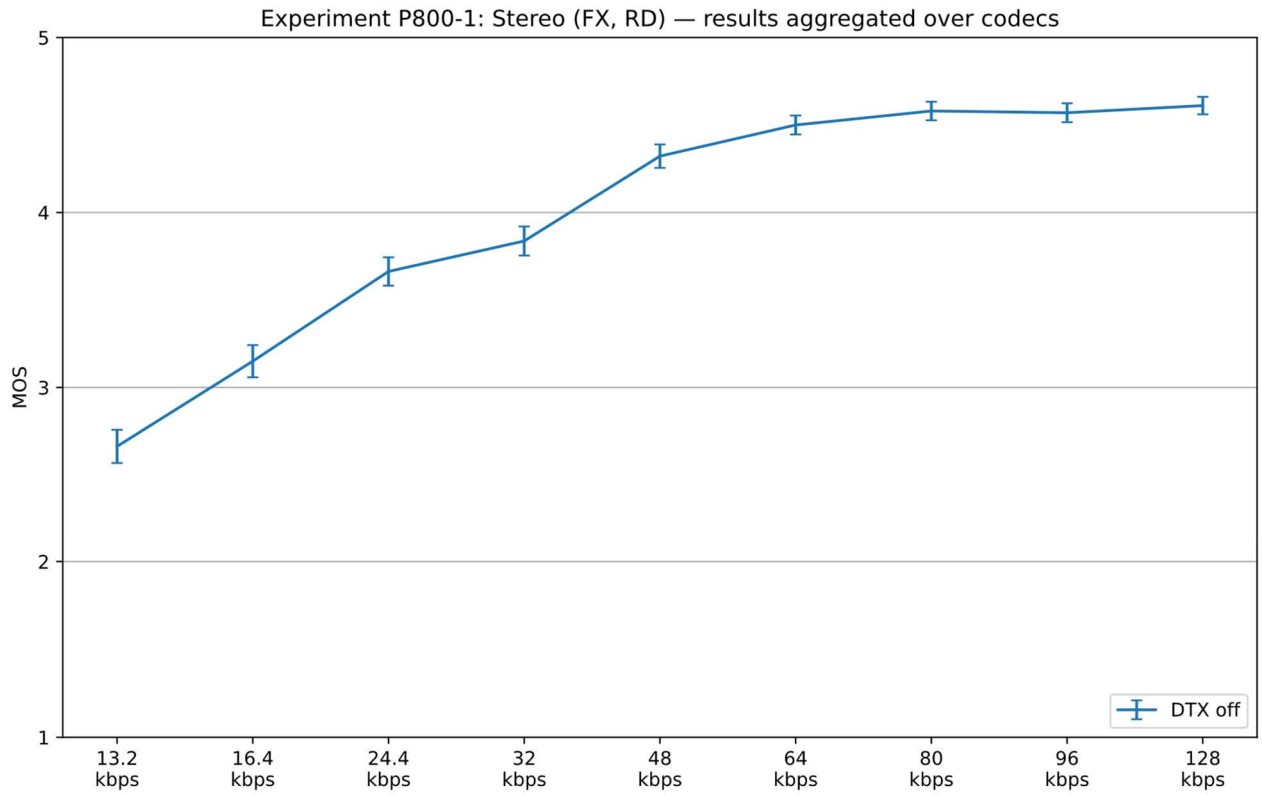


Figure 9.2-7: results aggregated over FX/FL for experiment P800-1

Table 9.2-12: statistical significance analysis for experiment P800-1

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	2.73	1.18	c19	2.66	1.09	0.07	0.558	0.288	NWT
c11	3.03	1.07	c20	3.23	1.11	-0.20	<0.001	0.954	WT
c12	3.61	0.95	c21	3.73	0.94	-0.12	<0.001	0.879	NWT
c13	3.92	0.94	c22	3.82	0.97	0.10	0.993	0.161	NWT
c14	4.35	0.80	c23	4.28	0.83	0.07	0.777	0.219	NWT
c15	4.57	0.64	c24	4.48	0.66	0.08	1.227	0.110	NWT
c16	4.61	0.60	c25	4.54	0.68	0.07	1.064	0.144	NWT
c17	4.56	0.66	c26	4.61	0.57	-0.05	<0.001	0.778	NWT
c18	4.62	0.60	c27	4.62	0.61	0.00	<0.001	0.500	NWT
c28	2.60	1.09	c19	2.66	1.09	-0.06	<0.001	0.702	NWT
c29	3.19	1.08	c20	3.23	1.11	-0.04	<0.001	0.632	NWT
c30	3.65	0.99	c21	3.73	0.94	-0.08	<0.001	0.778	NWT
c31	3.78	1.02	c22	3.82	0.97	-0.04	<0.001	0.645	NWT
c32	4.33	0.72	c23	4.28	0.83	0.05	0.607	0.272	NWT
c33	4.45	0.63	c24	4.48	0.66	-0.03	<0.001	0.686	NWT
c34	4.59	0.60	c25	4.54	0.68	0.05	0.742	0.229	NWT
c35	4.54	0.68	c26	4.61	0.57	-0.07	<0.001	0.861	NWT
c36	4.59	0.58	c27	4.62	0.61	-0.03	<0.001	0.701	NWT

9.2.8 Characterization Experiment P800-2 (Stereo, Speech and Mixed/Music, Impaired Channel, DTX on, Headphone Presentation)

Characterization Experiment P800-2 evaluates the performance of the IVAS fixed-point and floating-point implementation for Stereo speech and mixed/music, DTX on, under impaird channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-2 was conducted by NTT (language: JAP).

The aggregated results per condition are listed in Table 9.2-13 and are shown in Figure 9.2-8. The results aggregated over FX/FL are shown in Figure 9.2-9. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.2-14. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.2-13: results per condition for experiment P800-2

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.85	0.40	0.06
c02	MNRU Q = 27 dB	180	3.96	1.01	0.15
c03	MNRU Q = 22 dB	180	3.02	1.22	0.18
c04	MNRU Q = 17 dB	180	1.91	0.99	0.14
c05	MNRU Q = 12 dB	180	1.24	0.52	0.08
c06	ESDRU α=0.7	180	4.44	0.75	0.11
c07	ESDRU α=0.5	180	3.16	1.13	0.17
c08	ESDRU α=0.3	180	3.02	1.10	0.16
c09	ESDRU α=0.1	180	2.28	0.95	0.14
c10	IVAS FL enc / FX dec, 13.2 kbps, FER 5%, DTX on	180	2.61	1.03	0.15
c11	IVAS FX enc / FL dec, 16.4 kbps, FER 5%, DTX on	180	3.04	0.93	0.14
c12	IVAS FL enc / FX dec, 24.4 kbps, FER 5%, DTX on	180	3.33	0.93	0.14
c13	IVAS FX enc / FL dec, 32 kbps, FER 5%, DTX on	180	3.66	0.77	0.11
c14	IVAS FL enc / FX dec, 48 kbps, FER 5%, DTX on	180	3.76	0.84	0.12
c15	IVAS FX enc / FL dec, 64 kbps, FER 5%, DTX on	180	3.89	0.78	0.12
c16	IVAS FL enc / FX dec, 80 kbps, FER 5%, DTX on	180	3.89	0.80	0.12
c17	IVAS FX enc / FL dec, 96 kbps, FER 5%, DTX on	180	3.91	0.77	0.11
c18	IVAS FL enc / FX dec, 128 kbps, FER 5%, DTX on	180	3.82	0.76	0.11
c19	IVAS FL, 13.2 kbps, FER 5%, DTX on	180	2.64	1.01	0.15
c20	IVAS FL, 16.4 kbps, FER 5%, DTX on	180	3.08	1.03	0.15
c21	IVAS FL, 24.4 kbps, FER 5%, DTX on	180	3.38	1.00	0.15
c22	IVAS FL, 32 kbps, FER 5%, DTX on	180	3.68	0.84	0.12
c23	IVAS FL, 48 kbps, FER 5%, DTX on	180	3.77	0.83	0.12
c24	IVAS FL, 64 kbps, FER 5%, DTX on	180	3.84	0.78	0.11
c25	IVAS FL, 80 kbps, FER 5%, DTX on	180	3.97	0.75	0.11
c26	IVAS FL, 96 kbps, FER 5%, DTX on	180	3.94	0.74	0.11
c27	IVAS FL, 128 kbps, FER 5%, DTX on	180	3.92	0.78	0.12
c28	IVAS FX, 13.2 kbps, FER 5%, DTX on	180	2.62	1.02	0.15
c29	IVAS FX, 16.4 kbps, FER 5%, DTX on	180	3.12	0.95	0.14
c30	IVAS FX, 24.4 kbps, FER 5%, DTX on	180	3.33	0.99	0.14
c31	IVAS FX, 32 kbps, FER 5%, DTX on	180	3.58	0.84	0.12
c32	IVAS FX, 48 kbps, FER 5%, DTX on	180	3.78	0.76	0.11
c33	IVAS FX, 64 kbps, FER 5%, DTX on	180	3.82	0.76	0.11
c34	IVAS FX, 80 kbps, FER 5%, DTX on	180	3.91	0.76	0.11
c35	IVAS FX, 96 kbps, FER 5%, DTX on	180	3.88	0.75	0.11
c36	IVAS FX, 128 kbps, FER 5%, DTX on	180	3.87	0.81	0.12

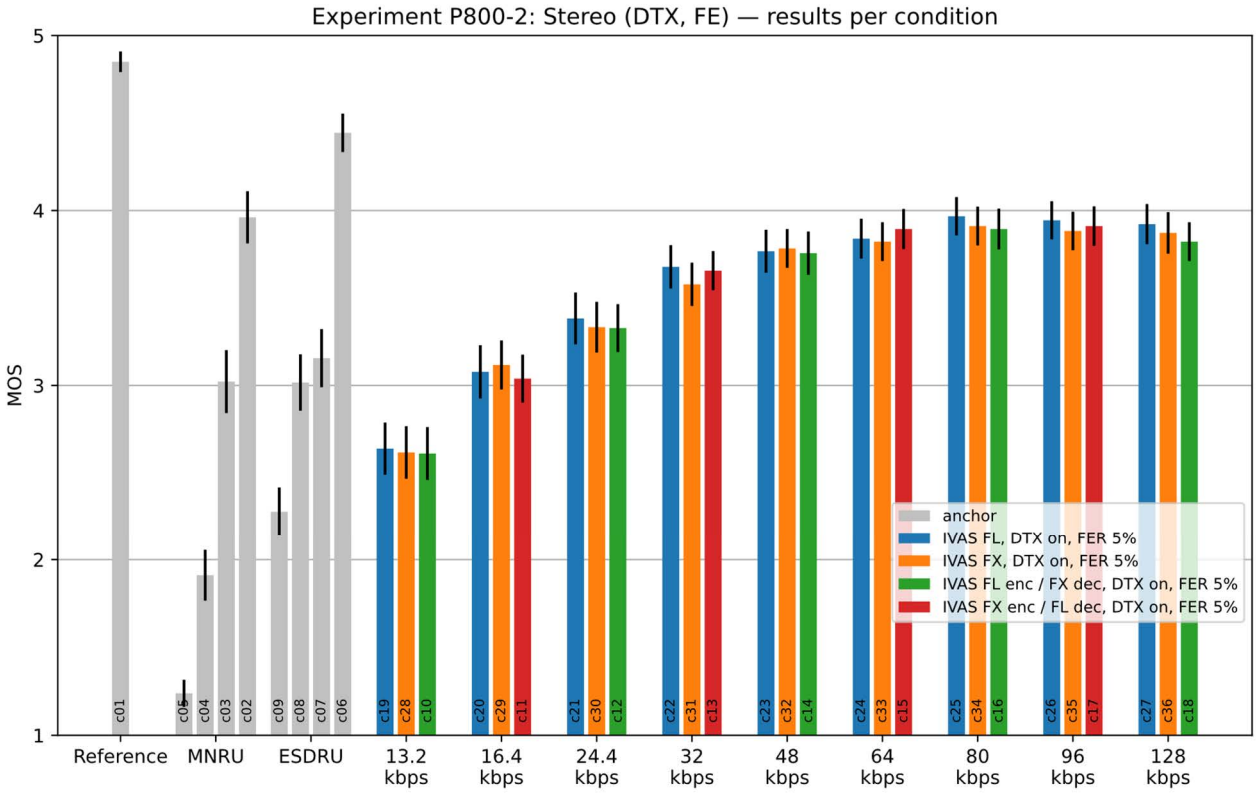


Figure 9.2-8: results per condition for experiment P800-2

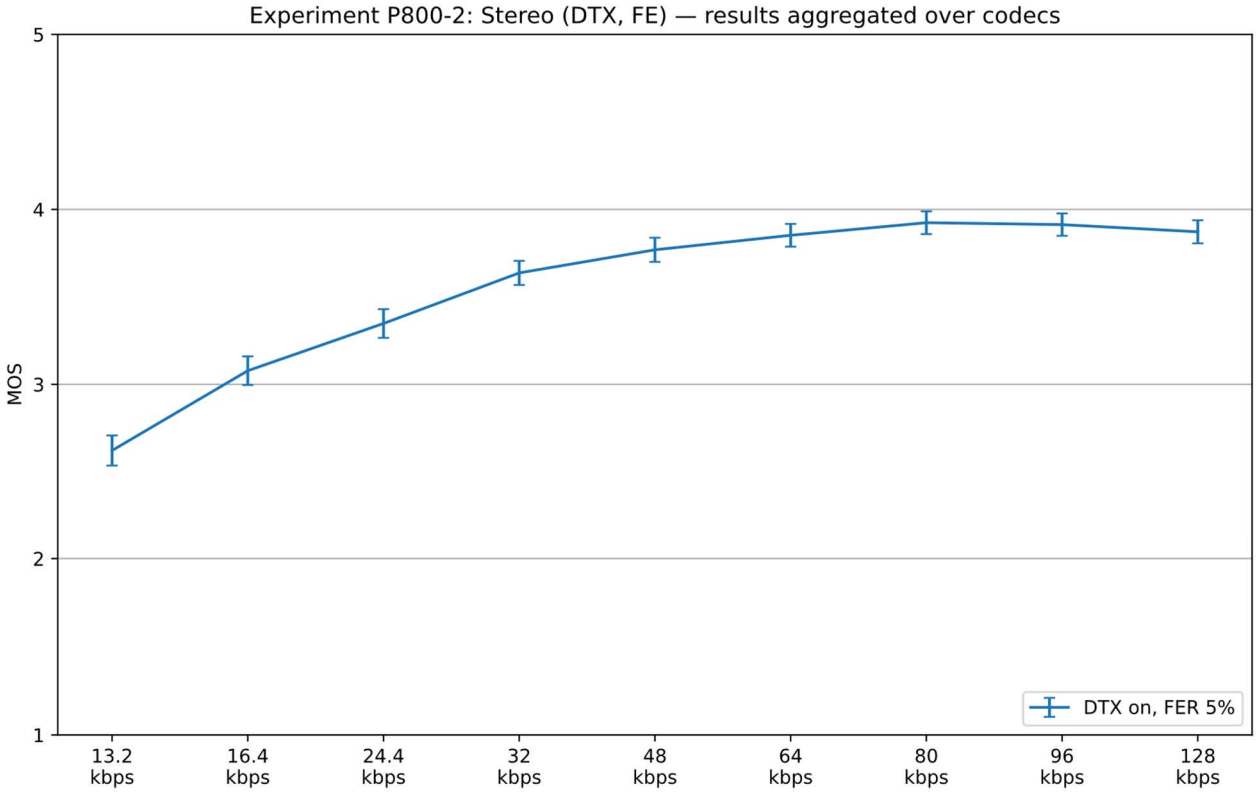


Figure 9.2-9: results aggregated over FX/FL for experiment P800-2

Table 9.2-14: statistical significance analysis for experiment P800-2

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	2.61	1.03	c19	2.64	1.01	-0.03	<0.001	0.603	NWT
c11	3.04	0.93	c20	3.08	1.03	-0.04	<0.001	0.647	NWT
c12	3.33	0.93	c21	3.38	1.00	-0.06	<0.001	0.705	NWT
c13	3.66	0.77	c22	3.68	0.84	-0.02	<0.001	0.602	NWT
c14	3.76	0.84	c23	3.77	0.83	-0.01	<0.001	0.550	NWT
c15	3.89	0.78	c24	3.84	0.78	0.06	0.670	0.252	NWT
c16	3.89	0.80	c25	3.97	0.75	-0.07	<0.001	0.815	NWT
c17	3.91	0.77	c26	3.94	0.74	-0.03	<0.001	0.661	NWT
c18	3.82	0.76	c27	3.92	0.78	-0.10	<0.001	0.891	NWT
c28	2.62	1.02	c19	2.64	1.01	-0.02	<0.001	0.581	NWT
c29	3.12	0.95	c20	3.08	1.03	0.04	0.372	0.355	NWT
c30	3.33	0.99	c21	3.38	1.00	-0.05	<0.001	0.683	NWT
c31	3.58	0.84	c22	3.68	0.84	-0.10	<0.001	0.869	NWT
c32	3.78	0.76	c23	3.77	0.83	0.02	0.191	0.424	NWT
c33	3.82	0.76	c24	3.84	0.78	-0.02	<0.001	0.583	NWT
c34	3.91	0.76	c25	3.97	0.75	-0.06	<0.001	0.760	NWT
c35	3.88	0.75	c26	3.94	0.74	-0.06	<0.001	0.782	NWT
c36	3.87	0.81	c27	3.92	0.78	-0.05	<0.001	0.724	NWT

9.2.9 Characterization Experiment P800-21 (Stereo, Speech and Mixed/Music, JBM, DTX on, Headphone Presentation)

Characterization Experiment P800-21 evaluates the performance of the IVAS JBM solution for fixed-point and floating-point implementation for Stereo speech and mixed/music, DTX on, under impaird channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-21 was conducted by Mesaqin.com (language: CZ).

The aggregated results per condition are listed in Table 9.2-15 and are shown in Figure 9.2-10. The results aggregated over FX/FL are shown in Figure 9.2-11. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.2-16. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.2-15: results per condition for experiment P800-21

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.77	0.43	0.06
c02	MNRU Q = 27 dB	180	3.92	0.99	0.15
c03	MNRU Q = 22 dB	180	3.16	1.16	0.17
c04	MNRU Q = 17 dB	180	2.07	1.05	0.15
c05	MNRU Q = 12 dB	180	1.28	0.56	0.08
c06	ESDRU α=0.7	180	4.54	0.70	0.10
c07	ESDRU α=0.5	180	3.95	1.05	0.15
c08	ESDRU α=0.3	180	3.73	1.13	0.17
c09	ESDRU α=0.1	180	2.97	1.17	0.17
c10	IVAS FL enc / FX dec, 24.4 kbps, No error, DTX on	180	4.32	0.74	0.11
c11	IVAS FL enc / FX dec, 48 kbps, No error, DTX on	180	4.59	0.59	0.09
c12	IVAS FL enc / FX dec, 96 kbps, No error, DTX on	180	4.75	0.48	0.07
c13	IVAS FL enc / FX dec, 24.4 kbps, Error I1. O1, DTX on	180	3.27	1.02	0.15
c14	IVAS FL enc / FX dec, 48 kbps, Error I1. O1, DTX on	180	3.42	1.05	0.15
c15	IVAS FL enc / FX dec, 24.4 kbps, Error I1. O2, DTX on	180	4.03	1.04	0.15
c16	IVAS FL enc / FX dec, 48 kbps, Error I1. O2, DTX on	180	4.30	0.89	0.13
c17	IVAS FL enc / FX dec, 24.4 kbps, I1. O1, DTX on	180	3.36	1.15	0.17
c18	IVAS FL enc / FX dec, 48 kbps, I1. O1, DTX on	180	3.58	0.93	0.14
c19	IVAS FL enc / FX dec, 96 kbps, I1. O1, DTX on	180	3.71	0.96	0.14

c20	IVAS FL enc / FX dec, 24.4 kbps, I1. O2, DTX on	180	3.84	1.01	0.15
c21	IVAS FL enc / FX dec, 48 kbps, I1. O2, DTX on	180	4.17	1.02	0.15
c22	IVAS FL enc / FX dec, 96 kbps, I1. O2, DTX on	180	4.24	1.04	0.15
c23	IVAS FL enc / FL dec, 24.4 kbps, No error, DTX on	180	4.28	0.82	0.12
c24	IVAS FL enc / FL dec, 48 kbps, No error, DTX on	180	4.60	0.59	0.09
c25	IVAS FL enc / FL dec, 96 kbps, No error, DTX on	180	4.67	0.53	0.08
c26	IVAS FL enc / FL dec, 24.4 kbps, Error I1. O1, DTX on	180	3.35	1.00	0.15
c27	IVAS FL enc / FL dec, 48 kbps, Error I1. O1, DTX on	180	3.51	0.99	0.14
c28	IVAS FL enc / FL dec, 24.4 kbps, Error I1. O2, DTX on	180	3.97	0.98	0.14
c29	IVAS FL enc / FL dec, 48 kbps, Error I1. O2, DTX on	180	4.32	0.94	0.14
c30	IVAS FL enc / FL dec, 96 kbps, Error I1. O2, DTX on	180	4.47	0.85	0.13
c31	IVAS FL enc / FL dec, 24.4 kbps, I1. O1, DTX on	180	3.38	1.03	0.15
c32	IVAS FL enc / FL dec, 48 kbps, I1. O1, DTX on	180	3.56	0.96	0.14
c33	IVAS FL enc / FL dec, 96 kbps, I1. O1, DTX on	180	3.68	0.92	0.14
c34	IVAS FL enc / FL dec, 24.4 kbps, I1. O2, DTX on	180	3.89	1.00	0.15
c35	IVAS FL enc / FL dec, 48 kbps, I1. O2, DTX on	180	4.15	1.05	0.15
c36	IVAS FL enc / FL dec, 96 kbps, I1. O2, DTX on	180	4.22	1.01	0.15

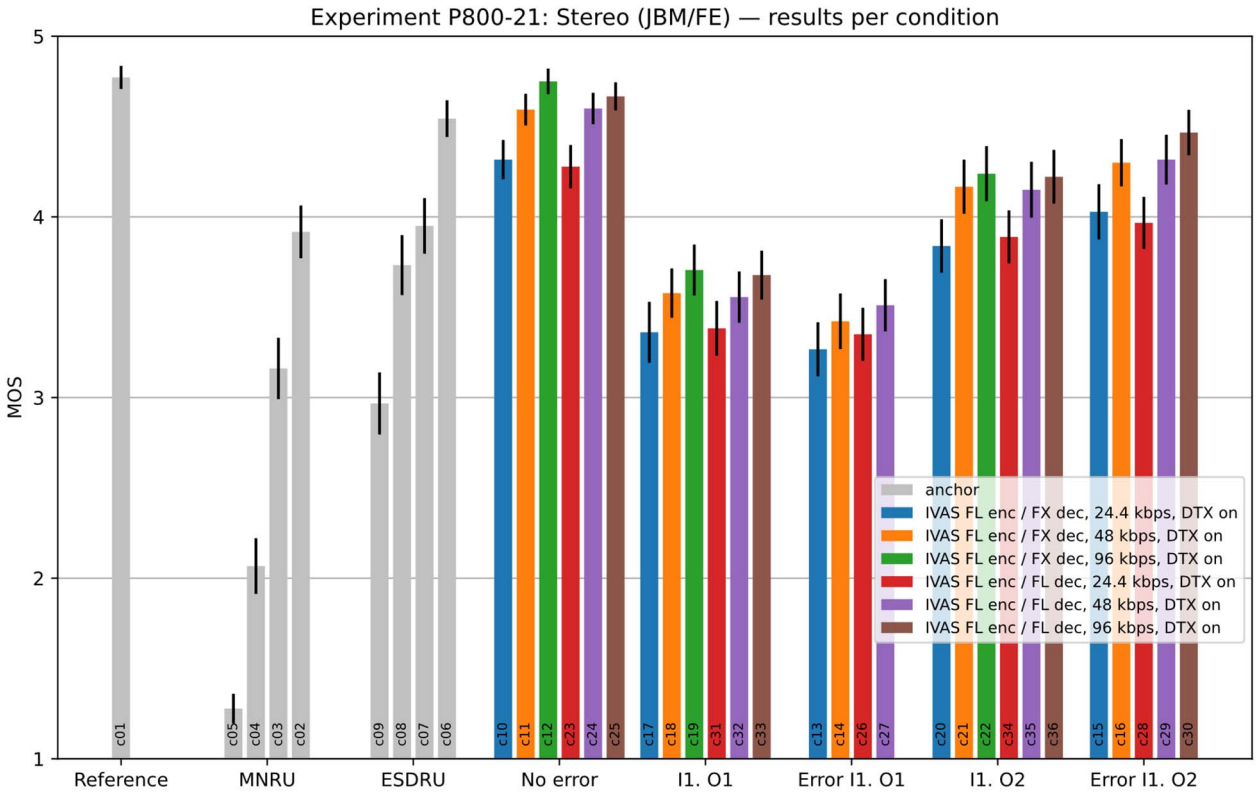


Figure 9.2-10: results per condition for experiment P800-21

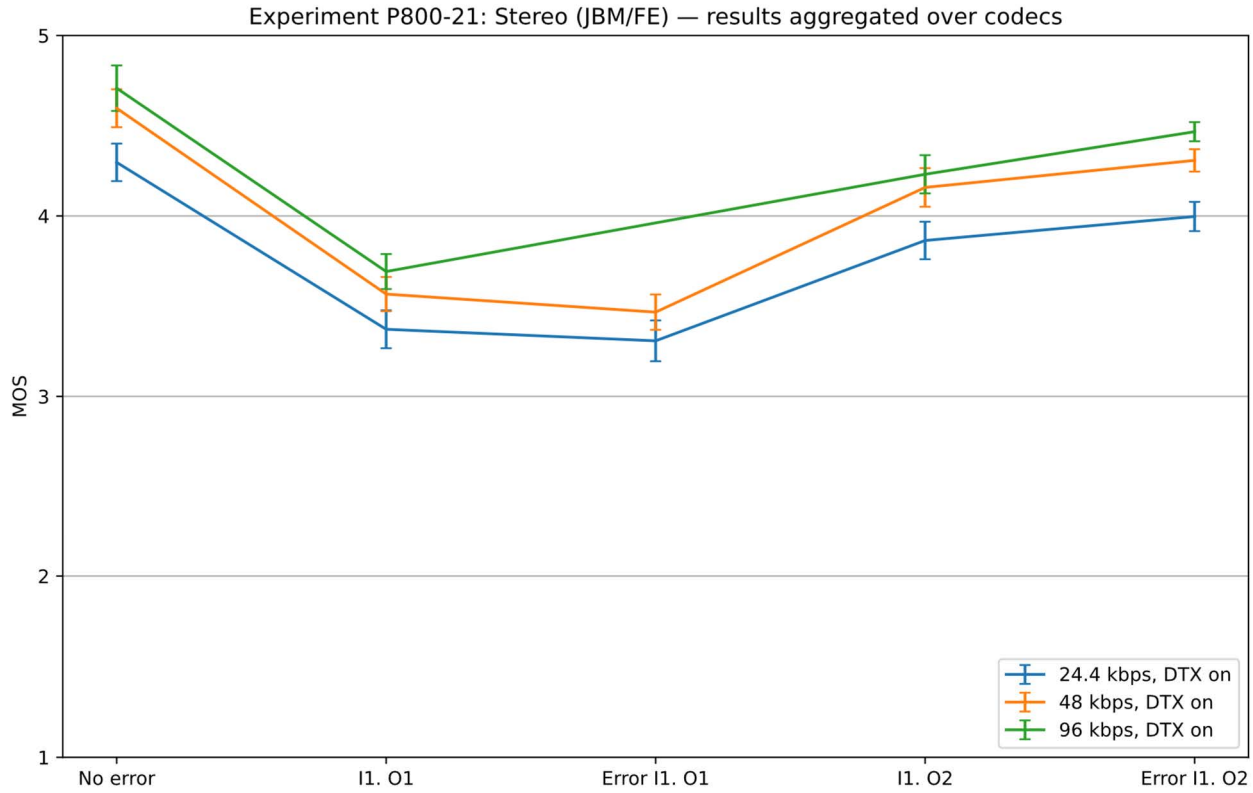


Figure 9.2-11: results aggregated over FX/FL for experiment P800-21

Table 9.2-16: statistical significance analysis for experiment P800-21

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
No terms of references mandated for experiment .									

9.2.10 Characterization Experiment BS1534-1 (Stereo, Generic Audio, 16.4 – 48 kbps, Headphone Presentation)

Characterization Experiment BS1534-1 evaluates the performance of the IVAS fixed-point implementation for Stereo generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-1 was conducted by FORCE Technology.

The aggregated results per condition are listed in Table 9.2-17 and are shown in Figure 9.2-12. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.2-18. Scores with result WT are highlighted in red.

Table 9.2-17: results per condition for experiment BS1534-1

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.67	2.17	0.33
c02	LP 3.5 kHz	168	7.69	12.34	1.88
c03	Mono EVS, 16.4 kbps	168	33.09	19.83	3.02
c04	Mono EVS, 32 kbps	168	41.30	20.94	3.19
c05	IVAS, 16.4 kbps	168	50.09	22.54	3.43
c06	IVAS, 24.4 kbps	168	65.46	20.24	3.08
c07	IVAS, 32 kbps	168	72.36	19.59	2.98
c08	IVAS, 48 kbps	168	88.07	12.01	1.83

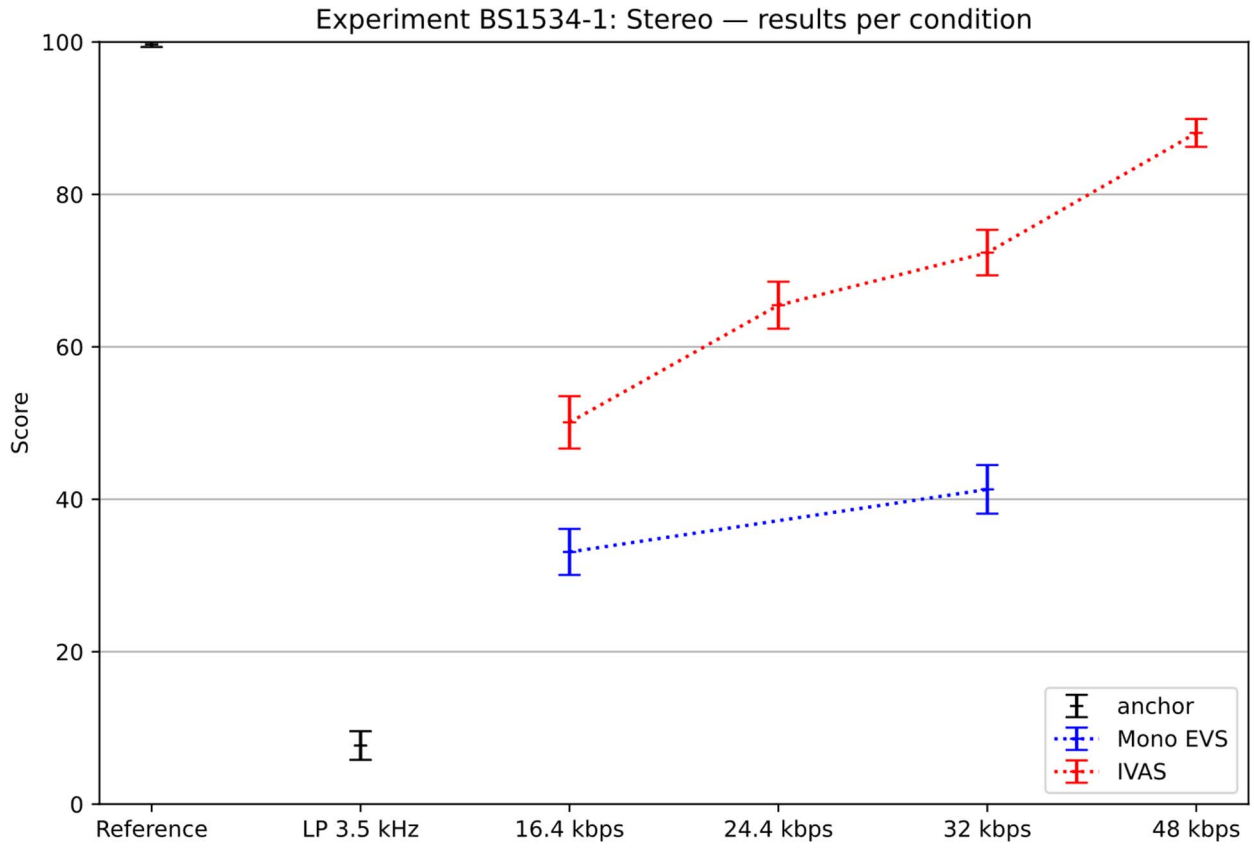


Figure 9.2-12: results per condition for experiment BS1534-1

Table 9.2-18: statistical significance analysis for experiment BS1534-1

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	50.09	22.54	c03	33.09	19.83	17.00	10.085	<0.001	BT
c07	72.36	19.59	c04	41.30	20.94	31.07	18.319	<0.001	BT

9.2.11 Characterization Experiment BS1534-2 (Stereo, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-2 evaluates the performance of the IVAS fixed-point implementation for Stereo generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-2 was conducted by Huawei Technologies Co Ltd.

The aggregated results per condition are listed in Table 9.2-19 and are shown in Figure 9.2-13. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.2-20. Scores with result WT are highlighted in red.

Table 9.2-19: results per condition for experiment BS1534-2

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	98.60	4.40	0.67
c02	LP 3.5 kHz	168	42.26	19.17	2.92
c03	Mono EVS, 64 kbps	168	75.81	15.30	2.33
c04	Mono EVS, 128 kbps	168	76.19	16.72	2.55
c05	IVAS, 64 kbps	168	95.26	8.50	1.29
c06	IVAS, 96 kbps	168	97.82	4.49	0.68
c07	IVAS, 128 kbps	168	97.19	5.67	0.86
c08	IVAS, 256 kbps	168	98.05	5.02	0.77

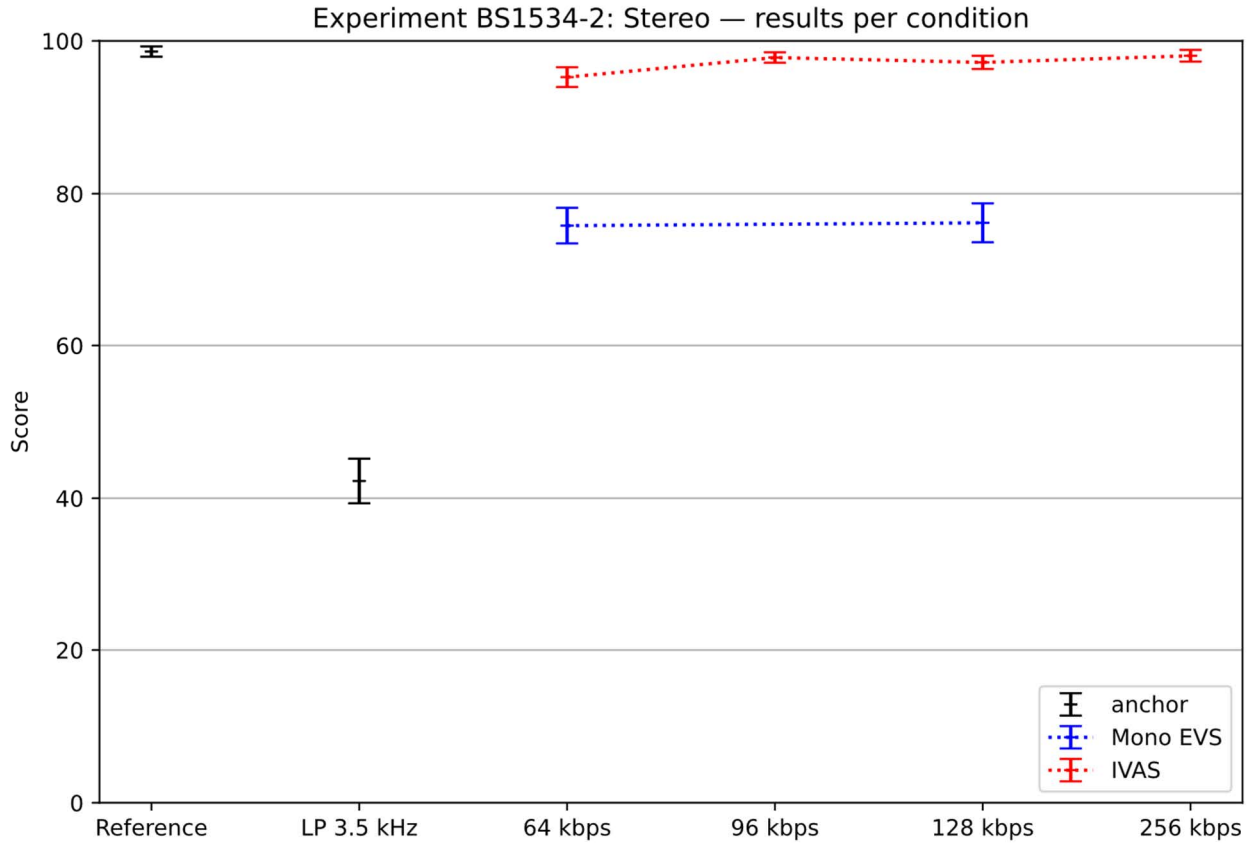


Figure 9.2-13: results per condition for experiment BS1534-2

Table 9.2-20: statistical significance analysis for experiment BS1534-2

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	95.26	8.50	c03	75.81	15.30	19.44	17.289	<0.001	BT
c07	97.19	5.67	c04	76.19	16.72	21.00	17.098	<0.001	BT

9.2.12 Conclusions on IVAS Performance for Stereo

Taking the results of the previous clauses in 9.2 into account, the following conclusions can be drawn for IVAS Stereo operating modes:

During selection phase, IVAS stereo coding was extensively compared against EVS dual-mono. For all tested content types (clean speech, noisy speech, mixed/music) and conditions incl. DTX and FER, IVAS has fulfilled the selection criterion: Better than EVS dual-mono at the same bitrate or not worse than EVS dual-mono at the next higher bitrate. Noisy speech and mixed/music as content is especially relevant for immersive communication. For clean speech, IVAS outperforms EVS dual-mono for bitrates up to and including 16.4 kbps. For noisy speech, IVAS outperforms EVS dual-mono for bitrates up to and including 24.4 kbps. For mixed/music and generic audio, IVAS significantly outperforms EVS dual-mono at the same bitrate for bitrates up to 64 kbps. For bitrates from 96 kbps and above, IVAS performance is equivalent to EVS dual-mono and reaches transparency.

During characterization phase, IVAS stereo was evaluated against EVS mono in a series of BS.1534 tests. Across all tested bitrates, IVAS significantly outperformed EVS mono at the same bitrate, with a mean improvement across subjects and items of up to 31 MUSHRA points (32 kbps); maintaining quality in the MUSHRA “good” range at starting from 24.4kbps and the “excellent” range starting from 48 kbps on general audio content

Additionally, the performance of the IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase. After adjustment of the significance level to account for multiple comparisons (Bonferroni method), all 36 tested operating points were statistically equivalent.

Overall, the IVAS stereo mode maintains a “good” MOS rating (>= 4.0) from bitrates starting from 24.4 kbps and above for clean speech, noisy speech and music/mixed content. A level 1 IVAS encoder already reaches an “excellent” MOS rating of approx. 4.5.

9.3 Scene-based Audio (SBA, Ambisonics)

9.3.1 Overview

In Selection phase, six experiments have been conducted to evaluate the performance of the IVAS codec with Ambisonics content. While the experiments P800-4 and P800-5 were conducted as P.800 DCR tests on FOA content, the experiments BS1534-4a, BS1534-4b, BS1534-5a and BS1534-5b were conducted as BS.1534 tests for different Ambisonics orders. The experiments P800-4, P800-5, BS1534-4a, BS1534-4b, BS1534-5a were conducted using headphone presentation, whereas experiment BS1534-5b was conducted using 7.1+4 loudspeaker presentation.

- Selection Experiment P800-4: Clean speech, FOA under clean and impaired channel conditions, headphone presentation
- Selection Experiment P800-5: Speech + background, FOA under clean channel conditions, DTX off and on, headphone presentation
- Selection Experiment BS1534-4a: Generic audio, FOA input, HOA3 output, 96 kbps, 128 kbps and 160 kbps, headphone presentation
- Selection Experiment BS1534-4b: Generic audio, HOA2 input, HOA3 output, 160 and 192 kbps, headphone presentation
- Selection Experiment BS1534-5a: Generic audio, HOA3 input, HOA3 output, 192 and 256 kbps, headphone presentation
- Selection Experiment BS1534-5b: Generic audio, HOA3 input, HOA3 output, 384 and 512 kbps, 7.1+4 loudspeaker presentation

In Characterization phase, seven additional experiments have been conducted. While the experiments P800-3, P800-4, P800-5 and P800-23 were conducted as P.800 DCR tests, the experiments BS1534-3, BS1534-4 and BS1534-5 were conducted as BS.1534 tests.

- Characterization Experiment P800-3: Speech + background and mixed/music, FOA under clean channel conditions, headphone presentation

- Characterization Experiment P800-4: Speech + background and mixed/music, HOA2 under impaired channel conditions, DTX off and on, headphone presentation
- Characterization Experiment P800-5: Speech + background and mixed/music, HOA3 under clean channel conditions, headphone presentation
- Characterization Experiment P800-23: Speech + background and mixed/music, FOA under impaired channel conditions using JBM, headphone presentation
- Characterization Experiment BS1534-3: Generic audio, FOA input, binaural output, 16.4 – 48 kbps, headphone presentation
- Characterization Experiment BS1534-4: Generic audio, FOA input, binaural output, 64 – 256 kbps, headphone presentation
- Characterization Experiment BS1534-5: Generic audio, HOA3 input, 7.1+4 output, 64 – 256 kbps, loudspeaker presentation

9.3.2 Selection Experiment P800-4 (FOA, Clean Speech, Headphone Presentation)

Selection Experiment P800-4 evaluates IVAS for FOA clean speech under clean and impaired channel conditions using headphone presentation. See Annex C.4 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-4 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.3-1: Statistical overview on the results of P800-4

		Type	CuT						EVS Reference				Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#													
a	c24	1	16.4	off		NWT	3.13	1.06	c09	3x7.2	2.47	1.14	5.66	BT	EXCEED
	c25	1	24.4	off		NWT	3.49	0.95	c10	4x7.2	2.54	1.07	8.85	BT	EXCEED
	c26	1	32	off		NWT	3.74	0.98	c12	4x9.6	3.4	0.98	3.33	BT	EXCEED
		2	32	off		BT	3.74	0.98	c11	4x8	2.73	1.11	9.24	BT	PASS
	c27	1	48	off		NWT	4.31	0.76	c14	4x16.4	4.34	0.77	-0.41	NWT	PASS
		2	48	off		BT	4.31	0.76	c13	4x13.2	4.1	0.88	2.43	BT	PASS
	c28	1	64	off		NWT	4.61	0.59	c15	4x24.4	4.52	0.66	1.35	NWT	PASS
		2	64	off		BT	4.61	0.59	c14	4x16.4	4.34	0.77	3.68	BT	PASS
	c29	1	80	off		NWT	4.66	0.51	c15	4x24.4	4.52	0.66	2.17	BT	EXCEED
		2	80	off		BT	4.66	0.51	c14	4x16.4	4.34	0.77	4.52	BT	PASS
	c30	1	96	off		NWT	4.72	0.47	c16	4x32	4.52	0.64	3.46	BT	EXCEED
		2	96	off		BT	4.72	0.47	c15	4x24.4	4.52	0.66	3.32	BT	PASS
	c31	1	24.4	off	5%	NWT	3.07	1.08	c17	4x7.2	2.27	1.05	7.07	BT	EXCEED
	c32	1	32	off	5%	NWT	3.31	0.97	c19	4x9.6	2.93	1.09	3.53	BT	EXCEED
		2	32	off	5%	BT	3.31	0.97	c18	4x8	2.38	1.14	8.34	BT	PASS
	c33	1	48	off	5%	NWT	3.81	0.95	c21	4x16.4	3.74	0.88	0.64	NWT	PASS
		2	48	off	5%	BT	3.81	0.95	c20	4x13.2	3.52	0.99	2.83	BT	PASS
	c34	1	64	off	5%	NWT	4.1	0.96	c22	4x24.4	3.96	0.87	1.44	NWT	PASS
		2	64	off	5%	BT	4.1	0.96	c21	4x16.4	3.74	0.88	3.66	BT	PASS
	c35	1	80	off	5%	NWT	4.15	0.91	c22	4x24.4	3.96	0.87	2.01	BT	EXCEED
		2	80	off	5%	BT	4.15	0.91	c21	4x16.4	3.74	0.88	4.3	BT	PASS
	c36	1	96	off	5%	NWT	4.02	0.92	c23	4x32	3.85	0.87	1.83	BT	EXCEED
		2	96	off	5%	BT	4.02	0.92	c22	4x24.4	3.96	0.87	0.65	NWT	FAIL
c	c24	1	16.4	off		NWT	3.53	0.92	c09	3x7.2	2.44	1	10.82	BT	EXCEED
	c25	1	24.4	off		NWT	3.68	0.86	c10	4x7.2	2.44	0.89	13.45	BT	EXCEED
	c26	1	32	off		NWT	3.77	0.83	c12	4x9.6	3.51	0.89	2.88	BT	EXCEED
		2	32	off		BT	3.77	0.83	c11	4x8	2.59	0.96	12.44	BT	PASS
	c27	1	48	off		NWT	4.39	0.63	c14	4x16.4	4.29	0.67	1.46	NWT	PASS
		2	48	off		BT	4.39	0.63	c13	4x13.2	4.1	0.73	4.01	BT	PASS
	c28	1	64	off		NWT	4.61	0.54	c15	4x24.4	4.5	0.62	1.81	BT	EXCEED
		2	64	off		BT	4.61	0.54	c14	4x16.4	4.29	0.67	5	BT	PASS
	c29	1	80	off		NWT	4.64	0.57	c15	4x24.4	4.5	0.62	2.22	BT	EXCEED
		2	80	off		BT	4.64	0.57	c14	4x16.4	4.29	0.67	5.34	BT	PASS
	c30	1	96	off		NWT	4.66	0.51	c16	4x32	4.46	0.62	3.35	BT	EXCEED
		2	96	off		BT	4.66	0.51	c15	4x24.4	4.5	0.62	2.69	BT	PASS
	c31	1	24.4	off	5%	NWT	3.14	0.99	c17	4x7.2	2.18	0.96	9.34	BT	EXCEED
	c32	1	32	off	5%	NWT	3.51	0.89	c19	4x9.6	3.08	1.02	4.19	BT	EXCEED
		2	32	off	5%	BT	3.51	0.89	c18	4x8	2.26	0.92	13.1	BT	PASS
	c33	1	48	off	5%	NWT	3.94	0.88	c21	4x16.4	3.63	0.92	3.32	BT	EXCEED
		2	48	off	5%	BT	3.94	0.88	c20	4x13.2	3.49	0.97	4.65	BT	PASS
	c34	1	64	off	5%	NWT	4.13	0.76	c22	4x24.4	4.09	0.75	0.49	NWT	PASS
		2	64	off	5%	BT	4.13	0.76	c21	4x16.4	3.63	0.92	5.61	BT	PASS
	c35	1	80	off	5%	NWT	4.18	0.7	c22	4x24.4	4.09	0.75	1.23	NWT	PASS
		2	80	off	5%	BT	4.18	0.7	c21	4x16.4	3.63	0.92	6.44	BT	PASS
	c36	1	96	off	5%	NWT	4.14	0.8	c23	4x32	3.81	0.76	4.05	BT	EXCEED
		2	96	off	5%	BT	4.14	0.8	c22	4x24.4	4.09	0.75	0.61	NWT	FAIL

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-2: Summary of the results of P800-4

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c24	16.4	off		NWT c09	EXCEED
	c25	24.4	off		NWT c10	EXCEED
	c26	32	off		NWT c12 or BT c11	EXCEED
	c27	48	off		NWT c14 or BT c13	PASS
	c28	64	off		NWT c15 or BT c14	PASS
	c29	80	off		NWT c15 or BT c14	EXCEED
	c30	96	off		NWT c16 or BT c15	EXCEED
	c31	24.4	off	5%	NWT c17	EXCEED
	c32	32	off	5%	NWT c19 or BT c18	EXCEED
	c33	48	off	5%	NWT c21 or BT c20	PASS
	c34	64	off	5%	NWT c22 or BT c21	PASS
	c35	80	off	5%	NWT c22 or BT c21	EXCEED
	c36	96	off	5%	NWT c23 or BT c22	PASS
c	c24	16.4	off		NWT c09	EXCEED
	c25	24.4	off		NWT c10	EXCEED
	c26	32	off		NWT c12 or BT c11	EXCEED
	c27	48	off		NWT c14 or BT c13	PASS
	c28	64	off		NWT c15 or BT c14	EXCEED
	c29	80	off		NWT c15 or BT c14	EXCEED

	c30	96	off		NWT c16 or BT c15	EXCEED
	c31	24.4	off	5%	NWT c17	EXCEED
	c32	32	off	5%	NWT c19 or BT c18	EXCEED
	c33	48	off	5%	NWT c21 or BT c20	EXCEED
	c34	64	off	5%	NWT c22 or BT c21	PASS
	c35	80	off	5%	NWT c22 or BT c21	PASS
	c36	96	off	5%	NWT c23 or BT c22	PASS

The following diagrams show the results for a range of conditions from experiment P800-4 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c1 – c16 for EVS conditions and c24 – c30 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c17 – c23 for EVS conditions and c31 – c36 for IVAS conditions.

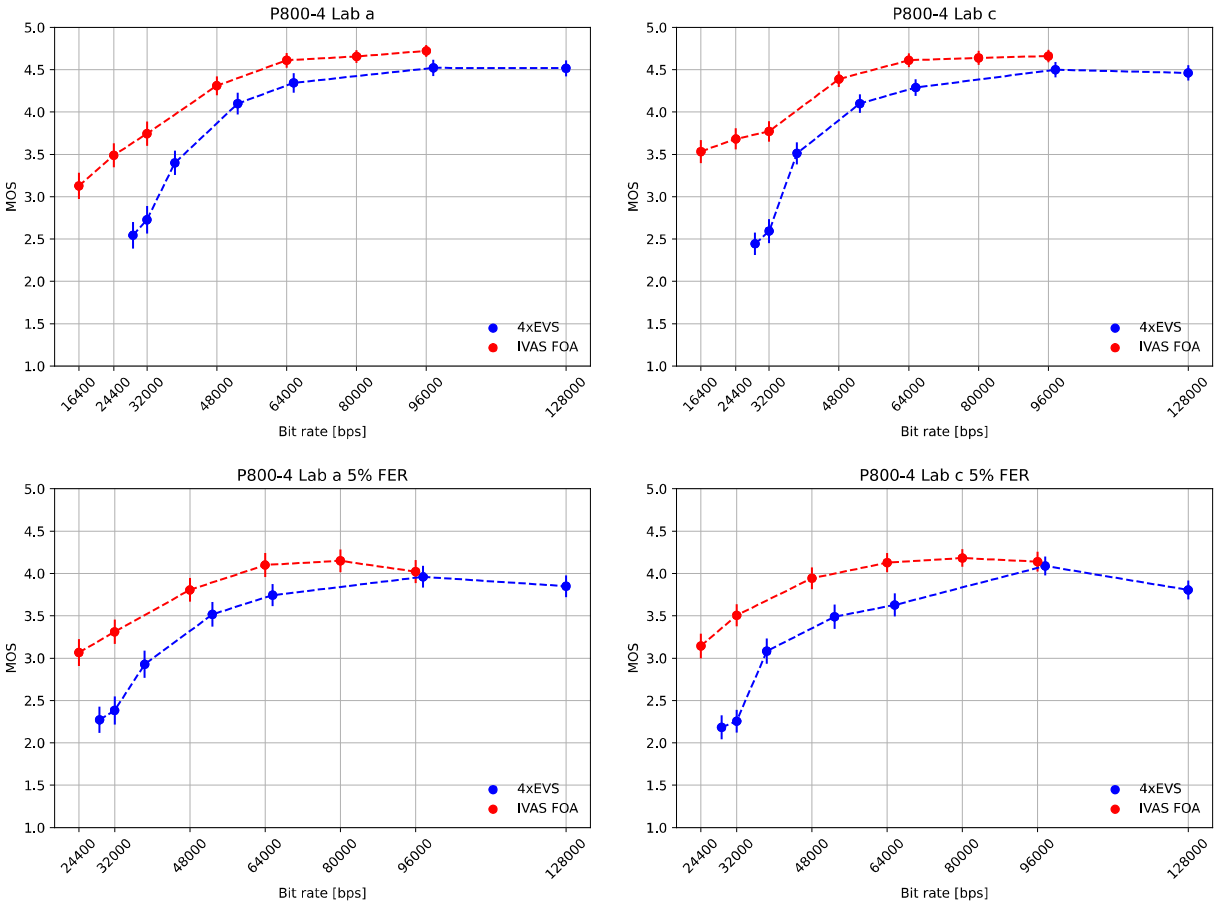


Figure 9.3-1: P800-4 (Ambisonics FOA input/output, clean speech, headphone presentation) rate distortion curves for clean and impaired channel conditions

9.3.3 Selection Experiment P800-5 (FOA, Speech+Background, Headphone Presentation)

Selection Experiment P800-5 evaluates IVAS for FOA speech + background under clean channel conditions, DTX off and on using headphone presentation. See Annex C.5 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-5 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.3-3: Statistical overview on the results of P800-5

		Type	CuT					EVS Reference				Evaluation		
		Value	Bitrate	DTX	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#												
a	c24	1	16.4	off	NWT	3.26	0.96	c09	3x7.2	2.93	1.04	3.06	BT	EXCEED
	c25	1	24.4	off	NWT	3.68	1	c10	4x7.2	3.04	1.08	5.82	BT	EXCEED
	c26	1	32	off	NWT	3.84	0.96	c12	4x9.6	3.74	0.88	1.08	NWT	PASS
		2	32	off	BT	3.84	0.96	c11	4x8	3.15	1.03	6.6	BT	PASS
	c27	1	48	off	NWT	4.45	0.66	c14	4x16.4	4.27	0.68	2.58	BT	EXCEED
		2	48	off	BT	4.45	0.66	c13	4x13.2	4	0.85	5.62	BT	PASS
	c28	1	64	off	NWT	4.54	0.59	c15	4x24.4	4.47	0.71	1.12	NWT	PASS
		2	64	off	BT	4.54	0.59	c14	4x16.4	4.27	0.68	4.12	BT	PASS
	c29	1	80	off	NWT	4.56	0.65	c15	4x24.4	4.47	0.71	1.31	NWT	PASS
		2	80	off	BT	4.56	0.65	c14	4x16.4	4.27	0.68	4.18	BT	PASS
	c30	1	96	off	NWT	4.54	0.65	c16	4x32	4.38	0.68	2.38	BT	EXCEED
		2	96	off	BT	4.54	0.65	c15	4x24.4	4.47	0.71	1.07	NWT	FAIL
	c31	1	16.4	on	NWT	3.26	1	c17	3x7.2	2.78	1.09	4.28	BT	EXCEED
	c32	1	24.4	on	NWT	3.49	0.98	c18	4x7.2	2.96	1.01	5.09	BT	EXCEED
	c33	1	32	on	NWT	3.8	0.88	c20	4x9.6	3.64	0.9	1.72	BT	EXCEED
		2	32	on	BT	3.8	0.88	c19	4x8	2.97	0.98	8.42	BT	PASS
	c34	1	48	on	NWT	4.26	0.7	c22	4x16.4	4.13	0.66	1.79	BT	EXCEED
		2	48	on	BT	4.26	0.7	c21	4x13.2	3.87	0.88	4.69	BT	PASS
	c35	1	64	on	NWT	4.25	0.76	c23	4x24.4	4.2	0.86	0.58	NWT	PASS
		2	64	on	BT	4.25	0.76	c22	4x16.4	4.13	0.66	1.56	NWT	FAIL
	c36	1	80	on	NWT	4.3	0.75	c23	4x24.4	4.2	0.86	1.18	NWT	PASS
		2	80	on	BT	4.3	0.75	c22	4x16.4	4.13	0.66	2.26	BT	PASS
b	c24	1	16.4	off	NWT	3.53	0.97	c09	3x7.2	2.82	0.99	6.91	BT	EXCEED
	c25	1	24.4	off	NWT	3.78	1.01	c10	4x7.2	2.93	1.04	7.88	BT	EXCEED
	c26	1	32	off	NWT	3.96	0.86	c12	4x9.6	3.81	0.99	1.53	NWT	PASS
		2	32	off	BT	3.96	0.86	c11	4x8	2.87	1.05	10.75	BT	PASS
	c27	1	48	off	NWT	4.61	0.57	c14	4x16.4	4.45	0.64	2.44	BT	EXCEED
		2	48	off	BT	4.61	0.57	c13	4x13.2	4.16	0.8	6.15	BT	PASS
	c28	1	64	off	NWT	4.77	0.46	c15	4x24.4	4.63	0.55	2.6	BT	EXCEED
		2	64	off	BT	4.77	0.46	c14	4x16.4	4.45	0.64	5.4	BT	PASS
	c29	1	80	off	NWT	4.71	0.53	c15	4x24.4	4.63	0.55	1.45	NWT	PASS
		2	80	off	BT	4.71	0.53	c14	4x16.4	4.45	0.64	4.21	BT	PASS
	c30	1	96	off	NWT	4.75	0.45	c16	4x32	4.6	0.59	2.71	BT	EXCEED
		2	96	off	BT	4.75	0.45	c15	4x24.4	4.63	0.55	2.31	BT	PASS
	c31	1	16.4	on	NWT	3.52	1.04	c17	3x7.2	2.72	0.99	7.5	BT	EXCEED
	c32	1	24.4	on	NWT	3.76	0.94	c18	4x7.2	2.73	1.02	9.97	BT	EXCEED
	c33	1	32	on	NWT	3.79	0.99	c20	4x9.6	3.83	0.88	-0.34	NWT	PASS
		2	32	on	BT	3.79	0.99	c19	4x8	2.88	1.06	8.49	BT	PASS
	c34	1	48	on	NWT	4.52	0.58	c22	4x16.4	4.33	0.72	2.73	BT	EXCEED
		2	48	on	BT	4.52	0.58	c21	4x13.2	4.02	0.84	6.65	BT	PASS
	c35	1	64	on	NWT	4.63	0.53	c23	4x24.4	4.51	0.68	1.9	BT	EXCEED
		2	64	on	BT	4.63	0.53	c22	4x16.4	4.33	0.72	4.41	BT	PASS
	c36	1	80	on	NWT	4.73	0.47	c23	4x24.4	4.51	0.68	3.69	BT	EXCEED
		2	80	on	BT	4.73	0.47	c22	4x16.4	4.33	0.72	6.22	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-4: Summary of the results of P800-5

Lab	Cond.	Bitrate	DTX	ToR	Status
a	c24	16.4	Off	NWT c09	EXCEED
	c25	24.4	Off	NWT c10	EXCEED
	c26	32	Off	NWT c12 or BT c11	PASS
	c27	48	Off	NWT c14 or BT c13	EXCEED
	c28	64	Off	NWT c15 or BT c14	PASS
	c29	80	off	NWT c15 or BT c14	PASS
	c30	96	off	NWT c16 or BT c15	PASS
	c31	16.4	on	NWT c17	EXCEED
	c32	24.4	on	NWT c18	EXCEED
	c33	32	on	NWT c20 or BT c19	EXCEED
	c34	48	on	NWT c22 or BT c21	EXCEED
	c35	64	on	NWT c23 or BT c22	PASS
	c36	80	on	NWT c23 or BT c22	PASS
b	c24	16.4	off	NWT c09	EXCEED
	c25	24.4	off	NWT c10	EXCEED
	c26	32	off	NWT c12 or BT c11	PASS
	c27	48	off	NWT c14 or BT c13	EXCEED
	c28	64	off	NWT c15 or BT c14	EXCEED
	c29	80	off	NWT c15 or BT c14	PASS
	c30	96	off	NWT c16 or BT c15	EXCEED
	c31	16.4	on	NWT c17	EXCEED

	c32	24.4	on	NWT c18	EXCEED
	c33	32	on	NWT c20 or BT c19	PASS
	c34	48	on	NWT c22 or BT c21	EXCEED
	c35	64	on	NWT c23 or BT c22	EXCEED
	c36	80	on	NWT c23 or BT c22	EXCEED

The following diagrams show the results for a range of conditions from experiment P800-5 as rate-distortion curves. The first two diagrams only show results for active coding conditions (DTX off), i.e. conditions c10 – c16 for EVS conditions and c24 – c30 for IVAS conditions. The second two diagrams show results for DTX conditions (DTX on), i.e. conditions c18 – c23 for EVS conditions and c31 – c36 for IVAS conditions.

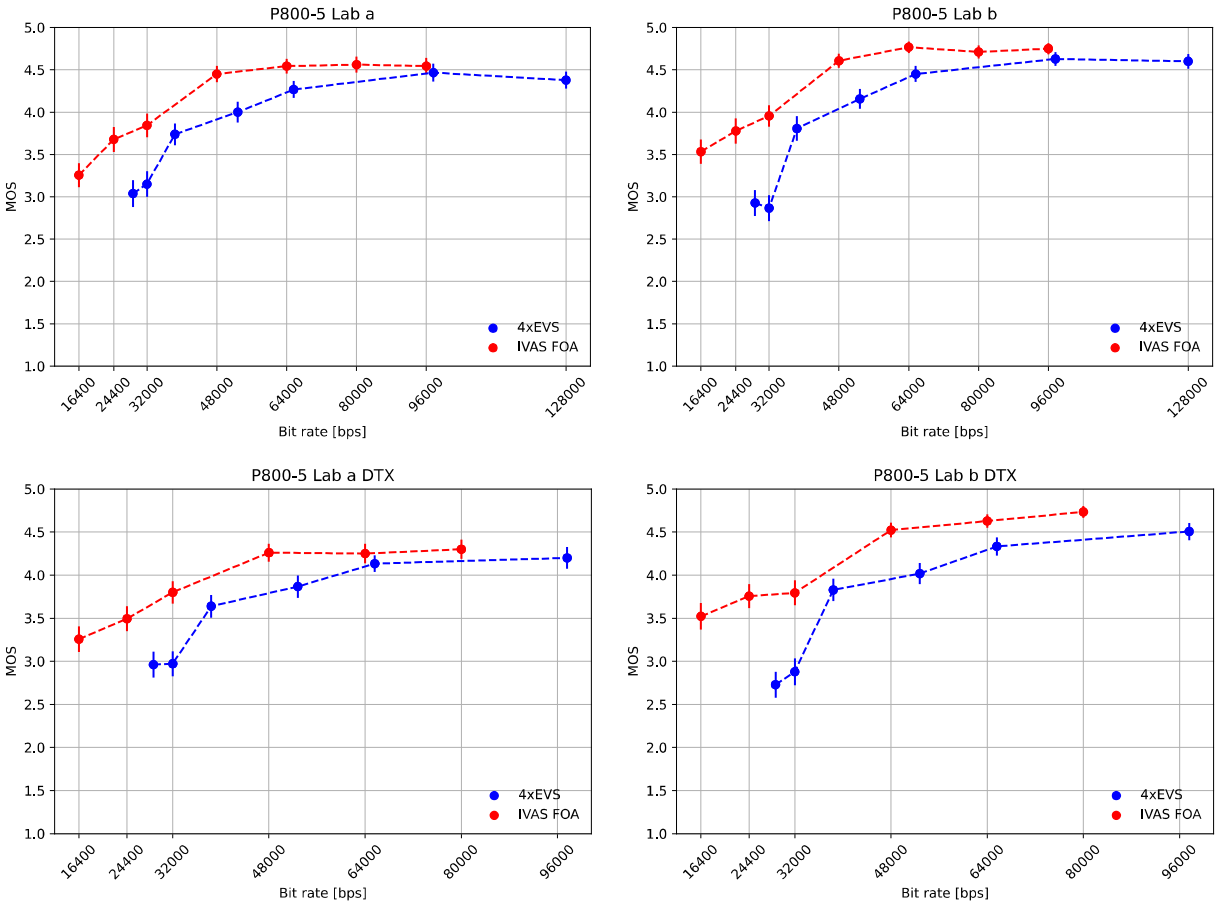


Figure 9.3-2: P800-5 (Ambisonics FOA input/output, speech + background, headphone presentation) rate distortion curves for conditions with DTX off and on

9.3.4 Selection Experiment BS1534-4a (FOA, Generic Audio, 96, 128 and 160 kbps, Headphone Presentation)

Selection Experiment BS1534-4a evaluates IVAS for Ambisonics generic audio for FOA input and HOA3 output at 96 kbps, 128 kbps and 160 kbps using headphone presentation. See , Annex C.16 for details.

The averaged results per condition for experiment BS1534-4a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c08).

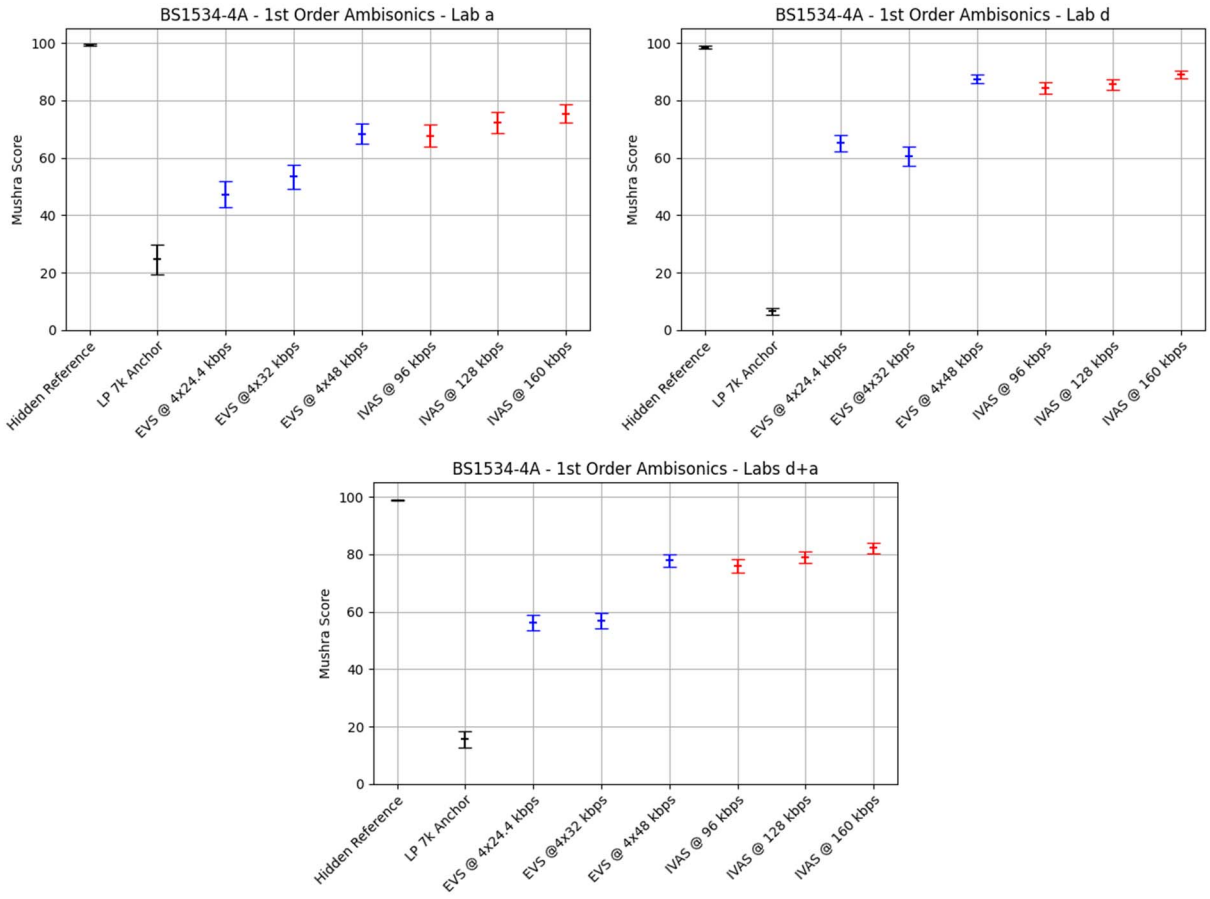


Figure 9.3-3: BS1534-4a (Ambisonics FOA input and HOA3 output, generic Audio, 96, 128 and 160 kbps, headphone presentation) MUSHRA plots for labs a and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-4a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.3-5: Statistical overview on the results of BS1534-4a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	96	NWT	67.7	25	c04	4x32	53.4	27.2	5	BT	EXCEED
		2	96	BT	67.7	25	c03	4x24.4	47.3	28.7	6.93	BT	PASS
	c07	1	128	NWT	72.3	23.6	c05	4x48	68.4	23	1.53	NWT	PASS
		2	128	BT	72.3	23.6	c04	4x32	53.4	27.2	6.78	BT	PASS
	c08	1	160	NWT	75.4	21	c05	4x48	68.4	23	2.93	BT	EXCEED
		2	160	BT	75.4	21	c04	4x32	53.4	27.2	8.28	BT	PASS
d	c06	1	96	NWT	84.3	13.4	c04	4x32	60.5	22.6	11.74	BT	EXCEED
		2	96	BT	84.3	13.4	c03	4x24.4	65.1	19	10.68	BT	PASS
	c07	1	128	NWT	85.7	12.3	c05	4x48	87.5	10.3	-1.48	NWT	PASS
		2	128	BT	85.7	12.3	c04	4x32	60.5	22.6	12.67	BT	PASS
	c08	1	160	NWT	89.2	8.7	c05	4x48	87.5	10.3	1.6	NWT	PASS
		2	160	BT	89.2	8.7	c04	4x32	60.5	22.6	15.33	BT	PASS
d+a	c06	1	96	NWT	76	21.7	c04	4x32	57	25.2	10.48	BT	EXCEED
		2	96	BT	76	21.7	c03	4x24.4	56.2	25.9	10.72	BT	PASS
	c07	1	128	NWT	79	19.9	c05	4x48	77.9	20.2	0.67	NWT	PASS
		2	128	BT	79	19.9	c04	4x32	57	25.2	12.53	BT	PASS
	c08	1	160	NWT	82.3	17.5	c05	4x48	77.9	20.2	2.99	BT	EXCEED
		2	160	BT	82.3	17.5	c04	4x32	57	25.2	15.12	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-6: Summary of the results of BS1534-4a

Lab	Cond.	Bitrate	ToR	Status
a	c06	96	NWT c04 OR BT c03	EXCEED
	c07	128	NWT c05 OR BT c04	PASS
	c08	160	NWT c05 OR BT c04	EXCEED
d	c06	96	NWT c04 OR BT c03	EXCEED
	c07	128	NWT c05 OR BT c04	PASS
	c08	160	NWT c05 OR BT c04	PASS
d+a	c06	96	NWT c04 OR BT c03	EXCEED
	c07	128	NWT c05 OR BT c04	PASS
	c08	160	NWT c05 OR BT c04	EXCEED

9.3.5 Selection Experiment BS1534-4b (HOA2, Generic Audio, 160 and 192 kbps, Headphone Presentation)

Selection Experiment BS1534-4b evaluates IVAS for Ambisonics generic audio for HOA2 input and HOA3 output at 160 and 192 kbps using headphone presentation. See Annex C.17 for details.

The averaged results per condition for experiment BS1534-4b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

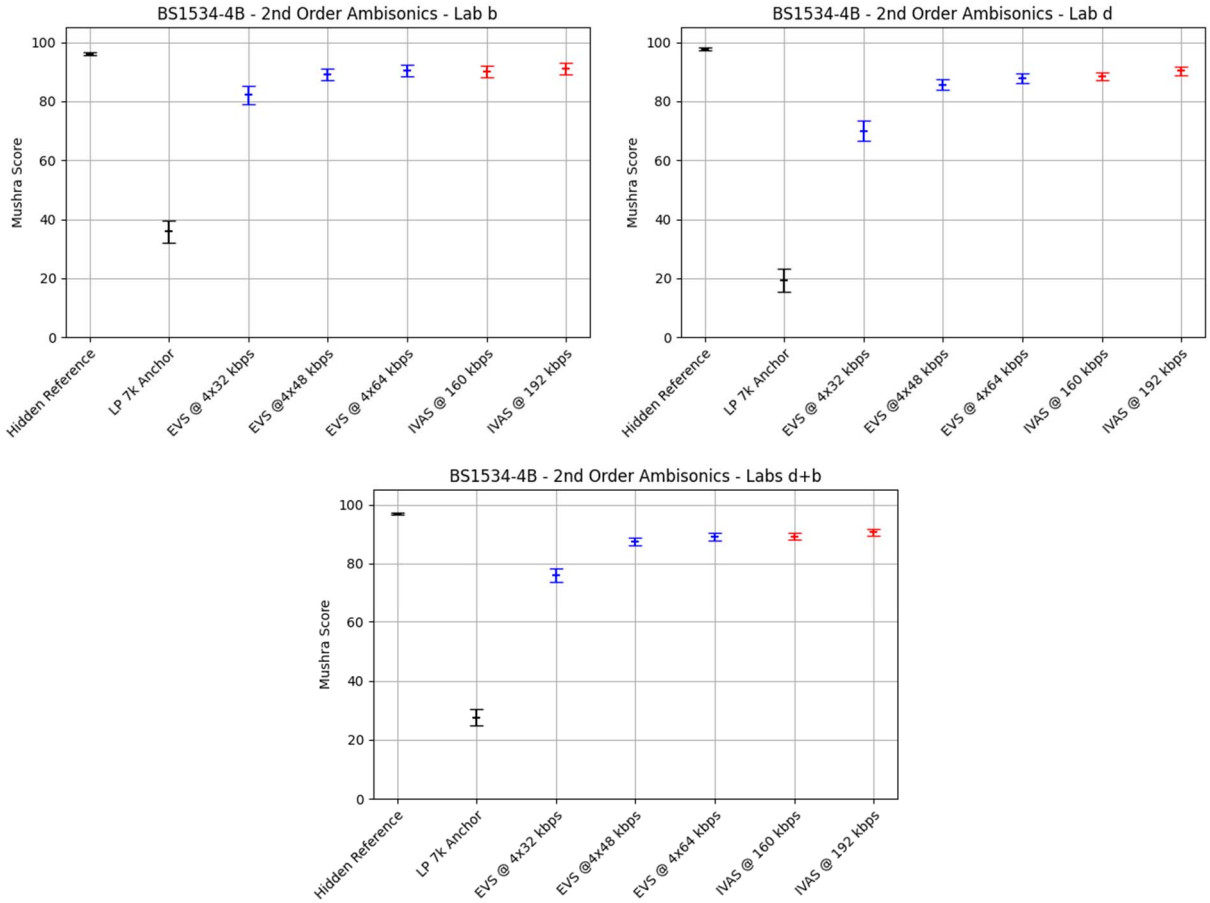


Figure 9.3-4: BS1534-4b (Ambisonics HOA2 input and HOA3 output, generic Audio, 160 and 192 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-4b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.3-7: Statistical overview on the results of BS1534-4b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c06	1	160	NWT	90	13.1	c04	4x48	89.1	13.3	0.63	NWT	PASS
		2	160	BT	90	13.1	c03	4x32	82.2	20.2	4.21	BT	PASS
	c07	1	192	NWT	91	13.1	c05	4x64	90.4	13.3	0.4	NWT	PASS
		2	192	BT	91	13.1	c04	4x48	89.1	13.3	1.29	NWT	FAIL
d	c06	1	160	NWT	88.4	8.3	c04	4x48	85.7	11.2	2.54	BT	EXCEED
		2	160	BT	88.4	8.3	c03	4x32	70	21.6	10.29	BT	PASS
	c07	1	192	NWT	90.3	8.9	c05	4x64	87.8	10.1	2.4	BT	EXCEED
		2	192	BT	90.3	8.9	c04	4x48	85.7	11.2	4.16	BT	PASS
d+b	c06	1	160	NWT	89.2	11	c04	4x48	87.4	12.4	2.01	BT	EXCEED
		2	160	BT	89.2	11	c03	4x32	76.1	21.8	9.85	BT	PASS
	c07	1	192	NWT	90.6	11.2	c05	4x64	89.1	11.9	1.72	BT	EXCEED
		2	192	BT	90.6	11.2	c04	4x48	87.4	12.4	3.54	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-8: Summary of the results of BS1534-4b

Lab	Cond.	Bitrate	ToR	Status
b	c06	160	NWT c04 OR BT c03	PASS
	c07	192	NWT c05 OR BT c04	PASS
d	c06	160	NWT c04 OR BT c03	EXCEED
	c07	192	NWT c05 OR BT c04	EXCEED
d+b	c06	160	NWT c04 OR BT c03	EXCEED
	c07	192	NWT c05 OR BT c04	EXCEED

9.3.6 Selection Experiment BS1534-5a (HOA3, Generic Audio, 192 and 256 kbps, Headphone Presentation)

Selection Experiment BS1534-5a evaluates IVAS for Ambisonics generic audio, HOA3 input and HOA3 output at 192 and 256 kbps using headphone presentation. See Annex C.18 for details.

The averaged results per condition for experiment BS1534-5a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

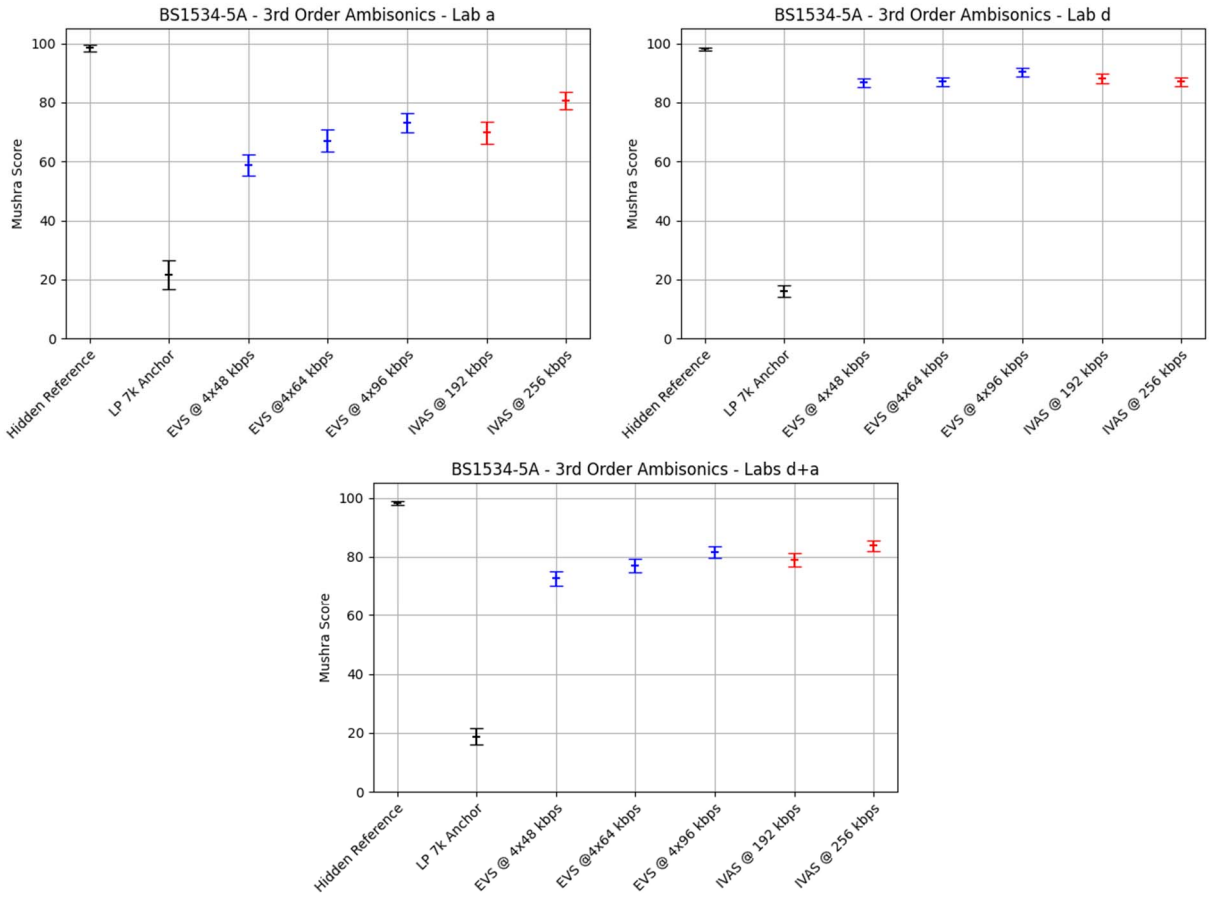


Figure 9.3-5: BS1534-5a (Ambisonics, HOA3 input/output generic audio, 192 and 256 kbps, headphone presentation) MUSHRA plots for labs a and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-5a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.3-9: Statistical overview on the results of BS1534-5a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	192	NWT	69.8	24.4	c04	4x64	67.1	24.5	1.03	NWT	PASS
		2	192	BT	69.8	24.4	c03	4x48	58.8	23.6	4.22	BT	PASS
	c07	1	256	NWT	80.6	20.1	c05	4x96	73	21.6	3.34	BT	EXCEED
		2	256	BT	80.6	20.1	c04	4x64	67.1	24.5	5.54	BT	PASS
d	c06	1	192	NWT	88.1	10.4	c04	4x64	87.1	10	0.86	NWT	PASS
		2	192	BT	88.1	10.4	c03	4x48	86.7	10.1	1.22	NWT	FAIL
	c07	1	256	NWT	87	9.7	c05	4x96	90.4	10	-3.13	WT	FAIL
		2	256	BT	87	9.7	c04	4x64	87.1	10	-0.08	NWT	FAIL
d+a	c06	1	192	NWT	78.9	20.8	c04	4x64	77.1	21.2	1.14	NWT	PASS
		2	192	BT	78.9	20.8	c03	4x48	72.7	22.9	3.67	BT	PASS
	c07	1	256	NWT	83.8	16.1	c05	4x96	81.7	18.9	1.57	NWT	PASS
		2	256	BT	83.8	16.1	c04	4x64	77.1	21.2	4.64	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-10: Summary of the results of BS1534-5a

Lab	Cond.	Bitrate	ToR	Status
a	c06	192	NWT c04 OR BT c03	PASS
	c07	256	NWT c05 OR BT c04	EXCEED
d	c06	192	NWT c04 OR BT c03	PASS
	c07	256	NWT c05 OR BT c04	FAIL

d+a	c06	192	NWT c04 OR BT c03	PASS
	c07	256	NWT c05 OR BT c04	PASS

9.3.7 Selection Experiment BS1534-5b (HOA3, Generic Audio, 384 and 512 kbps, 7.1+4 Loudspeaker Presentation)

Selection Experiment BS1534-5b evaluates IVAS for Ambisonics generic audio, HOA3 input and HOA3 output at 384 and 512 kbps using 7.1+4 loudspeaker presentation. See Annex C.19 for details.

The averaged results per condition for experiment BS1534-5b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c04) and IVAS conditions with increasing bitrate (c05 – c06).

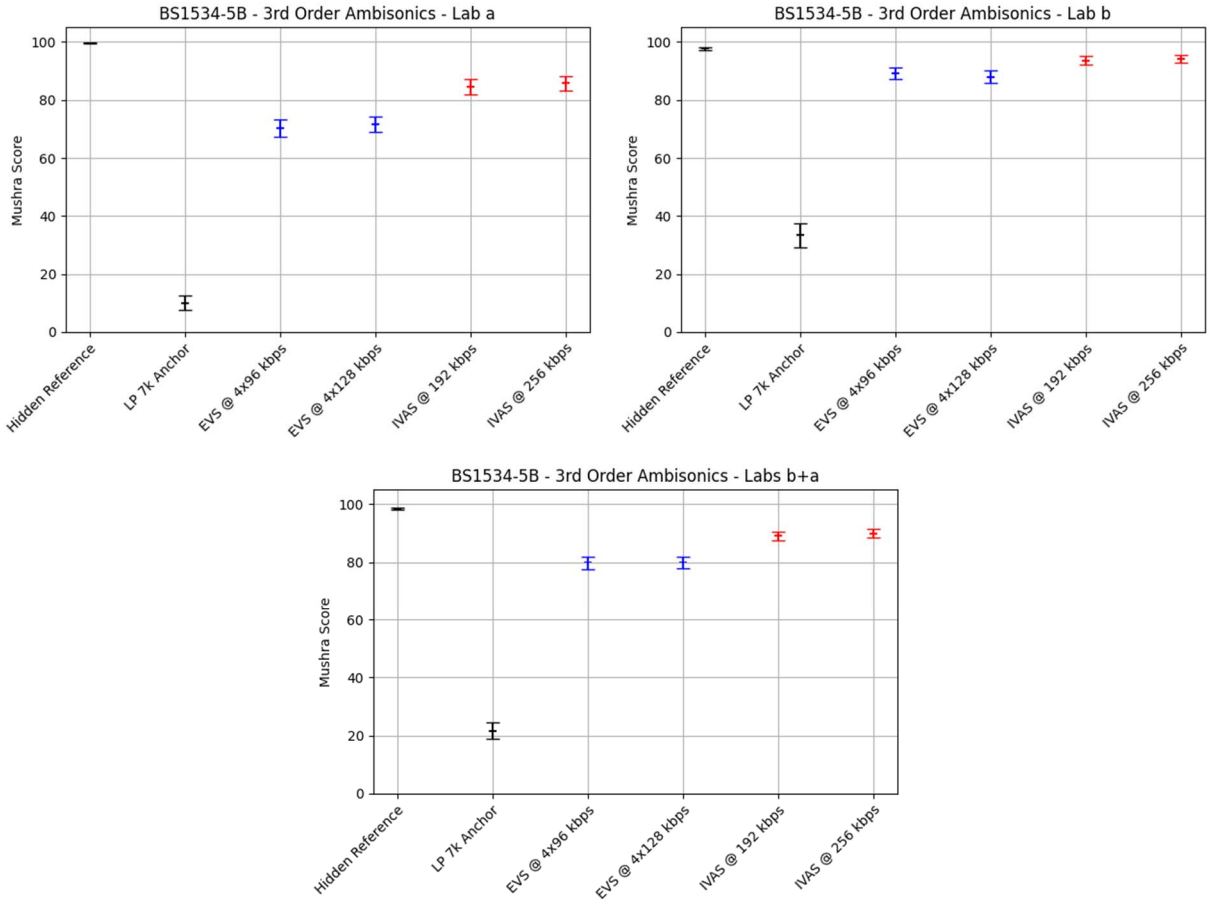


Figure 9.3-6: BS1534-5b (Ambisonics, HOA3 input/output, generic audio, 384 and 512 kbps, 7.1+4 loudspeaker presentation) MUSHRA plots for labs a and b, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-5b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.3-11: Statistical overview on the results of BS1534-5b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c05	1	384	NWT	84.5	16.8	c04	4x128	71.6	18.1	6.78	BT	EXCEED
		2	384	BT	84.5	16.8	c03	4x96	70.4	19.3	7.15	BT	PASS
	c06	1	512	NWT	85.8	16.1	c04	4x128	71.6	18.1	7.62	BT	EXCEED
b	c05	1	384	NWT	93.6	9.5	c04	4x128	88	14.4	4.24	BT	EXCEED
		2	384	BT	93.6	9.5	c03	4x96	89.1	13.3	3.55	BT	PASS
	c06	1	512	NWT	94.2	8.5	c04	4x128	88	14.4	4.8	BT	EXCEED
b+a	c05	1	384	NWT	89.1	14.4	c04	4x128	79.8	18.3	7.32	BT	EXCEED
		2	384	BT	89.1	14.4	c03	4x96	79.8	19	7.15	BT	PASS
	c06	1	512	NWT	90	13.5	c04	4x128	79.8	18.3	8.23	BT	EXCEED

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.3-12: Summary of the results of BS1534-5b

Lab	Cond.	Bitrate	ToR	Status
a	c05	384	NWT c04 OR BT c03	EXCEED
	c06	512	NWT c04	EXCEED
b	c05	384	NWT c04 OR BT c03	EXCEED
	c06	512	NWT c04	EXCEED
b+a	c05	384	NWT c04 OR BT c03	EXCEED
	c06	512	NWT c04	EXCEED

9.3.8 Characterization Experiment P800-3 (FOA, Speech + background and mixed/music, Headphone Presentation)

Characterization Experiment P800-3 evaluates the performance of the IVAS fixed-point and floating-point implementation for FOA speech + background and mixed/music under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-3 was conducted by Orange (language: FR).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.2-13 and are shown in Figure 9.3-7. The results aggregated over FX/FL are shown in Figure 9.3-8. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-14. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.3-13: results per condition for experiment P800-3

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.64	0.60	0.09
c02	MNRU Q = 32 dB	180	4.17	0.88	0.13
c03	MNRU Q = 27 dB	180	3.18	1.20	0.18
c04	MNRU Q = 22 dB	180	2.19	1.25	0.18
c05	MNRU Q = 17 dB	180	1.62	0.93	0.14
c06	ESDRU α = 0.8	180	4.62	0.64	0.09
c07	ESDRU α = 0.6	180	3.58	0.93	0.14
c08	ESDRU α = 0.4	180	2.73	1.11	0.16
c09	ESDRU α = 0.2	180	2.34	1.15	0.17
c10	IVAS FL enc / FX dec, 13.2 kbps, DTX off	179	3.13	1.10	0.16
c11	IVAS FL enc / FX dec, 16.4 kbps, DTX off	180	3.53	1.05	0.15
c12	IVAS FX enc / FL dec, 24.4 kbps, DTX off	180	3.87	0.89	0.13
c13	IVAS FL enc / FX dec, 32 kbps, DTX off	179	4.06	0.79	0.12
c14	IVAS FX enc / FL dec, 48 kbps, DTX off	180	4.36	0.77	0.11
c15	IVAS FL enc / FX dec, 64 kbps, DTX off	176	4.49	0.68	0.10
c16	IVAS FL enc / FX dec, 96 kbps, DTX off	180	4.64	0.60	0.09
c17	IVAS FX enc / FL dec, 128 kbps, DTX off	180	4.64	0.57	0.08

c18	IVAS FL enc / FX dec, 256 kbps, DTX off	180	4.64	0.60	0.09
c19	IVAS FL, 13.2 kbps, DTX off	180	3.22	1.11	0.16
c20	IVAS FL, 16.4 kbps, DTX off	180	3.64	0.95	0.14
c21	IVAS FL, 24.4 kbps, DTX off	180	3.96	0.89	0.13
c22	IVAS FL, 32 kbps, DTX off	180	4.12	0.83	0.12
c23	IVAS FL, 48 kbps, DTX off	180	4.30	0.73	0.11
c24	IVAS FL, 64 kbps, DTX off	180	4.43	0.68	0.10
c25	IVAS FL, 96 kbps, DTX off	180	4.66	0.54	0.08
c26	IVAS FL, 128 kbps, DTX off	180	4.68	0.56	0.08
c27	IVAS FL, 256 kbps, DTX off	180	4.62	0.60	0.09
c28	IVAS FX, 13.2 kbps, DTX off	180	3.18	1.11	0.16
c29	IVAS FX, 16.4 kbps, DTX off	180	3.62	0.96	0.14
c30	IVAS FX, 24.4 kbps, DTX off	180	3.81	0.90	0.13
c31	IVAS FX, 32 kbps, DTX off	180	4.13	0.78	0.12
c32	IVAS FX, 48 kbps, DTX off	180	4.40	0.71	0.10
c33	IVAS FX, 64 kbps, DTX off	180	4.38	0.76	0.11
c34	IVAS FX, 96 kbps, DTX off	180	4.56	0.63	0.09
c35	IVAS FX, 128 kbps, DTX off	180	4.64	0.58	0.09
c36	IVAS FX, 256 kbps, DTX off	180	4.72	0.48	0.07

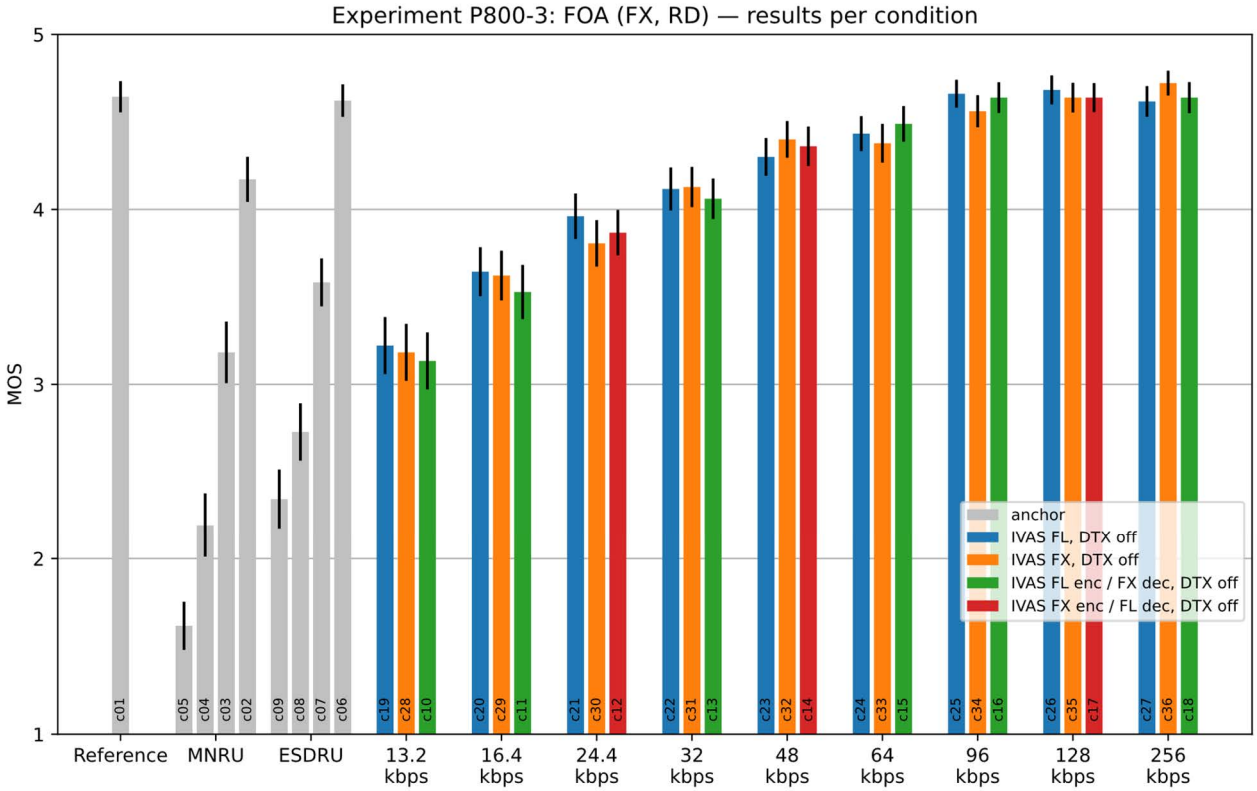


Figure 9.3-7: results per condition for experiment P800-3

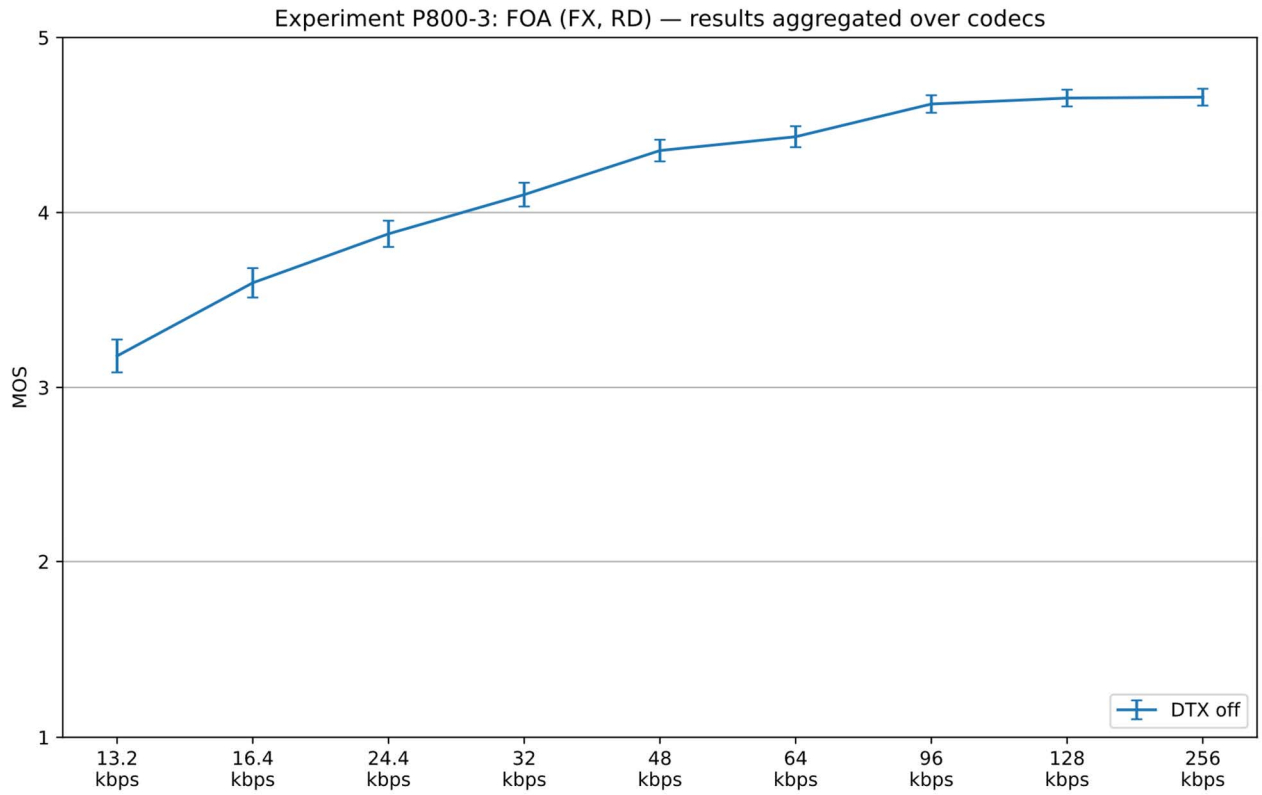


Figure 9.3-8: results aggregated over FX/FL for experiment P800-3

Table 9.3-14: statistical significance analysis for experiment P800-3

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.13	1.10	c19	3.22	1.11	-0.09	<0.001	0.774	NWT
c11	3.53	1.05	c20	3.64	0.95	-0.12	<0.001	0.863	NWT
c12	3.87	0.89	c21	3.96	0.89	-0.09	<0.001	0.842	NWT
c13	4.06	0.79	c22	4.12	0.83	-0.06	<0.001	0.743	NWT
c14	4.36	0.77	c23	4.30	0.73	0.06	0.771	0.220	NWT
c15	4.49	0.68	c24	4.43	0.68	0.06	0.776	0.219	NWT
c16	4.64	0.60	c25	4.66	0.54	-0.02	<0.001	0.643	NWT
c17	4.64	0.57	c26	4.68	0.56	-0.04	<0.001	0.770	NWT
c18	4.64	0.60	c27	4.62	0.60	0.02	0.346	0.365	NWT
c28	3.18	1.11	c19	3.22	1.11	-0.04	<0.001	0.630	NWT
c29	3.62	0.96	c20	3.64	0.95	-0.02	<0.001	0.586	NWT
c30	3.81	0.90	c21	3.96	0.89	-0.15	<0.001	0.949	NWT
c31	4.13	0.78	c22	4.12	0.83	0.01	0.129	0.449	NWT
c32	4.40	0.71	c23	4.30	0.73	0.10	1.312	0.095	NWT
c33	4.38	0.76	c24	4.43	0.68	-0.05	<0.001	0.766	NWT
c34	4.56	0.63	c25	4.66	0.54	-0.10	<0.001	0.947	NWT
c35	4.64	0.58	c26	4.68	0.56	-0.04	<0.001	0.768	NWT
c36	4.72	0.48	c27	4.62	0.60	0.11	1.824	0.035	BT

9.3.9 Characterization Experiment P800-4 (HOA2, Speech + background and mixed/music, Impaired Channel, DTX, Headphone Presentation)

Characterization Experiment P800-4 evaluates the performance of the IVAS fixed-point and floating-point implementation for FOA speech + background and mixed/music, DTX on, under impaired channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-4 was conducted by Orange (language: FR).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.3-15 and are shown in Figure 9.3-9. The results aggregated over FX/FL are shown in Figure 9.3-10. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-16. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.3-15: results per condition for experiment P800-4

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.69	0.55	0.08
c02	MNRU Q = 32 dB	180	3.65	1.05	0.15
c03	MNRU Q = 27 dB	180	2.73	1.22	0.18
c04	MNRU Q = 22 dB	179	1.98	1.15	0.17
c05	MNRU Q = 17 dB	179	1.51	0.90	0.13
c06	ESDRU α = 0.8	180	4.58	0.63	0.09
c07	ESDRU α = 0.6	180	3.63	1.01	0.15
c08	ESDRU α = 0.4	179	2.96	1.14	0.17
c09	ESDRU α = 0.2	180	2.58	1.17	0.17
c10	IVAS FL enc / FX dec, 16.4 kbps, FER 5%, DTX On	180	3.23	1.15	0.17
c11	IVAS FX enc / FL dec, 24.4 kbps, FER 5%, DTX On	180	3.36	1.06	0.16
c12	IVAS FL enc / FX dec, 32 kbps, FER 5%, DTX On	180	3.62	1.00	0.15
c13	IVAS FX enc / FL dec, 48 kbps, FER 5%, DTX On	179	4.11	0.90	0.13
c14	IVAS FL enc / FX dec, 64 kbps, FER 5%, DTX On	180	4.14	0.85	0.12
c15	IVAS FX enc / FL dec, 96 kbps, FER 5%, DTX Off	180	4.29	0.79	0.12
c16	IVAS FL enc / FX dec, 128 kbps, FER 5%, DTX Off	179	4.28	0.81	0.12
c17	IVAS FX enc / FL dec, 256 kbps, FER 5%, DTX Off	180	4.43	0.65	0.10
c18	IVAS FL enc / FX dec, 384 kbps, FER 5%, DTX Off	180	4.31	0.80	0.12
c19	IVAS FL, 16.4 kbps, FER 5%, DTX On	180	3.39	1.09	0.16
c20	IVAS FL, 24.4 kbps, FER 5%, DTX On	180	3.29	1.05	0.15
c21	IVAS FL, 32 kbps, FER 5%, DTX On	180	3.56	1.10	0.16
c22	IVAS FL, 48 kbps, FER 5%, DTX On	180	3.97	0.87	0.13
c23	IVAS FL, 64 kbps, FER 5%, DTX On	180	4.17	0.82	0.12
c24	IVAS FL, 96 kbps, FER 5%, DTX Off	180	4.34	0.71	0.10
c25	IVAS FL, 128 kbps, FER 5%, DTX Off	180	4.36	0.67	0.10
c26	IVAS FL, 256 kbps, FER 5%, DTX Off	180	4.32	0.78	0.12
c27	IVAS FL, 384 kbps, FER 5%, DTX Off	180	4.32	0.76	0.11
c28	IVAS FX, 16.4 kbps, FER 5%, DTX On	180	3.22	1.15	0.17
c29	IVAS FX, 24.4 kbps, FER 5%, DTX On	180	3.41	1.06	0.15
c30	IVAS FX, 32 kbps, FER 5%, DTX On	180	3.59	0.95	0.14
c31	IVAS FX, 48 kbps, FER 5%, DTX On	180	3.98	0.90	0.13
c32	IVAS FX, 64 kbps, FER 5%, DTX On	180	4.18	0.84	0.12
c33	IVAS FX, 96 kbps, FER 5%, DTX Off	180	4.27	0.77	0.11
c34	IVAS FX, 128 kbps, FER 5%, DTX Off	180	4.37	0.80	0.12
c35	IVAS FX, 256 kbps, FER 5%, DTX Off	180	4.40	0.65	0.10
c36	IVAS FX, 384 kbps, FER 5%, DTX Off	179	4.41	0.67	0.10

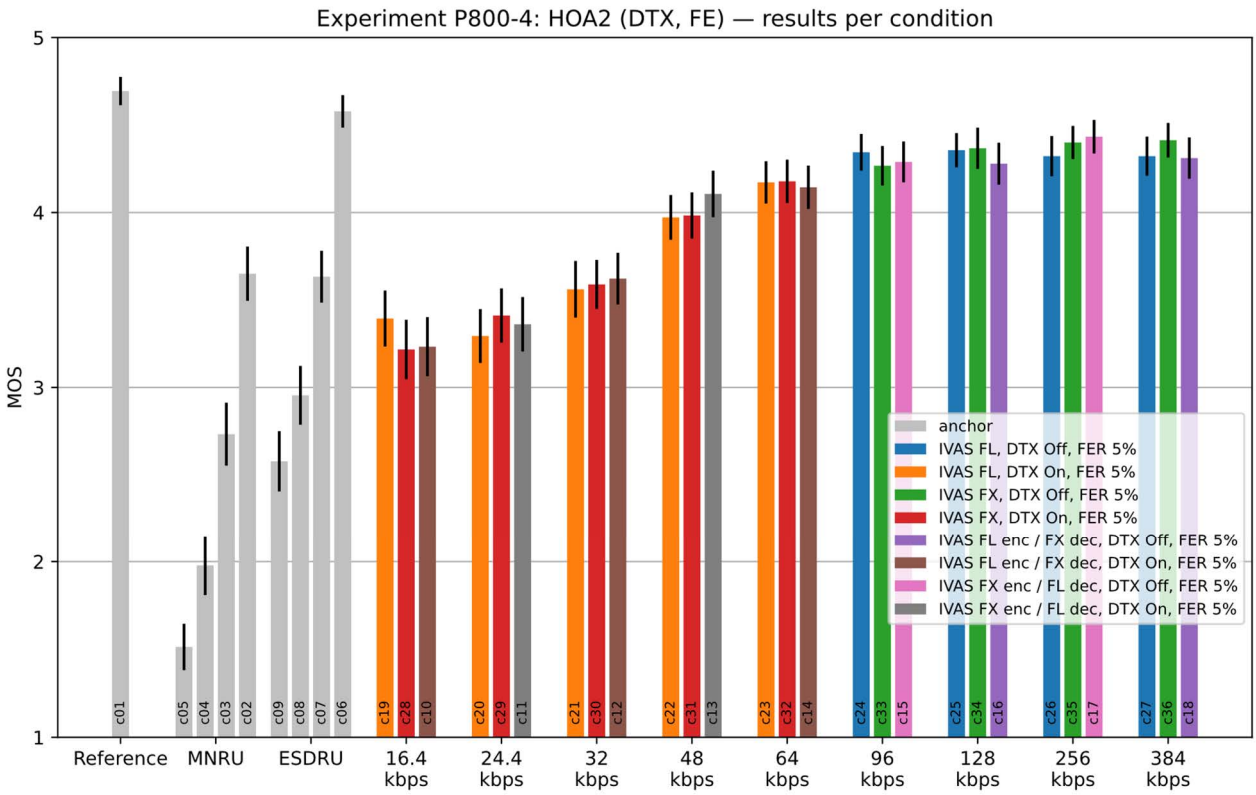


Figure 9.3-9: results per condition for experiment P800-4

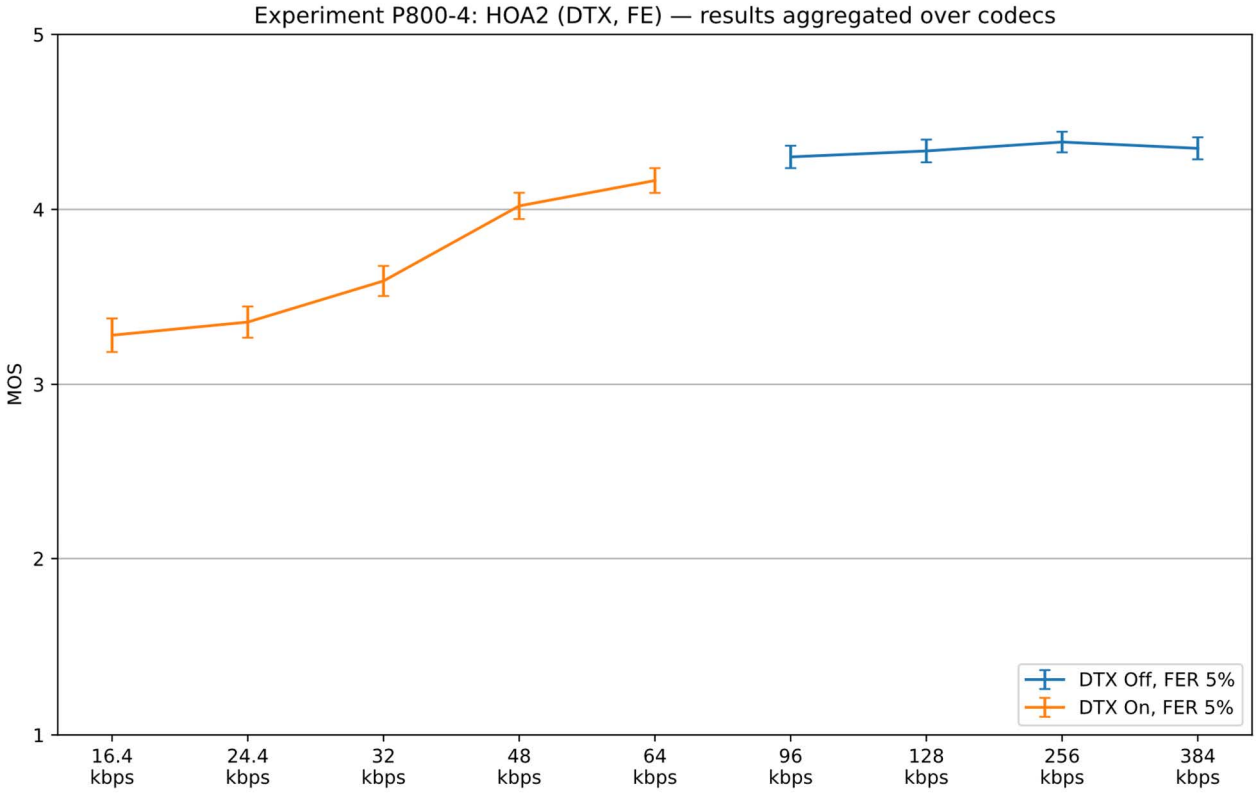


Figure 9.3-10: results aggregated over FX/FL for experiment P800-4

Table 9.3-16: statistical significance analysis for experiment P800-4

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.23	1.15	c19	3.39	1.09	-0.16	<0.001	0.913	NWT
c11	3.36	1.06	c20	3.29	1.05	0.07	0.602	0.274	NWT
c12	3.62	1.00	c21	3.56	1.10	0.06	0.550	0.291	NWT
c13	4.11	0.90	c22	3.97	0.87	0.13	1.433	0.076	NWT
c14	4.14	0.85	c23	4.17	0.82	-0.03	<0.001	0.625	NWT
c15	4.29	0.79	c24	4.34	0.71	-0.06	<0.001	0.755	NWT
c16	4.28	0.81	c25	4.36	0.67	-0.08	<0.001	0.836	NWT
c17	4.43	0.65	c26	4.32	0.78	0.11	1.464	0.072	NWT
c18	4.31	0.80	c27	4.32	0.76	-0.01	<0.001	0.553	NWT
c28	3.22	1.15	c19	3.39	1.09	-0.18	<0.001	0.932	NWT
c29	3.41	1.06	c20	3.29	1.05	0.12	1.054	0.146	NWT
c30	3.59	0.95	c21	3.56	1.10	0.03	0.259	0.398	NWT
c31	3.98	0.90	c22	3.97	0.87	0.01	0.118	0.453	NWT
c32	4.18	0.84	c23	4.17	0.82	0.01	0.068	0.473	NWT
c33	4.27	0.77	c24	4.34	0.71	-0.08	<0.001	0.838	NWT
c34	4.37	0.80	c25	4.36	0.67	0.01	0.141	0.444	NWT
c35	4.40	0.65	c26	4.32	0.78	0.08	1.031	0.152	NWT
c36	4.41	0.67	c27	4.32	0.76	0.09	1.205	0.114	NWT

9.3.10 Characterization Experiment P800-5 (HOA3, Speech + background and mixed/music, Headphone Presentation)

Characterization Experiment P800-5 evaluates the performance of the IVAS fixed-point and floating-point implementation for HOA3 speech + background and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-5 was conducted by Dolby Laboratories, Inc. (language: ENG).

The aggregated results per condition are listed in Table 9.3-17 and are shown in Figure 9.3-11. The results aggregated over FX/FL are shown in Figure 9.3-12. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-18. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.3-17: results per condition for experiment P800-5

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.72	0.51	0.07
c02	MNRU Q = 32 dB	180	3.69	1.09	0.16
c03	MNRU Q = 27 dB	180	2.83	1.17	0.17
c04	MNRU Q = 22 dB	180	2.07	1.19	0.17
c05	MNRU Q = 17 dB	180	1.68	1.03	0.15
c06	ESDRU α = 0.8	180	4.51	0.63	0.09
c07	ESDRU α = 0.6	180	3.70	1.02	0.15
c08	ESDRU α = 0.4	180	3.11	1.23	0.18
c09	ESDRU α = 0.2	180	2.49	1.15	0.17
c10	IVAS FL enc / FX dec, 32 kbps, DTX off	180	4.26	0.81	0.12
c11	IVAS FX enc / FL dec, 48 kbps, DTX off	180	4.41	0.70	0.10
c12	IVAS FL enc / FX dec, 64 kbps, DTX off	180	4.45	0.69	0.10
c13	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.60	0.68	0.10
c14	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.65	0.57	0.08
c15	IVAS FX enc / FL dec, 160 kbps, DTX off	180	4.67	0.58	0.09
c16	IVAS FL enc / FX dec, 256 kbps, DTX off	180	4.67	0.60	0.09
c17	IVAS FX enc / FL dec, 384 kbps, DTX off	180	4.67	0.53	0.08
c18	IVAS FL enc / FX dec, 512 kbps, DTX off	180	4.66	0.61	0.09
c19	IVAS FL, 32 kbps, DTX off	180	4.16	0.81	0.12

c20	IVAS FL, 48 kbps, DTX off	180	4.35	0.74	0.11
c21	IVAS FL, 64 kbps, DTX off	180	4.47	0.66	0.10
c22	IVAS FL, 96 kbps, DTX off	180	4.56	0.68	0.10
c23	IVAS FL, 128 kbps, DTX off	180	4.60	0.65	0.10
c24	IVAS FL, 160 kbps, DTX off	180	4.67	0.58	0.09
c25	IVAS FL, 256 kbps, DTX off	180	4.72	0.56	0.08
c26	IVAS FL, 384 kbps, DTX off	180	4.66	0.58	0.09
c27	IVAS FL, 512 kbps, DTX off	180	4.59	0.61	0.09
c28	IVAS FX, 32 kbps, DTX off	180	4.16	0.80	0.12
c29	IVAS FX, 48 kbps, DTX off	180	4.36	0.81	0.12
c30	IVAS FX, 64 kbps, DTX off	180	4.46	0.72	0.11
c31	IVAS FX, 96 kbps, DTX off	180	4.66	0.53	0.08
c32	IVAS FX, 128 kbps, DTX off	180	4.72	0.48	0.07
c33	IVAS FX, 160 kbps, DTX off	180	4.69	0.52	0.08
c34	IVAS FX, 256 kbps, DTX off	180	4.59	0.65	0.10
c35	IVAS FX, 384 kbps, DTX off	180	4.58	0.66	0.10
c36	IVAS FX, 512 kbps, DTX off	180	4.70	0.55	0.08

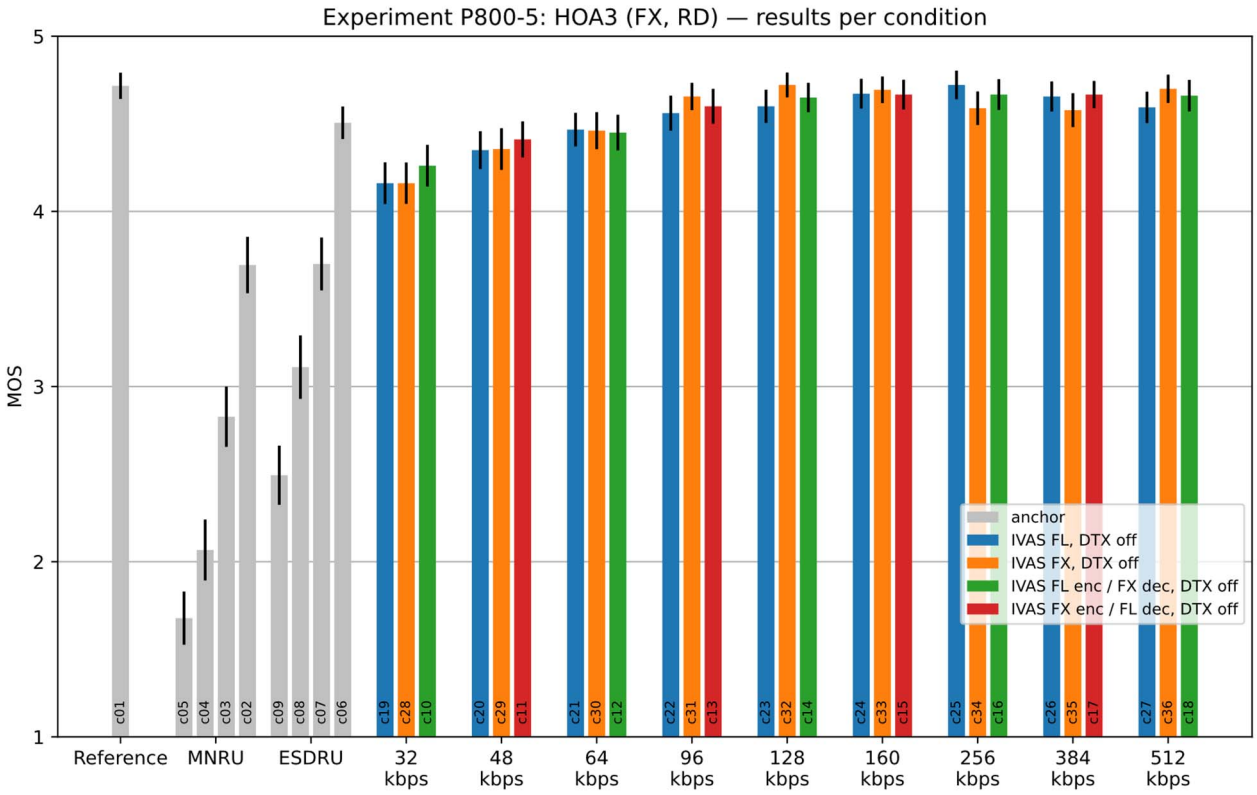


Figure 9.3-11: results per condition for experiment P800-5

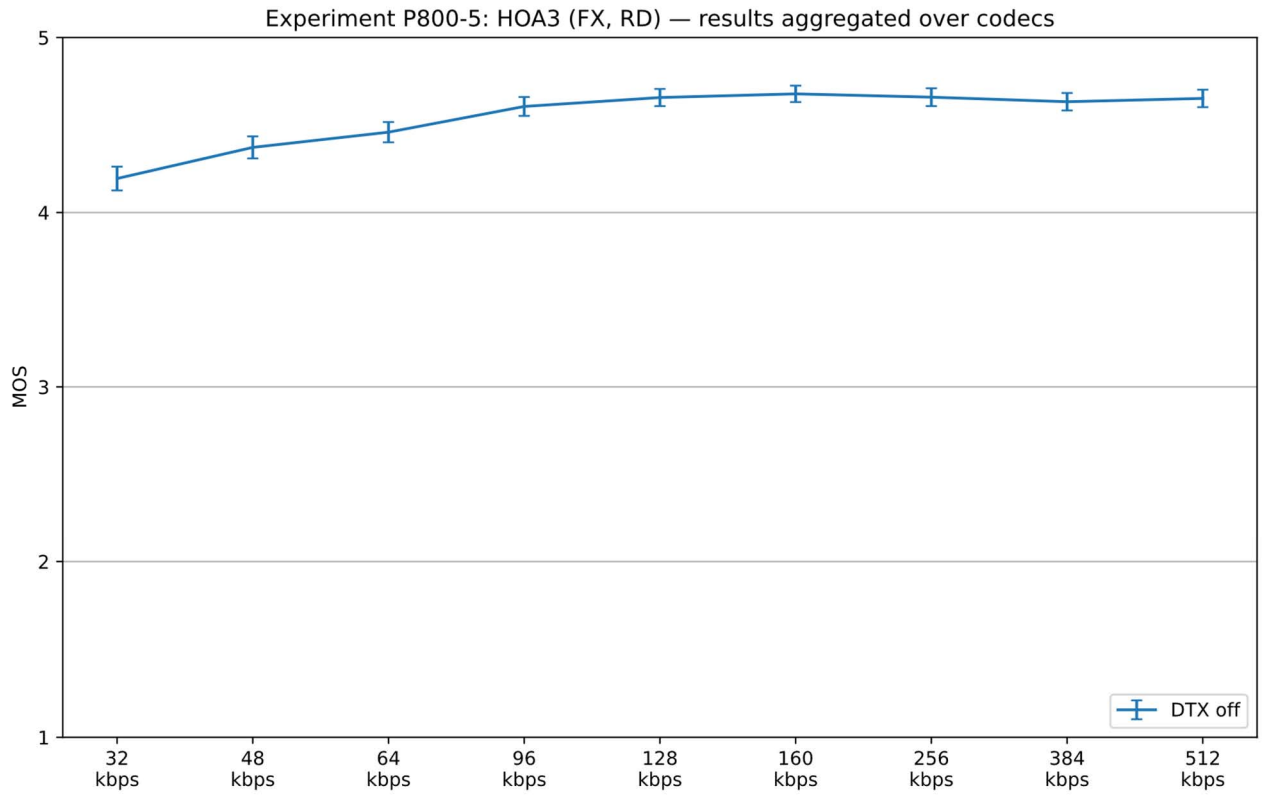


Figure 9.3-12: results aggregated over FX/FL for experiment P800-5

Table 9.3-18: statistical significance analysis for experiment P800-5

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	4.26	0.81	c19	4.16	0.81	0.10	1.176	0.120	NWT
c11	4.41	0.70	c20	4.35	0.74	0.06	0.806	0.210	NWT
c12	4.45	0.69	c21	4.47	0.66	-0.02	<0.001	0.594	NWT
c13	4.60	0.68	c22	4.56	0.68	0.04	0.544	0.293	NWT
c14	4.65	0.57	c23	4.60	0.65	0.05	0.775	0.219	NWT
c15	4.67	0.58	c24	4.67	0.58	-0.00	<0.001	0.533	NWT
c16	4.67	0.60	c25	4.72	0.56	-0.06	<0.001	0.816	NWT
c17	4.67	0.53	c26	4.66	0.58	0.01	0.188	0.426	NWT
c18	4.66	0.61	c27	4.59	0.61	0.07	1.039	0.150	NWT
c28	4.16	0.80	c19	4.16	0.81	0.00	<0.001	0.500	NWT
c29	4.36	0.81	c20	4.35	0.74	0.01	0.074	0.471	NWT
c30	4.46	0.72	c21	4.47	0.66	-0.01	<0.001	0.533	NWT
c31	4.66	0.53	c22	4.56	0.68	0.09	1.479	0.070	NWT
c32	4.72	0.48	c23	4.60	0.65	0.12	2.022	0.022	BT
c33	4.69	0.52	c24	4.67	0.58	0.02	0.380	0.352	NWT
c34	4.59	0.65	c25	4.72	0.56	-0.13	<0.001	0.981	WT
c35	4.58	0.66	c26	4.66	0.58	-0.08	<0.001	0.882	NWT
c36	4.70	0.55	c27	4.59	0.61	0.11	1.728	0.042	BT

9.3.11 Characterization Experiment P800-23 (FOA, Speech + background and mixed/music, JBM, Headphone Presentation)

Characterization Experiment P800-23 evaluates the performance of the IVAS JBM solution for fixed-point and floating-point implementation for FOA speech + background and mixed/music, under impaird channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-23 was conducted by FORCE Technology (language: DAN).

The aggregated results per condition are listed in Table 9.3-19 and are shown in Figure 9.3-13. The results aggregated over FX/FL are shown in Figure 9.3-14. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-20. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.3-19: results per condition for experiment P800-23

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.55	0.68	0.10
c02	MNRU Q = 32 dB	180	4.09	0.87	0.13
c03	MNRU Q = 27 dB	180	3.22	1.12	0.17
c04	MNRU Q = 22 dB	180	2.41	1.18	0.17
c05	MNRU Q = 17 dB	180	1.64	0.90	0.13
c06	ESDRU α = 0.8	180	4.55	0.69	0.10
c07	ESDRU α = 0.6	180	4.05	0.87	0.13
c08	ESDRU α = 0.4	180	3.36	1.07	0.16
c09	ESDRU α = 0.2	180	2.97	1.18	0.17
c10	IVAS FL enc / FX dec, No error, 24.4 kbps	180	3.84	0.85	0.12
c11	IVAS FL enc / FX dec, No error, 48 kbps	180	4.12	0.88	0.13
c12	IVAS FL enc / FX dec, No error, 96 kbps	180	4.44	0.76	0.11
c13	IVAS FL enc / FX dec, Error I1. O1, 24.4 kbps	180	2.69	1.10	0.16
c14	IVAS FL enc / FX dec, Error I1. O1, 48 kbps	180	2.97	1.10	0.16
c15	IVAS FL enc / FX dec, Error I1. O2, 24.4 kbps	180	3.47	1.08	0.16
c16	IVAS FL enc / FX dec, Error I1. O2, 48 kbps	180	3.87	1.14	0.17
c17	IVAS FL enc / FX dec, I1. O1, 24.4 kbps	180	2.65	1.14	0.17
c18	IVAS FL enc / FX dec, I1. O1, 48 kbps	180	2.93	1.05	0.15
c19	IVAS FL enc / FX dec, I1. O1, 96 kbps	180	3.30	1.05	0.15
c20	IVAS FL enc / FX dec, I1. O2, 24.4 kbps	180	3.38	1.10	0.16
c21	IVAS FL enc / FX dec, I1. O2, 48 kbps	180	3.86	1.12	0.17
c22	IVAS FL enc / FX dec, I1. O2, 96 kbps	180	3.95	1.15	0.17
c23	IVAS FL, No error, 24.4 kbps	180	3.79	0.82	0.12
c24	IVAS FL, No error, 48 kbps	180	4.14	0.93	0.14
c25	IVAS FL, No error, 96 kbps	180	4.46	0.74	0.11
c26	IVAS FL, Error I1. O1, 24.4 kbps	180	2.62	1.10	0.16
c27	IVAS FL, Error I1. O1, 48 kbps	180	2.99	1.15	0.17
c28	IVAS FL, Error I1. O2, 24.4 kbps	180	3.51	1.07	0.16
c29	IVAS FL, Error I1. O2, 48 kbps	180	3.80	1.16	0.17
c30	IVAS FL, Error I1. O2, 96 kbps	180	4.13	1.12	0.16
c31	IVAS FL, I1. O1, 24.4 kbps	180	2.51	1.04	0.15
c32	IVAS FL, I1. O1, 48 kbps	180	2.91	1.09	0.16
c33	IVAS FL, I1. O1, 96 kbps	180	3.26	1.05	0.15
c34	IVAS FL, I1. O2, 24.4 kbps	180	3.35	0.98	0.14
c35	IVAS FL, I1. O2, 48 kbps	180	3.79	1.13	0.17
c36	IVAS FL, I1. O2, 96 kbps	180	4.06	1.08	0.16

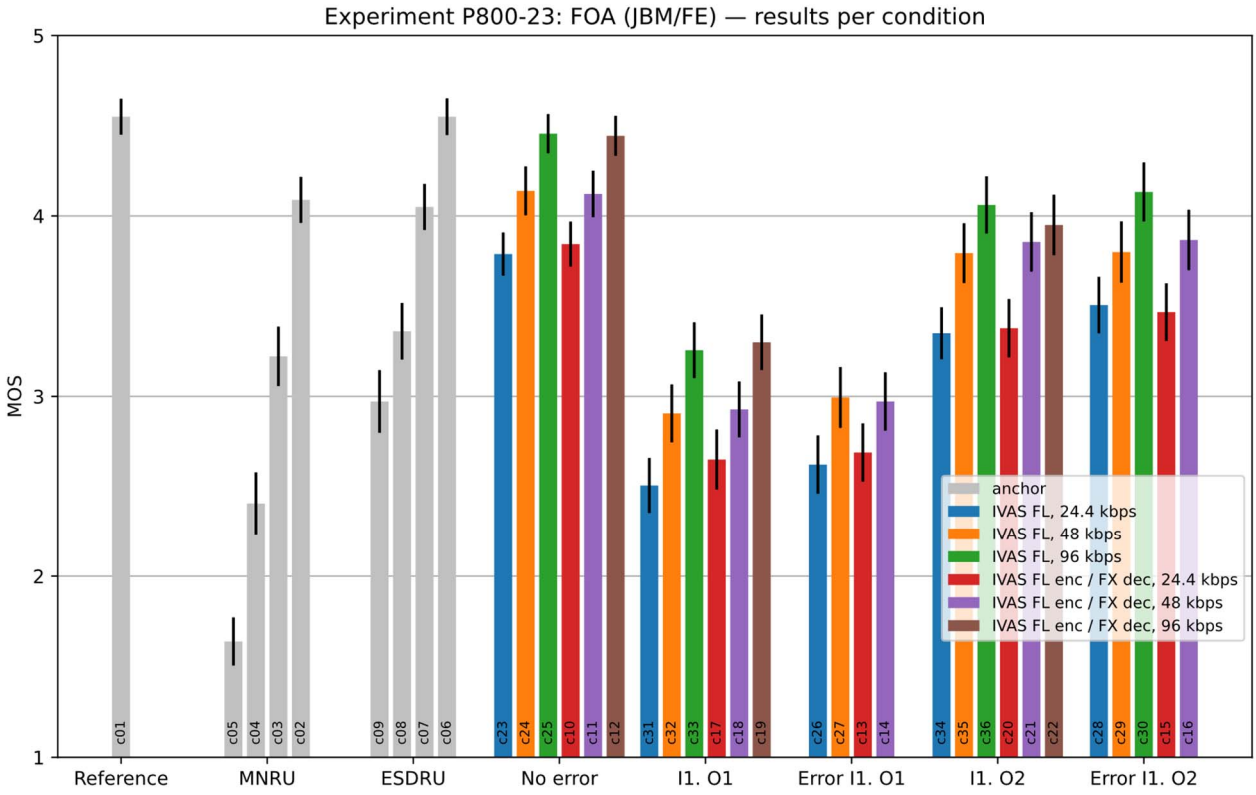


Figure 9.3-13: results per condition for experiment P800-23

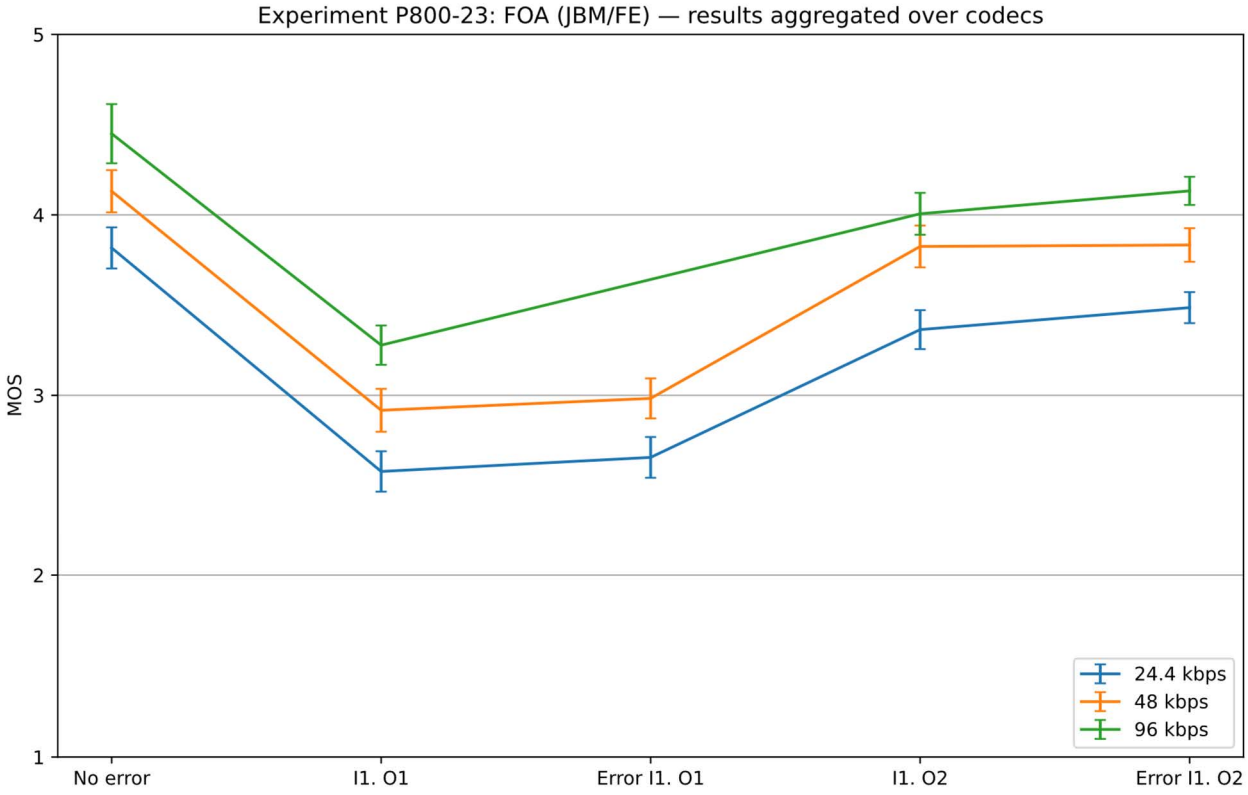


Figure 9.3-14: results aggregated over FX/FL for experiment P800-23

Table 9.3-20: statistical significance analysis for experiment P800-23

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
No terms of references mandated for experiment.									

9.3.12 Characterization Experiment BS1534-3 (FOA, Generic Audio, 16.4 - 48 kbps, Headphone Presentation)

Characterization Experiment BS1534-3 evaluates the performance of the IVAS fixed-point implementation for FOA generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-3 was conducted by Huawei Technologies Co Ltd.

The aggregated results per condition are listed in Table 9.3-21 and are shown in Figure 9.3-15. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-22. Scores with result WT are highlighted in red.

Table 9.3-21: results per condition for experiment BS1534-3

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.29	3.72	0.57
c02	LP 3.5 kHz	168	48.10	20.95	3.19
c03	Mono EVS, 16.4 kbps	168	68.49	17.83	2.72
c04	Mono EVS, 32 kbps	168	70.41	17.73	2.70
c05	IVAS, 16.4 kbps	168	76.88	19.76	3.01
c06	IVAS, 24.4 kbps	168	82.06	16.66	2.54
c07	IVAS, 32 kbps	168	84.23	15.03	2.29
c08	IVAS, 48 kbps	168	89.92	11.70	1.78

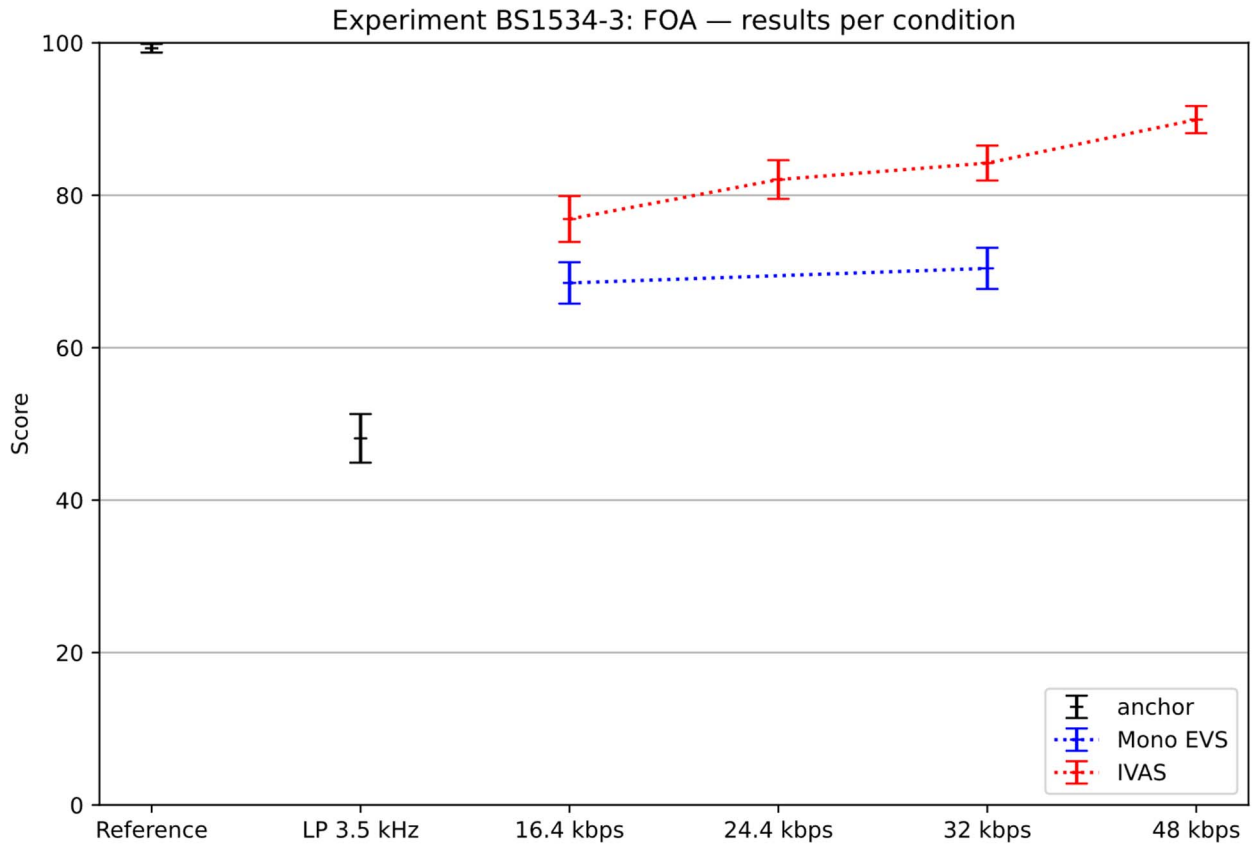


Figure 9.3-15: results per condition for experiment BS1534-3

Table 9.3-22: statistical significance analysis for experiment BS1534-3

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	76.88	19.76	c03	68.49	17.83	8.39	6.273	<0.001	BT
c07	84.23	15.03	c04	70.41	17.73	13.82	10.457	<0.001	BT

9.3.13 Characterization Experiment BS1534-4 (FOA, Generic Audio, 64 - 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-4 evaluates the performance of the IVAS fixed-point implementation for FOA generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-4 was conducted by Panasonic Holdings Corporation.

The aggregated results per condition are listed in Table 9.3-23 and are shown in Figure 9.3-16. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-24. Scores with result WT are highlighted in red.

Table 9.3-23: results per condition for experiment BS1534-4

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	96.23	12.30	1.87
c02	LP 3.5 kHz	168	30.53	24.66	3.76
c03	Mono EVS, 64 kbps	168	55.52	23.86	3.63
c04	Mono EVS, 128 kbps	168	55.44	22.89	3.49
c05	IVAS, 64 kbps	168	79.27	17.92	2.73
c06	IVAS, 96 kbps	168	85.74	17.37	2.65
c07	IVAS, 128 kbps	168	88.51	15.37	2.34
c08	IVAS, 256 kbps	168	88.51	16.24	2.47

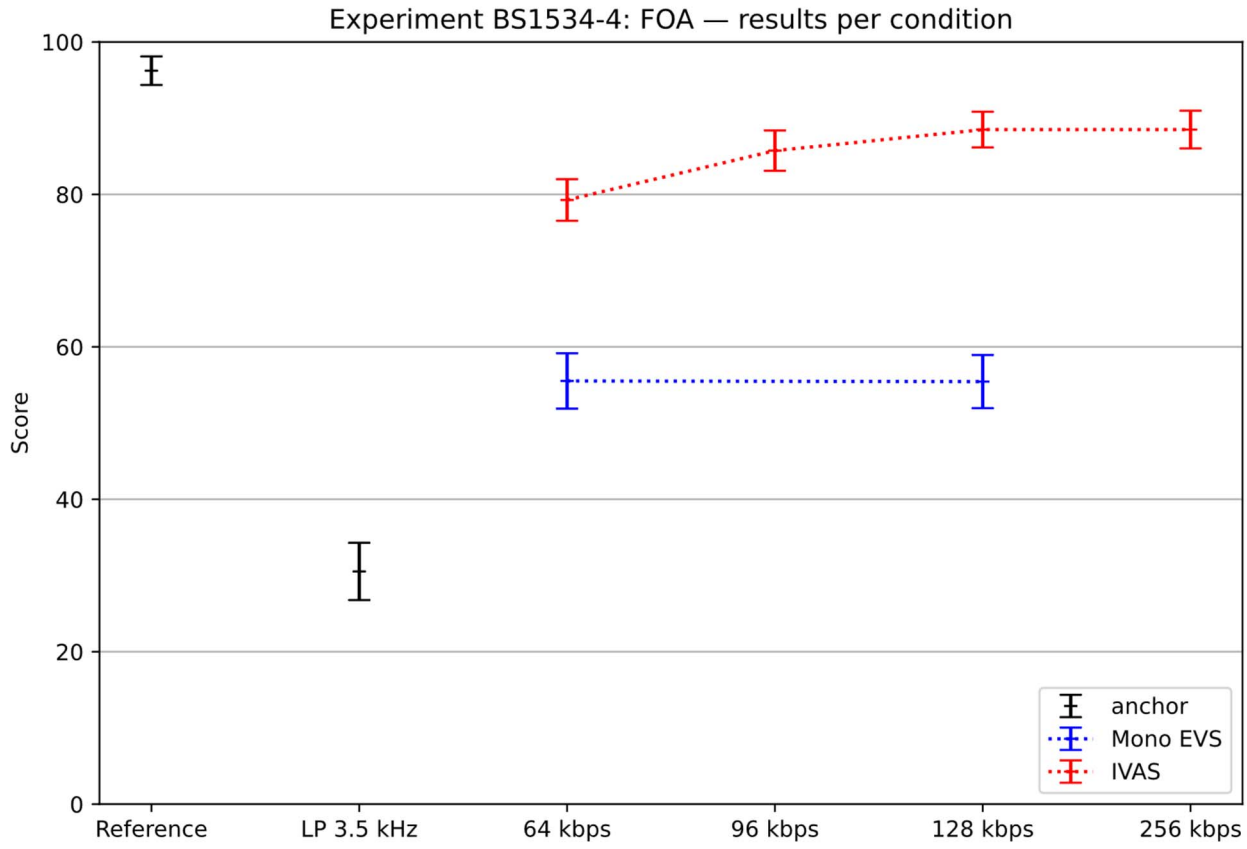


Figure 9.3-16: results per condition for experiment BS1534-4

Table 9.3-24: statistical significance analysis for experiment BS1534-4

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	79.27	17.92	c03	55.52	23.86	23.74	12.408	<0.001	BT
c07	88.51	15.37	c04	55.44	22.89	33.07	17.972	<0.001	BT

9.3.14 Characterization Experiment BS1534-5 (HOA3, Generic Audio, 64 - 256 kbps, 7.1+4 Loudspeaker Presentation)

Characterization Experiment BS1534-5 evaluates the performance of the IVAS fixed-point implementation for HOA3 generic audio content with 7.1+4 loudspeaker presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-5 was conducted by Dolby Laboratories, Inc.

The aggregated results per condition are listed in Table 9.3-25 and are shown in Figure 9.3-17. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.3-26. Scores with result WT are highlighted in red.

Table 9.3-25: results per condition for experiment BS1534-5

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.89	0.84	0.13
c02	LP 3.5 kHz	168	35.19	9.75	1.49
c03	Mono EVS, 64 kbps	168	58.90	10.00	1.52
c04	Mono EVS, 128 kbps	168	58.79	9.62	1.47
c05	IVAS, 64 kbps	168	68.99	9.25	1.41
c06	IVAS, 96 kbps	168	70.70	9.01	1.37
c07	IVAS, 128 kbps	168	72.16	8.06	1.23
c08	IVAS, 256 kbps	168	76.05	7.87	1.20

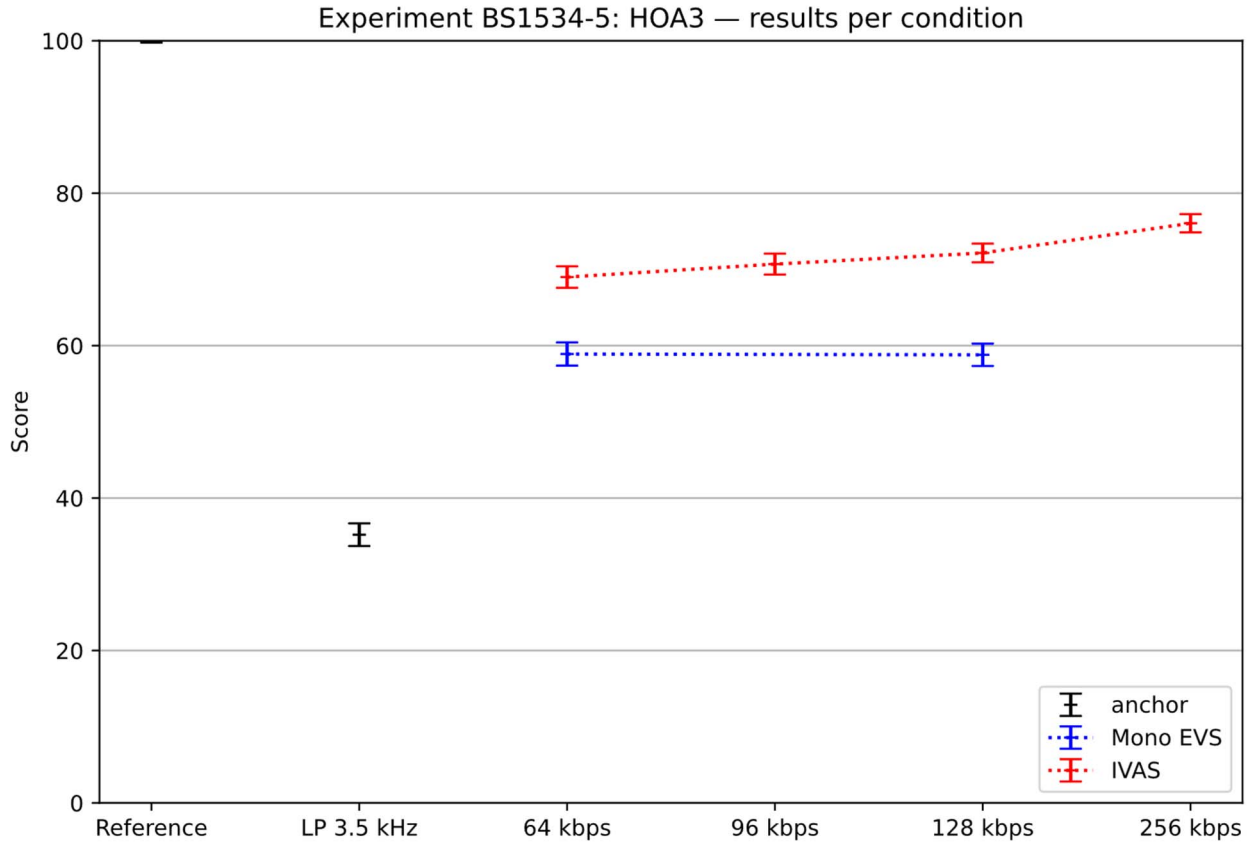


Figure 9.3-17: results per condition for experiment BS1534-5

Table 9.3-26: statistical significance analysis for experiment BS1534-5

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	68.99	9.25	c03	58.90	10.00	10.10	19.312	<0.001	BT
c07	72.16	8.06	c04	58.79	9.62	13.36	24.389	<0.001	BT

9.3.15 Conclusions on IVAS Performance for Scene Based Audio (SBA, Ambisonics)

Taking the results of the previous clauses in 9.3 into account, the following conclusions can be drawn for IVAS scene-based audio operating modes:

During selection phase, IVAS SBA coding was extensively compared against EVS multi-mono. For all tested content types (clean speech, noisy speech, generic audio) and conditions incl. DTX and FER, IVAS Ambisonics has fulfilled the selection criterion: Better than EVS multi-mono at the same bitrate or not worse than EVS multi-mono at the next higher bitrate. For clean and noisy speech content, IVAS FOA outperforms 4x EVS significantly for bitrates up to (including) 64 kbps. For generic audio, this range extends up to 128 kbps. For bitrates above 128 kbps, the performance of IVAS Ambisonics is equivalent to EVS multi-mono and reaches transparency.

During characterization phase, IVAS Ambisonics was evaluated for FOA and HOA3 against EVS mono in a series of BS.1534 tests. Across all tested bitrates, IVAS significantly outperformed EVS mono at the same bitrate, with a mean improvement across items and listeners of up to 33 MUSHRA points (128 kbps) and maintaining the “excellent” quality range from 24kbps to 256kbps on headphone playback (with the possible exception of 64 kbps, which yielded a result slightly below 80, but with a 95% confidence interval crossing 80 MUSHRA points). On the 7.1.4 playback test, IVAS maintained a “good” quality range on the MUSHRA scale from 64 to 256 kbps, against an HOA3 uncoded reference.

Additionally, the performance of the IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for FOA, HOA2 and HOA3 in a series of P.800 tests. After adjustment of the significance level to account for multiple comparisons (Bonferroni method), all 54 tested operating points were statistically equivalent.

Overall, IVAS Ambisonics reached and maintained “good” MOS scores (>= 4.0) starting from 32 kbps for both noisy speech and mixed/music content. A level 1 IVAS encoder already reaches an “very good” MOS rating of approx. 4.4.

9.4 Objects (Independent Streams with Metadata, ISM)

9.4.1 Overview

In Selection phase, four experiments have been conducted to evaluate the performance of the IVAS codec with object content. While the experiments P800-6 and P800-7 were conducted as P.800 DCR tests on 1 or 2 object content, the experiments BS1534-6a and BS1534-6b were conducted as BS.1534 test on 3 or 4 object content. All experiments were conducted using headphone presentation.

- Selection Experiment P800-6: Clean speech, 1 object under clean and impaired channel conditions, DTX off and on, headphone presentation
- Selection Experiment P800-7: Clean speech, 2 objects under clean and impaired channel conditions, DTX off and on, headphone presentation
- Selection Experiment BS1534-6a: Generic audio, 3 objects, 48, 64 and 96 kbps, headphone presentation
- Selection Experiment BS1534-6b: Generic audio, 4 objects, 96, 128 and 256 kbps, headphone presentation

In Characterization phase, seven additional experiments have been conducted. While the experiments P800-9, P800-10, P800-11 and P800-22 were conducted as P.800 DCR tests, the experiments BS1534-10, BS1534-11 and BS1534-12 were conducted as BS.1534 tests.

- Characterization Experiment P800-9: Speech and mixed/music, 1-2 objects under clean channel conditions, headphone presentation
- Characterization Experiment P800-10: Speech and mixed/music, 3-4 objects under clean channel conditions, headphone presentation
- Characterization Experiment P800-11: Speech and mixed/music, 1-4 objects under impaired channel conditions, DTX on, headphone presentation
- Characterization Experiment P800-22: Speech and mixed/music, 1-2 objects under impaired channel conditions using JBM, headphone presentation

- Characterization Experiment BS1534-10: Generic audio, 1-2 objects, 64 – 256 kbps, headphone presentation
- Characterization Experiment BS1534-11: Generic audio, 3-4 objects, 24.4 – 64 kbps, headphone presentation
- Characterization Experiment BS1534-12: Generic audio, 3-4 objects, 64 – 256 kbps, headphone presentation

9.4.2 Selection Experiment P800-6 (1 Object, Clean speech, Headphone Presentation)

Selection Experiment P800-6 evaluates IVAS for 1 object clean speech, under clean and impaired channel conditions, DTX off and on using headphone presentation. See Annex C.6 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-6 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.4-1: Statistical overview on the results of P800-6

		Type	CuT						EVS Reference			Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#												
a	c24	1	13.2	Off		NWT	4.49	0.77	c11	4.46	0.75	0.48	NWT	PASS
	c25	1	16.4	Off		NWT	4.44	0.73	c12	4.53	0.72	-1.1	NWT	PASS
	c26	1	24.4	Off		NWT	4.58	0.67	c13	4.61	0.6	-0.49	NWT	PASS
	c27	1	32	Off		NWT	4.62	0.68	c14	4.59	0.61	0.48	NWT	PASS
	c28	1	48	Off		NWT	4.58	0.65	c15	4.63	0.62	-0.67	NWT	PASS
	c29	1	64	Off		NWT	4.57	0.6	c16	4.7	0.56	-2.18	WT	FAIL
	c30	1	13.2	Off	5%	NWT	3.67	1.14	c17	3.56	1.15	0.96	NWT	PASS
	c31	1	16.4	Off	5%	NWT	3.49	1.13	c18	3.74	1.07	-2.15	WT	FAIL
	c32	1	24.4	Off	5%	NWT	3.82	1.07	c19	3.77	1.06	0.49	NWT	PASS
	c33	1	32	Off	5%	NWT	3.8	1.08	c20	3.85	1.04	-0.45	NWT	PASS
	c34	1	13.2	On		NWT	4.4	0.86	c21	4.44	0.73	-0.52	NWT	PASS
	c35	1	16.4	On		NWT	4.45	0.74	c22	4.51	0.69	-0.81	NWT	PASS
	c36	1	24.4	On		NWT	4.54	0.63	c23	4.66	0.56	-1.94	WT	FAIL
c	c24	1	13.2	Off		NWT	4.23	0.71	c11	4.14	0.76	1.09	NWT	PASS
	c25	1	16.4	Off		NWT	4.34	0.69	c12	4.44	0.62	-1.45	NWT	PASS
	c26	1	24.4	Off		NWT	4.49	0.66	c13	4.52	0.6	-0.42	NWT	PASS
	c27	1	32	Off		NWT	4.52	0.59	c14	4.51	0.55	0.1	NWT	PASS
	c28	1	48	Off		NWT	4.48	0.64	c15	4.59	0.6	-1.69	WT	FAIL
	c29	1	64	Off		NWT	4.57	0.59	c16	4.58	0.61	-0.17	NWT	PASS
	c30	1	13.2	Off	5%	NWT	3.4	0.82	c17	3.43	0.89	-0.37	NWT	PASS
	c31	1	16.4	Off	5%	NWT	3.44	0.89	c18	3.48	0.96	-0.35	NWT	PASS
	c32	1	24.4	Off	5%	NWT	3.58	0.89	c19	3.68	0.89	-1.07	NWT	PASS
	c33	1	32	Off	5%	NWT	3.65	0.88	c20	3.55	0.88	1.08	NWT	PASS
	c34	1	13.2	On		NWT	4.26	0.69	c21	4.29	0.68	-0.52	NWT	PASS
	c35	1	16.4	On		NWT	4.46	0.63	c22	4.31	0.7	2.21	BT	EXCEED
	c36	1	24.4	On		NWT	4.56	0.58	c23	4.6	0.61	-0.7	NWT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.4-2: Summary of the results of P800-6

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c24	13.2	Off		NWT C11	PASS
	c25	16.4	Off		NWT C12	PASS
	c26	24.4	Off		NWT C13	PASS
	c27	32	Off		NWT C14	PASS
	c28	48	Off		NWT C15	PASS
	c29	64	Off		NWT C16	FAIL
	c30	13.2	Off	5%	NWT C17	PASS
	c31	16.4	Off	5%	NWT C18	FAIL
	c32	24.4	Off	5%	NWT C19	PASS
	c33	32	Off	5%	NWT C20	PASS
	c34	13.2	On		NWT C21	PASS
	c35	16.4	On		NWT C22	PASS
	c36	24.4	On		NWT C23	FAIL
c	c24	13.2	Off		NWT C11	PASS
	c25	16.4	Off		NWT C12	PASS
	c26	24.4	Off		NWT C13	PASS
	c27	32	Off		NWT C14	PASS
	c28	48	Off		NWT C15	FAIL

	c29	64	Off		NWT C16	PASS
	c30	13.2	Off	5%	NWT C17	PASS
	c31	16.4	Off	5%	NWT C18	PASS
	c32	24.4	Off	5%	NWT C19	PASS
	c33	32	Off	5%	NWT C20	PASS
	c34	13.2	On		NWT C21	PASS
	c35	16.4	On		NWT C22	EXCEED
	c36	24.4	On		NWT C23	PASS

The following diagrams show the results for a range of conditions from experiment P800-6 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c11 – c16 for EVS conditions and c24 – c29 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c17 – c21 for EVS conditions and c30 – c34 for IVAS conditions.

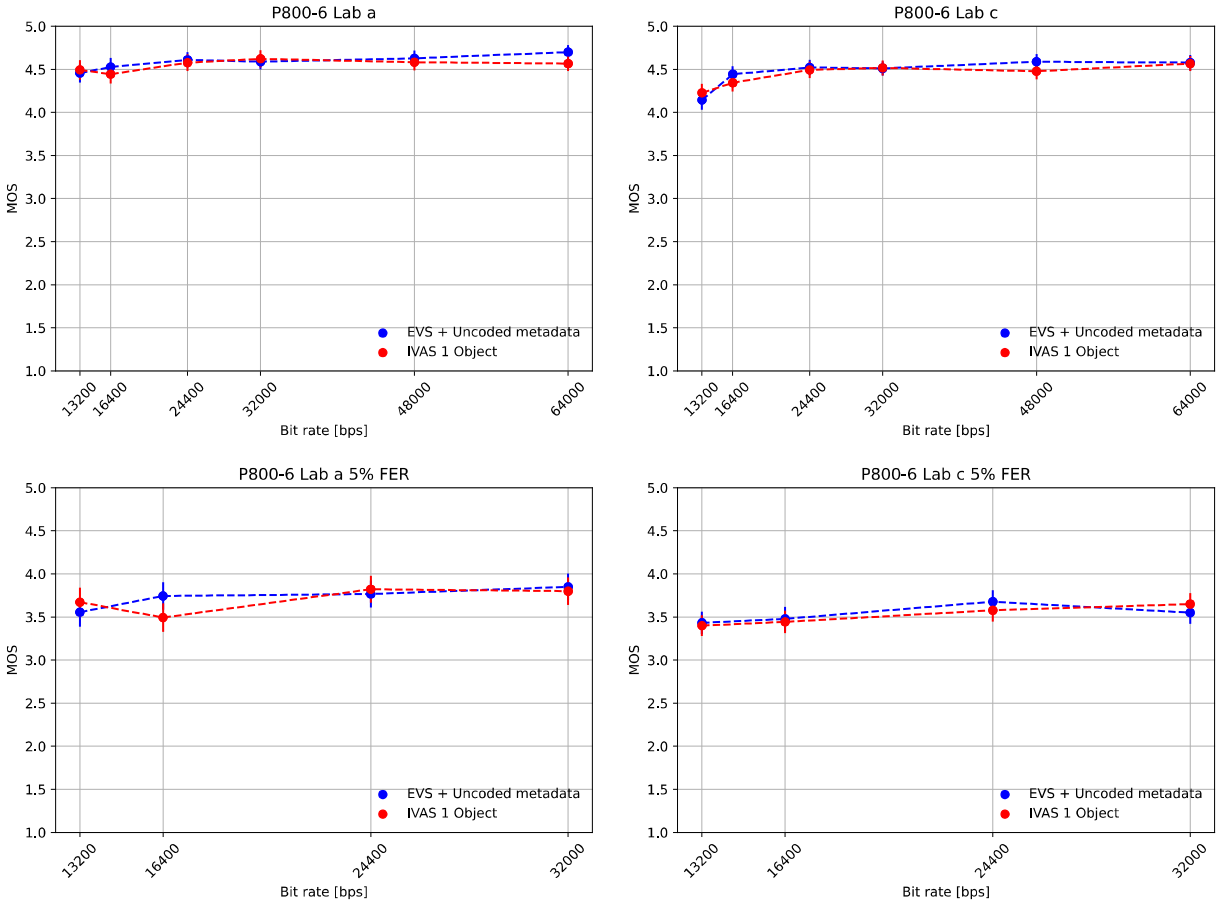


Figure 9.4-1: P800-6 (1 object, clean speech, headphone presentation) rate distortion curves for clean and impaired channel conditions

9.4.3 Selection Experiment P800-7 (2 Objects, Clean speech, Headphone Presentation)

Selection Experiment P800-7 evaluates IVAS for 2 objects, clean speech, under clean and impaired channel conditions, DTX off and on using headphone presentation. See Annex C.7 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-7 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.4-3: Statistical overview on the results of P800-7

		Type	CuT						EVS Reference				Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#													
a	c24	1	16.4	Off		NWT	3.08	1.14	c11	2 x 8.0	2.63	1.06	3.87	BT	EXCEED
	c25	1	24.4	Off		NWT	4.42	0.68	c12	2 x 13.2	4.41	0.73	0.15	NWT	PASS
	c26	1	32	Off		NWT	4.59	0.59	c13	2 x 16.4	4.51	0.72	1.2	NWT	PASS
	c27	1	48	Off		NWT	4.57	0.58	c14	2 x 24.4	4.58	0.66	-0.17	NWT	PASS
	c28	1	64	Off		NWT	4.55	0.66	c15	2 x 32	4.5	0.74	0.67	NWT	PASS
	c29	1	16.4	Off	5%	NWT	2.79	1.16	c16	2 x 8.0	2.46	1.09	2.76	BT	EXCEED
	c30	1	24.4	Off	5%	NWT	3.92	1.04	c17	2 x 13.2	3.87	0.91	0.54	NWT	PASS
	c31	1	32	Off	5%	NWT	4.05	0.84	c18	2 x 16.4	3.94	0.91	1.2	NWT	PASS
	c32	1	48	Off	5%	NWT	4.07	0.83	c19	2 x 24.4	4.07	0.89	0.06	NWT	PASS
	c33	1	16.4	On		NWT	3.03	1.13	c20	2 x 8.0	2.76	1.12	2.3	BT	EXCEED
	c34	1	24.4	On		NWT	4.45	0.69	c21	2 x 13.2	4.36	0.74	1.25	NWT	PASS
	c35	1	32	On		NWT	4.58	0.67	c22	2 x 16.4	4.47	0.7	1.47	NWT	PASS
	c36	1	48	On		NWT	4.54	0.65	c23	2 x 24.4	4.54	0.6	0.08	NWT	PASS
d	c24	1	16.4	Off		NWT	3.96	0.89	c11	2 x 8.0	3.19	0.96	7.79	BT	EXCEED
	c25	1	24.4	Off		NWT	4.66	0.54	c12	2 x 13.2	4.64	0.55	0.38	NWT	PASS
	c26	1	32	Off		NWT	4.7	0.52	c13	2 x 16.4	4.63	0.59	1.23	NWT	PASS
	c27	1	48	Off		NWT	4.7	0.51	c14	2 x 24.4	4.76	0.44	-1.12	NWT	PASS
	c28	1	64	Off		NWT	4.74	0.49	c15	2 x 32	4.71	0.49	0.54	NWT	PASS
	c29	1	16.4	Off	5%	NWT	3.47	0.97	c16	2 x 8.0	2.89	1.1	5.34	BT	EXCEED
	c30	1	24.4	Off	5%	NWT	4.26	0.81	c17	2 x 13.2	4.23	0.83	0.27	NWT	PASS
	c31	1	32	Off	5%	NWT	4.32	0.75	c18	2 x 16.4	4.33	0.79	-0.14	NWT	PASS
	c32	1	48	Off	5%	NWT	4.32	0.78	c19	2 x 24.4	4.43	0.71	-1.41	NWT	PASS
	c33	1	16.4	On		NWT	3.86	0.91	c20	2 x 8.0	3.32	0.91	5.56	BT	EXCEED
	c34	1	24.4	On		NWT	4.63	0.6	c21	2 x 13.2	4.64	0.53	-0.1	NWT	PASS
	c35	1	32	On		NWT	4.64	0.57	c22	2 x 16.4	4.62	0.57	0.36	NWT	PASS
	c36	1	48	On		NWT	4.66	0.54	c23	2 x 24.4	4.73	0.47	-1.35	NWT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.4-4: Summary of the results of P800-7

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c24	16.4	Off		NWT C11	EXCEED
	c25	24.4	Off		NWT C12	PASS
	c26	32	Off		NWT C13	PASS
	c27	48	Off		NWT C14	PASS
	c28	64	Off		NWT C15	PASS
	c29	16.4	Off	5%	NWT C16	EXCEED
	c30	24.4	Off	5%	NWT C17	PASS
	c31	32	Off	5%	NWT C18	PASS
	c32	48	Off	5%	NWT C19	PASS
	c33	16.4	On		NWT C20	EXCEED
	c34	24.4	On		NWT C21	PASS
	c35	32	On		NWT C22	PASS
	c36	48	On		NWT C23	PASS
d	c24	16.4	Off		NWT C11	EXCEED
	c25	24.4	Off		NWT C12	PASS
	c26	32	Off		NWT C13	PASS
	c27	48	Off		NWT C14	PASS
	c28	64	Off		NWT C15	PASS
	c29	16.4	Off	5%	NWT C16	EXCEED
	c30	24.4	Off	5%	NWT C17	PASS
	c31	32	Off	5%	NWT C18	PASS
	c32	48	Off	5%	NWT C19	PASS
	c33	16.4	On		NWT C20	EXCEED
	c34	24.4	On		NWT C21	PASS
	c35	32	On		NWT C22	PASS
	c36	48	On		NWT C23	PASS

The following diagrams show the results for a range of conditions from experiment P800-7 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c11 – c15 for EVS conditions and c24 – c29 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c16 – c19 for EVS conditions and c29 – c32 for IVAS conditions.

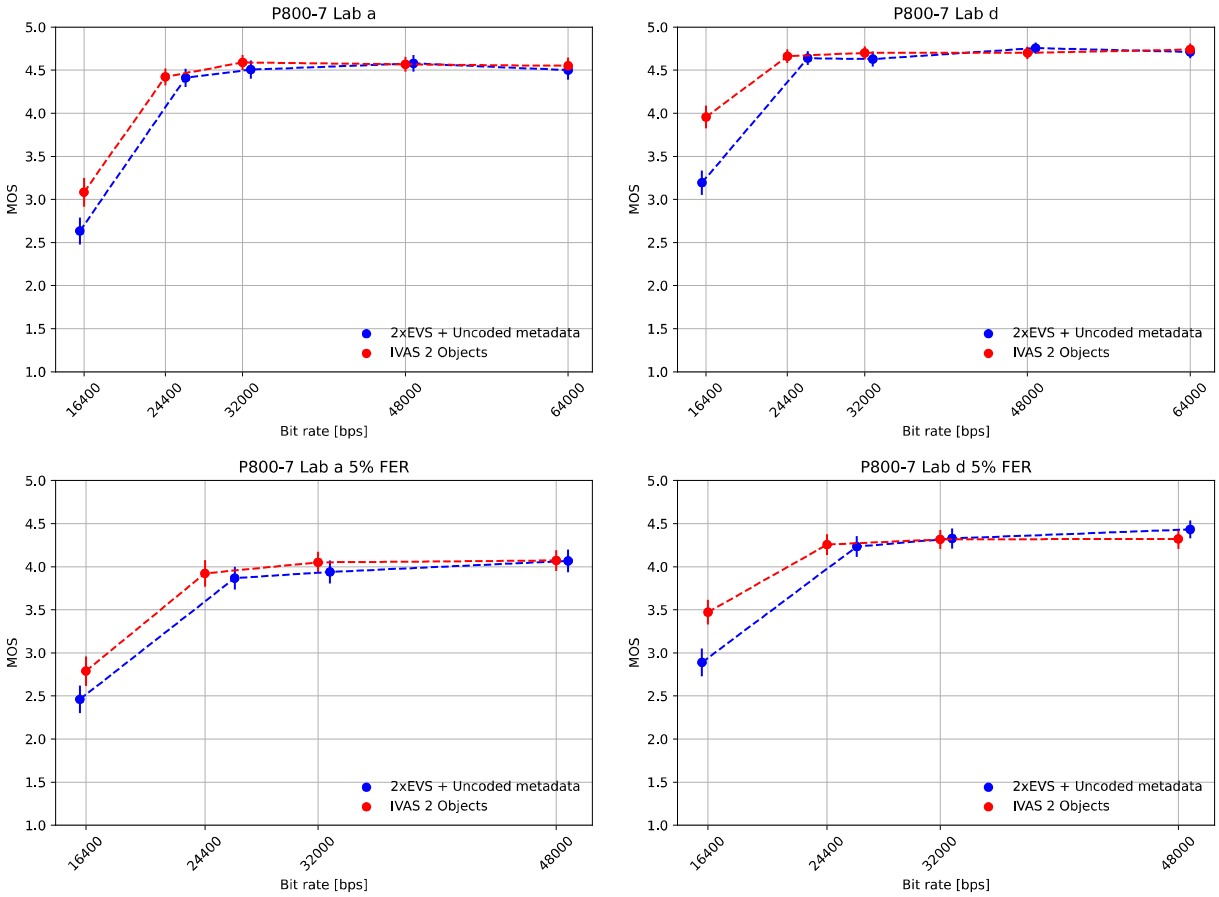


Figure 9.4-2: P800-7 (2 objects, clean speech, headphone presentation) rate distortion curves for clean and impaired channel conditions

9.4.4 Selection Experiment BS1534-6a (3 Objects, Generic Audio, 48, 64 and 96 kbps, Headphone Presentation)

Selection Experiment BS1534-6a evaluates IVAS for 3 objects, generic audio at 48, 64 and 96 kbps using headphone presentation. See Annex C.20 for details.

The averaged results per condition for experiment BS1534-6a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c08).

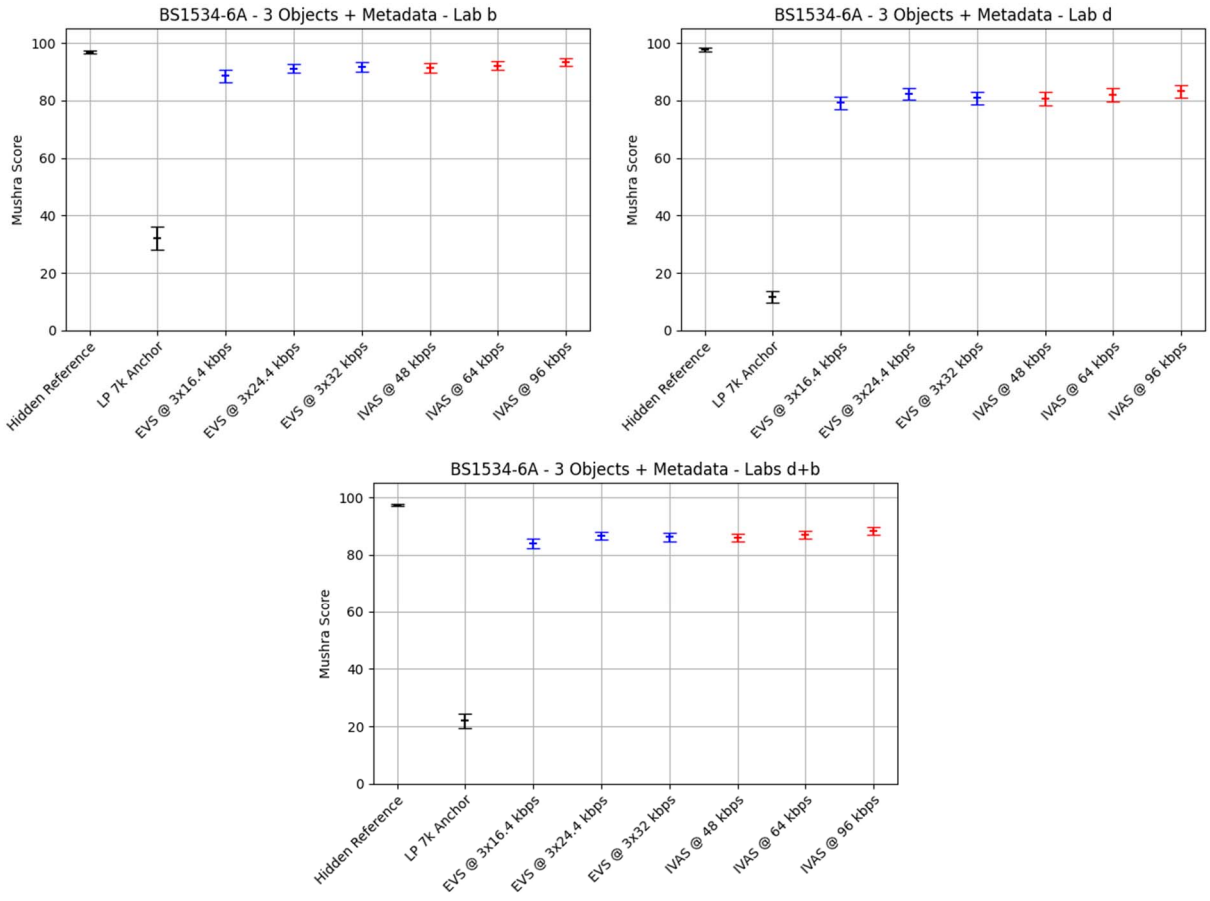


Figure 9.4-3: BS1534-6a (3 objects, generic audio, 48, 64 and 96 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-6a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.4-5: Statistical overview on the results of BS1534-6a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c06	1	48	NWT	91.4	10.2	c03	3x16.4	88.6	14.6	1.97	BT	EXCEED
	c07	1	64	NWT	92.2	9.2	c04	3x24.4	91.2	10.8	0.91	NWT	PASS
	c08	1	96	NWT	93.4	8.4	c05	3x32	91.7	11.1	1.59	NWT	PASS
d	c06	1	48	NWT	80.7	15.2	c03	3x16.4	79.2	13.9	0.94	NWT	PASS
	c07	1	64	NWT	82	15.1	c04	3x24.4	82.4	13.7	-0.26	NWT	PASS
	c08	1	96	NWT	83.2	14.4	c05	3x32	80.8	14.3	1.5	NWT	PASS
d+b	c06	1	48	NWT	86	14	c03	3x16.4	83.9	15	1.88	BT	EXCEED
	c07	1	64	NWT	87.1	13.5	c04	3x24.4	86.8	13.1	0.28	NWT	PASS
	c08	1	96	NWT	88.3	12.9	c05	3x32	86.2	13.9	1.97	BT	EXCEED

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.4-6: Summary of the results of BS1534-6a

Lab	Cond.	Bitrate	ToR	Status
b	c06	48	NWT c03	EXCEED
	c07	64	NWT c04	PASS
	c08	96	NWT c05	PASS
d	c06	48	NWT c03	PASS
	c07	64	NWT c04	PASS
	c08	96	NWT c05	PASS

d+b	c06	48	NWT c03	EXCEED
	c07	64	NWT c04	PASS
	c08	96	NWT c05	EXCEED

9.4.5 Selection Experiment BS1534-6b (4 Objects, Generic Audio, 96, 128 and 256 kbps, Headphone Presentation)

Selection Experiment BS1534-6b evaluates IVAS for 4 objects, generic audio, at 96, 128 and 256 kbps using headphone presentation. See Annex C.21. for details.

The averaged results per condition for experiment BS1534-6b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c08).

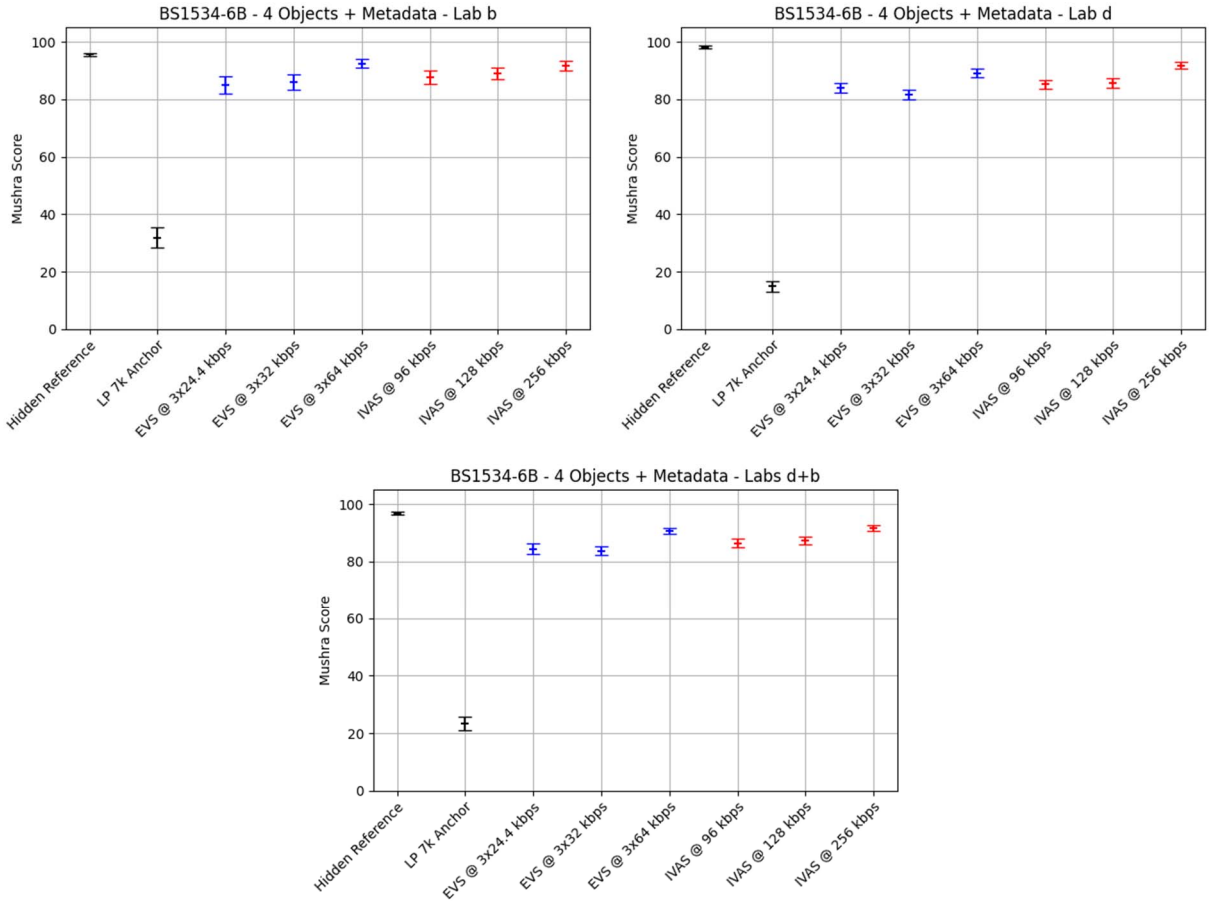


Figure 9.4-4: BS1534-6b (4 objects, generic audio, 96, 128 and 192 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-6b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.4-7: Statistical overview on the results of BS1534-6b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c06	1	96	NWT	87.7	15.7	c03	4x24.4	84.9	19.4	1.44	NWT	PASS
	c07	1	128	NWT	89	14	c04	4x32	86	17.8	1.72	BT	EXCEED
	c08	1	256	NWT	91.7	11	c05	4x64	92.4	9.9	-0.64	NWT	PASS
d	c06	1	96	NWT	85.2	10.7	c03	4x24.4	84.1	11.3	0.97	NWT	PASS
	c07	1	128	NWT	85.7	10.4	c04	4x32	81.6	11	3.51	BT	EXCEED
	c08	1	256	NWT	91.8	8.1	c05	4x64	89.1	9.3	2.82	BT	EXCEED
d+b	c06	1	96	NWT	86.5	13.5	c03	4x24.4	84.5	15.9	1.73	BT	EXCEED
	c07	1	128	NWT	87.4	12.4	c04	4x32	83.8	14.9	3.35	BT	EXCEED
	c08	1	256	NWT	91.7	9.6	c05	4x64	90.8	9.7	1.31	NWT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.4-8: Summary of the results of BS1534-6b

Lab	Cond.	Bitrate	ToR	Status
b	c06	96	NWT c03	PASS
	c07	128	NWT c04	EXCEED
	c08	256	NWT c05	PASS
d	c06	96	NWT c03	PASS
	c07	128	NWT c04	EXCEED
	c08	256	NWT c05	EXCEED
d+b	c06	96	NWT c03	EXCEED
	c07	128	NWT c04	EXCEED
	c08	256	NWT c05	PASS

9.4.6 Characterization Experiment P800-9 (1-2 Objects, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-9 evaluates the performance of the IVAS fixed-point and floating-point implementation for 1-2 objects for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-9 was conducted by Mesaqin.com (language: MAN).

The listening lab has reported that for each panel, up to seven instead of five participants participated. The five subjects with the highest PCC were selected for result submission.

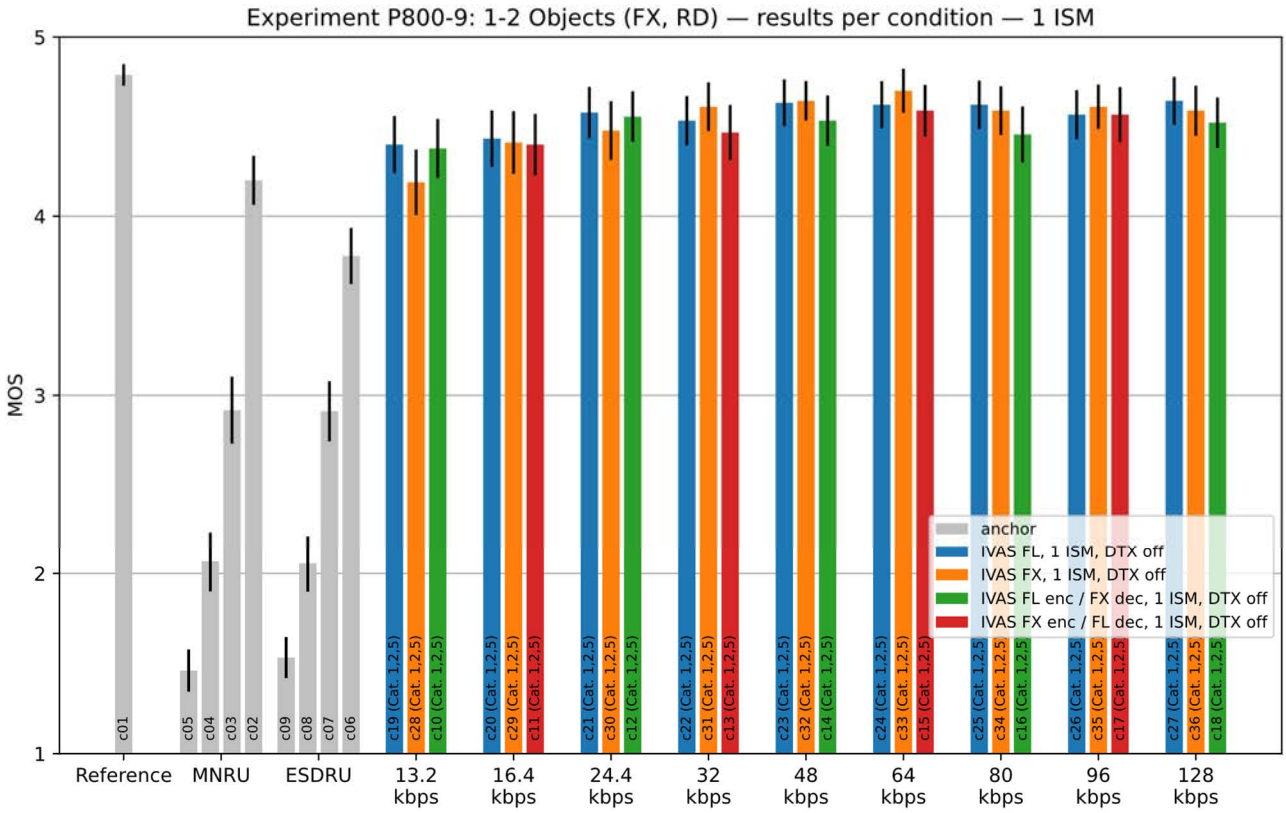
The aggregated results per condition are listed in Table 9.4-9 and are shown in Figure 9.4-5. The results aggregated over FX/FL are shown in Figure 9.4-6. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-10. Significant different MOS values (BT or WT) are highlighted in red.

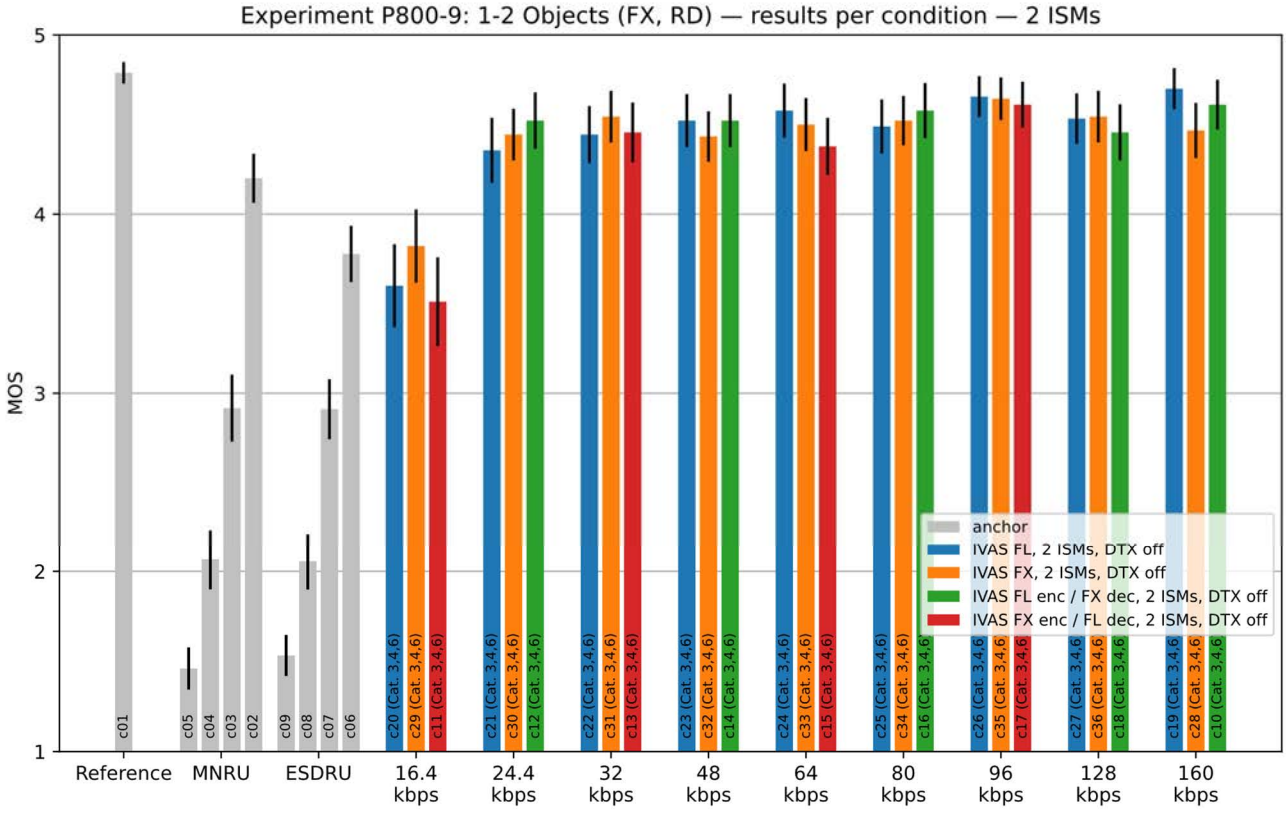
Table 9.4-9: results per condition for experiment P800-9

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.79	0.41	0.06
c02	MNRU Q = 36 dB	180	4.20	0.93	0.14
c03	MNRU Q = 30 dB	180	2.92	1.27	0.19
c04	MNRU Q = 24 dB	180	2.07	1.14	0.17
c05	MNRU Q = 18 dB	180	1.46	0.79	0.12
c06	ESDRU α=0.7	180	3.78	1.07	0.16
c07	ESDRU α=0.5	180	2.91	1.14	0.17
c08	ESDRU α=0.3	180	2.06	1.07	0.16
c09	ESDRU α=0.1	180	1.53	0.77	0.11
c10 (Cat. 1,2,5)	IVAS FL enc / FX dec, 13.2 kbps, 1 ISM, DTX off	90	4.38	0.79	0.17
c10 (Cat. 3,4,6)	IVAS FL enc / FX dec, 160 kbps, 2 ISMs, DTX off	90	4.61	0.67	0.14
c11 (Cat. 1,2,5)	IVAS FX enc / FL dec, 16.4 kbps, 1 ISM, DTX off	90	4.40	0.82	0.17
c11 (Cat. 3,4,6)	IVAS FX enc / FL dec, 16.4 kbps, 2 ISMs, DTX off	90	3.51	1.18	0.25
c12 (Cat. 1,2,5)	IVAS FL enc / FX dec, 24.4 kbps, 1 ISM, DTX off	90	4.56	0.67	0.14

c12 (Cat. 3,4,6)	IVAS FL enc / FX dec, 24.4 kbps, 2 ISMs, DTX off	90	4.52	0.75	0.16
c13 (Cat. 1,2,5)	IVAS FX enc / FL dec, 32 kbps, 1 ISM, DTX off	90	4.47	0.74	0.15
c13 (Cat. 3,4,6)	IVAS FX enc / FL dec, 32 kbps, 2 ISMs, DTX off	90	4.46	0.80	0.17
c14 (Cat. 1,2,5)	IVAS FL enc / FX dec, 48 kbps, 1 ISM, DTX off	90	4.53	0.67	0.14
c14 (Cat. 3,4,6)	IVAS FL enc / FX dec, 48 kbps, 2 ISMs, DTX off	90	4.52	0.71	0.15
c15 (Cat. 1,2,5)	IVAS FX enc / FL dec, 64 kbps, 1 ISM, DTX off	90	4.59	0.69	0.14
c15 (Cat. 3,4,6)	IVAS FX enc / FL dec, 64 kbps, 2 ISMs, DTX off	90	4.38	0.76	0.16
c16 (Cat. 1,2,5)	IVAS FL enc / FX dec, 80 kbps, 1 ISM, DTX off	90	4.46	0.75	0.16
c16 (Cat. 3,4,6)	IVAS FL enc / FX dec, 80 kbps, 2 ISMs, DTX off	90	4.58	0.73	0.15
c17 (Cat. 1,2,5)	IVAS FX enc / FL dec, 96 kbps, 1 ISM, DTX off	90	4.57	0.73	0.15
c17 (Cat. 3,4,6)	IVAS FX enc / FL dec, 96 kbps, 2 ISMs, DTX off	90	4.61	0.61	0.13
c18 (Cat. 1,2,5)	IVAS FL enc / FX dec, 128 kbps, 1 ISM, DTX off	90	4.52	0.67	0.14
c18 (Cat. 3,4,6)	IVAS FL enc / FX dec, 128 kbps, 2 ISMs, DTX off	90	4.46	0.75	0.16
c19 (Cat. 1,2,5)	IVAS FL, 13.2 kbps, 1 ISM, DTX off	90	4.40	0.76	0.16
c19 (Cat. 3,4,6)	IVAS FL, 160 kbps, 2 ISMs, DTX off	90	4.70	0.55	0.12
c20 (Cat. 1,2,5)	IVAS FL, 16.4 kbps, 1 ISM, DTX off	90	4.43	0.75	0.16
c20 (Cat. 3,4,6)	IVAS FL, 16.4 kbps, 2 ISMs, DTX off	90	3.60	1.11	0.23
c21 (Cat. 1,2,5)	IVAS FL, 24.4 kbps, 1 ISM, DTX off	90	4.58	0.69	0.14
c21 (Cat. 3,4,6)	IVAS FL, 24.4 kbps, 2 ISMs, DTX off	90	4.36	0.86	0.18
c22 (Cat. 1,2,5)	IVAS FL, 32 kbps, 1 ISM, DTX off	90	4.53	0.66	0.14
c22 (Cat. 3,4,6)	IVAS FL, 32 kbps, 2 ISMs, DTX off	90	4.44	0.77	0.16
c23 (Cat. 1,2,5)	IVAS FL, 48 kbps, 1 ISM, DTX off	90	4.63	0.63	0.13
c23 (Cat. 3,4,6)	IVAS FL, 48 kbps, 2 ISMs, DTX off	90	4.52	0.71	0.15
c24 (Cat. 1,2,5)	IVAS FL, 64 kbps, 1 ISM, DTX off	90	4.62	0.63	0.13
c24 (Cat. 3,4,6)	IVAS FL, 64 kbps, 2 ISMs, DTX off	90	4.58	0.72	0.15
c25 (Cat. 1,2,5)	IVAS FL, 80 kbps, 1 ISM, DTX off	90	4.62	0.65	0.14
c25 (Cat. 3,4,6)	IVAS FL, 80 kbps, 2 ISMs, DTX off	90	4.49	0.72	0.15
c26 (Cat. 1,2,5)	IVAS FL, 96 kbps, 1 ISM, DTX off	90	4.57	0.65	0.14
c26 (Cat. 3,4,6)	IVAS FL, 96 kbps, 2 ISMs, DTX off	90	4.66	0.54	0.11
c27 (Cat. 1,2,5)	IVAS FL, 128 kbps, 1 ISM, DTX off	90	4.64	0.64	0.13
c27 (Cat. 3,4,6)	IVAS FL, 128 kbps, 2 ISMs, DTX off	90	4.53	0.67	0.14
c28 (Cat. 1,2,5)	IVAS FX, 13.2 kbps, 1 ISM, DTX off	90	4.19	0.87	0.18
c28 (Cat. 3,4,6)	IVAS FX, 160 kbps, 2 ISMs, DTX off	90	4.47	0.74	0.15
c29 (Cat. 1,2,5)	IVAS FX, 16.4 kbps, 1 ISM, DTX off	90	4.41	0.83	0.17
c29 (Cat. 3,4,6)	IVAS FX, 16.4 kbps, 2 ISMs, DTX off	90	3.82	0.98	0.20
c30 (Cat. 1,2,5)	IVAS FX, 24.4 kbps, 1 ISM, DTX off	90	4.48	0.78	0.16
c30 (Cat. 3,4,6)	IVAS FX, 24.4 kbps, 2 ISMs, DTX off	90	4.44	0.69	0.14
c31 (Cat. 1,2,5)	IVAS FX, 32 kbps, 1 ISM, DTX off	90	4.61	0.65	0.14
c31 (Cat. 3,4,6)	IVAS FX, 32 kbps, 2 ISMs, DTX off	90	4.54	0.69	0.14
c32 (Cat. 1,2,5)	IVAS FX, 48 kbps, 1 ISM, DTX off	90	4.64	0.53	0.11
c32 (Cat. 3,4,6)	IVAS FX, 48 kbps, 2 ISMs, DTX off	90	4.43	0.67	0.14
c33 (Cat. 1,2,5)	IVAS FX, 64 kbps, 1 ISM, DTX off	90	4.70	0.59	0.12
c33 (Cat. 3,4,6)	IVAS FX, 64 kbps, 2 ISMs, DTX off	90	4.50	0.71	0.15
c34 (Cat. 1,2,5)	IVAS FX, 80 kbps, 1 ISM, DTX off	90	4.59	0.65	0.14
c34 (Cat. 3,4,6)	IVAS FX, 80 kbps, 2 ISMs, DTX off	90	4.52	0.66	0.14
c35 (Cat. 1,2,5)	IVAS FX, 96 kbps, 1 ISM, DTX off	90	4.61	0.59	0.12
c35 (Cat. 3,4,6)	IVAS FX, 96 kbps, 2 ISMs, DTX off	90	4.64	0.57	0.12
c36 (Cat. 1,2,5)	IVAS FX, 128 kbps, 1 ISM, DTX off	90	4.59	0.67	0.14
c36 (Cat. 3,4,6)	IVAS FX, 128 kbps, 2 ISMs, DTX off	90	4.54	0.69	0.14



(a)



(b)

Figure 9.4-5: results per condition for experiment P800-9

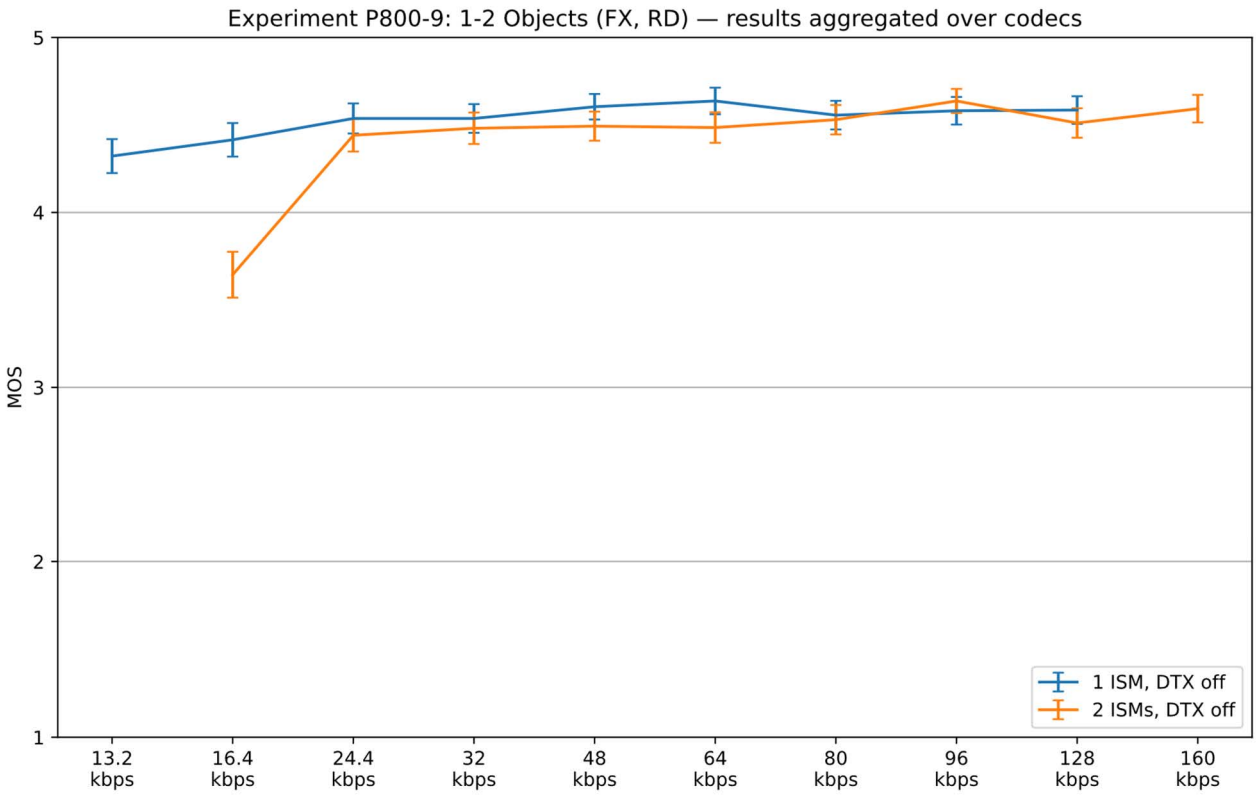


Figure 9.4-6: results aggregated over FX/FL for experiment P800-9

Table 9.4-10: statistical significance analysis for experiment P800-9

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10 (Cat. 1,2,5)	4.38	0.79	c19 (Cat. 1,2,5)	4.40	0.76	-0.02	<0.001	0.575	NWT
c10 (Cat. 3,4,6)	4.61	0.67	c19 (Cat. 3,4,6)	4.70	0.55	-0.09	<0.001	0.835	NWT
c11 (Cat. 1,2,5)	4.40	0.82	c20 (Cat. 1,2,5)	4.43	0.75	-0.03	<0.001	0.611	NWT
c11 (Cat. 3,4,6)	3.51	1.18	c20 (Cat. 3,4,6)	3.60	1.11	-0.09	<0.001	0.698	NWT
c12 (Cat. 1,2,5)	4.56	0.67	c21 (Cat. 1,2,5)	4.58	0.69	-0.02	<0.001	0.586	NWT
c12 (Cat. 3,4,6)	4.52	0.75	c21 (Cat. 3,4,6)	4.36	0.86	0.17	1.373	0.086	NWT
c13 (Cat. 1,2,5)	4.47	0.74	c22 (Cat. 1,2,5)	4.53	0.66	-0.07	<0.001	0.737	NWT
c13 (Cat. 3,4,6)	4.46	0.80	c22 (Cat. 3,4,6)	4.44	0.77	0.01	0.103	0.459	NWT
c14 (Cat. 1,2,5)	4.53	0.67	c23 (Cat. 1,2,5)	4.63	0.63	-0.10	<0.001	0.848	NWT
c14 (Cat. 3,4,6)	4.52	0.71	c23 (Cat. 3,4,6)	4.52	0.71	0.00	<0.001	0.500	NWT
c15 (Cat. 1,2,5)	4.59	0.69	c24 (Cat. 1,2,5)	4.62	0.63	-0.03	<0.001	0.632	NWT
c15 (Cat. 3,4,6)	4.38	0.76	c24 (Cat. 3,4,6)	4.58	0.72	-0.20	<0.001	0.964	WT
c16 (Cat. 1,2,5)	4.46	0.75	c25 (Cat. 1,2,5)	4.62	0.65	-0.17	<0.001	0.943	NWT

c16 (Cat. 3,4,6)	4.58	0.73	c25 (Cat. 3,4,6)	4.49	0.72	0.09	0.820	0.207	NWT
c17 (Cat. 1,2,5)	4.57	0.73	c26 (Cat. 1,2,5)	4.57	0.65	0.00	<0.001	0.500	NWT
c17 (Cat. 3,4,6)	4.61	0.61	c26 (Cat. 3,4,6)	4.66	0.54	-0.04	<0.001	0.699	NWT
c18 (Cat. 1,2,5)	4.52	0.67	c27 (Cat. 1,2,5)	4.64	0.64	-0.12	<0.001	0.892	NWT
c18 (Cat. 3,4,6)	4.46	0.75	c27 (Cat. 3,4,6)	4.53	0.67	-0.08	<0.001	0.765	NWT
c28 (Cat. 1,2,5)	4.19	0.87	c19 (Cat. 1,2,5)	4.40	0.76	-0.21	<0.001	0.957	WT
c28 (Cat. 3,4,6)	4.47	0.74	c19 (Cat. 3,4,6)	4.70	0.55	-0.23	<0.001	0.991	WT
c29 (Cat. 1,2,5)	4.41	0.83	c20 (Cat. 1,2,5)	4.43	0.75	-0.02	<0.001	0.574	NWT
c29 (Cat. 3,4,6)	3.82	0.98	c20 (Cat. 3,4,6)	3.60	1.11	0.22	1.424	0.078	NWT
c30 (Cat. 1,2,5)	4.48	0.78	c21 (Cat. 1,2,5)	4.58	0.69	-0.10	<0.001	0.818	NWT
c30 (Cat. 3,4,6)	4.44	0.69	c21 (Cat. 3,4,6)	4.36	0.86	0.09	0.755	0.226	NWT
c31 (Cat. 1,2,5)	4.61	0.65	c22 (Cat. 1,2,5)	4.53	0.66	0.08	0.802	0.212	NWT
c31 (Cat. 3,4,6)	4.54	0.69	c22 (Cat. 3,4,6)	4.44	0.77	0.10	0.920	0.179	NWT
c32 (Cat. 1,2,5)	4.64	0.53	c23 (Cat. 1,2,5)	4.63	0.63	0.01	0.128	0.449	NWT
c32 (Cat. 3,4,6)	4.43	0.67	c23 (Cat. 3,4,6)	4.52	0.71	-0.09	<0.001	0.806	NWT
c33 (Cat. 1,2,5)	4.70	0.59	c24 (Cat. 1,2,5)	4.62	0.63	0.08	0.859	0.196	NWT
c33 (Cat. 3,4,6)	4.50	0.71	c24 (Cat. 3,4,6)	4.58	0.72	-0.08	<0.001	0.768	NWT
c34 (Cat. 1,2,5)	4.59	0.65	c25 (Cat. 1,2,5)	4.62	0.65	-0.03	<0.001	0.633	NWT
c34 (Cat. 3,4,6)	4.52	0.66	c25 (Cat. 3,4,6)	4.49	0.72	0.03	0.320	0.374	NWT
c35 (Cat. 1,2,5)	4.61	0.59	c26 (Cat. 1,2,5)	4.57	0.65	0.04	0.472	0.319	NWT
c35 (Cat. 3,4,6)	4.64	0.57	c26 (Cat. 3,4,6)	4.66	0.54	-0.01	<0.001	0.558	NWT
c36 (Cat. 1,2,5)	4.59	0.67	c27 (Cat. 1,2,5)	4.64	0.64	-0.05	<0.001	0.713	NWT
c36 (Cat. 3,4,6)	4.54	0.69	c27 (Cat. 3,4,6)	4.53	0.67	0.01	0.108	0.457	NWT

9.4.7 Characterization Experiment P800-10 (3-4 Objects, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-10 evaluates the performance of the IVAS fixed-point and floating-point implementation for 3-4 objects for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-10 was conducted by VoiceAge Corporation (language: FR).

The aggregated results per condition are listed in Table 9.4-11 and are shown in Figure 9.4-7. The results aggregated over FX/FL are shown in Figure 9.4-8. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-12. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.4-11: results per condition for experiment P800-10

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.57	0.71	0.10
c02	MNRU Q = 36 dB	180	4.35	0.83	0.12
c03	MNRU Q = 30 dB	180	3.72	1.15	0.17
c04	MNRU Q = 24 dB	180	2.26	1.24	0.18
c05	MNRU Q = 18 dB	180	1.31	0.69	0.10
c06	ESDRU α=0.7	180	3.86	1.02	0.15
c07	ESDRU α=0.5	180	2.69	1.12	0.17
c08	ESDRU α=0.3	180	2.25	1.16	0.17
c09	ESDRU α=0.1	180	1.58	0.90	0.13
c10	IVAS FL enc / FX dec, 24.4 kbps, DTX off	180	2.29	1.10	0.16
c11	IVAS FX enc / FL dec, 32 kbps, DTX off	180	3.14	1.12	0.17
c12	IVAS FL enc / FX dec, 48 kbps, DTX off	180	4.14	0.90	0.13
c13	IVAS FX enc / FL dec, 64 kbps, DTX off	180	4.24	0.82	0.12
c14	IVAS FL enc / FX dec, 80 kbps, DTX off	180	4.36	0.74	0.11
c15	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.16	0.93	0.14
c16	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.57	0.69	0.10
c17	IVAS FX enc / FL dec, 160 kbps, DTX off	180	4.66	0.58	0.09
c18	IVAS FL enc / FX dec, 192 kbps, DTX off	180	4.56	0.68	0.10
c19	IVAS FL, 24.4 kbps, DTX off	180	2.27	1.14	0.17
c20	IVAS FL, 32 kbps, DTX off	180	3.08	1.12	0.16
c21	IVAS FL, 48 kbps, DTX off	180	4.17	0.85	0.12
c22	IVAS FL, 64 kbps, DTX off	180	4.35	0.79	0.12
c23	IVAS FL, 80 kbps, DTX off	180	4.44	0.66	0.10
c24	IVAS FL, 96 kbps, DTX off	180	4.35	0.76	0.11
c25	IVAS FL, 128 kbps, DTX off	180	4.48	0.74	0.11
c26	IVAS FL, 160 kbps, DTX off	180	4.48	0.73	0.11
c27	IVAS FL, 192 kbps, DTX off	180	4.52	0.69	0.10
c28	IVAS FX, 24.4 kbps, DTX off	180	2.23	1.09	0.16
c29	IVAS FX, 32 kbps, DTX off	180	3.07	1.13	0.17
c30	IVAS FX, 48 kbps, DTX off	180	4.20	0.80	0.12
c31	IVAS FX, 64 kbps, DTX off	180	4.29	0.87	0.13
c32	IVAS FX, 80 kbps, DTX off	180	4.21	0.79	0.12
c33	IVAS FX, 96 kbps, DTX off	180	4.20	0.87	0.13
c34	IVAS FX, 128 kbps, DTX off	180	4.40	0.71	0.10
c35	IVAS FX, 160 kbps, DTX off	180	4.58	0.65	0.10
c36	IVAS FX, 192 kbps, DTX off	180	4.62	0.59	0.09

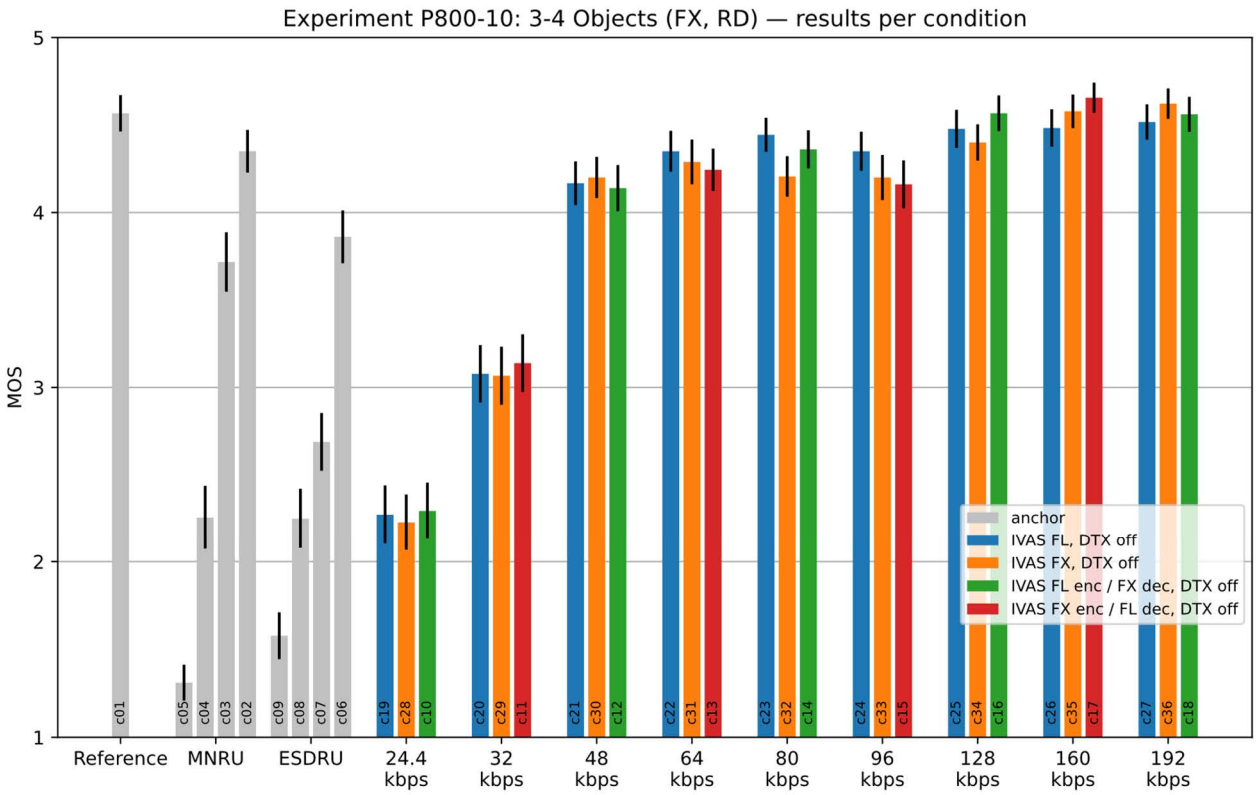


Figure 9.4-7: results per condition for experiment P800-10

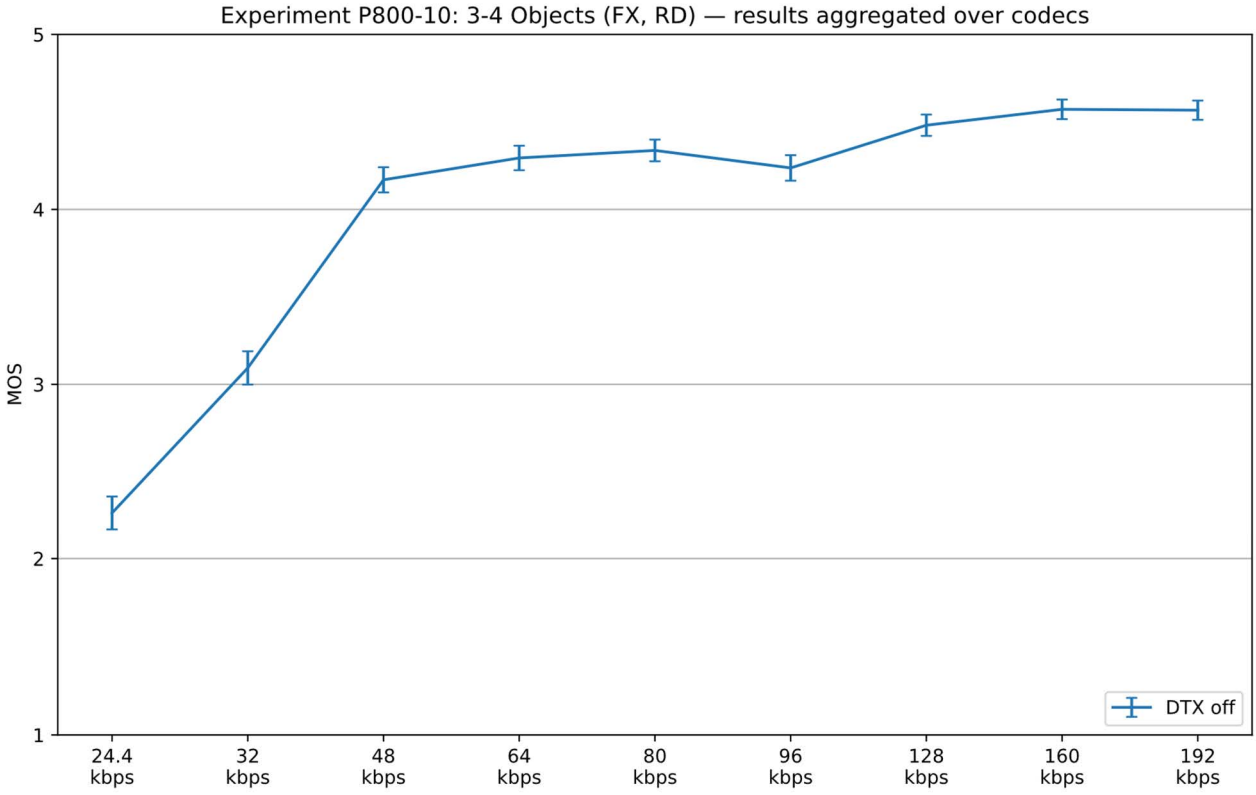


Figure 9.4-8: results aggregated over FX/FL for experiment P800-10

Table 9.4-12: statistical significance analysis for experiment P800-10

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	2.29	1.10	c19	2.27	1.14	0.02	0.186	0.426	NWT
c11	3.14	1.12	c20	3.08	1.12	0.06	0.517	0.303	NWT
c12	4.14	0.90	c21	4.17	0.85	-0.03	<0.001	0.620	NWT
c13	4.24	0.82	c22	4.35	0.79	-0.11	<0.001	0.893	NWT
c14	4.36	0.74	c23	4.44	0.66	-0.08	<0.001	0.869	NWT
c15	4.16	0.93	c24	4.35	0.76	-0.19	<0.001	0.982	WT
c16	4.57	0.69	c25	4.48	0.74	0.09	1.174	0.121	NWT
c17	4.66	0.58	c26	4.48	0.73	0.17	2.490	0.007	BT
c18	4.56	0.68	c27	4.52	0.69	0.04	0.611	0.271	NWT
c28	2.23	1.09	c19	2.27	1.14	-0.04	<0.001	0.646	NWT
c29	3.07	1.13	c20	3.08	1.12	-0.01	<0.001	0.537	NWT
c30	4.20	0.80	c21	4.17	0.85	0.03	0.379	0.352	NWT
c31	4.29	0.87	c22	4.35	0.79	-0.06	<0.001	0.756	NWT
c32	4.21	0.79	c23	4.44	0.66	-0.24	<0.001	0.999	WT
c33	4.20	0.87	c24	4.35	0.76	-0.15	<0.001	0.959	WT
c34	4.40	0.71	c25	4.48	0.74	-0.08	<0.001	0.846	NWT
c35	4.58	0.65	c26	4.48	0.73	0.10	1.305	0.096	NWT
c36	4.62	0.59	c27	4.52	0.69	0.10	1.553	0.061	NWT

9.4.8 Characterization Experiment P800-10 (1-4 Objects, Speech and Mixed/Music, Impaired Channel, DTX on, Headphone Presentation)

Characterization Experiment P800-11 evaluates the performance of the IVAS fixed-point and floating-point implementation for 1-4 objects for speech and mixed/music, DTX on, under impaired channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-11 was conducted by FORCE Technology (language: DAN).

The aggregated results per condition are listed in Table 9.4-13 and are shown in Figure 9.4-9. The results aggregated over FX/FL are shown in Figure 9.4-10. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

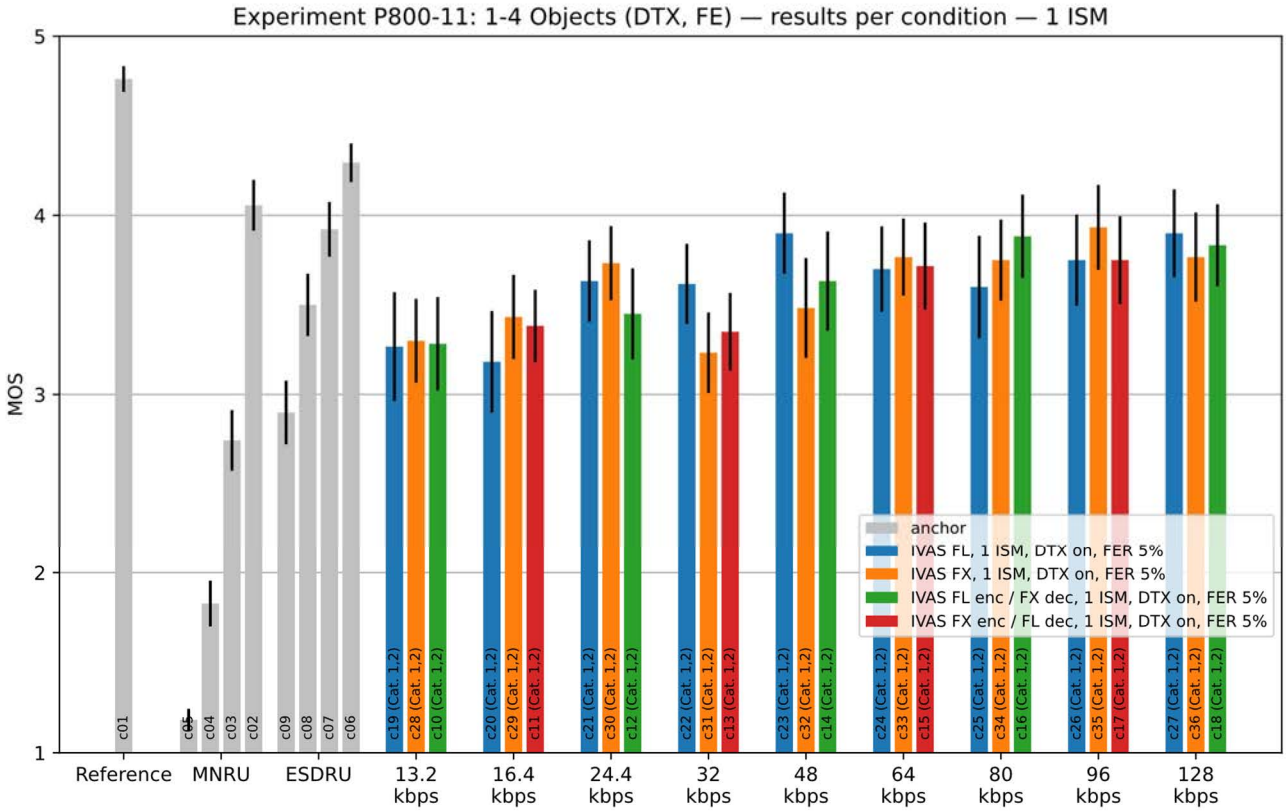
The statistical significance analysis is shown in Table 9.4-14. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.4-13: results per condition for experiment P800-11

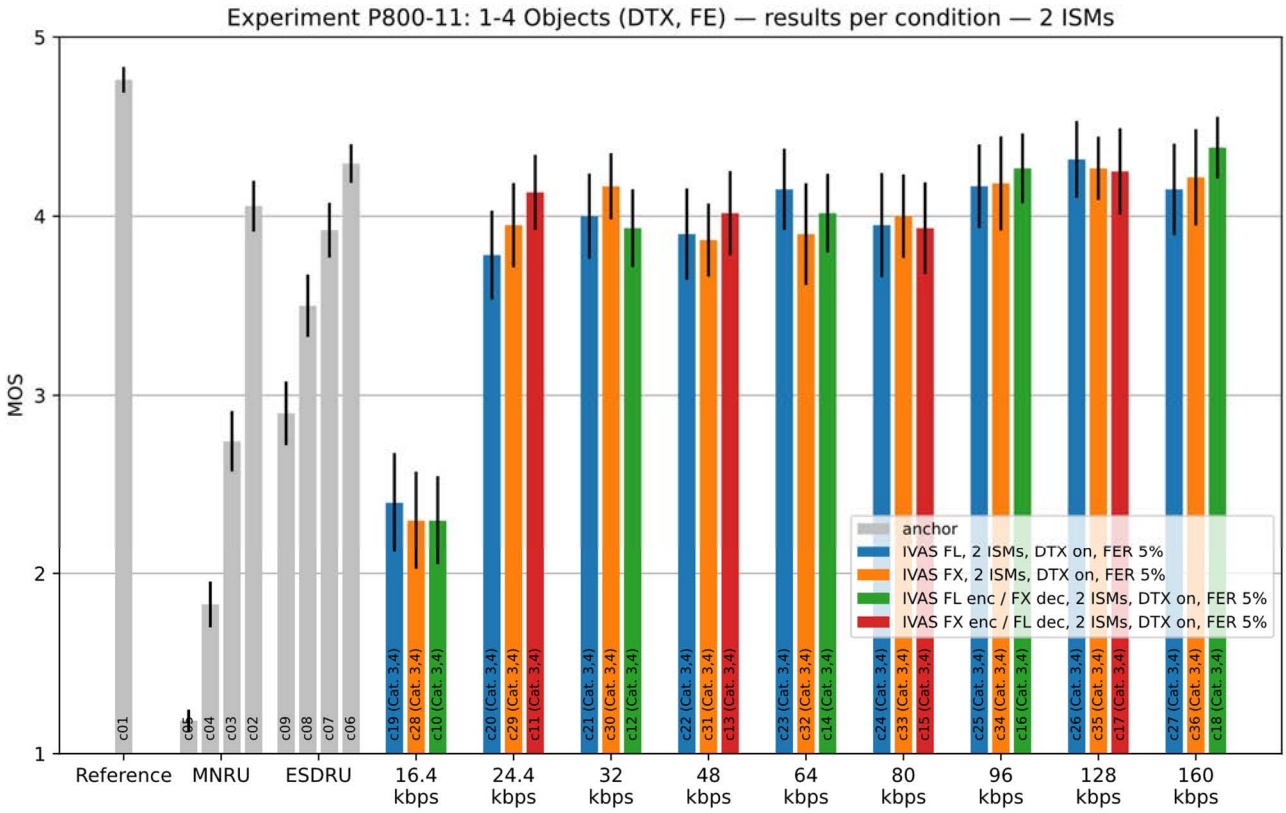
Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.76	0.49	0.07
c02	MNRU Q = 36 dB	180	4.06	0.97	0.14
c03	MNRU Q = 30 dB	180	2.74	1.15	0.17
c04	MNRU Q = 24 dB	180	1.83	0.86	0.13
c05	MNRU Q = 18 dB	180	1.18	0.42	0.06
c06	ESDRU α=0.7	180	4.29	0.74	0.11
c07	ESDRU α=0.5	180	3.92	1.04	0.15
c08	ESDRU α=0.3	180	3.50	1.18	0.17
c09	ESDRU α=0.1	180	2.90	1.21	0.18
c10 (Cat. 1,2)	IVAS FL enc / FX dec, FER 5%, 13.2 kbps, 1 ISM, DTX on	60	3.28	1.01	0.26
c10 (Cat. 3,4)	IVAS FL enc / FX dec, FER 5%, 16.4 kbps, 2 ISMs, DTX on	60	2.30	0.96	0.25
c10 (Cat. 5,6)	IVAS FL enc / FX dec, FER 5%, 24.4 kbps, 3-4 ISMs, DTX on	60	1.83	0.87	0.22
c11 (Cat. 1,2)	IVAS FX enc / FL dec, FER 5%, 16.4 kbps, 1 ISM, DTX on	60	3.38	0.78	0.20
c11 (Cat. 3,4)	IVAS FX enc / FL dec, FER 5%, 24.4 kbps, 2 ISMs, DTX on	60	4.13	0.81	0.21
c11 (Cat. 5,6)	IVAS FX enc / FL dec, FER 5%, 32 kbps, 3-4 ISMs, DTX on	60	2.35	1.07	0.28
c12 (Cat. 1,2)	IVAS FL enc / FX dec, FER 5%, 24.4 kbps, 1 ISM, DTX on	60	3.45	0.98	0.25
c12 (Cat. 3,4)	IVAS FL enc / FX dec, FER 5%, 32 kbps, 2 ISMs, DTX on	60	3.93	0.84	0.22
c12 (Cat. 5,6)	IVAS FL enc / FX dec, FER 5%, 48 kbps, 3-4 ISMs, DTX on	60	3.55	0.96	0.25
c13 (Cat. 1,2)	IVAS FX enc / FL dec, FER 5%, 32 kbps, 1 ISM, DTX on	60	3.35	0.84	0.22

c13 (Cat. 3,4)	IVAS FX enc / FL dec, FER 5%, 48 kbps, 2 ISMs, DTX on	60	4.02	0.91	0.23
c13 (Cat. 5,6)	IVAS FX enc / FL dec, FER 5%, 64 kbps, 3-4 ISMs, DTX on	60	3.60	0.98	0.25
c14 (Cat. 1,2)	IVAS FL enc / FX dec, FER 5%, 48 kbps, 1 ISM, DTX on	60	3.63	1.07	0.28
c14 (Cat. 3,4)	IVAS FL enc / FX dec, FER 5%, 64 kbps, 2 ISMs, DTX on	60	4.02	0.85	0.22
c14 (Cat. 5,6)	IVAS FL enc / FX dec, FER 5%, 80 kbps, 3-4 ISMs, DTX on	60	3.48	1.02	0.26
c15 (Cat. 1,2)	IVAS FX enc / FL dec, FER 5%, 64 kbps, 1 ISM, DTX on	60	3.72	0.94	0.24
c15 (Cat. 3,4)	IVAS FX enc / FL dec, FER 5%, 80 kbps, 2 ISMs, DTX on	60	3.93	0.99	0.26
c15 (Cat. 5,6)	IVAS FX enc / FL dec, FER 5%, 96 kbps, 3-4 ISMs, DTX on	60	3.57	0.98	0.25
c16 (Cat. 1,2)	IVAS FL enc / FX dec, FER 5%, 80 kbps, 1 ISM, DTX on	60	3.88	0.90	0.23
c16 (Cat. 3,4)	IVAS FL enc / FX dec, FER 5%, 96 kbps, 2 ISMs, DTX on	60	4.27	0.76	0.20
c16 (Cat. 5,6)	IVAS FL enc / FX dec, FER 5%, 128 kbps, 3-4 ISMs, DTX on	60	3.78	0.92	0.24
c17 (Cat. 1,2)	IVAS FX enc / FL dec, FER 5%, 96 kbps, 1 ISM, DTX on	60	3.75	0.95	0.24
c17 (Cat. 3,4)	IVAS FX enc / FL dec, FER 5%, 128 kbps, 2 ISMs, DTX on	60	4.25	0.93	0.24
c17 (Cat. 5,6)	IVAS FX enc / FL dec, FER 5%, 160 kbps, 3-4 ISMs, DTX on	60	3.80	0.76	0.20
c18 (Cat. 1,2)	IVAS FL enc / FX dec, FER 5%, 128 kbps, 1 ISM, DTX on	60	3.83	0.89	0.23
c18 (Cat. 3,4)	IVAS FL enc / FX dec, FER 5%, 160 kbps, 2 ISMs, DTX on	60	4.38	0.67	0.17
c18 (Cat. 5,6)	IVAS FL enc / FX dec, FER 5%, 192 kbps, 3-4 ISMs, DTX on	60	3.90	1.05	0.27
c19 (Cat. 1,2)	IVAS FL, FER 5%, 13.2 kbps, 1 ISM, DTX on	60	3.27	1.18	0.30
c19 (Cat. 3,4)	IVAS FL, FER 5%, 16.4 kbps, 2 ISMs, DTX on	60	2.40	1.08	0.28
c19 (Cat. 5,6)	IVAS FL, FER 5%, 24.4 kbps, 3-4 ISMs, DTX on	60	1.92	0.85	0.22
c20 (Cat. 1,2)	IVAS FL, FER 5%, 16.4 kbps, 1 ISM, DTX on	60	3.18	1.10	0.28
c20 (Cat. 3,4)	IVAS FL, FER 5%, 24.4 kbps, 2 ISMs, DTX on	60	3.78	0.96	0.25
c20 (Cat. 5,6)	IVAS FL, FER 5%, 32 kbps, 3-4 ISMs, DTX on	60	2.35	0.92	0.24
c21 (Cat. 1,2)	IVAS FL, FER 5%, 24.4 kbps, 1 ISM, DTX on	60	3.63	0.88	0.23
c21 (Cat. 3,4)	IVAS FL, FER 5%, 32 kbps, 2 ISMs, DTX on	60	4.00	0.92	0.24
c21 (Cat. 5,6)	IVAS FL, FER 5%, 48 kbps, 3-4 ISMs, DTX on	60	3.53	0.98	0.25
c22 (Cat. 1,2)	IVAS FL, FER 5%, 32 kbps, 1 ISM, DTX on	60	3.62	0.86	0.22
c22 (Cat. 3,4)	IVAS FL, FER 5%, 48 kbps, 2 ISMs, DTX on	60	3.90	0.99	0.26
c22 (Cat. 5,6)	IVAS FL, FER 5%, 64 kbps, 3-4 ISMs, DTX on	60	3.62	1.01	0.26
c23 (Cat. 1,2)	IVAS FL, FER 5%, 48 kbps, 1 ISM, DTX on	60	3.90	0.88	0.23
c23 (Cat. 3,4)	IVAS FL, FER 5%, 64 kbps, 2 ISMs, DTX on	60	4.15	0.88	0.23
c23 (Cat. 5,6)	IVAS FL, FER 5%, 80 kbps, 3-4 ISMs, DTX on	60	3.40	1.04	0.27
c24 (Cat. 1,2)	IVAS FL, FER 5%, 64 kbps, 1 ISM, DTX on	60	3.70	0.93	0.24
c24 (Cat. 3,4)	IVAS FL, FER 5%, 80 kbps, 2 ISMs, DTX on	60	3.95	1.13	0.29
c24 (Cat. 5,6)	IVAS FL, FER 5%, 96 kbps, 3-4 ISMs, DTX on	60	3.48	1.08	0.28
c25 (Cat. 1,2)	IVAS FL, FER 5%, 80 kbps, 1 ISM, DTX on	60	3.60	1.11	0.29
c25 (Cat. 3,4)	IVAS FL, FER 5%, 96 kbps, 2 ISMs, DTX on	60	4.17	0.91	0.23
c25 (Cat. 5,6)	IVAS FL, FER 5%, 128 kbps, 3-4 ISMs, DTX on	60	3.83	0.85	0.22
c26 (Cat. 1,2)	IVAS FL, FER 5%, 96 kbps, 1 ISM, DTX on	60	3.75	0.98	0.25
c26 (Cat. 3,4)	IVAS FL, FER 5%, 128 kbps, 2 ISMs, DTX on	60	4.32	0.83	0.21
c26 (Cat. 5,6)	IVAS FL, FER 5%, 160 kbps, 3-4 ISMs, DTX on	60	3.93	0.92	0.24
c27 (Cat. 1,2)	IVAS FL, FER 5%, 128 kbps, 1 ISM, DTX on	60	3.90	0.95	0.25
c27 (Cat. 3,4)	IVAS FL, FER 5%, 160 kbps, 2 ISMs, DTX on	60	4.15	0.99	0.26
c27 (Cat. 5,6)	IVAS FL, FER 5%, 192 kbps, 3-4 ISMs, DTX on	60	3.93	0.80	0.21
c28 (Cat. 1,2)	IVAS FX, FER 5%, 13.2 kbps, 1 ISM, DTX on	60	3.30	0.91	0.23
c28 (Cat. 3,4)	IVAS FX, FER 5%, 16.4 kbps, 2 ISMs, DTX on	60	2.30	1.06	0.27
c28 (Cat. 5,6)	IVAS FX, FER 5%, 24.4 kbps, 3-4 ISMs, DTX on	60	1.93	0.92	0.24
c29 (Cat. 1,2)	IVAS FX, FER 5%, 16.4 kbps, 1 ISM, DTX on	60	3.43	0.91	0.23
c29 (Cat. 3,4)	IVAS FX, FER 5%, 24.4 kbps, 2 ISMs, DTX on	60	3.95	0.91	0.23
c29 (Cat. 5,6)	IVAS FX, FER 5%, 32 kbps, 3-4 ISMs, DTX on	60	2.43	1.03	0.27
c30 (Cat. 1,2)	IVAS FX, FER 5%, 24.4 kbps, 1 ISM, DTX on	60	3.73	0.80	0.21
c30 (Cat. 3,4)	IVAS FX, FER 5%, 32 kbps, 2 ISMs, DTX on	60	4.17	0.72	0.18
c30 (Cat. 5,6)	IVAS FX, FER 5%, 48 kbps, 3-4 ISMs, DTX on	60	3.38	1.03	0.27
c31 (Cat. 1,2)	IVAS FX, FER 5%, 32 kbps, 1 ISM, DTX on	60	3.23	0.87	0.23
c31 (Cat. 3,4)	IVAS FX, FER 5%, 48 kbps, 2 ISMs, DTX on	60	3.87	0.79	0.20
c31 (Cat. 5,6)	IVAS FX, FER 5%, 64 kbps, 3-4 ISMs, DTX on	60	3.65	0.90	0.23
c32 (Cat. 1,2)	IVAS FX, FER 5%, 48 kbps, 1 ISM, DTX on	60	3.48	1.08	0.28
c32 (Cat. 3,4)	IVAS FX, FER 5%, 64 kbps, 2 ISMs, DTX on	60	3.90	1.10	0.28
c32 (Cat. 5,6)	IVAS FX, FER 5%, 80 kbps, 3-4 ISMs, DTX on	60	3.32	1.00	0.26
c33 (Cat. 1,2)	IVAS FX, FER 5%, 64 kbps, 1 ISM, DTX on	60	3.77	0.83	0.21
c33 (Cat. 3,4)	IVAS FX, FER 5%, 80 kbps, 2 ISMs, DTX on	60	4.00	0.90	0.23
c33 (Cat. 5,6)	IVAS FX, FER 5%, 96 kbps, 3-4 ISMs, DTX on	60	3.48	1.02	0.26
c34 (Cat. 1,2)	IVAS FX, FER 5%, 80 kbps, 1 ISM, DTX on	60	3.75	0.88	0.23
c34 (Cat. 3,4)	IVAS FX, FER 5%, 96 kbps, 2 ISMs, DTX on	60	4.18	1.02	0.26
c34 (Cat. 5,6)	IVAS FX, FER 5%, 128 kbps, 3-4 ISMs, DTX on	60	3.60	1.01	0.26

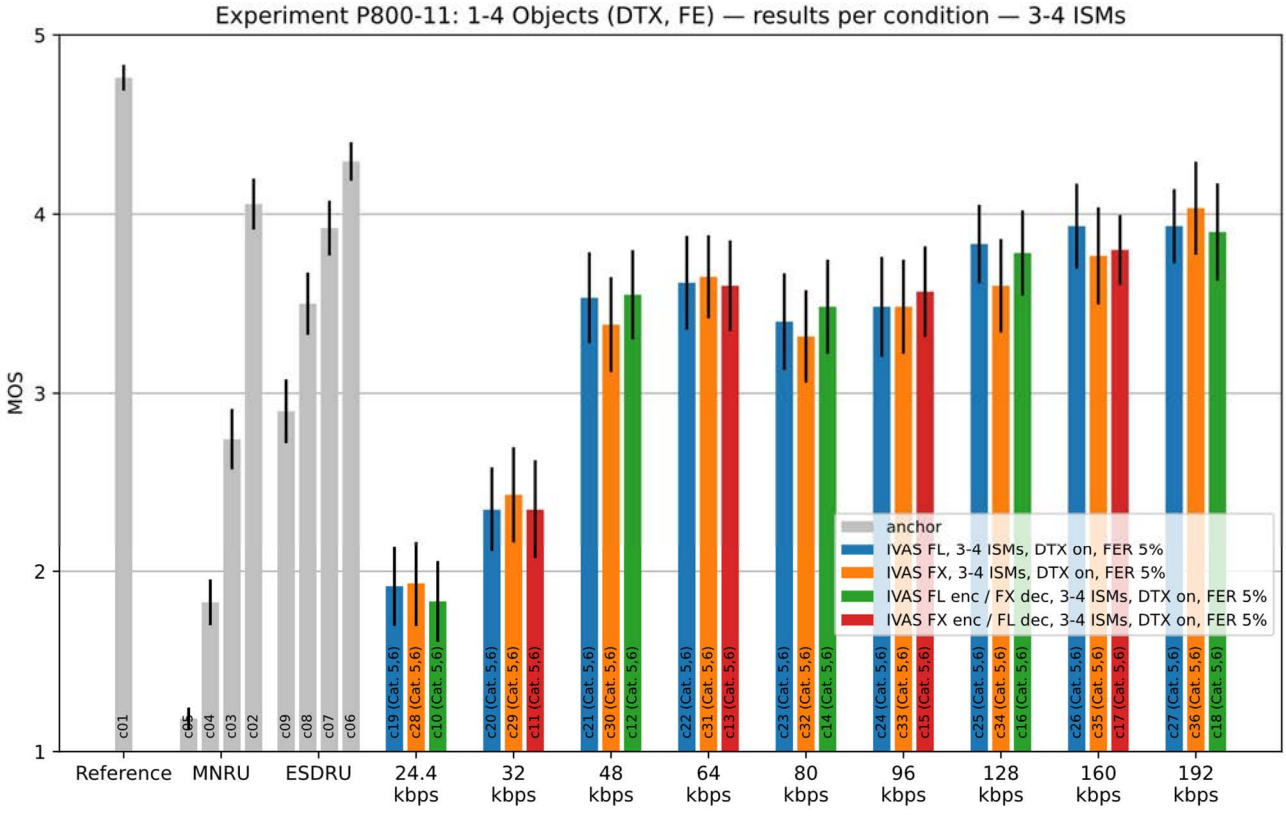
c35 (Cat. 1,2)	IVAS FX, FER 5%, 96 kbps, 1 ISM, DTX on	60	3.93	0.92	0.24
c35 (Cat. 3,4)	IVAS FX, FER 5%, 128 kbps, 2 ISMs, DTX on	60	4.27	0.69	0.18
c35 (Cat. 5,6)	IVAS FX, FER 5%, 160 kbps, 3-4 ISMs, DTX on	60	3.77	1.05	0.27
c36 (Cat. 1,2)	IVAS FX, FER 5%, 128 kbps, 1 ISM, DTX on	60	3.77	0.96	0.25
c36 (Cat. 3,4)	IVAS FX, FER 5%, 160 kbps, 2 ISMs, DTX on	60	4.22	1.04	0.27
c36 (Cat. 5,6)	IVAS FX, FER 5%, 192 kbps, 3-4 ISMs, DTX on	60	4.03	1.01	0.26



(a)



(b)



(c)

Figure 9.4-9: results per condition for experiment P800-11

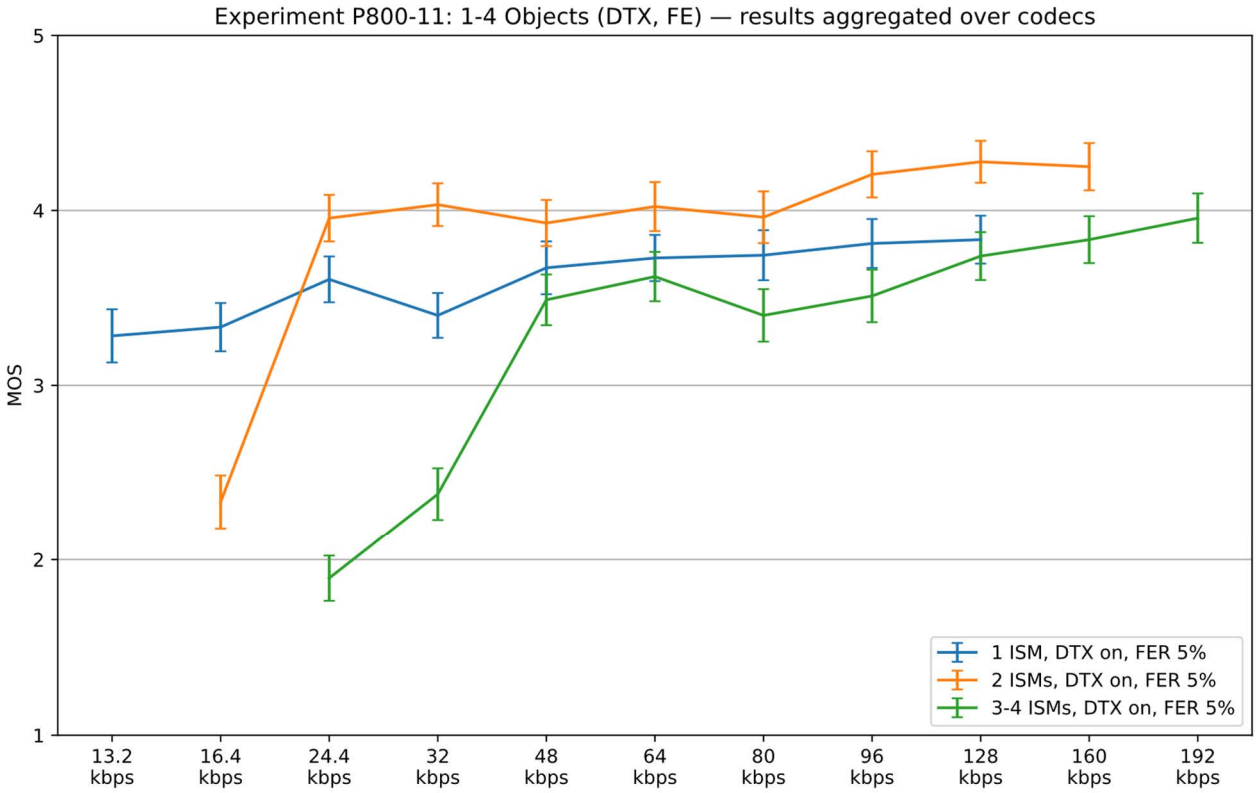


Figure 9.4-10: results aggregated over FX/FL for experiment P800-11

Table 9.4-14: statistical significance analysis for experiment P800-11

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10 (Cat. 1,2)	3.28	1.01	c19 (Cat. 1,2)	3.27	1.18	0.02	0.080	0.468	NWT
c10 (Cat. 3,4)	2.30	0.96	c19 (Cat. 3,4)	2.40	1.08	-0.10	<0.001	0.704	NWT
c10 (Cat. 5,6)	1.83	0.87	c19 (Cat. 5,6)	1.92	0.85	-0.08	<0.001	0.703	NWT
c11 (Cat. 1,2)	3.38	0.78	c20 (Cat. 1,2)	3.18	1.10	0.20	1.149	0.126	NWT
c11 (Cat. 3,4)	4.13	0.81	c20 (Cat. 3,4)	3.78	0.96	0.35	2.159	0.016	BT
c11 (Cat. 5,6)	2.35	1.07	c20 (Cat. 5,6)	2.35	0.92	0.00	<0.001	0.500	NWT
c12 (Cat. 1,2)	3.45	0.98	c21 (Cat. 1,2)	3.63	0.88	-0.18	<0.001	0.857	NWT
c12 (Cat. 3,4)	3.93	0.84	c21 (Cat. 3,4)	4.00	0.92	-0.07	<0.001	0.661	NWT
c12 (Cat. 5,6)	3.55	0.96	c21 (Cat. 5,6)	3.53	0.98	0.02	0.096	0.462	NWT
c13 (Cat. 1,2)	3.35	0.84	c22 (Cat. 1,2)	3.62	0.86	-0.27	<0.001	0.956	WT
c13 (Cat. 3,4)	4.02	0.91	c22 (Cat. 3,4)	3.90	0.99	0.12	0.675	0.250	NWT
c13 (Cat. 5,6)	3.60	0.98	c22 (Cat. 5,6)	3.62	1.01	-0.02	<0.001	0.537	NWT
c14 (Cat. 1,2)	3.63	1.07	c23 (Cat. 1,2)	3.90	0.88	-0.27	<0.001	0.931	NWT

c14 (Cat. 3,4)	4.02	0.85	c23 (Cat. 3,4)	4.15	0.88	-0.13	<0.001	0.799	NWT
c14 (Cat. 5,6)	3.48	1.02	c23 (Cat. 5,6)	3.40	1.04	0.08	0.441	0.330	NWT
c15 (Cat. 1,2)	3.72	0.94	c24 (Cat. 1,2)	3.70	0.93	0.02	0.100	0.460	NWT
c15 (Cat. 3,4)	3.93	0.99	c24 (Cat. 3,4)	3.95	1.13	-0.02	<0.001	0.535	NWT
c15 (Cat. 5,6)	3.57	0.98	c24 (Cat. 5,6)	3.48	1.08	0.08	0.446	0.328	NWT
c16 (Cat. 1,2)	3.88	0.90	c25 (Cat. 1,2)	3.60	1.11	0.28	1.533	0.064	NWT
c16 (Cat. 3,4)	4.27	0.76	c25 (Cat. 3,4)	4.17	0.91	0.10	0.657	0.256	NWT
c16 (Cat. 5,6)	3.78	0.92	c25 (Cat. 5,6)	3.83	0.85	-0.05	<0.001	0.621	NWT
c17 (Cat. 1,2)	3.75	0.95	c26 (Cat. 1,2)	3.75	0.98	0.00	<0.001	0.500	NWT
c17 (Cat. 3,4)	4.25	0.93	c26 (Cat. 3,4)	4.32	0.83	-0.07	<0.001	0.661	NWT
c17 (Cat. 5,6)	3.80	0.76	c26 (Cat. 5,6)	3.93	0.92	-0.13	<0.001	0.806	NWT
c18 (Cat. 1,2)	3.83	0.89	c27 (Cat. 1,2)	3.90	0.95	-0.07	<0.001	0.655	NWT
c18 (Cat. 3,4)	4.38	0.67	c27 (Cat. 3,4)	4.15	0.99	0.23	1.515	0.066	NWT
c18 (Cat. 5,6)	3.90	1.05	c27 (Cat. 5,6)	3.93	0.80	-0.03	<0.001	0.576	NWT
c28 (Cat. 1,2)	3.30	0.91	c19 (Cat. 1,2)	3.27	1.18	0.03	0.172	0.432	NWT
c28 (Cat. 3,4)	2.30	1.06	c19 (Cat. 3,4)	2.40	1.08	-0.10	<0.001	0.695	NWT
c28 (Cat. 5,6)	1.93	0.92	c19 (Cat. 5,6)	1.92	0.85	0.02	0.099	0.461	NWT
c29 (Cat. 1,2)	3.43	0.91	c20 (Cat. 1,2)	3.18	1.10	0.25	1.359	0.088	NWT
c29 (Cat. 3,4)	3.95	0.91	c20 (Cat. 3,4)	3.78	0.96	0.17	0.979	0.165	NWT
c29 (Cat. 5,6)	2.43	1.03	c20 (Cat. 5,6)	2.35	0.92	0.08	0.466	0.321	NWT
c30 (Cat. 1,2)	3.73	0.80	c21 (Cat. 1,2)	3.63	0.88	0.10	0.651	0.258	NWT
c30 (Cat. 3,4)	4.17	0.72	c21 (Cat. 3,4)	4.00	0.92	0.17	1.108	0.135	NWT
c30 (Cat. 5,6)	3.38	1.03	c21 (Cat. 5,6)	3.53	0.98	-0.15	<0.001	0.792	NWT
c31 (Cat. 1,2)	3.23	0.87	c22 (Cat. 1,2)	3.62	0.86	-0.38	<0.001	0.992	WT
c31 (Cat. 3,4)	3.87	0.79	c22 (Cat. 3,4)	3.90	0.99	-0.03	<0.001	0.580	NWT
c31 (Cat. 5,6)	3.65	0.90	c22 (Cat. 5,6)	3.62	1.01	0.03	0.189	0.425	NWT
c32 (Cat. 1,2)	3.48	1.08	c23 (Cat. 1,2)	3.90	0.88	-0.42	<0.001	0.989	WT
c32 (Cat. 3,4)	3.90	1.10	c23 (Cat. 3,4)	4.15	0.88	-0.25	<0.001	0.914	NWT
c32 (Cat. 5,6)	3.32	1.00	c23 (Cat. 5,6)	3.40	1.04	-0.08	<0.001	0.671	NWT
c33 (Cat. 1,2)	3.77	0.83	c24 (Cat. 1,2)	3.70	0.93	0.07	0.417	0.339	NWT
c33 (Cat. 3,4)	4.00	0.90	c24 (Cat. 3,4)	3.95	1.13	0.05	0.268	0.394	NWT
c33 (Cat. 5,6)	3.48	1.02	c24 (Cat. 5,6)	3.48	1.08	0.00	<0.001	0.500	NWT
c34 (Cat. 1,2)	3.75	0.88	c25 (Cat. 1,2)	3.60	1.11	0.15	0.823	0.206	NWT

c34 (Cat. 3,4)	4.18	1.02	c25 (Cat. 3,4)	4.17	0.91	0.02	0.091	0.464	NWT
c34 (Cat. 5,6)	3.60	1.01	c25 (Cat. 5,6)	3.83	0.85	-0.23	<0.001	0.913	NWT
c35 (Cat. 1,2)	3.93	0.92	c26 (Cat. 1,2)	3.75	0.98	0.18	1.053	0.147	NWT
c35 (Cat. 3,4)	4.27	0.69	c26 (Cat. 3,4)	4.32	0.83	-0.05	<0.001	0.640	NWT
c35 (Cat. 5,6)	3.77	1.05	c26 (Cat. 5,6)	3.93	0.92	-0.17	<0.001	0.821	NWT
c36 (Cat. 1,2)	3.77	0.96	c27 (Cat. 1,2)	3.90	0.95	-0.13	<0.001	0.776	NWT
c36 (Cat. 3,4)	4.22	1.04	c27 (Cat. 3,4)	4.15	0.99	0.07	0.361	0.359	NWT
c36 (Cat. 5,6)	4.03	1.01	c27 (Cat. 5,6)	3.93	0.80	0.10	0.602	0.274	NWT

9.4.9 Characterization Experiment P800-22 (1-2 Objects, Speech and Mixed/Music, JBM, Headphone Presentation)

Characterization Experiment P800-23 evaluates the performance of the IVAS JBM solution for fixed-point and floating-point implementation for objects for speech and mixed/music, under impaird channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-22 was conducted by FORCE Technology (language: DAN).

The aggregated results per condition are listed in Table 9.4-15 and are shown in Figure 9.4-11. The results aggregated over FX/FL are shown in Figure 9.4-12. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-16. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.4-15: results per condition for experiment P800-22

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.69	0.54	0.08
c02	MNRU Q = 36 dB	180	4.11	0.87	0.13
c03	MNRU Q = 30 dB	180	3.06	1.15	0.17
c04	MNRU Q = 24 dB	180	2.03	0.94	0.14
c05	MNRU Q = 18 dB	180	1.40	0.65	0.10
c06	ESDRU α=0.7	180	4.38	0.69	0.10
c07	ESDRU α=0.5	180	4.04	0.87	0.13
c08	ESDRU α=0.3	180	3.50	1.11	0.16
c09	ESDRU α=0.1	180	2.99	1.22	0.18
c10	IVAS FL enc / FX dec, No error, 24.4 kbps	180	4.36	0.77	0.11
c11	IVAS FL enc / FX dec, No error, 48 kbps	180	4.51	0.68	0.10
c12	IVAS FL enc / FX dec, No error, 96 kbps	180	4.68	0.55	0.08
c13	IVAS FL enc / FX dec, Error I1. O1, 24.4 kbps	180	3.28	1.06	0.16
c14	IVAS FL enc / FX dec, Error I1. O1, 48 kbps	180	3.33	1.09	0.16
c15	IVAS FL enc / FX dec, Error I1. O2, 24.4 kbps	180	4.19	1.01	0.15
c16	IVAS FL enc / FX dec, Error I1. O2, 48 kbps	180	4.42	0.88	0.13
c17	IVAS FL enc / FX dec, I1. O1, 24.4 kbps	180	3.09	1.22	0.18
c18	IVAS FL enc / FX dec, I1. O1, 48 kbps	180	3.44	1.06	0.16
c19	IVAS FL enc / FX dec, I1. O1, 96 kbps	180	3.53	0.99	0.14
c20	IVAS FL enc / FX dec, I1. O2, 24.4 kbps	180	4.26	0.96	0.14
c21	IVAS FL enc / FX dec, I1. O2, 48 kbps	180	4.37	0.86	0.13
c22	IVAS FL enc / FX dec, I1. O2, 96 kbps	180	4.46	0.85	0.13
c23	IVAS FL enc / FL dec, No error, 24.4 kbps	180	4.32	0.84	0.12
c24	IVAS FL enc / FL dec, No error, 48 kbps	180	4.59	0.61	0.09
c25	IVAS FL enc / FL dec, No error, 96 kbps	180	4.59	0.59	0.09
c26	IVAS FL enc / FL dec, Error I1. O1, 24.4 kbps	180	3.37	1.13	0.17
c27	IVAS FL enc / FL dec, Error I1. O1, 48 kbps	180	3.54	1.07	0.16

c28	IVAS FL enc / FL dec, Error I1. O2, 24.4 kbps	180	4.14	0.96	0.14
c29	IVAS FL enc / FL dec, Error I1. O2, 48 kbps	180	4.37	0.91	0.13
c30	IVAS FL enc / FL dec, Error I1. O2, 96 kbps	180	4.49	0.90	0.13
c31	IVAS FL enc / FL dec, I1. O1, 24.4 kbps	180	3.26	1.08	0.16
c32	IVAS FL enc / FL dec, I1. O1, 48 kbps	180	3.42	1.08	0.16
c33	IVAS FL enc / FL dec, I1. O1, 96 kbps	180	3.45	1.06	0.16
c34	IVAS FL enc / FL dec, I1. O2, 24.4 kbps	180	4.21	0.91	0.13
c35	IVAS FL enc / FL dec, I1. O2, 48 kbps	180	4.34	0.94	0.14
c36	IVAS FL enc / FL dec, I1. O2, 96 kbps	180	4.52	0.82	0.12

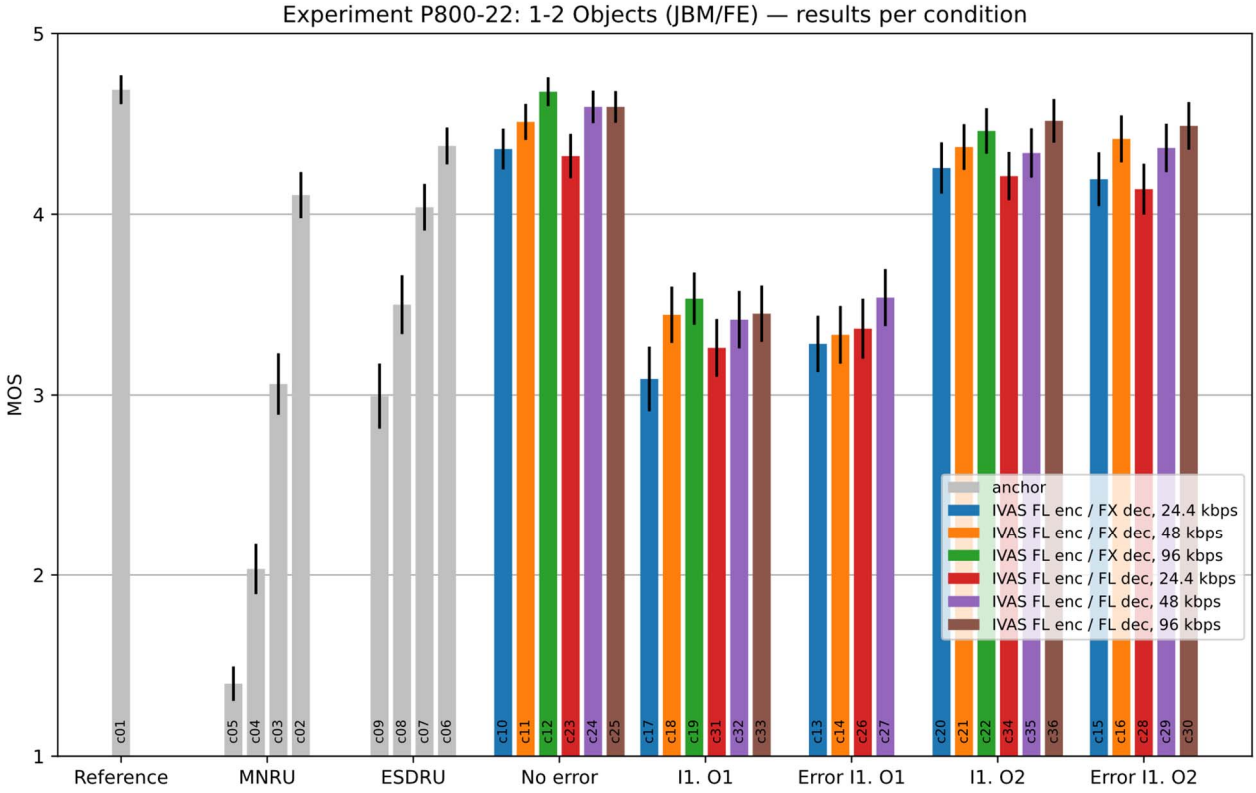


Figure 9.4-11: results per condition for experiment P800-22

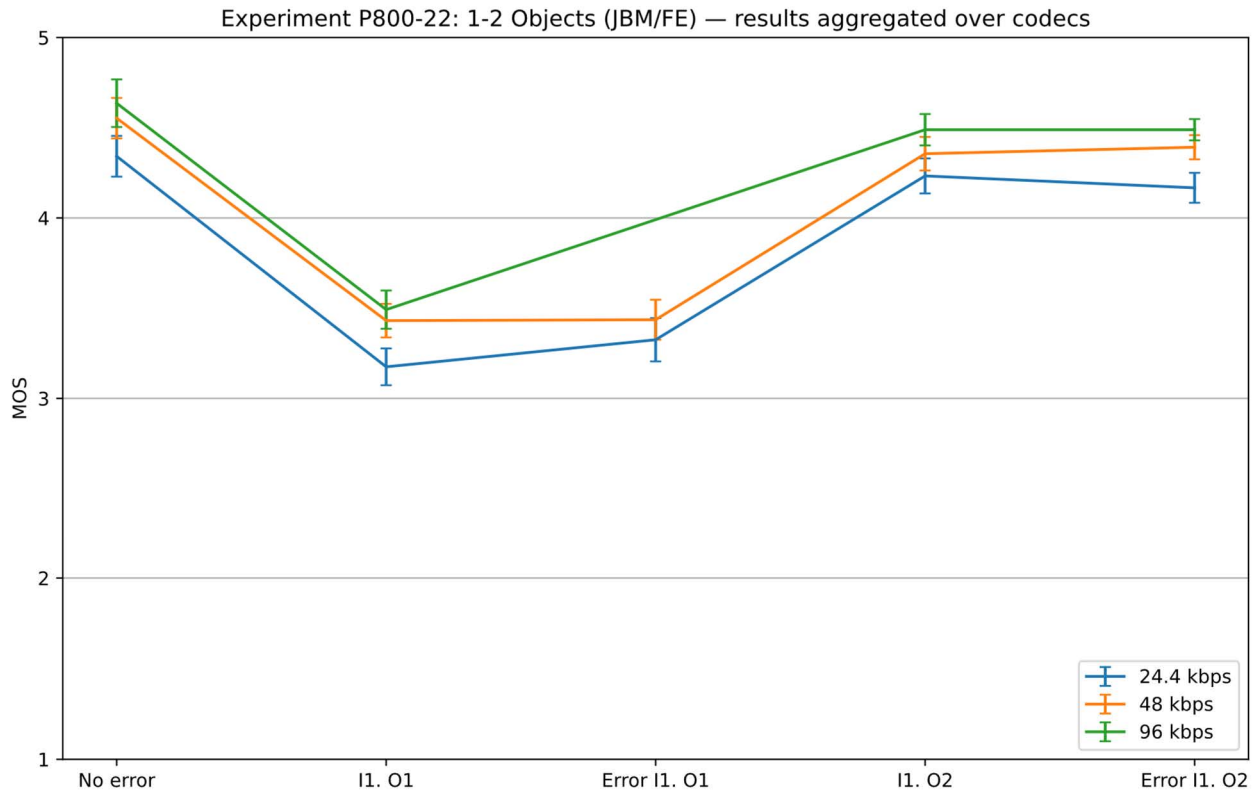


Figure 9.4-12: results aggregated over FX/FL for experiment P800-22

Table 9.4-16: statistical significance analysis for experiment P800-22

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
No terms of references mandated for experiment .									

9.4.10 Characterization Experiment BS1534-10 (1-2 Objects, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-10 evaluates the performance of the IVAS fixed-point implementation for 1-2 objects with generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-10 was conducted by Qualcomm Incorporated.

The aggregated results per condition are listed in Table 9.4-17 and are shown in Figure 9.4-13. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-18. Scores with result WT are highlighted in red.

Table 9.4-17: results per condition for experiment BS1534-10

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	97.54	5.29	0.81
c02	LP 3.5 kHz	168	28.43	12.56	1.91
c03	Mono EVS, 48 kbps	168	58.14	14.03	2.14
c04	Mono EVS, 96 kbps	168	58.32	13.85	2.11
c05	IVAS, 48 kbps	168	90.23	11.86	1.81
c06	IVAS, 64 kbps	168	92.47	10.65	1.62
c07	IVAS, 96 kbps	168	95.03	8.40	1.28
c08	IVAS, 128 kbps	168	97.59	4.97	0.76

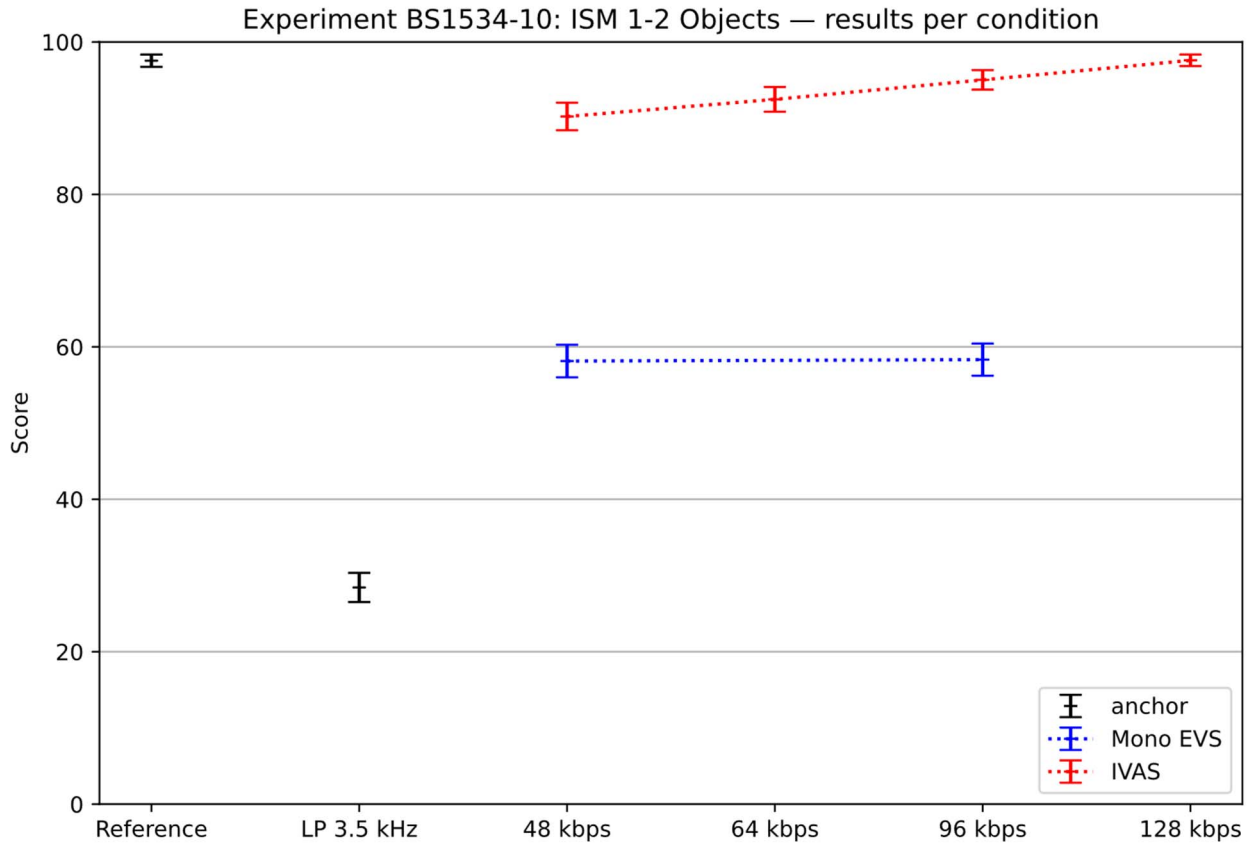


Figure 9.4-13: results per condition for experiment BS1534-10

Table 9.4-18: statistical significance analysis for experiment BS1534-10

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	90.23	11.86	c03	58.14	14.03	32.09	23.951	<0.001	BT
c07	95.03	8.40	c04	58.32	13.85	36.71	29.071	<0.001	BT

9.4.11 Characterization Experiment BS1534-11 (3-4 Objects, Generic Audio, 24.4 – 64 kbps, Headphone Presentation)

Characterization Experiment BS1534-11 evaluates the performance of the IVAS fixed-point implementation for 3-4 objects with generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-11 was conducted by Dolby Laboratories, Inc.

The aggregated results per condition are listed in Table 9.4-19 and are shown in Figure 9.4-14. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-20. Scores with result WT are highlighted in red.

Table 9.4-19: results per condition for experiment BS1534-11

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	100.00	0.00	0.00
c02	LP 3.5 kHz	168	33.83	6.77	1.03
c03	Mono EVS, 24.4 kbps	168	61.81	10.79	1.64
c04	Mono EVS, 48 kbps	168	62.39	10.14	1.54
c05	IVAS, 24.4 kbps	168	60.78	12.26	1.87
c06	IVAS, 32 kbps	168	64.11	12.07	1.84
c07	IVAS, 48 kbps	168	78.29	7.25	1.10
c08	IVAS, 64 kbps	168	79.31	7.58	1.15

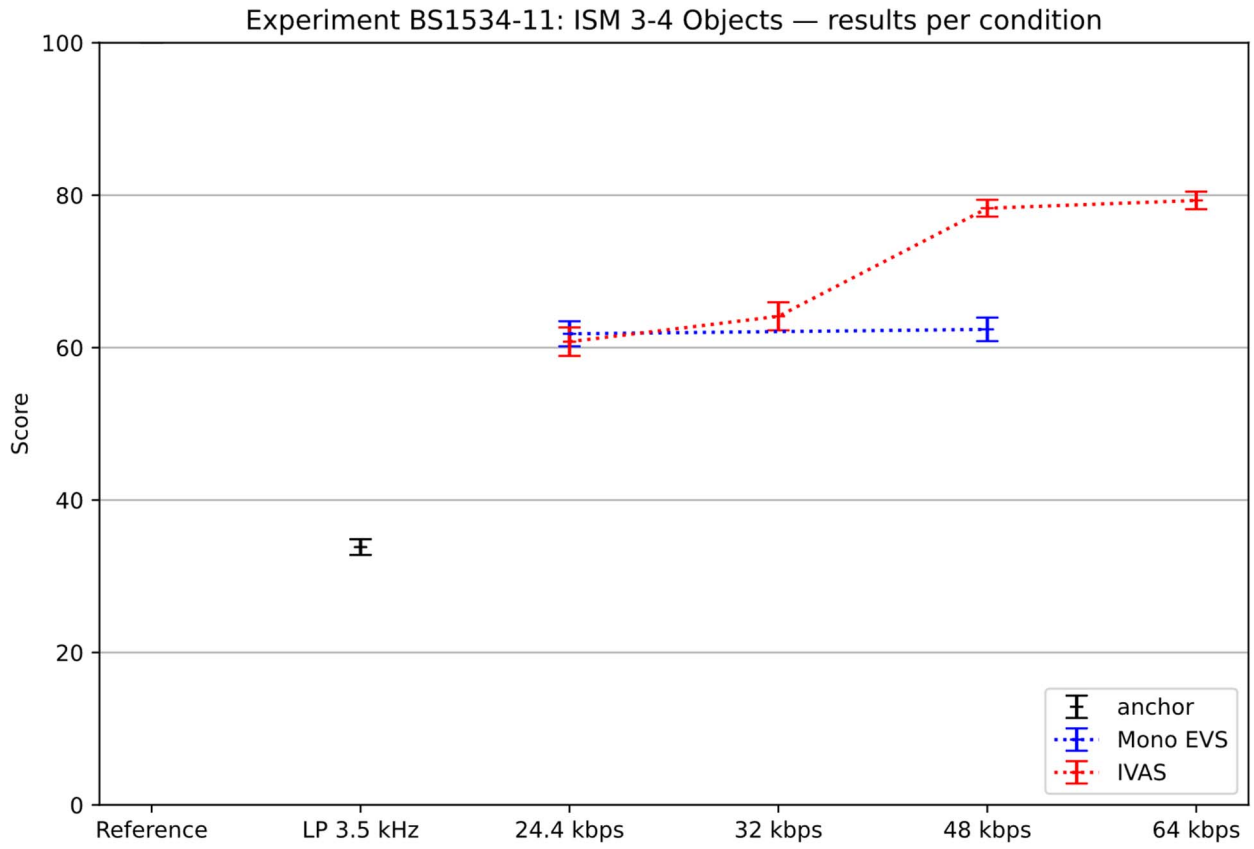


Figure 9.4-14: results per condition for experiment BS1534-11

Table 9.4-20: statistical significance analysis for experiment BS1534-11

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	60.78	12.26	c03	61.81	10.79	-1.04	<0.001	0.172	NWT
c07	78.29	7.25	c04	62.39	10.14	15.90	20.547	<0.001	BT

9.4.12 Characterization Experiment BS1534-12 (3-4 Objects, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-12 evaluates the performance of the IVAS fixed-point implementation for 3-4 objects with generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-12 was conducted by Philips International B.V.

The aggregated results per condition are listed in Table 9.4-21 and are shown in Figure 9.4-15. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.4-22. Scores with result WT are highlighted in red.

Table 9.4-21: results per condition for experiment BS1534-12

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	97.99	4.83	0.74
c02	LP 3.5 kHz	168	17.28	17.11	2.61
c03	Mono EVS, 64 kbps	168	41.79	23.57	3.59
c04	Mono EVS, 128 kbps	168	42.76	24.35	3.71
c05	IVAS, 64 kbps	168	90.99	13.98	2.13
c06	IVAS, 96 kbps	168	91.35	13.40	2.04
c07	IVAS, 128 kbps	168	92.04	13.04	1.99
c08	IVAS, 256 kbps	168	94.05	11.16	1.70

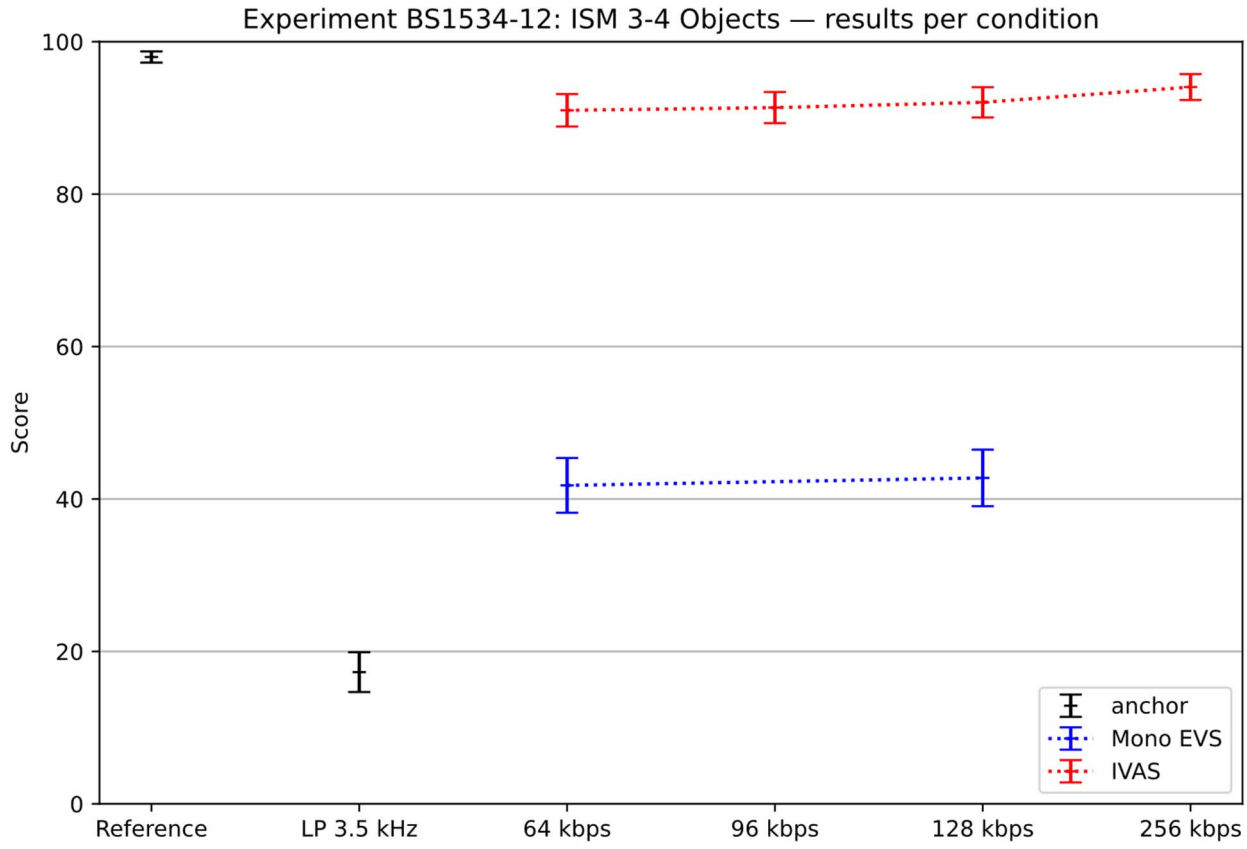


Figure 9.4-15: results per condition for experiment BS1534-12

Table 9.4-22: statistical significance analysis for experiment BS1534-12

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	90.99	13.98	c03	41.79	23.57	49.21	26.034	<0.001	BT
c07	92.04	13.04	c04	42.76	24.35	49.28	23.804	<0.001	BT

9.4.13 Conclusions on IVAS Performance for Objects (Independent Streams with Metadata, ISM)

Taking the results of the previous clauses in 9.4 into account, the following conclusions can be drawn for IVAS object-based audio operating modes:

During selection phase, IVAS object coding was extensively compared against EVS multi-mono. EVS doesn't support the transmission of object metadata. Thus, it should be noted that for the EVS reference conditions, the object metadata was not quantized and not considered in terms of bitrate for the EVS conditions. Configurations from 1-4 objects were tested. Across all content categories, IVAS performs at least equivalent to EVS, even though for EVS the object metadata bitrate isn't considered. For 2 objects at 16.4 kbps, there is even a significant advantage over EVS observed.

During characterization phase, IVAS object coding was evaluated against EVS mono in a series of BS.1534 tests. Across most tested bitrates, IVAS was significantly outperforming EVS mono at the same bitrate. For the transmission of 3-4 objects at 24 kbps, IVAS was found to be on par with EVS mono.

Additionally, the performance of the IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for 1 – 4 objects. Out of 92 tested operating points, 82 operating points were statistically equivalent.

Overall, the IVAS object coding for 1-2 objects maintains a "good" MOS rating (>= 4.0) from bitrates starting from 16.4 kbps and above. A level 1 IVAS encoder alreads reaches an "excellent" MOS rating of approx. 4.6 for this configuration. IVAS object coding for 3-4 objects maintains a "good" MOS rating (>= 4.0) from bitrates starting from 48 kbps and above. A level 1 IVAS encoder alreads reaches an "very good" MOS rating of approx. 4.4 for this configuration.

9.5 Metadata-assisted Spatial Audio (MASA)

9.5.1 Overview

In Selection phase, four experiments have been conducted to evaluate the performance of the IVAS codec with 2TC MASA content. While the experiments P800-8 and P800-9 were conducted as P.800 DCR tests, the experiments BS1534-7a and BS1534-7b were conducted as BS.1534 tests. All experiments were conducted using headphone presentation.

- Selection Experiment P800-8: Clean speech, MASA 2TC under clean and impaired channel conditions, headphone presentation
- Selection Experiment P800-9: Speech + background, MASA 2TC under clean channel conditions, DTX off and on, headphone presentation
- Selection Experiment BS1534-7a: Generic audio, MASA 2TC, 96 and 128 kbps, headphone presentation
- Selection Experiment BS1534-7b: Generic audio, MASA 2TC, 192 and 256 kbps, headphone presentation

In Characterization phase, six additional experiments have been conducted, with MASA 1TC and MASA 2TC content. While the experiments P800-12, P800-13, P800-14 were conducted as P.800 DCR tests, the experiments BS1534-13, BS1534-14 and BS1534-15 were conducted as BS.1534 tests.

- Characterization Experiment P800-12: Speech + background and mixed/music, MASA 1TC under clean channel conditions, headphone presentation

- Characterization Experiment P800-13: Speech + background and mixed/music, MASA 2TC under clean channel conditions, headphone presentation
- Characterization Experiment P800-14: Speech + background and mixed/music, MASA 1-2TC under impaired channel conditions, DTX on, headphone presentation
- Characterization Experiment BS1534-13: Generic audio, MASA 1TC, 16.4 – 48 kbps, headphone presentation
- Characterization Experiment BS1534-14: Generic audio, MASA 1TC, 64 – 256 kbps, headphone presentation
- Characterization Experiment BS1534-15: Generic audio, MASA 2TC, 64 – 256 kbps, headphone presentation

9.5.2 Selection Experiment P800-8 (MASA, Clean speech, Headphone Presentation)

Selection Experiment P800-8 evaluates IVAS for MASA clean speech under clean and impaired channel conditions using headphone presentation. See Annex C.8 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-8 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.5-1: Statistical overview on the results of P800-8

		Type	CuT						EVS Reference				Evaluation		
		Value	Bitrate	DTX	FER	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#													
a	c25	1	13.2	Off		NWT	3.61	0.91	c09	3x7.2	2.91	1.06	6.7	BT	EXCEED
		2	13.2	Off		NWT	3.61	0.91	c15	2x7.2	2.93	1.24	5.91	BT	EXCEED
	c26	1	16.4	Off		NWT	3.91	0.83	c10	4x7.2	2.9	1.09	9.89	BT	EXCEED
		2	16.4	Off		NWT	3.91	0.83	c15	2x7.2	2.93	1.24	8.78	BT	EXCEED
	c27	1	24.4	Off		NWT	4.08	0.78	c11	4x8	2.96	1.08	11.28	BT	EXCEED
		2	24.4	Off		NWT	4.08	0.78	c16	2x8	3.12	1.23	8.94	BT	EXCEED
	c28	1	32	Off		NWT	4.28	0.77	c12	4x9.6	3.62	1	6.96	BT	EXCEED
		2	32	Off		NWT	4.28	0.77	c17	2x9.6	3.67	1.1	6.09	BT	EXCEED
	c29	1	48	Off		NWT	4.46	0.67	c13	4x16.4	4.28	0.76	2.36	BT	EXCEED
		2	48	Off		NWT	4.46	0.67	c18	2x16.4	4.45	0.68	0.08	NWT	PASS
	c30	1	64	Off		NWT	4.59	0.6	c14	4x24.4	4.48	0.65	1.76	BT	EXCEED
		2	64	Off		NWT	4.59	0.6	c19	2x24.4	4.67	0.57	-1.26	NWT	PASS
	c31	1	80	Off		NWT	4.68	0.5	c14	4x24.4	4.48	0.65	3.27	BT	EXCEED
		2	80	Off		NWT	4.68	0.5	c19	2x24.4	4.67	0.57	0.11	NWT	PASS
	c32	1	13.2	Off	5%	NWT	3.36	0.98	c20	3x7.2	2.57	1.04	7.38	BT	EXCEED
	c33	1	16.4	Off	5%	NWT	3.68	1	c21	4x7.2	2.51	1.16	10.22	BT	EXCEED
	c34	1	24.4	Off	5%	NWT	3.83	0.92	c22	4x8	2.73	1.07	10.5	BT	EXCEED
	c35	1	48	Off	5%	NWT	4.22	0.8	c23	4x16.4	4	0.9	2.48	BT	EXCEED
	c36	1	64	Off	5%	NWT	4.37	0.76	c24	4x24.4	4.29	0.74	0.99	NWT	PASS
b	c25	1	13.2	Off		NWT	3.36	1.03	c09	3x7.2	2.37	0.98	9.32	BT	EXCEED
		2	13.2	Off		NWT	3.36	1.03	c15	2x7.2	2.41	0.94	9.11	BT	EXCEED
	c26	1	16.4	Off		NWT	3.78	0.92	c10	4x7.2	2.36	1	14.09	BT	EXCEED
		2	16.4	Off		NWT	3.78	0.92	c15	2x7.2	2.41	0.94	13.97	BT	EXCEED
	c27	1	24.4	Off		NWT	3.98	0.89	c11	4x8	2.43	0.91	16.43	BT	EXCEED
		2	24.4	Off		NWT	3.98	0.89	c16	2x8	2.57	1.05	13.84	BT	EXCEED
	c28	1	32	Off		NWT	4.04	0.92	c12	4x9.6	3.54	1	4.94	BT	EXCEED
		2	32	Off		NWT	4.04	0.92	c17	2x9.6	3.62	1.16	3.78	BT	EXCEED
	c29	1	48	Off		NWT	4.46	0.74	c13	4x16.4	4.13	0.92	3.73	BT	EXCEED
		2	48	Off		NWT	4.46	0.74	c18	2x16.4	4.46	0.68	0.07	NWT	PASS
	c30	1	64	Off		NWT	4.75	0.48	c14	4x24.4	4.42	0.72	5.17	BT	EXCEED
		2	64	Off		NWT	4.75	0.48	c19	2x24.4	4.58	0.66	2.74	BT	EXCEED
	c31	1	80	Off		NWT	4.74	0.51	c14	4x24.4	4.42	0.72	5	BT	EXCEED
		2	80	Off		NWT	4.74	0.51	c19	2x24.4	4.58	0.66	2.6	BT	EXCEED
	c32	1	13.2	Off	5%	NWT	3.09	1.07	c20	3x7.2	2.05	0.86	10.16	BT	EXCEED
	c33	1	16.4	Off	5%	NWT	3.27	1.04	c21	4x7.2	2.11	0.9	11.3	BT	EXCEED
	c34	1	24.4	Off	5%	NWT	3.51	1.01	c22	4x8	2.26	0.88	12.51	BT	EXCEED
	c35	1	48	Off	5%	NWT	4.23	0.86	c23	4x16.4	3.9	0.94	3.51	BT	EXCEED
	c36	1	64	Off	5%	NWT	4.34	0.76	c24	4x24.4	4.14	0.89	2.36	BT	EXCEED

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.5-2: Summary of the results of P800-8

Lab	Cond.	Bitrate	DTX	FER	ToR	Status
a	c25	13.2	Off		NWT c09 OR NWT c15	EXCEED
	c26	16.4	Off		NWT c10 OR NWT c15	EXCEED
	c27	24.4	Off		NWT c11 OR NWT c16	EXCEED

	c28	32	Off		NWT c12 OR NWT c17	EXCEED
	c29	48	Off		NWT c13 OR NWT c18	PASS
	c30	64	Off		NWT c14 OR NWT c19	PASS
	c31	80	Off		NWT c14 OR NWT c19	PASS
	c32	13.2	Off	5%	NWT c20	EXCEED
	c33	16.4	Off	5%	NWT c21	EXCEED
	c34	24.4	Off	5%	NWT c22	EXCEED
	c35	48	Off	5%	NWT c23	EXCEED
	c36	64	Off	5%	NWT c24	PASS
b	c25	13.2	Off		NWT c09 OR NWT c15	EXCEED
	c26	16.4	Off		NWT c10 OR NWT c15	EXCEED
	c27	24.4	Off		NWT c11 OR NWT c16	EXCEED
	c28	32	Off		NWT c12 OR NWT c17	EXCEED
	c29	48	Off		NWT c13 OR NWT c18	PASS
	c30	64	Off		NWT c14 OR NWT c19	EXCEED
	c31	80	Off		NWT c14 OR NWT c19	EXCEED
	c32	13.2	Off	5%	NWT c20	EXCEED
	c33	16.4	Off	5%	NWT c21	EXCEED
	c34	24.4	Off	5%	NWT c22	EXCEED
	c35	48	Off	5%	NWT c23	EXCEED
	c36	64	Off	5%	NWT c24	EXCEED

The following diagrams show the results for a range of conditions from experiment P800-8 as rate-distortion curves. The first two diagrams only show results for clean channel conditions, i.e. conditions c10 – c14 for EVS conditions and c25 – c31 for IVAS conditions. The second two diagrams show results for conditions with 5% simulated frame loss, i.e. conditions c21 – c24 for EVS conditions and c32 – c36 for IVAS conditions.

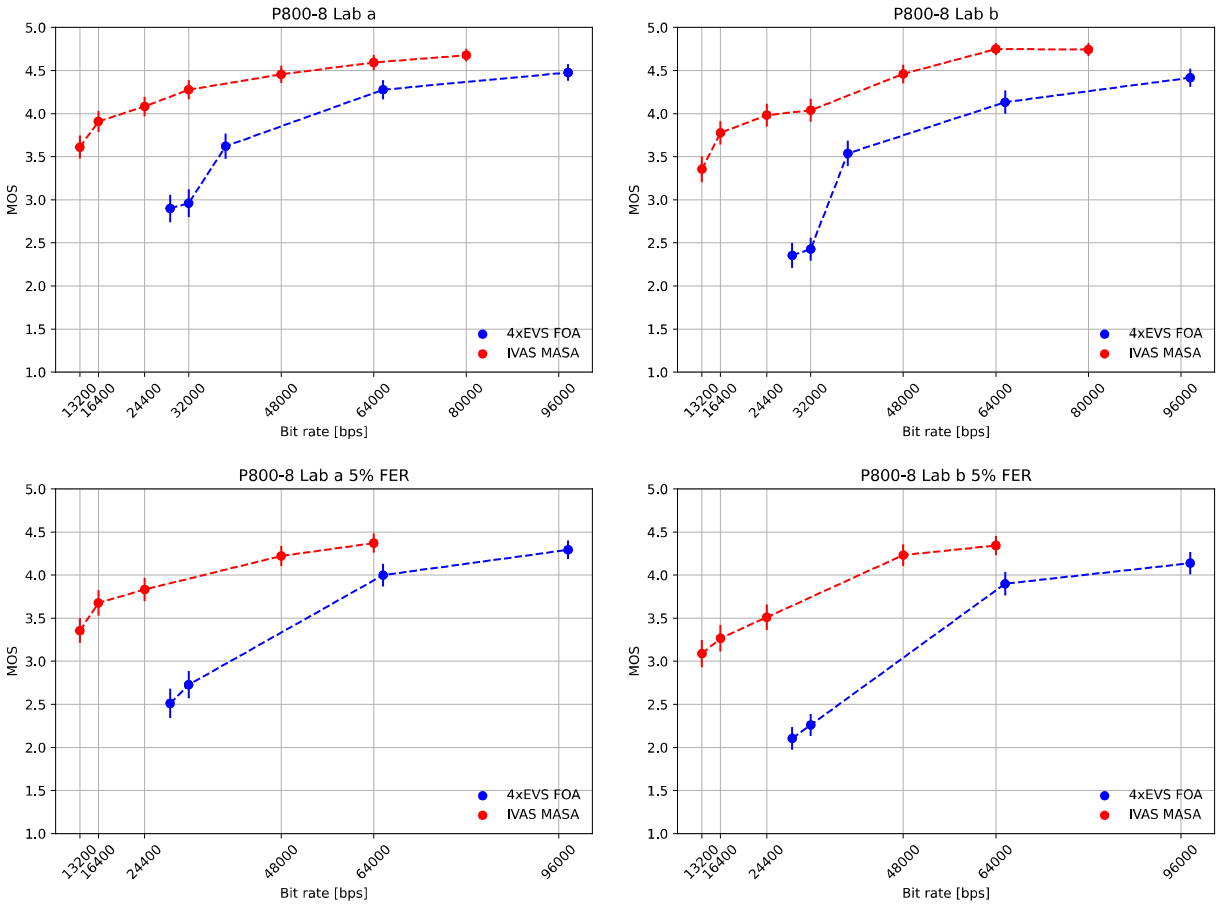


Figure 9.5-1: P800-8 (MASA, clean speech, headphone presentation) rate distortion curves for clean and impaired channel conditions

9.5.3 Selection Experiment P800-9 (MASA, Speech+Background, Headphone Presentation)

Selection Experiment P800-9 evaluates IVAS for MASA fors peech + background conditions under clean channel conditions, DTX off and on using headphone presentation. See Annex C.9 for details.

The complete statistical evaluation of the requirement ToR tests for experiment P800-9 is given in the following table. The evaluation is done separately for the data from the two listening laboratories.

Table 9.5-3: Statistical overview on the results of P800-9

		Type	CuT					EVS Reference				Evaluation		
		Value	Bitrate	DTX	Req.	MOS	Std.	Cond.	Bitrate	MOS	Std.	T-Stat	Result	State
Lab	Cond.	ToR#												
a	c24	1	13.2	off	NWT	3.46	1.1	c09	3x7.2	3.44	1.1	0.15	NWT	PASS
	c25	1	16.4	off	NWT	3.69	0.97	c10	4x7.2	3.47	1.11	2.07	BT	EXCEED
	c26	1	24.4	off	NWT	4.23	0.86	c11	4x8	3.68	1.01	5.5	BT	EXCEED
	c27	1	32	off	NWT	4.31	0.79	c12	4x9.6	4.04	0.9	3.04	BT	EXCEED
	c28	1	48	off	NWT	4.39	0.83	c13	4x16.4	4.38	0.75	0.07	NWT	PASS
	c29	1	64	off	NWT	4.44	0.82	c14	4x24.4	4.49	0.71	-0.62	NWT	PASS
	c30	1	80	off	NWT	4.44	0.78	c14	4x24.4	4.49	0.71	-0.63	NWT	PASS
	c31	1	13.2	on	NWT	3.37	1.12	c18	3x7.2	3.33	1.09	0.33	NWT	PASS
		2	13.2	on	NWT	3.37	1.12	c15	2x7.2	3.35	1.06	0.15	NWT	PASS
	c32	1	16.4	on	NWT	3.8	0.95	c19	4x7.2	3.33	1.09	4.36	BT	EXCEED
		2	16.4	on	NWT	3.8	0.95	c15	2x7.2	3.35	1.06	4.22	BT	EXCEED
	c33	1	24.4	on	NWT	4.09	0.91	c20	4x8	3.33	1.05	7.33	BT	EXCEED
	c34	1	32	on	NWT	4.04	0.96	c21	4x9.6	4.06	0.92	-0.17	NWT	PASS
		2	32	on	NWT	4.04	0.96	c16	2x9.6	4.16	0.91	-1.19	NWT	PASS
	c35	1	48	on	NWT	4.29	0.8	c22	4x16.4	4.31	0.77	-0.21	NWT	PASS
		2	48	on	NWT	4.29	0.8	c17	2x16.4	4.33	0.78	-0.53	NWT	PASS
	c36	1	64	on	NWT	4.4	0.79	c23	4x24.4	4.34	0.8	0.67	NWT	PASS
d	c24	1	13.2	off	NWT	3.44	0.89	c09	3x7.2	2.49	0.8	10.57	BT	EXCEED
	c25	1	16.4	off	NWT	3.78	0.92	c10	4x7.2	2.66	0.88	11.79	BT	EXCEED
	c26	1	24.4	off	NWT	3.71	0.86	c11	4x8	2.67	0.87	11.35	BT	EXCEED
	c27	1	32	off	NWT	3.84	0.76	c12	4x9.6	3.73	0.78	1.36	NWT	PASS
	c28	1	48	off	NWT	4.28	0.64	c13	4x16.4	4.15	0.78	1.71	BT	EXCEED
	c29	1	64	off	NWT	4.24	0.71	c14	4x24.4	4.32	0.74	-1.02	NWT	PASS
	c30	1	80	off	NWT	4.29	0.67	c14	4x24.4	4.32	0.74	-0.38	NWT	PASS
	c31	1	13.2	on	NWT	3.25	0.88	c18	3x7.2	2.52	0.8	8.24	BT	EXCEED
		2	13.2	on	NWT	3.25	0.88	c15	2x7.2	2.46	0.88	8.49	BT	EXCEED
	c32	1	16.4	on	NWT	3.58	0.86	c19	4x7.2	2.57	0.86	11.13	BT	EXCEED
		2	16.4	on	NWT	3.58	0.86	c15	2x7.2	2.46	0.88	12.25	BT	EXCEED
	c33	1	24.4	on	NWT	3.65	0.84	c20	4x8	2.61	0.87	11.63	BT	EXCEED
	c34	1	32	on	NWT	3.86	0.74	c21	4x9.6	3.73	0.91	1.46	NWT	PASS
		2	32	on	NWT	3.86	0.74	c16	2x9.6	3.83	0.83	0.34	NWT	PASS
	c35	1	48	on	NWT	4.06	0.78	c22	4x16.4	3.89	0.91	1.88	BT	EXCEED
		2	48	on	NWT	4.06	0.78	c17	2x16.4	3.96	0.81	1.14	NWT	PASS
	c36	1	64	on	NWT	4.22	0.73	c23	4x24.4	4.09	0.79	1.6	NWT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.5-4: Summary of the results of P800-9

Lab	Cond.	Bitrate	DTX	ToR	Status
a	c24	13.2	off	NWT c09	PASS
	c25	16.4	off	NWT c10	EXCEED
	c26	24.4	off	NWT c11	EXCEED
	c27	32	off	NWT c12	EXCEED
	c28	48	off	NWT c13	PASS
	c29	64	off	NWT c14	PASS
	c30	80	off	NWT c14	PASS
	c31	13.2	on	NWT c18 or NWT c15	PASS
	c32	16.4	on	NWT c19 or NWT c15	EXCEED
	c33	24.4	on	NWT c20	EXCEED
	c34	32	on	NWT c21 or NWT c16	PASS
	c35	48	on	NWT c22 or NWT c17	PASS
	c36	64	on	NWT c23	PASS
d	c24	13.2	off	NWT c09	EXCEED
	c25	16.4	off	NWT c10	EXCEED
	c26	24.4	off	NWT c11	EXCEED
	c27	32	off	NWT c12	PASS
	c28	48	off	NWT c13	EXCEED

	c29	64	off	NWT c14	PASS
	c30	80	off	NWT c14	PASS
	c31	13.2	on	NWT c18 or NWT c15	EXCEED
	c32	16.4	on	NWT c19 or NWT c15	EXCEED
	c33	24.4	on	NWT c20	EXCEED
	c34	32	on	NWT c21 or NWT c16	PASS
	c35	48	on	NWT c22 or NWT c17	PASS
	c36	64	on	NWT c23	PASS

The following diagrams show the results for a range of conditions from experiment P800-9 as rate-distortion curves. The first two diagrams only show results for active coding conditions (DTX off), i.e. conditions c10 – c14 for EVS conditions and c24 – c30 for IVAS conditions. The second two diagrams show results for DTX conditions (DTX on), i.e. conditions c19 – c23 for EVS conditions and c31 – c36 for IVAS conditions.

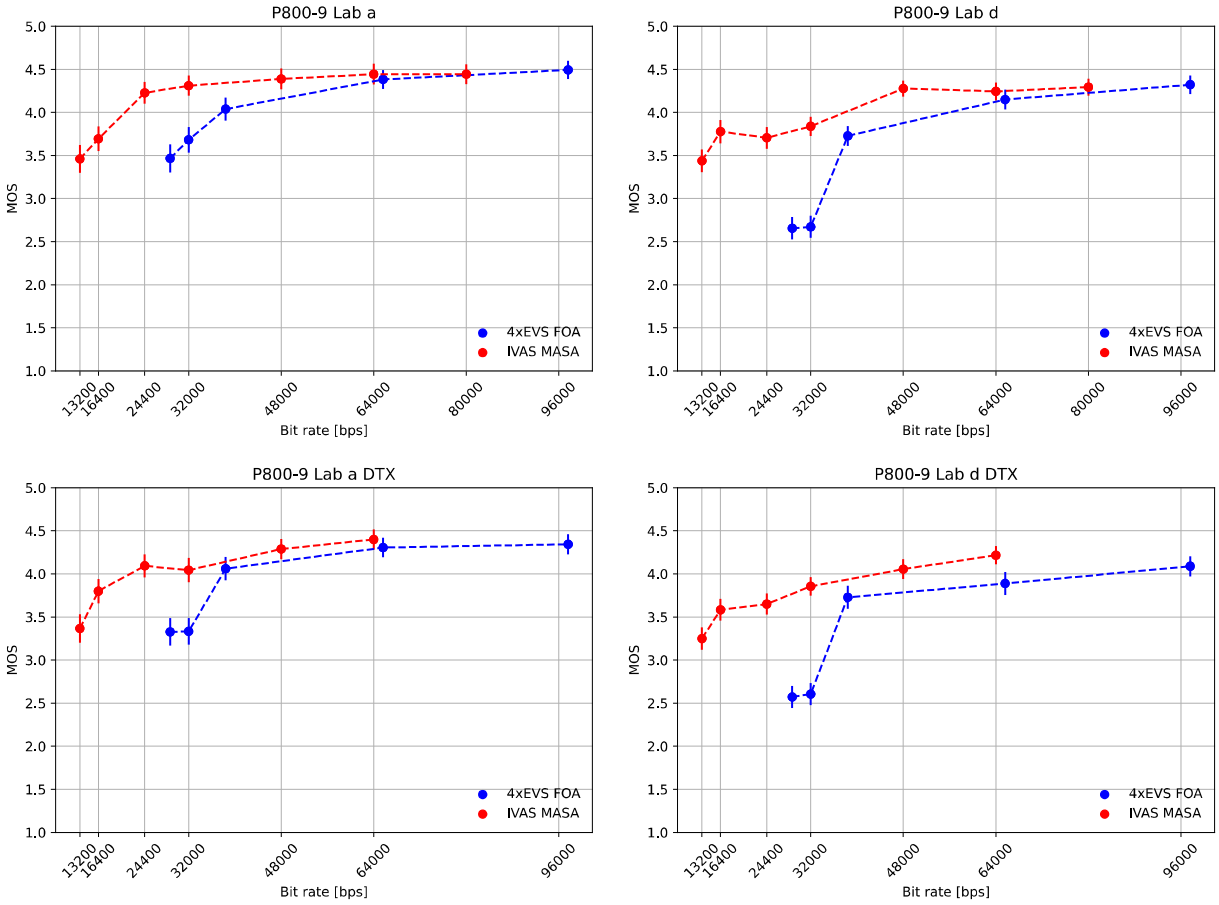


Figure 9.5-2: Figure 16: P800-9 (MASA, speech + background, headphone presentation) rate distortion curves for conditions with DTX off and on

9.5.4 Selection Experiment BS1534-7a (MASA, Generic Audio, 96 and 128 kbps, Headphone Presentation)

Selection Experiment BS1534-7a evaluates IVAS for MASA, generic audio, at 96 and 128 kbps using headphone presentation. See Annex C.22 for details.

The averaged results per condition for experiment BS1534-7a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS multi-mono FOA conditions with increasing bitrate (c03 – c04), EVS dual-mono MASA conditions with increasing bitrate (c05 – c06) and IVAS conditions with increasing bitrate (c07 – c08).

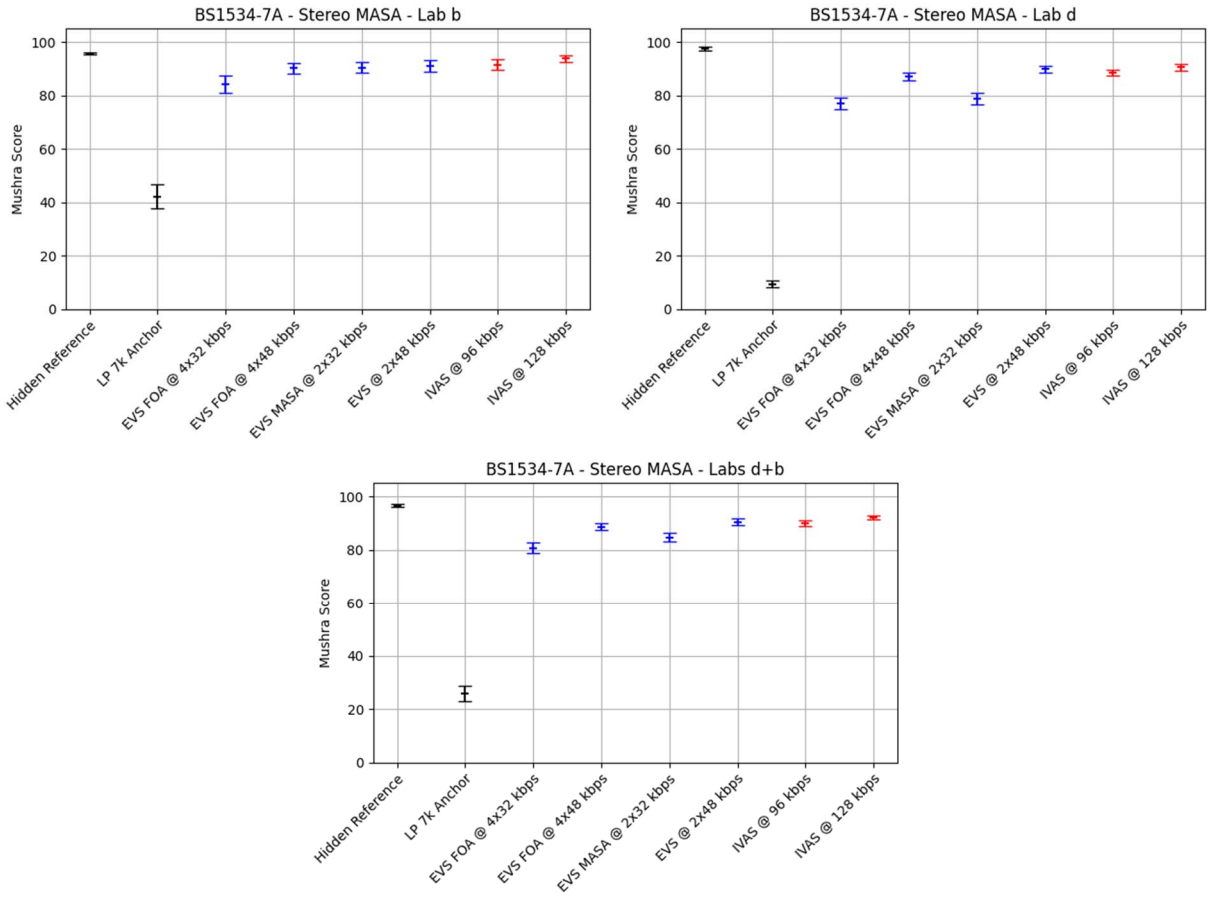


Figure 9.5-3: BS1534-7a (MASA, generic audio, 96, 128 and 192 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-7a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.5-5: Statistical overview on the results of BS1534-7a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c07	1	96	NWT	91.6	11.9	c03	4x32	84.1	20.9	4.01	BT	EXCEED
		2	96	NWT	91.6	11.9	c05	2x32, unq. MD	90.5	12.6	0.85	NWT	PASS
	c08	1	128	NWT	93.9	7.7	c04	4x48	90.2	13.1	3.14	BT	EXCEED
		2	128	NWT	93.9	7.7	c06	2x48, unq. MD	91.2	14.1	2.14	BT	EXCEED
d	c07	1	96	NWT	88.5	7.9	c03	4x32	77.2	14.3	9.05	BT	EXCEED
		2	96	NWT	88.5	7.9	c05	2x32, unq. MD	78.8	15	7.45	BT	EXCEED
	c08	1	128	NWT	90.6	7.5	c04	4x48	87	8.9	3.91	BT	EXCEED
		2	128	NWT	90.6	7.5	c06	2x48, unq. MD	89.8	8.8	0.82	NWT	PASS
d+b	c07	1	96	NWT	90.1	10.2	c03	4x32	80.7	18.2	8.27	BT	EXCEED
		2	96	NWT	90.1	10.2	c05	2x32, unq. MD	84.6	15	5.49	BT	EXCEED
	c08	1	128	NWT	92.2	7.7	c04	4x48	88.6	11.3	4.81	BT	EXCEED
		2	128	NWT	92.2	7.7	c06	2x48, unq. MD	90.5	11.8	2.2	BT	EXCEED

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.5-6: Summary of the results of BS1534-7a

Lab	Cond.	Bitrate	ToR	Status
b	c07	96	NWT c03 OR NWT c05	PASS
	c08	128	NWT c04 OR NWT c06	EXCEED

d	c07	96	NWT c03 OR NWT c05	EXCEED
	c08	128	NWT c04 OR NWT c06	PASS
d+b	c07	96	NWT c03 OR NWT c05	EXCEED
	c08	128	NWT c04 OR NWT c06	EXCEED

9.5.5 Selection Experiment BS1534-7b (MASA, Generic Audio, 192 and 256 kbps, Headphone Presentation)

Selection Experiment BS1534-7b evaluates IVAS for MASA generic audio, at 192 and 256 kbps using headphone presentation. See Annex C.23 for details.

The averaged results per condition for experiment BS1534-7b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS multi-mono FOA conditions with increasing bitrate (c03 – c04), EVS dual-mono MASA conditions with increasing bitrate (c05 – c06) and IVAS conditions with increasing bitrate (c07 – c08).

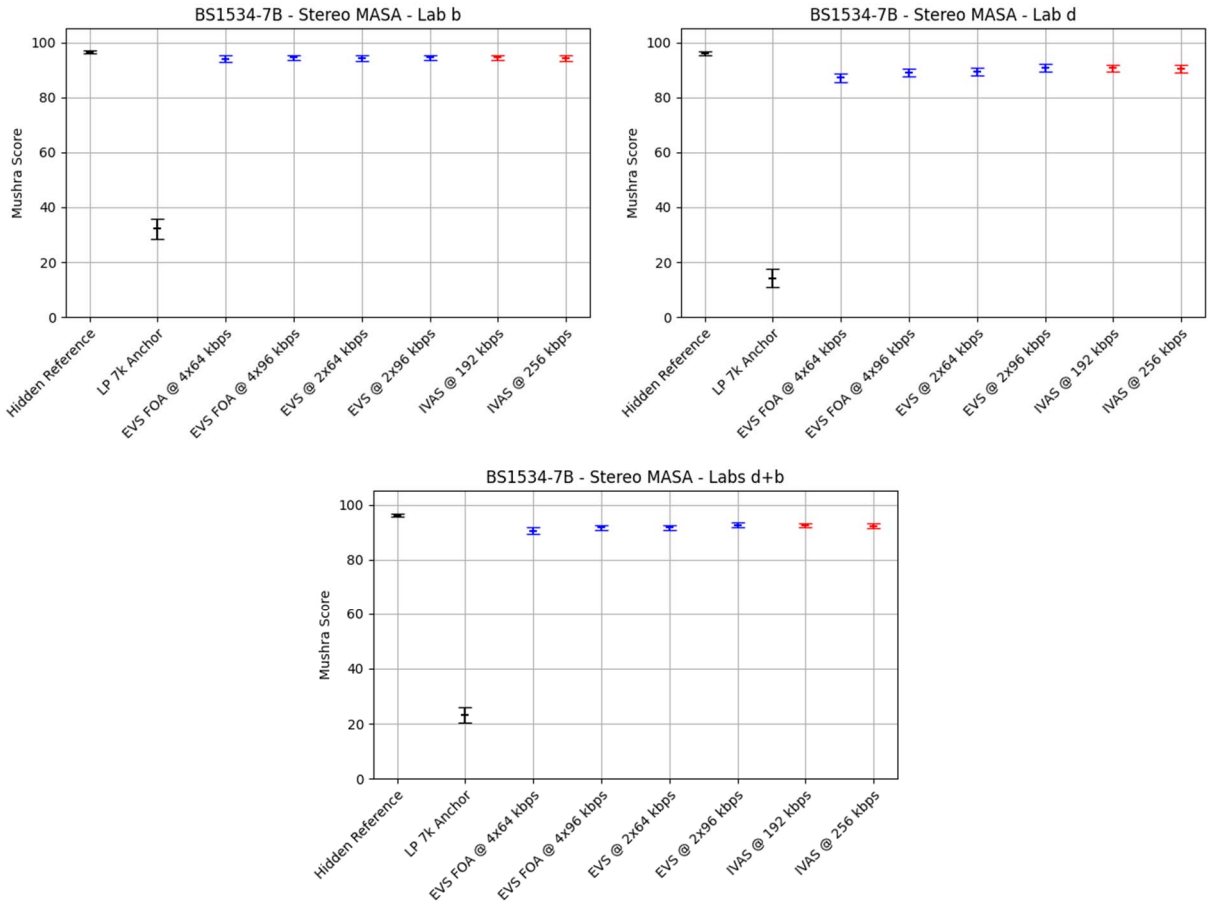


Figure 9.5-4: BS1534-7b (MASA, generic audio, 192 and 256 kbps, headphone presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-7b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.5-7: Statistical overview on the results of BS1534-7b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
b	c07	1	192	NWT	94.5	6.5	c03	4x64	93.9	8	0.71	NWT	PASS
		2	192	NWT	94.5	6.5	c05	2x64, unq. MD	94.2	6.6	0.37	NWT	PASS
	c08	1	256	NWT	94.2	6.1	c04	4x96	94.5	6	-0.41	NWT	PASS
		2	256	NWT	94.2	6.1	c06	2x96, unq. MD	94.5	6.3	-0.37	NWT	PASS
d	c07	1	192	NWT	90.6	8.4	c03	4x64	87.2	10.5	3.26	BT	EXCEED
		2	192	NWT	90.6	8.4	c05	2x64, unq. MD	89.4	8.8	1.33	NWT	PASS
	c08	1	256	NWT	90.4	8.6	c04	4x96	88.8	9.3	1.64	NWT	PASS
		2	256	NWT	90.4	8.6	c06	2x96, unq. MD	90.7	9.1	-0.23	NWT	PASS
d+b	c07	1	192	NWT	92.5	7.7	c03	4x64	90.6	9.9	2.87	BT	EXCEED
		2	192	NWT	92.5	7.7	c05	2x64, unq. MD	91.8	8.1	1.23	NWT	PASS
	c08	1	256	NWT	92.3	7.7	c04	4x96	91.7	8.3	1.07	NWT	PASS
		2	256	NWT	92.3	7.7	c06	2x96, unq. MD	92.6	8.1	-0.39	NWT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.5-8: Summary of the results of BS1534-7b

Lab	Cond.	Bitrate	ToR	Status
b	c07	192	NWT c03 OR NWT c05	PASS
	c08	256	NWT c04 OR NWT c06	PASS
d	c07	192	NWT c03 OR NWT c05	PASS
	c08	256	NWT c04 OR NWT c06	PASS
d+b	c07	192	NWT c03 OR NWT c05	PASS
	c08	256	NWT c04 OR NWT c06	PASS

9.5.6 Characterization Experiment P800-12 (MASA 1TC, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-12 evaluates the performance of the IVAS fixed-point and floating-point implementation for MASA 1TC for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-12 was conducted by Mesaqin.com (language: MAN).

The listening lab has reported that for each panel, up to seven instead of five participants participated. The five subjects with the highest PCC were selected for result submission.

The aggregated results per condition are listed in Table 9.5-9 and are shown in Figure 9.5-5. The results aggregated over FX/FL are shown in Figure 9.5-6. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-10. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.5-9: results per condition for experiment P800-12

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.79	0.42	0.06
c02	MNRU Q = 32 dB	180	4.24	0.88	0.13
c03	MNRU Q = 27 dB	180	3.53	1.27	0.19
c04	MNRU Q = 22 dB	180	2.37	1.18	0.17
c05	MNRU Q = 17 dB	180	1.57	0.75	0.11
c06	ESDRU α = 0.8	180	4.52	0.74	0.11
c07	ESDRU α = 0.6	180	3.86	1.11	0.16
c08	ESDRU α = 0.4	180	2.89	1.18	0.17
c09	ESDRU α = 0.2	180	2.25	1.17	0.17
c10	IVAS FL enc / FX dec, 13.2 kbps, DTX off	180	3.52	1.22	0.18
c11	IVAS FX enc / FL dec, 16.4 kbps, DTX off	180	3.96	1.04	0.15
c12	IVAS FL enc / FX dec, 24.4 kbps, DTX off	180	4.19	0.95	0.14
c13	IVAS FX enc / FL dec, 32 kbps, DTX off	180	4.25	0.88	0.13

c14	IVAS FL enc / FX dec, 48 kbps, DTX off	180	4.43	0.73	0.11
c15	IVAS FX enc / FL dec, 64 kbps, DTX off	180	4.46	0.80	0.12
c16	IVAS FL enc / FX dec, 80 kbps, DTX off	180	4.59	0.69	0.10
c17	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.56	0.68	0.10
c18	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.56	0.69	0.10
c19	IVAS FL, 13.2 kbps, DTX off	180	3.51	1.18	0.17
c20	IVAS FL, 16.4 kbps, DTX off	180	4.18	1.00	0.15
c21	IVAS FL, 24.4 kbps, DTX off	180	4.18	0.92	0.14
c22	IVAS FL, 32 kbps, DTX off	180	4.26	0.89	0.13
c23	IVAS FL, 48 kbps, DTX off	180	4.40	0.84	0.12
c24	IVAS FL, 64 kbps, DTX off	180	4.53	0.73	0.11
c25	IVAS FL, 80 kbps, DTX off	180	4.48	0.77	0.11
c26	IVAS FL, 96 kbps, DTX off	180	4.54	0.73	0.11
c27	IVAS FL, 128 kbps, DTX off	180	4.57	0.72	0.11
c28	IVAS FX, 13.2 kbps, DTX off	180	3.79	1.18	0.17
c29	IVAS FX, 16.4 kbps, DTX off	180	3.97	1.07	0.16
c30	IVAS FX, 24.4 kbps, DTX off	180	4.24	0.89	0.13
c31	IVAS FX, 32 kbps, DTX off	180	4.28	0.92	0.14
c32	IVAS FX, 48 kbps, DTX off	180	4.28	0.93	0.14
c33	IVAS FX, 64 kbps, DTX off	180	4.47	0.78	0.12
c34	IVAS FX, 80 kbps, DTX off	180	4.56	0.69	0.10
c35	IVAS FX, 96 kbps, DTX off	180	4.47	0.77	0.11
c36	IVAS FX, 128 kbps, DTX off	180	4.48	0.75	0.11

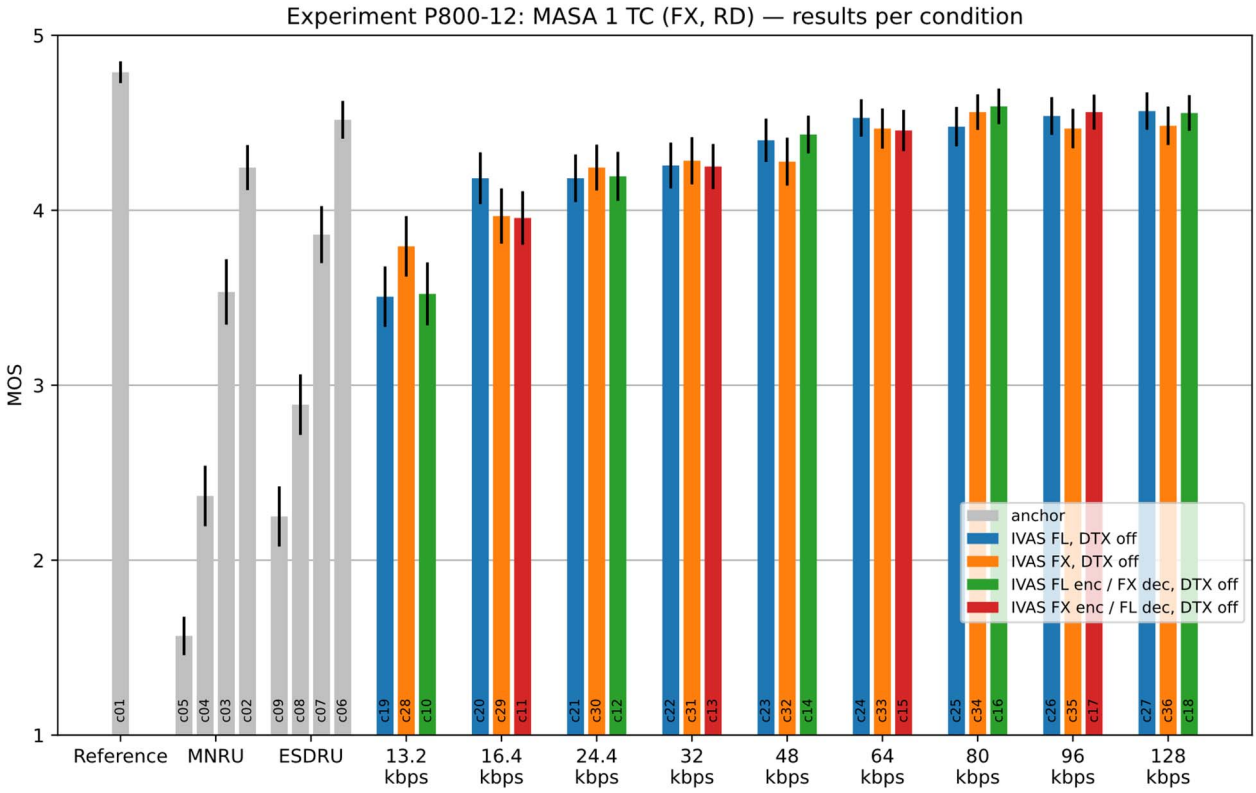


Figure 9.5-5: results per condition for experiment P800-12

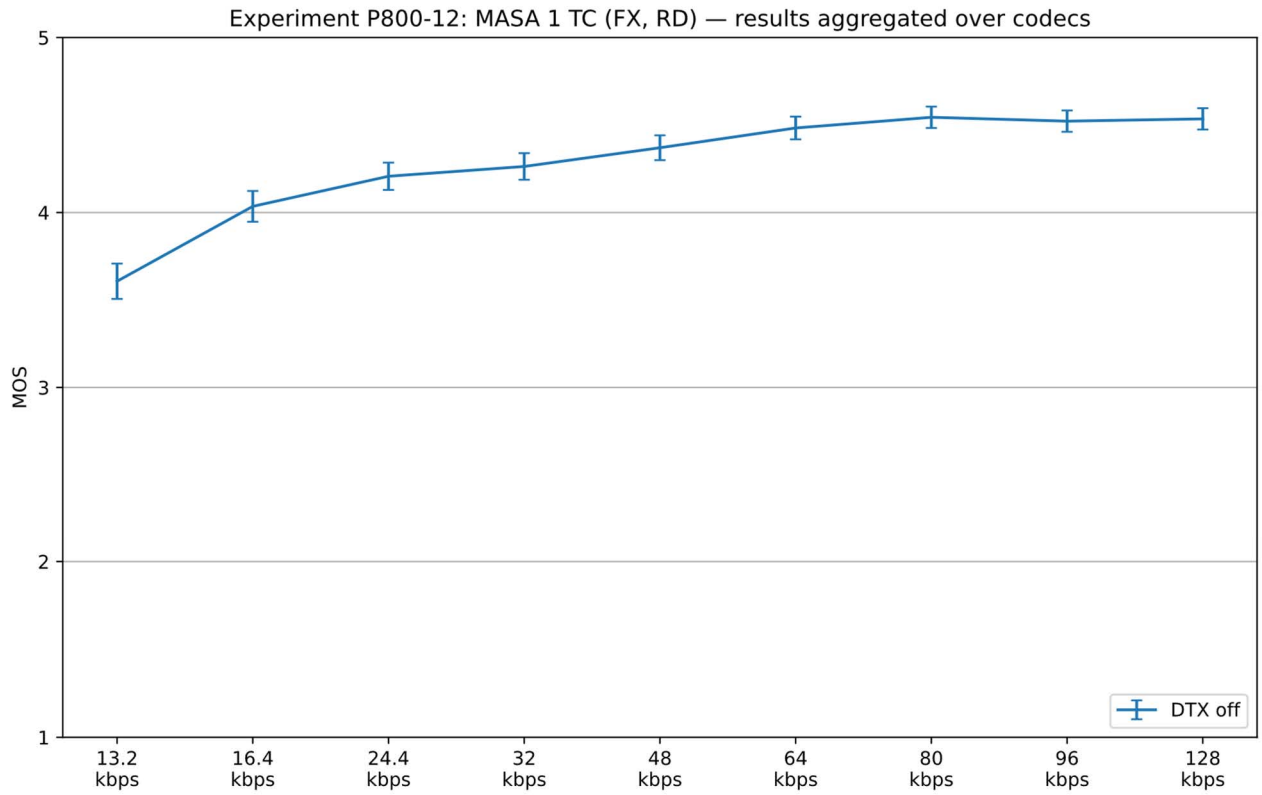


Figure 9.5-6: results aggregated over FX/FL for experiment P800-12

Table 9.5-10: statistical significance analysis for experiment P800-12

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.52	1.22	c19	3.51	1.18	0.02	0.126	0.450	NWT
c11	3.96	1.04	c20	4.18	1.00	-0.23	<0.001	0.982	WT
c12	4.19	0.95	c21	4.18	0.92	0.01	0.111	0.456	NWT
c13	4.25	0.88	c22	4.26	0.89	-0.01	<0.001	0.526	NWT
c14	4.43	0.73	c23	4.40	0.84	0.03	0.396	0.346	NWT
c15	4.46	0.80	c24	4.53	0.73	-0.07	<0.001	0.814	NWT
c16	4.59	0.69	c25	4.48	0.77	0.12	1.509	0.066	NWT
c17	4.56	0.68	c26	4.54	0.73	0.02	0.295	0.384	NWT
c18	4.56	0.69	c27	4.57	0.72	-0.01	<0.001	0.558	NWT
c28	3.79	1.18	c19	3.51	1.18	0.29	2.320	0.010	BT
c29	3.97	1.07	c20	4.18	1.00	-0.22	<0.001	0.975	WT
c30	4.24	0.89	c21	4.18	0.92	0.06	0.639	0.262	NWT
c31	4.28	0.92	c22	4.26	0.89	0.03	0.283	0.389	NWT
c32	4.28	0.93	c23	4.40	0.84	-0.12	<0.001	0.903	NWT
c33	4.47	0.78	c24	4.53	0.73	-0.06	<0.001	0.778	NWT
c34	4.56	0.69	c25	4.48	0.77	0.08	1.077	0.141	NWT
c35	4.47	0.77	c26	4.54	0.73	-0.07	<0.001	0.818	NWT
c36	4.48	0.75	c27	4.57	0.72	-0.08	<0.001	0.859	NWT

9.5.7 Characterization Experiment P800-13 (MASA 2TC, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-13 evaluates the performance of the IVAS fixed-point and floating-point implementation for MASA 2TC for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-13 was conducted by FORCE Technology (language: DAN).

The aggregated results per condition are listed in Table 9.5-11 and are shown in Figure 9.5-7. The results aggregated over FX/FL are shown in Figure 9.5-8. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-12. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.5-11: results per condition for experiment P800-13

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.78	0.45	0.07
c02	MNRU Q = 32 dB	180	3.79	1.07	0.16
c03	MNRU Q = 27 dB	180	2.85	1.16	0.17
c04	MNRU Q = 22 dB	180	1.92	0.94	0.14
c05	MNRU Q = 17 dB	180	1.25	0.51	0.07
c06	ESDRU α = 0.8	180	4.66	0.59	0.09
c07	ESDRU α = 0.6	180	4.11	0.86	0.13
c08	ESDRU α = 0.4	180	3.39	1.07	0.16
c09	ESDRU α = 0.2	180	3.00	1.17	0.17
c10	IVAS FL enc / FX dec, 16.4 kbps, DTX off	180	3.67	1.11	0.16
c11	IVAS FX enc / FL dec, 24.4 kbps, DTX off	180	3.59	1.00	0.15
c12	IVAS FL enc / FX dec, 32 kbps, DTX off	180	3.88	0.95	0.14
c13	IVAS FX enc / FL dec, 48 kbps, DTX off	180	4.27	0.74	0.11
c14	IVAS FL enc / FX dec, 64 kbps, DTX off	180	4.35	0.74	0.11
c15	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.59	0.63	0.09
c16	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.54	0.64	0.09
c17	IVAS FX enc / FL dec, 192 kbps, DTX off	180	4.59	0.65	0.10
c18	IVAS FL enc / FX dec, 384 kbps, DTX off	180	4.58	0.62	0.09
c19	IVAS FL, 16.4 kbps, DTX off	180	3.46	1.14	0.17
c20	IVAS FL, 24.4 kbps, DTX off	180	3.69	0.93	0.14
c21	IVAS FL, 32 kbps, DTX off	180	3.81	0.99	0.15
c22	IVAS FL, 48 kbps, DTX off	180	4.17	0.83	0.12
c23	IVAS FL, 64 kbps, DTX off	180	4.35	0.70	0.10
c24	IVAS FL, 96 kbps, DTX off	180	4.54	0.66	0.10
c25	IVAS FL, 128 kbps, DTX off	180	4.51	0.61	0.09
c26	IVAS FL, 192 kbps, DTX off	180	4.55	0.67	0.10
c27	IVAS FL, 384 kbps, DTX off	180	4.56	0.63	0.09
c28	IVAS FX, 16.4 kbps, DTX off	180	3.64	1.08	0.16
c29	IVAS FX, 24.4 kbps, DTX off	180	3.66	1.06	0.16
c30	IVAS FX, 32 kbps, DTX off	180	3.71	1.07	0.16
c31	IVAS FX, 48 kbps, DTX off	180	4.12	0.85	0.12
c32	IVAS FX, 64 kbps, DTX off	180	4.39	0.71	0.10
c33	IVAS FX, 96 kbps, DTX off	180	4.46	0.73	0.11
c34	IVAS FX, 128 kbps, DTX off	180	4.59	0.57	0.08
c35	IVAS FX, 192 kbps, DTX off	180	4.62	0.59	0.09
c36	IVAS FX, 384 kbps, DTX off	180	4.63	0.60	0.09

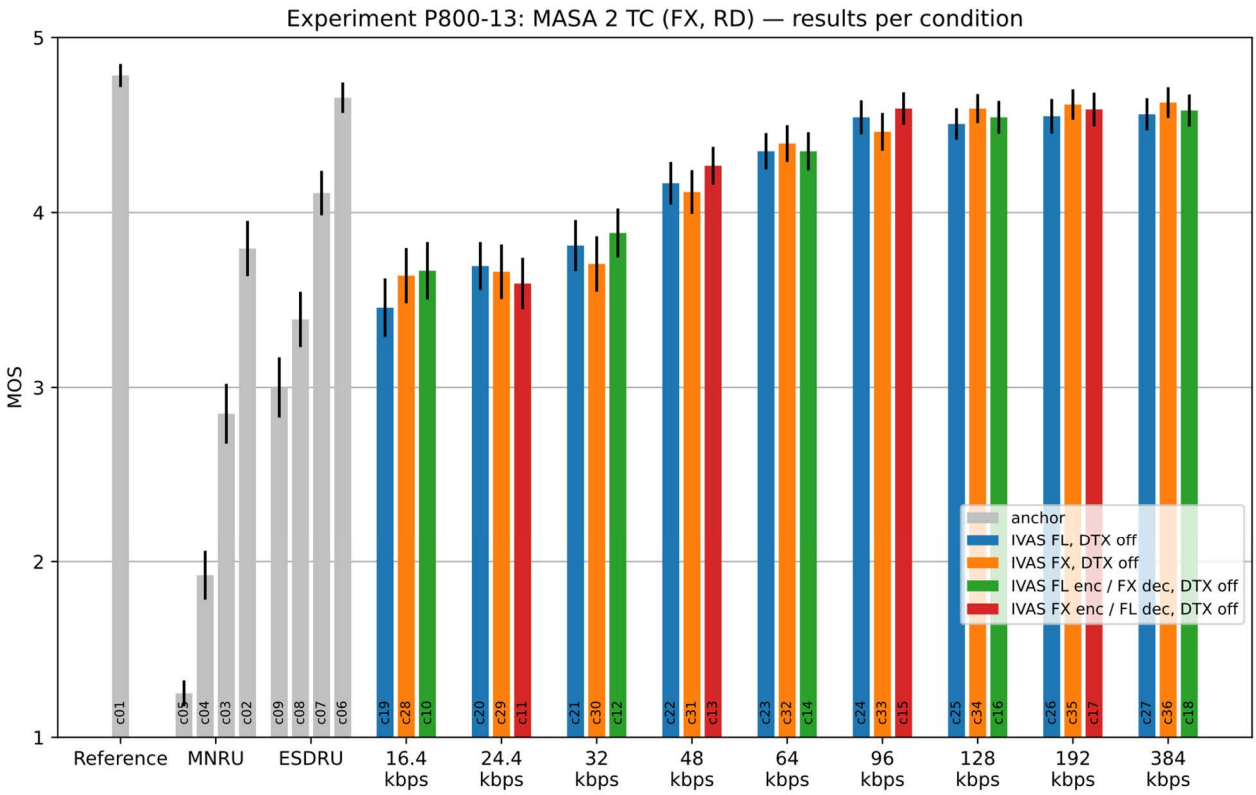


Figure 9.5-7: results per condition for experiment P800-13

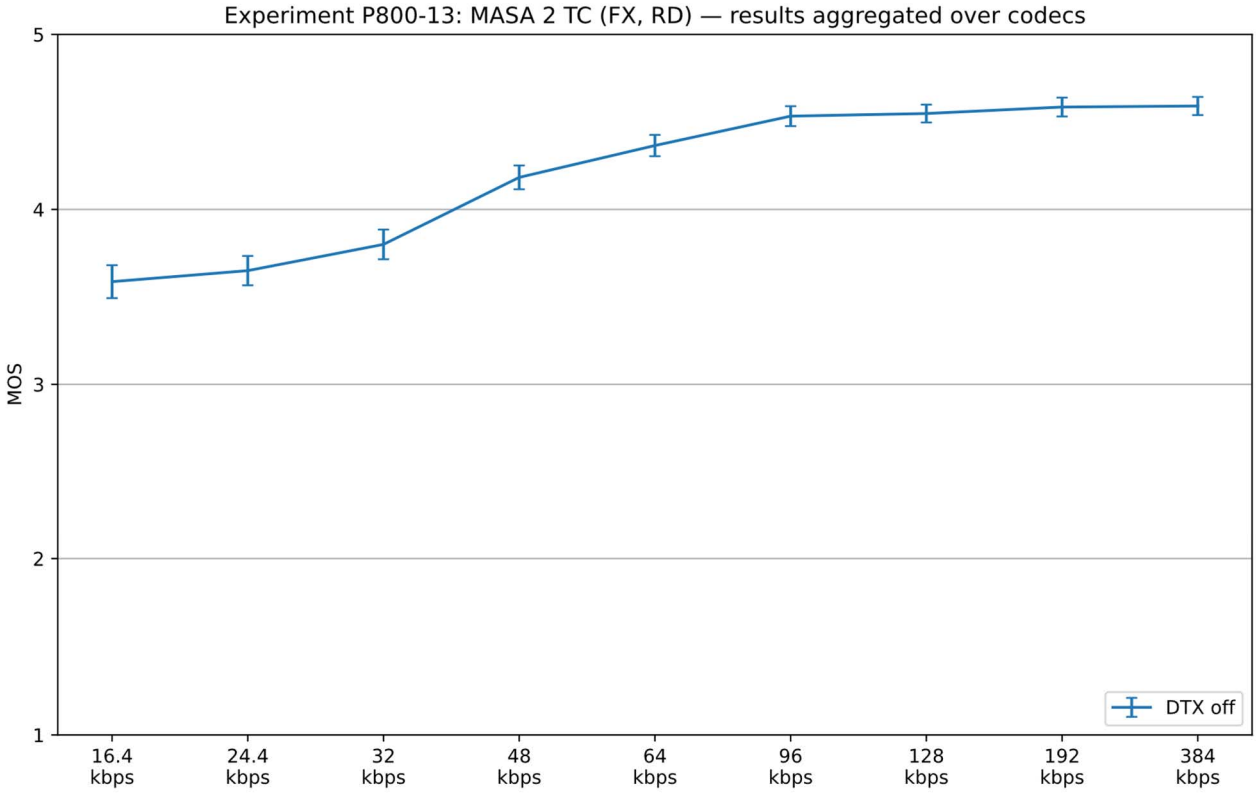


Figure 9.5-8: results aggregated over FX/FL for experiment P800-13

Table 9.5-12: statistical significance analysis for experiment P800-13

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.67	1.11	c19	3.46	1.14	0.21	1.780	0.038	BT
c11	3.59	1.00	c20	3.69	0.93	-0.10	<0.001	0.836	NWT
c12	3.88	0.95	c21	3.81	0.99	0.07	0.703	0.241	NWT
c13	4.27	0.74	c22	4.17	0.83	0.10	1.210	0.114	NWT
c14	4.35	0.74	c23	4.35	0.70	0.00	<0.001	0.500	NWT
c15	4.59	0.63	c24	4.54	0.66	0.05	0.732	0.232	NWT
c16	4.54	0.64	c25	4.51	0.61	0.04	0.577	0.282	NWT
c17	4.59	0.65	c26	4.55	0.67	0.04	0.560	0.288	NWT
c18	4.58	0.62	c27	4.56	0.63	0.02	0.334	0.369	NWT
c28	3.64	1.08	c19	3.46	1.14	0.18	1.569	0.059	NWT
c29	3.66	1.06	c20	3.69	0.93	-0.03	<0.001	0.623	NWT
c30	3.71	1.07	c21	3.81	0.99	-0.10	<0.001	0.833	NWT
c31	4.12	0.85	c22	4.17	0.83	-0.05	<0.001	0.714	NWT
c32	4.39	0.71	c23	4.35	0.70	0.04	0.589	0.278	NWT
c33	4.46	0.73	c24	4.54	0.66	-0.08	<0.001	0.869	NWT
c34	4.59	0.57	c25	4.51	0.61	0.09	1.416	0.079	NWT
c35	4.62	0.59	c26	4.55	0.67	0.07	1.005	0.158	NWT
c36	4.63	0.60	c27	4.56	0.63	0.07	1.037	0.150	NWT

9.5.8 Characterization Experiment P800-14 (MASA 1-2TC, Speech and Mixed/Music, Impaired Channel, DTX on, Headphone Presentation)

Characterization Experiment P800-14 evaluates the performance of the IVAS fixed-point and floating-point implementation for MASA 1-2TC for speech and mixed/music, DTX on, under impaired channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-14 was conducted by Nokia (language: FIN).

The listening lab has reported that the presentation order of the samples to the listeners was not according to the original randomization tables due to a configuration issue in the listening test software.

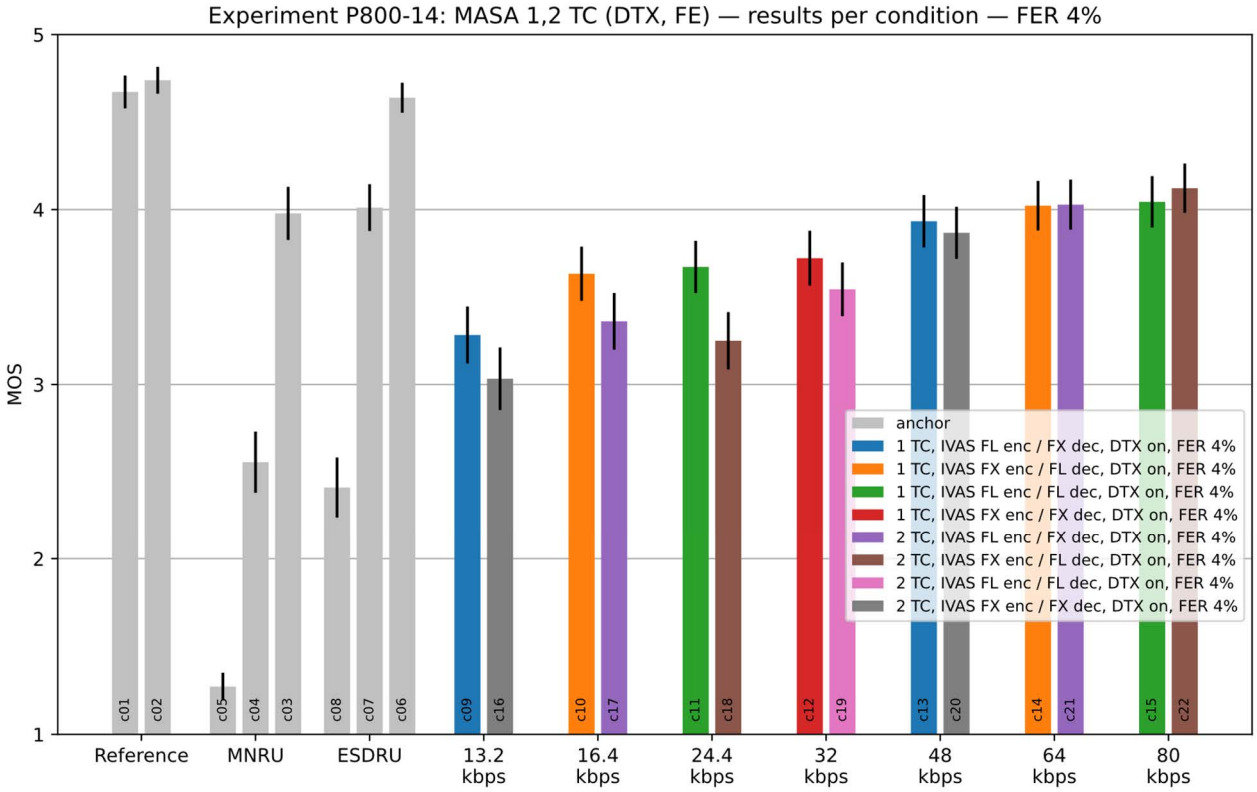
The aggregated results per condition are listed in Table 9.5-13 and are shown in Figure 9.5-9. The results aggregated over FX/FL are shown in Figure 9.5-10. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-14. Significant different MOS values (BT or WT) are highlighted in red.

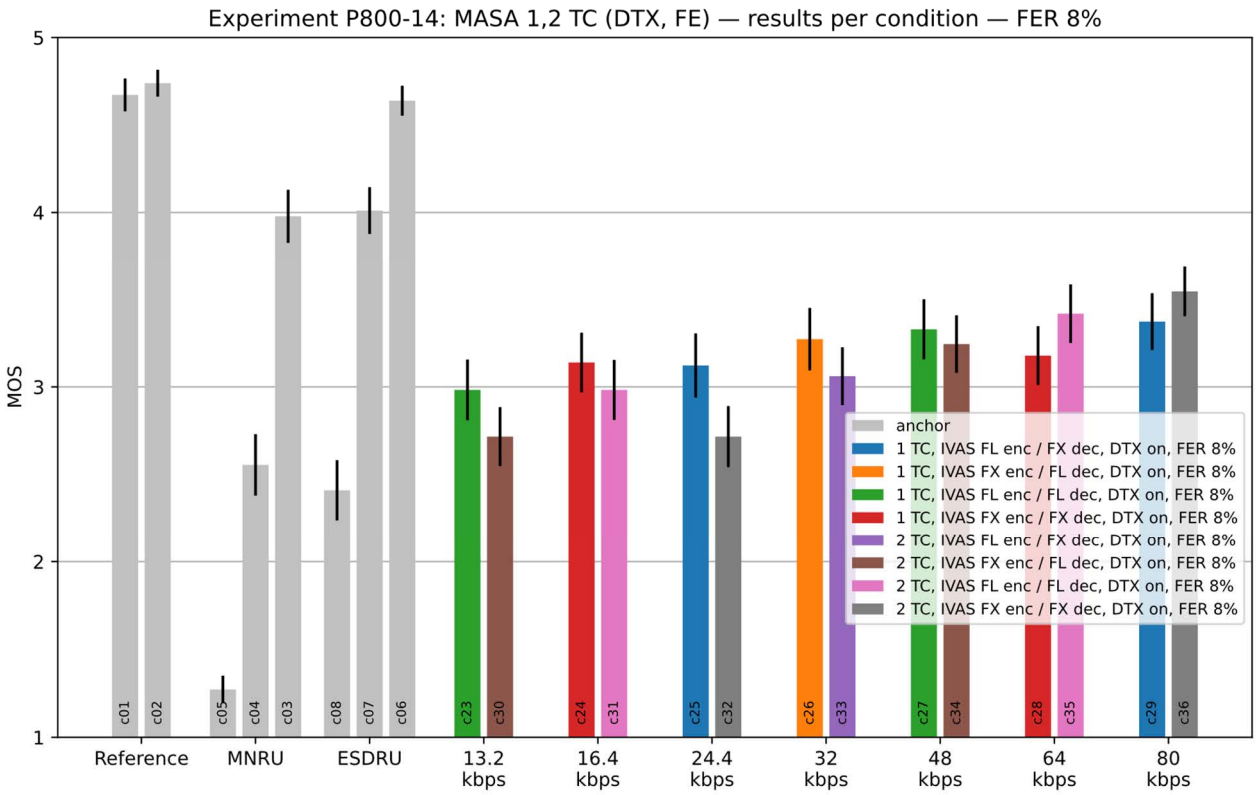
Table 9.5-13: results per condition for experiment P800-14

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference 1 TC, 1 TC	180	4.67	0.64	0.09
c02	Reference 2 TC, 2 TC	180	4.74	0.52	0.08
c03	MNRU Q = 32 dB	180	3.98	1.03	0.15
c04	MNRU Q = 25 dB	180	2.56	1.19	0.17
c05	MNRU Q = 18 dB	180	1.27	0.54	0.08
c06	ESDRU α = 0.8	180	4.64	0.59	0.09
c07	ESDRU α = 0.6	180	4.01	0.91	0.13
c08	ESDRU α = 0.4	180	2.41	1.17	0.17
c09	IVAS FL enc / FX dec, FER 4%, 1 TC, 13.2 kbps, DTX on	180	3.28	1.10	0.16
c10	IVAS FX enc / FL dec, FER 4%, 1 TC, 16.4 kbps, DTX on	180	3.63	1.06	0.15
c11	IVAS FL enc / FL dec, FER 4%, 1 TC, 24.4 kbps, DTX on	180	3.67	1.01	0.15
c12	IVAS FX enc / FX dec, FER 4%, 1 TC, 32 kbps, DTX on	180	3.72	1.07	0.16
c13	IVAS FL enc / FX dec, FER 4%, 1 TC, 48 kbps, DTX on	180	3.93	1.02	0.15
c14	IVAS FX enc / FL dec, FER 4%, 1 TC, 64 kbps, DTX on	180	4.02	0.97	0.14
c15	IVAS FL enc / FL dec, FER 4%, 1 TC, 80 kbps, DTX on	180	4.04	1.00	0.15
c16	IVAS FX enc / FX dec, FER 4%, 2 TC, 13.2 kbps, DTX on	180	3.03	1.21	0.18

c17	IVAS FL enc / FX dec, FER 4%, 2 TC, 16.4 kbps, DTX on	180	3.36	1.10	0.16
c18	IVAS FX enc / FL dec, FER 4%, 2 TC, 24.4 kbps, DTX on	180	3.25	1.11	0.16
c19	IVAS FL enc / FL dec, FER 4%, 2 TC, 32 kbps, DTX on	180	3.54	1.05	0.15
c20	IVAS FX enc / FX dec, FER 4%, 2 TC, 48 kbps, DTX on	180	3.87	1.02	0.15
c21	IVAS FL enc / FX dec, FER 4%, 2 TC, 64 kbps, DTX on	180	4.03	0.97	0.14
c22	IVAS FX enc / FL dec, FER 4%, 2 TC, 80 kbps, DTX on	180	4.12	0.96	0.14
c23	IVAS FL enc / FL dec, FER 8%, 1 TC, 13.2 kbps, DTX on	180	2.98	1.18	0.17
c24	IVAS FX enc / FX dec, FER 8%, 1 TC, 16.4 kbps, DTX on	180	3.14	1.16	0.17
c25	IVAS FL enc / FX dec, FER 8%, 1 TC, 24.4 kbps, DTX on	180	3.12	1.25	0.18
c26	IVAS FX enc / FL dec, FER 8%, 1 TC, 32 kbps, DTX on	180	3.27	1.21	0.18
c27	IVAS FL enc / FL dec, FER 8%, 1 TC, 48 kbps, DTX on	180	3.33	1.16	0.17
c28	IVAS FX enc / FX dec, FER 8%, 1 TC, 64 kbps, DTX on	180	3.18	1.14	0.17
c29	IVAS FL enc / FX dec, FER 8%, 1 TC, 80 kbps, DTX on	180	3.37	1.10	0.16
c30	IVAS FX enc / FL dec, FER 8%, 2 TC, 13.2 kbps, DTX on	180	2.72	1.14	0.17
c31	IVAS FL enc / FL dec, FER 8%, 2 TC, 16.4 kbps, DTX on	180	2.98	1.16	0.17
c32	IVAS FX enc / FX dec, FER 8%, 2 TC, 24.4 kbps, DTX on	180	2.72	1.18	0.17
c33	IVAS FL enc / FX dec, FER 8%, 2 TC, 32 kbps, DTX on	180	3.06	1.12	0.17
c34	IVAS FX enc / FL dec, FER 8%, 2 TC, 48 kbps, DTX on	180	3.24	1.11	0.16
c35	IVAS FL enc / FL dec, FER 8%, 2 TC, 64 kbps, DTX on	180	3.42	1.14	0.17
c36	IVAS FX enc / FX dec, FER 8%, 2 TC, 80 kbps, DTX on	180	3.54	0.96	0.14



(a)



(b)

Figure 9.5-9: results per condition for experiment P800-14

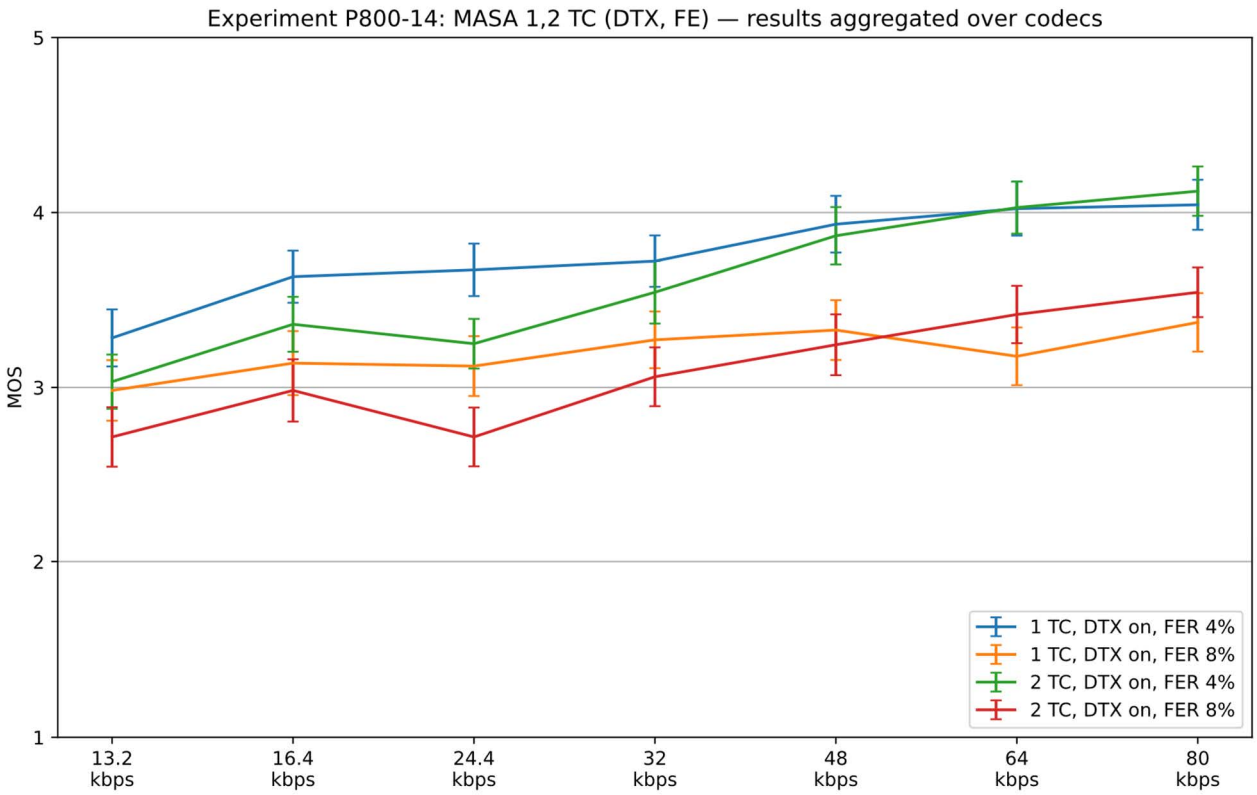


Figure 9.5-10: results aggregated over FX/FL for experiment P800-14

Table 9.5-14: statistical significance analysis for experiment P800-14

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
No terms of references mandated for experiment .									

9.5.9 Characterization Experiment BS1534-13 (MASA 1TC, Generic Audio, 16.4 – 48 kbps, Headphone Presentation)

Characterization Experiment BS1534-13 evaluates the performance of the IVAS fixed-point implementation for MASA 1TC for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-13 was conducted by Qualcomm Incorporated.

The aggregated results per condition are listed in Table 9.5-15 and are shown in Figure 9.5-11. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-16. Scores with result WT are highlighted in red.

Table 9.5-15: results per condition for experiment BS1534-13

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	98.84	4.19	0.64
c02	LP 3.5 kHz	168	31.63	11.22	1.71
c03	Mono EVS, 16.4 kbps	168	56.34	14.76	2.25
c04	Mono EVS, 32 kbps	168	58.44	14.43	2.20
c05	IVAS, 16.4 kbps	168	72.13	18.82	2.87
c06	IVAS, 24.4 kbps	168	80.51	15.57	2.37
c07	IVAS, 32 kbps	168	83.29	14.49	2.21
c08	IVAS, 48 kbps	168	87.05	16.39	2.50

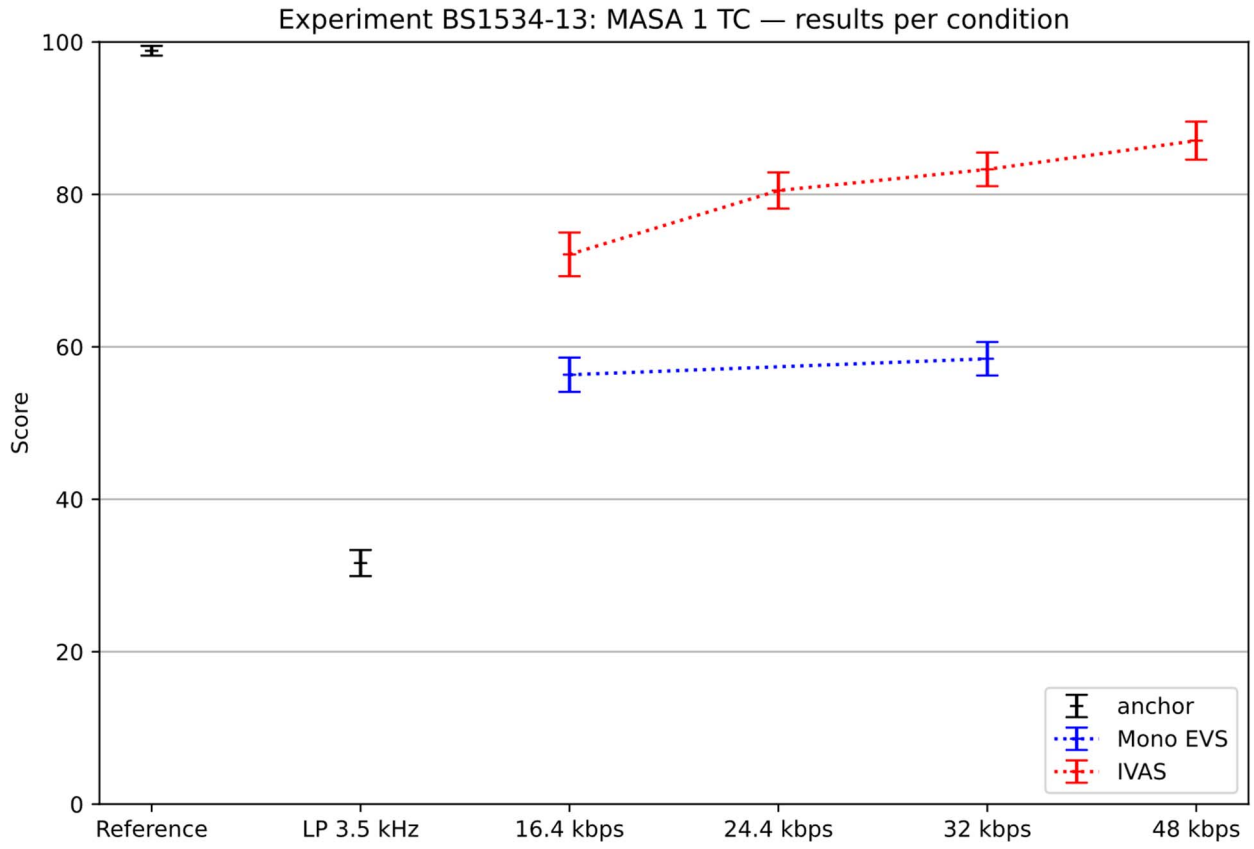


Figure 9.5-11: results per condition for experiment BS1534-13

Table 9.5-16: statistical significance analysis for experiment BS1534-13

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	72.13	18.82	c03	56.34	14.76	15.79	10.297	<0.001	BT
c07	83.29	14.49	c04	58.44	14.43	24.85	17.690	<0.001	BT

9.5.10 Characterization Experiment BS1534-14 (MASA 1TC, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-14 evaluates the performance of the IVAS fixed-point implementation for MASA 1TC for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-14 was conducted by Dolby Laboratories, Inc.

The aggregated results per condition are listed in Table 9.5-17 and are shown in Figure 9.5-12. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-18. Scores with result WT are highlighted in red.

Table 9.5-17: results per condition for experiment BS1534-14

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.79	1.25	0.19
c02	LP 3.5 kHz	168	32.71	6.44	0.98
c03	Mono EVS, 64 kbps	168	65.31	8.78	1.34
c04	Mono EVS, 128 kbps	168	65.82	8.69	1.32
c05	IVAS, 64 kbps	168	75.21	8.68	1.32
c06	IVAS, 96 kbps	168	77.26	8.80	1.34
c07	IVAS, 128 kbps	168	78.49	8.82	1.34
c08	IVAS, 256 kbps	168	79.63	8.96	1.36

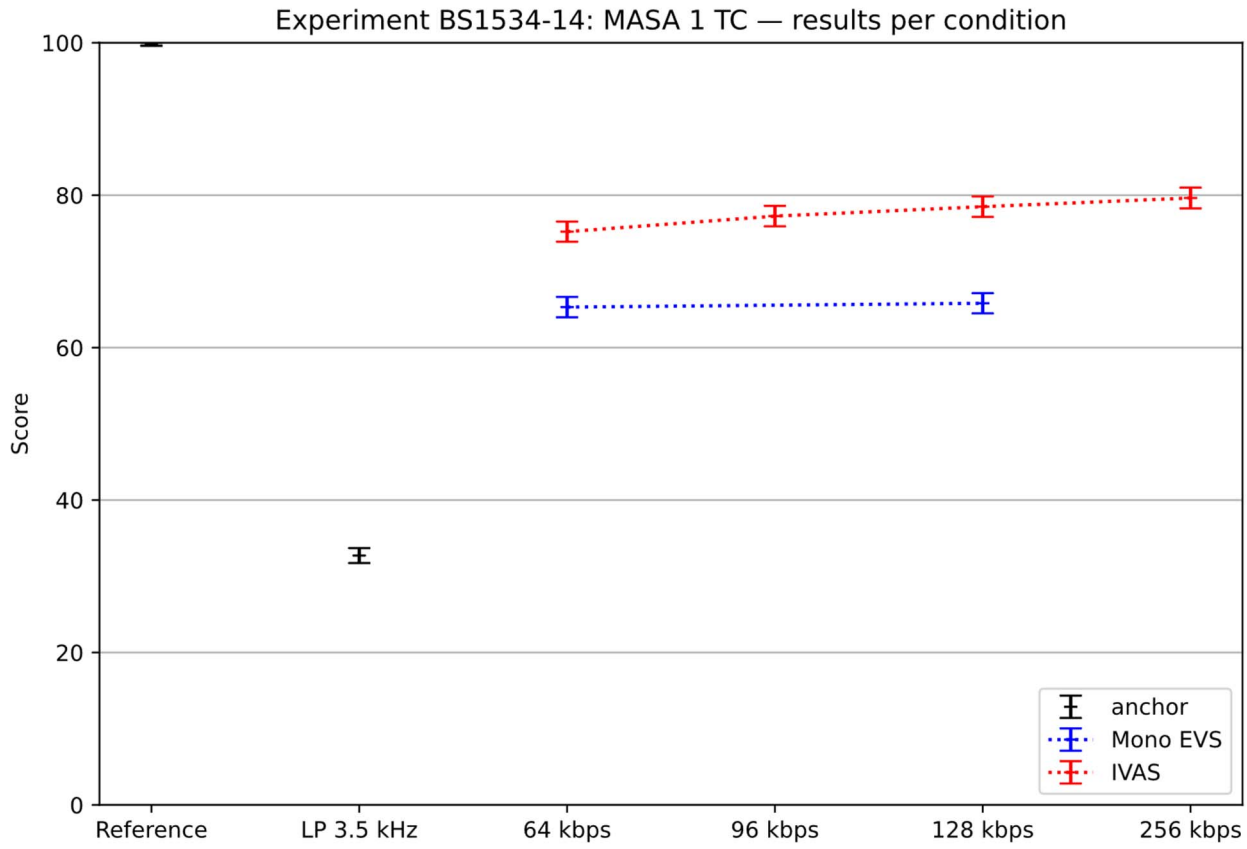


Figure 9.5-12: results per condition for experiment BS1534-14

Table 9.5-18: statistical significance analysis for experiment BS1534-14

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	75.21	8.68	c03	65.31	8.78	9.90	15.721	<0.001	BT
c07	78.49	8.82	c04	65.82	8.69	12.67	20.767	<0.001	BT

9.5.11 Characterization Experiment BS1534-15 (MASA 2TC, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-15 evaluates the performance of the IVAS fixed-point implementation for MASA 2TC for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-15 was conducted by Nokia.

The aggregated results per condition are listed in Table 9.5-19 and are shown in Figure 9.5-13. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.5-20. Scores with result WT are highlighted in red.

Table 9.5-19: results per condition for experiment BS1534-15

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	97.34	5.71	0.87
c02	LP 3.5 kHz	168	12.12	10.53	1.60
c03	Mono EVS, 64 kbps	168	35.27	13.52	2.06
c04	Mono EVS, 128 kbps	168	37.52	15.59	2.37
c05	IVAS, 64 kbps	168	66.74	17.32	2.64
c06	IVAS, 96 kbps	168	74.52	14.56	2.22
c07	IVAS, 128 kbps	168	83.95	11.72	1.79
c08	IVAS, 256 kbps	168	89.06	10.46	1.59

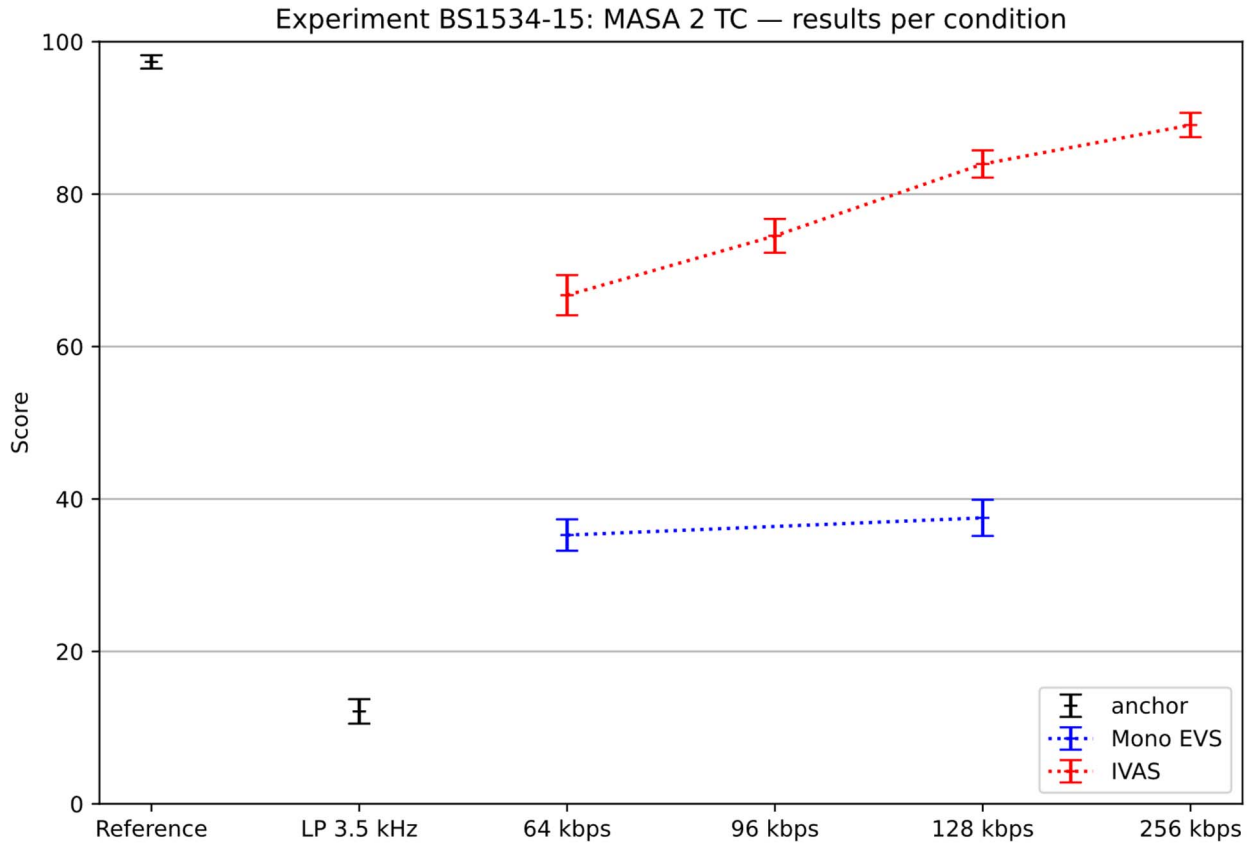


Figure 9.5-13: results per condition for experiment BS1534-15

Table 9.5-20: statistical significance analysis for experiment BS1534-15

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	66.74	17.32	c03	35.27	13.52	31.47	25.802	<0.001	BT
c07	83.95	11.72	c04	37.52	15.59	46.43	33.814	<0.001	BT

9.5.12 Conclusions on IVAS Performance for Metadata-assisted Spatial Audio (MASA)

Taking the results of the previous clauses in 9.5 into account, the following conclusions can be drawn for IVAS Metadata-assisted Spatial Audio (MASA) modes:

During selection phase, IVAS MASA 2TC coding was extensively compared against EVS multi-mono. For all tested content types (clean speech, noisy speech, generic audio) and conditions incl. DTX and FER, IVAS MASA has fulfilled the selection criterion: Better than EVS multi-mono at the same bitrate or not worse than EVS multi-mono at the next higher bitrate. For clean speech, IVAS MASA outperforms EVS multi-mono up to 64 kbps. For noisy speech, IVAS MASA outperforms EVS multi-mono up to 48 kbps. For bitrates above 96 kbps, the performance of IVAS MASA is equivalent to EVS multi-mono and reaches transparency.

During characterization phase, IVAS MASA 1TC/2TC was evaluated against EVS mono in a series of BS.1534 tests. Across all tested bitrates, IVAS was significantly outperforming EVS mono at the same bitrate. In addition, the characterization phase testing provides additional information on the performance of IVAS MASA 1TC, that was compared to IVAS MASA 2TC under impaired channel conditions.

Additionally, the performance of the IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for MASA. Out of 36 tested operating points, 32 operating points were statistically equivalent.

Overall, the IVAS MASA coding for maintains a “good” MOS rating (>= 4.0) from bitrates starting from 16.4 kbps and above (MASA 1TC). A level 1 IVAS encoder alreads reaches an “excellent” MOS rating of approx. 4.5 for this configuration.

9.6 Multi-Channel (MC)

9.6.1 Overview

In Selection phase, four experiments have been conducted to evaluate the performance of the IVAS codec with multi-channel content. All experiments were conducted as BS.1534 tests. While experiments BS1534-2a and BS1534-2b were conducted using 5.1 loudspeaker presentation, experiments BS1534-3a and BS1534-3b were conducted using 7.1+4 loudspeaker presentation.

- Selection Experiment BS1534-2a: Generic audio, multi-channel 5.1, 64 and 96 kbps, loudspeaker 5.1 presentation
- Selection Experiment BS1534-2b: Generic audio, multi-channel 5.1, 128 and 160 kbps, loudspeaker 5.1 presentation
- Selection Experiment BS1534-3a: Generic audio, multi-channel 7.1.4, 128 and 160 kbps, loudspeaker 7.1+4 presentation
- Selection Experiment BS1534-3b: Generic audio, multi-channel 7.1.4, 384 and 512 kbps, loudspeaker 7.1+4 presentation

In Characterization phase, seven additional experiments have been conducted. While the experiments P800-6, P800-7, P800-8 were conducted as P.800 DCR tests, the experiments BS1534-6, BS1534-7, BS1534-8 and BS1534-9 were conducted as BS.1534 tests.

- Characterization Experiment P800-6: Speech + background and mixed/music, multi-channel 5.1 and 7.1, under clean channel conditions, headphone presentation

- Characterization Experiment P800-7: Speech + background and mixed/music, multi-channel 5.1+4 and 7.1+4, under clean channel conditions, headphone presentation
- Characterization Experiment P800-8: Speech + background and mixed/music, mixed CICP layouts, under impaired channel conditions, headphone presentation
- Characterization Experiment BS1534-6: Generic audio, multi-channel 5.1, 16.4 – 48 kbps, loudspeaker 5.1 presentation
- Characterization Experiment BS1534-7: Generic audio, multi-channel 5.1 and 7.1, 64 – 256 kbps, headphone presentation
- Characterization Experiment BS1534-8: Generic audio, multi-channel 5.1+2 and 5.1+4, 64 – 256 kbps, headphone presentation
- Characterization Experiment BS1534-9: Generic audio, multi-channel 7.1+4, 64 – 256 kbps, loudspeaker 7.1+4 presentation

9.6.2 Selection Experiment BS1534-2a (MC 5.1, Generic Audio, 64 and 96 kbps, 5.1 Loudspeaker Presentation)

Selection Experiment BS1534-2a evaluates IVAS for multi-channel 5.1, generic audio at 64 and 96 kbps using loudspeaker 5.1 presentation. See Annex C.123 for details.

The averaged results per condition for experiment BS1534-2a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

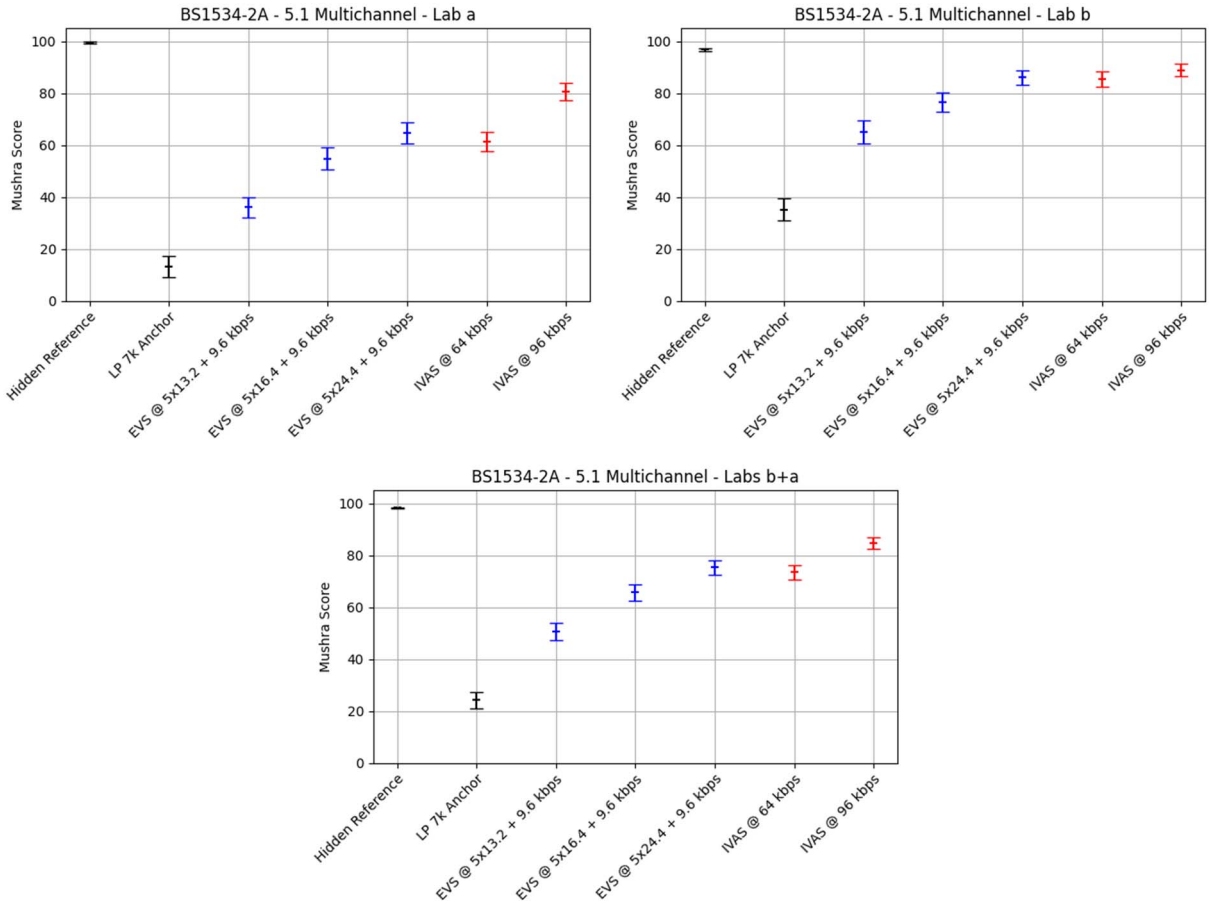


Figure 9.6-2: BS1534-2a (MC 5.1, generic audio, 64 and 96 kbps, 5.1 loudspeaker presentation) MUSHRA plots for labs a and b, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-2a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.6-1: Statistical overview on the results of BS1534-2a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	64	NWT	61.4	24.2	c04	5x16.4	54.8	28.3	2.29	BT	EXCEED
		2	64	BT	61.4	24.2	c03	5x13.2	36.1	25.5	9.35	BT	PASS
	c07	1	96	NWT	80.6	21.9	c05	5x24.4	64.7	26.7	5.95	BT	EXCEED
		2	96	BT	80.6	21.9	c04	5x16.4	54.8	28.3	9.32	BT	PASS
b	c06	1	64	NWT	85.6	19	c04	5x16.4	76.6	24.8	3.71	BT	EXCEED
		2	64	BT	85.6	19	c03	5x13.2	65.2	28.9	7.63	BT	PASS
	c07	1	96	NWT	88.9	15.5	c05	5x24.4	86.1	18.6	1.51	NWT	PASS
		2	96	BT	88.9	15.5	c04	5x16.4	76.6	24.8	5.45	BT	PASS
b+a	c06	1	64	NWT	73.5	24.9	c04	5x16.4	65.7	28.7	3.75	BT	EXCEED
		2	64	BT	73.5	24.9	c03	5x13.2	50.6	30.9	10.57	BT	PASS
	c07	1	96	NWT	84.7	19.4	c05	5x24.4	75.4	25.3	5.36	BT	EXCEED
		2	96	BT	84.7	19.4	c04	5x16.4	65.7	28.7	10.05	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.6-2: Summary of the results of BS1534-2a

Lab	Cond.	Bitrate	ToR	Status
a	c06	64	NWT c04 OR BT c03	EXCEED
	c07	96	NWT c05 OR BT c04	EXCEED
b	c06	64	NWT c04 OR BT c03	EXCEED
	c07	96	NWT c05 OR BT c04	PASS
b+a	c06	64	NWT c04 OR BT c03	EXCEED
	c07	96	NWT c05 OR BT c04	EXCEED

9.6.3 Selection Experiment BS1534-2b (MC 5.1, Generic Audio, 128 and 160 kbps, 5.1 Loudspeaker Presentation)

Selection Experiment BS1534-2b evaluates IVAS for multi-channel 5.1, generic audio at 128 and 160 kbps using loudspeaker 5.1 presentation. See Annex C.13 for details.

The averaged results per condition for experiment BS1534-2b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

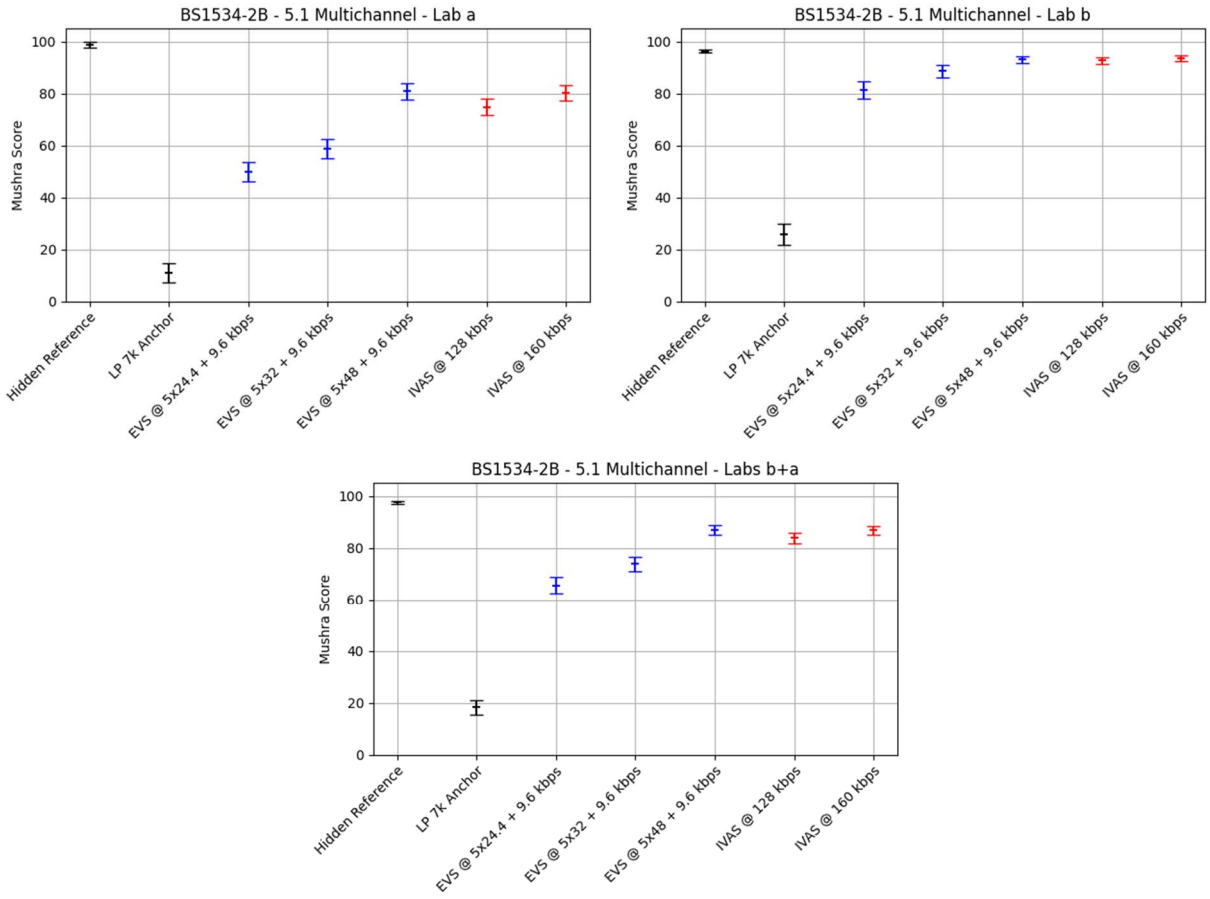


Figure 9.6-3: BS1534-2b (MC 5.1, generic audio, 128 and 160 kbps, 5.1 loudspeaker presentation) MUSHRA plots for labs b and d, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-2b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.6-3: Statistical overview on the results of BS1534-2b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	128	NWT	74.9	21.4	c04	5x32	58.9	24.6	6.36	BT	EXCEED
		2	128	BT	74.9	21.4	c03	5x24.4	49.9	24.1	10.07	BT	PASS
	c07	1	160	NWT	80.2	18.8	c05	5x48	81	20.2	-0.34	NWT	PASS
		2	160	BT	80.2	18.8	c04	5x32	58.9	24.6	8.93	BT	PASS
b	c06	1	128	NWT	92.8	8.2	c04	5x32	88.8	15.7	2.96	BT	EXCEED
		2	128	BT	92.8	8.2	c03	5x24.4	81.4	22.3	6.23	BT	PASS
	c07	1	160	NWT	93.6	7.4	c05	5x48	93.1	8.6	0.56	NWT	PASS
		2	160	BT	93.6	7.4	c04	5x32	88.8	15.7	3.59	BT	PASS
b+a	c06	1	128	NWT	83.9	18.5	c04	5x32	73.8	25.5	5.84	BT	EXCEED
		2	128	BT	83.9	18.5	c03	5x24.4	65.6	28	9.94	BT	PASS
	c07	1	160	NWT	86.9	15.7	c05	5x48	87	16.7	-0.09	NWT	PASS
		2	160	BT	86.9	15.7	c04	5x32	73.8	25.5	8.01	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.6-4: Summary of the results of BS1534-2b

Lab	Cond.	Bitrate	ToR	Status
a	c06	128	NWT c04 OR BT c03	EXCEED
	c07	160	NWT c05 OR BT c04	PASS
b	c06	128	NWT c04 OR BT c03	EXCEED
	c07	160	NWT c05 OR BT c04	PASS

b+a	c06	128	NWT c04 OR BT c03	EXCEED
	c07	160	NWT c05 OR BT c04	PASS

9.6.4 Selection Experiment BS1534-3a (MC 7.1.4, Generic Audio, 128 and 160 kbps, 7.1+4 Loudspeaker Presentation)

Selection Experiment BS1534-3a evaluates IVAS for multi-channel 7.1.4, generic audio at 128 and kbps using loudspeaker 7.1+4 presentation. See Annex C.15 for details.

The averaged results per condition for experiment BS1534-3a are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

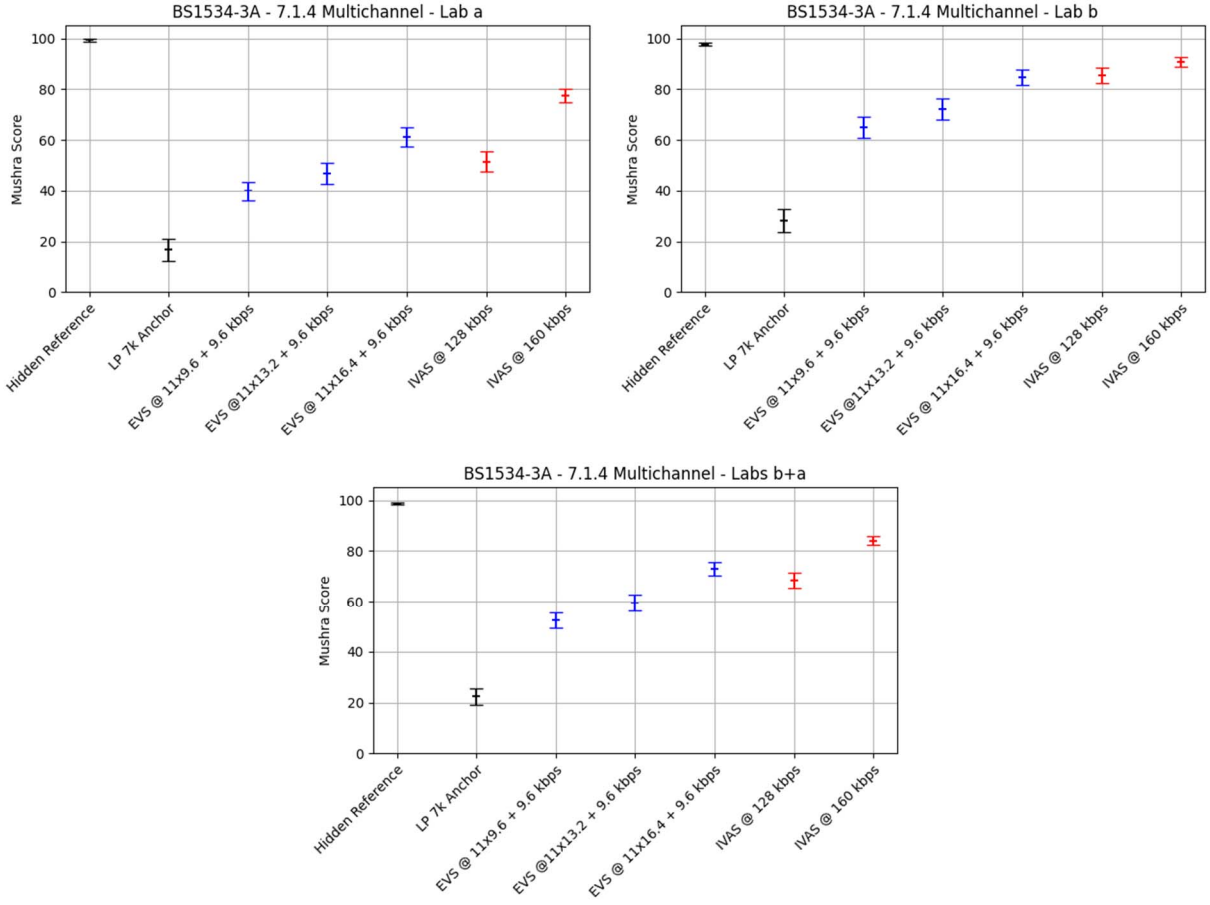


Figure 9.6-4: BS1534-3a (MC 7.1.4, generic audio, 128 and 160 kbps, 7.1+4 loudspeaker presentation) MUSHRA plots for labs a and b, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-3a is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.6-5: Statistical overview on the results of BS1534-3a

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	128	NWT	51.5	26.3	c04	11x13.2	46.8	26.7	1.61	NWT	PASS
		2	128	BT	51.5	26.3	c03	11x9.6	39.8	23.9	4.25	BT	PASS
	c07	1	160	NWT	77.5	17.7	c05	11x16.4	61.1	24.6	7.01	BT	EXCEED
		2	160	BT	77.5	17.7	c04	11x13.2	46.8	26.7	12.42	BT	PASS
b	c06	1	128	NWT	85.2	19.8	c04	11x13.2	72.2	27.3	4.99	BT	EXCEED
		2	128	BT	85.2	19.8	c03	11x9.6	64.9	28.2	7.64	BT	PASS
	c07	1	160	NWT	90.6	12.3	c05	11x16.4	84.7	18.9	3.41	BT	EXCEED
		2	160	BT	90.6	12.3	c04	11x13.2	72.2	27.3	7.95	BT	PASS
b+a	c06	1	128	NWT	68.4	28.8	c04	11x13.2	59.5	29.8	3.91	BT	EXCEED
		2	128	BT	68.4	28.8	c03	11x9.6	52.4	29	7.18	BT	PASS
	c07	1	160	NWT	84.1	16.6	c05	11x16.4	72.9	24.9	6.84	BT	EXCEED
		2	160	BT	84.1	16.6	c04	11x13.2	59.5	29.8	13.18	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.6-6: Summary of the results of BS1534-3a

Lab	Cond.	Bitrate	ToR	Status
a	c06	128	NWT c04 OR BT c03	PASS
	c07	160	NWT c05 OR BT c04	EXCEED
b	c06	128	NWT c04 OR BT c03	EXCEED
	c07	160	NWT c05 OR BT c04	EXCEED
b+a	c06	128	NWT c04 OR BT c03	EXCEED
	c07	160	NWT c05 OR BT c04	EXCEED

9.6.5 Selection Experiment BS1534-3b (MC 7.1.4, Generic Audio, 384 and 512 kbps, 7.1+4 Loudspeaker Presentation)

Selection Experiment BS1534-3b evaluates IVAS for multi-channel 7.1.4, generic audio at, 384 and 512 kbps, loudspeaker 7.1+4 presentation. See Annex C.15 for details.

The averaged results per condition for experiment BS1534-3b are depicted in the following figures. The three figures show the individual results for the two labs and the results for a joint evaluation, respectively. The conditions are shown grouped by Hidden Reference (c01), LP 7k anchor (c02), EVS conditions with increasing bitrate (c03 – c05) and IVAS conditions with increasing bitrate (c06 – c07).

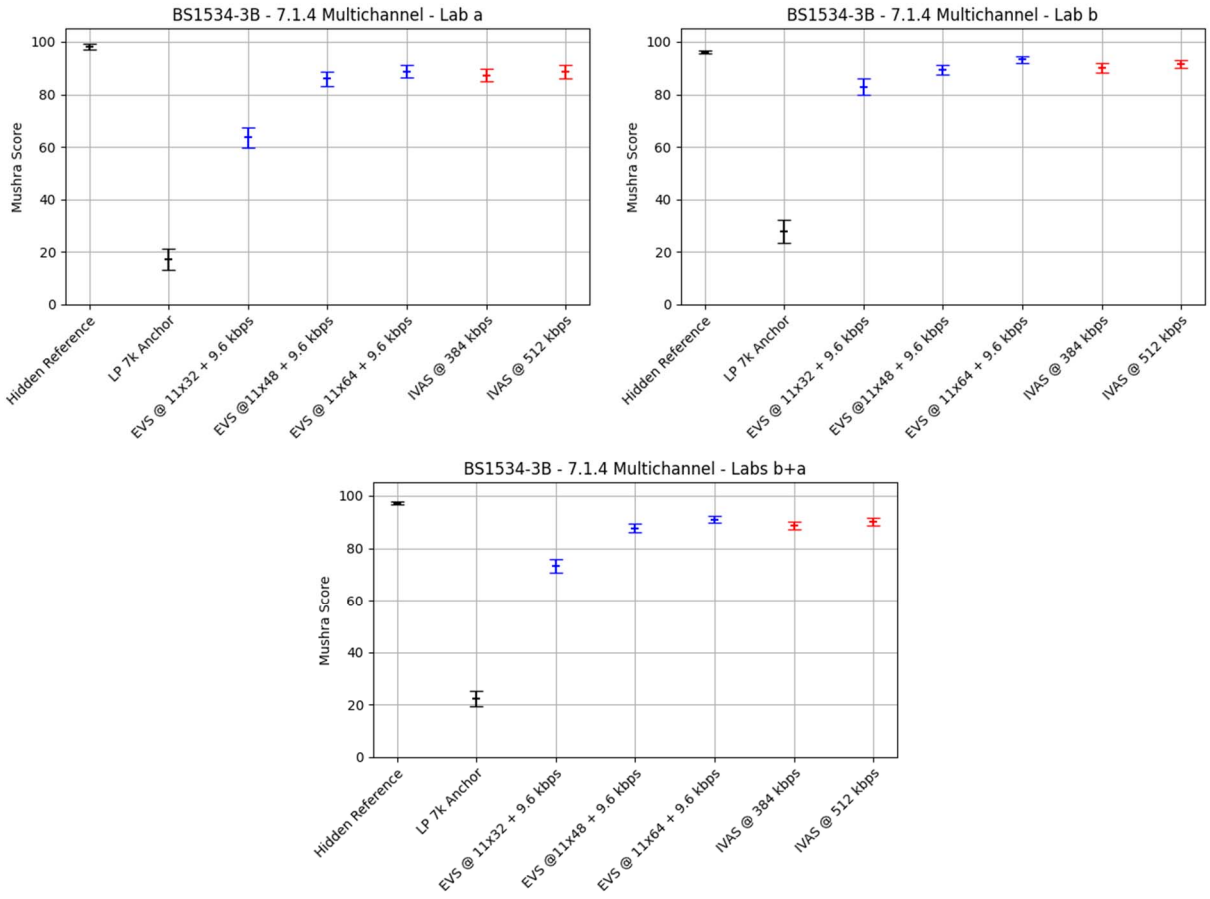


Figure 9.6-5: BS1534-3b (MC 7.1.4, generic audio, 384 and 512 kbps, 7.1+4 loudspeaker presentation) MUSHRA plots for labs a and b, both labs combined

The complete statistical evaluation of the requirement ToR tests for experiment BS1534-3b is given in the following table. The evaluation is done separately for the data from the two listening laboratories and for a combination of the two data sets.

Table 9.6-7: Statistical overview on the results of BS1534-3b

		Type	CuT				EVS Reference				Evaluation		
		Value	Bitrate	Req.	Score	Std.	Cond.	Bitrate	Score	Std.	T-Stat	Result	State
Lab	Cond.	ToR#											
a	c06	1	384	NWT	87.3	16	c04	11x48	85.9	17.6	0.72	NWT	PASS
		2	384	BT	87.3	16	c03	11x32	63.6	25	10.32	BT	PASS
	c07	1	512	NWT	88.6	15.8	c05	11x64	88.7	16	-0.04	NWT	PASS
		2	512	BT	88.6	15.8	c04	11x48	85.9	17.6	1.48	NWT	FAIL
b	c06	1	384	NWT	90.1	11.5	c04	11x48	89.3	12.4	0.6	NWT	PASS
		2	384	BT	90.1	11.5	c03	11x32	82.9	20.1	4.04	BT	PASS
	c07	1	512	NWT	91.5	10.1	c05	11x64	93.3	8	-1.76	WT	FAIL
		2	512	BT	91.5	10.1	c04	11x48	89.3	12.4	1.8	BT	PASS
b+a	c06	1	384	NWT	88.7	14	c04	11x48	87.6	15.3	0.94	NWT	PASS
		2	384	BT	88.7	14	c03	11x32	73.3	24.6	9.99	BT	PASS
	c07	1	512	NWT	90.1	13.3	c05	11x64	91	12.8	-0.9	NWT	PASS
		2	512	BT	90.1	13.3	c04	11x48	87.6	15.3	2.23	BT	PASS

The following table provides a summary of the results. For this summary, the requirements that are defined as a disjunction of two separate checks have been combined into an overall status for this requirement as described before.

Table 9.6-8: Summary of the results of BS1534-3b

Lab	Cond.	Bitrate	ToR	Status
a	c06	384	NWT c04 OR BT c03	PASS
	c07	512	NWT c05 OR BT c04	PASS
b	c06	384	NWT c04 OR BT c03	PASS
	c07	512	NWT c05 OR BT c04	PASS

b+a	c06	384	NWT c04 OR BT c03	PASS
	c07	512	NWT c05 OR BT c04	PASS

9.6.6 Characterization Experiment P800-6 (MC 5.1/7.1, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-6 evaluates the performance of the IVAS fixed-point and floating-point implementation for Multi-channel 5.1/7.1 for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-6 was conducted by Mesaqin.com (language: MAN).

The listening lab has reported that for each panel, up to seven instead of five participants participated. The five subjects with the highest PCC were selected for result submission.

The aggregated results per condition are listed in Table 9.6-9 and are shown in Figure 9.6-6. The results aggregated over FX/FL are shown in Figure 9.6-7. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-10. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.6-9: results per condition for experiment P800-6

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.73	0.52	0.08
c02	MNRU Q = 32 dB	180	3.82	1.11	0.16
c03	MNRU Q = 27 dB	180	3.14	1.25	0.18
c04	MNRU Q = 22 dB	180	2.22	1.18	0.17
c05	MNRU Q = 17 dB	180	1.54	0.83	0.12
c06	ESDRU α = 0.8	180	4.32	0.83	0.12
c07	ESDRU α = 0.6	180	2.95	1.17	0.17
c08	ESDRU α = 0.4	180	2.08	1.10	0.16
c09	ESDRU α = 0.2	180	1.63	0.86	0.13
c10	IVAS FL enc / FX dec, 24.4 kbps, DTX off	180	3.86	1.01	0.15
c11	IVAS FX enc / FL dec, 32 kbps, DTX off	180	4.09	0.95	0.14
c12	IVAS FL enc / FX dec, 48 kbps, DTX off	180	4.34	0.83	0.12
c13	IVAS FX enc / FL dec, 64 kbps, DTX off	180	4.46	0.77	0.11
c14	IVAS FL enc / FX dec, 80 kbps, DTX off	180	4.51	0.70	0.10
c15	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.43	0.78	0.11
c16	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.47	0.73	0.11
c17	IVAS FX enc / FL dec, 160 kbps, DTX off	180	4.48	0.70	0.10
c18	IVAS FL enc / FX dec, 256 kbps, DTX off	180	4.53	0.69	0.10
c19	IVAS FL, 24.4 kbps, DTX off	180	3.95	1.00	0.15
c20	IVAS FL, 32 kbps, DTX off	180	4.02	0.95	0.14
c21	IVAS FL, 48 kbps, DTX off	180	4.43	0.83	0.12
c22	IVAS FL, 64 kbps, DTX off	180	4.42	0.75	0.11
c23	IVAS FL, 80 kbps, DTX off	180	4.49	0.72	0.11
c24	IVAS FL, 96 kbps, DTX off	180	4.53	0.68	0.10
c25	IVAS FL, 128 kbps, DTX off	180	4.52	0.72	0.11
c26	IVAS FL, 160 kbps, DTX off	180	4.54	0.69	0.10
c27	IVAS FL, 256 kbps, DTX off	180	4.53	0.71	0.10
c28	IVAS FX, 24.4 kbps, DTX off	180	3.87	1.02	0.15
c29	IVAS FX, 32 kbps, DTX off	180	4.00	1.05	0.15
c30	IVAS FX, 48 kbps, DTX off	180	4.42	0.77	0.11
c31	IVAS FX, 64 kbps, DTX off	180	4.44	0.76	0.11
c32	IVAS FX, 80 kbps, DTX off	180	4.46	0.69	0.10
c33	IVAS FX, 96 kbps, DTX off	180	4.50	0.68	0.10
c34	IVAS FX, 128 kbps, DTX off	180	4.43	0.80	0.12
c35	IVAS FX, 160 kbps, DTX off	180	4.48	0.67	0.10
c36	IVAS FX, 256 kbps, DTX off	180	4.49	0.71	0.10

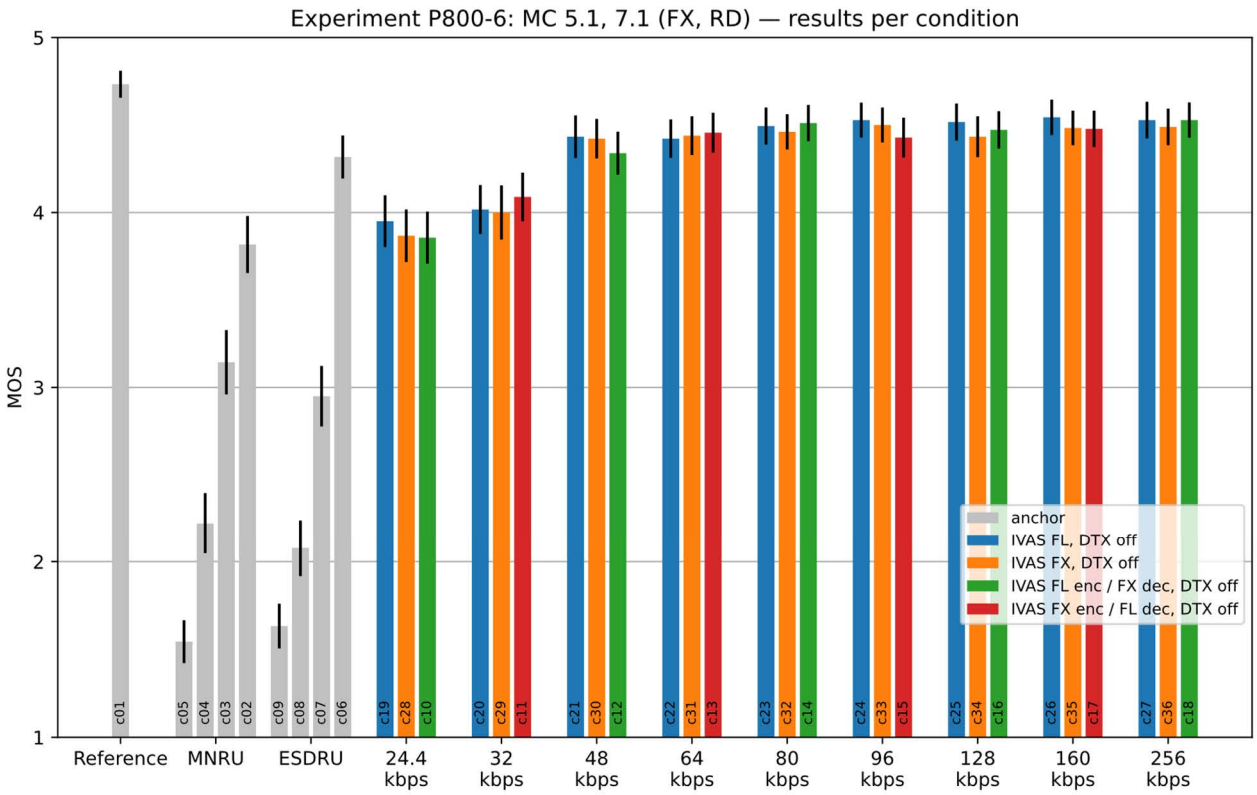


Figure 9.6-6: results per condition for experiment P800-6

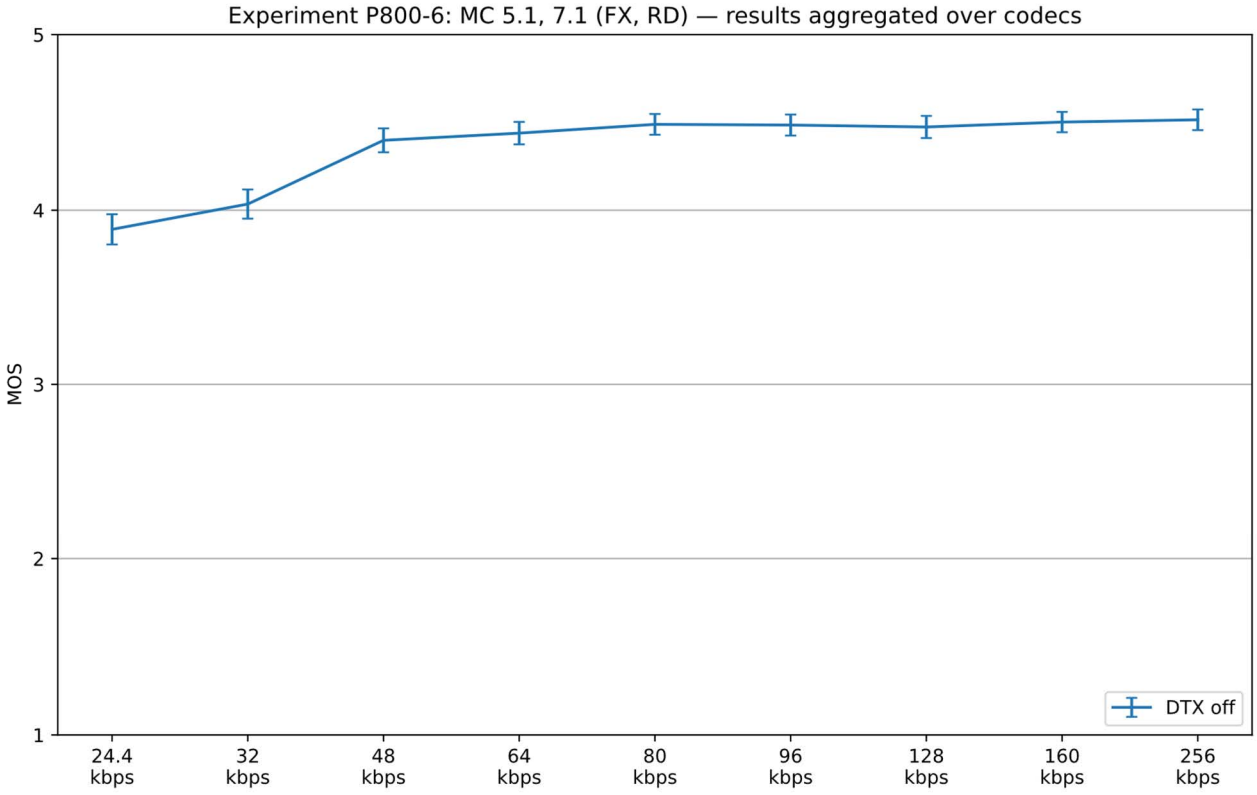


Figure 9.6-7: results aggregated over FX/FL for experiment P800-6

Table 9.6-10: statistical significance analysis for experiment P800-6

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.86	1.01	c19	3.95	1.00	-0.09	<0.001	0.811	NWT
c11	4.09	0.95	c20	4.02	0.95	0.07	0.719	0.236	NWT
c12	4.34	0.83	c21	4.43	0.83	-0.09	<0.001	0.858	NWT
c13	4.46	0.77	c22	4.42	0.75	0.03	0.425	0.336	NWT
c14	4.51	0.70	c23	4.49	0.72	0.02	0.226	0.411	NWT
c15	4.43	0.78	c24	4.53	0.68	-0.10	<0.001	0.903	NWT
c16	4.47	0.73	c25	4.52	0.72	-0.04	<0.001	0.722	NWT
c17	4.48	0.70	c26	4.54	0.69	-0.07	<0.001	0.815	NWT
c18	4.53	0.69	c27	4.53	0.71	0.00	<0.001	0.500	NWT
c28	3.87	1.02	c19	3.95	1.00	-0.08	<0.001	0.781	NWT
c29	4.00	1.05	c20	4.02	0.95	-0.02	<0.001	0.564	NWT
c30	4.42	0.77	c21	4.43	0.83	-0.01	<0.001	0.552	NWT
c31	4.44	0.76	c22	4.42	0.75	0.02	0.215	0.415	NWT
c32	4.46	0.69	c23	4.49	0.72	-0.03	<0.001	0.671	NWT
c33	4.50	0.68	c24	4.53	0.68	-0.03	<0.001	0.652	NWT
c34	4.43	0.80	c25	4.52	0.72	-0.08	<0.001	0.852	NWT
c35	4.48	0.67	c26	4.54	0.69	-0.06	<0.001	0.802	NWT
c36	4.49	0.71	c27	4.53	0.71	-0.04	<0.001	0.698	NWT

9.6.7 Characterization Experiment P800-7 (MC 5.1+4/7.1+4, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-7 evaluates the performance of the IVAS fixed-point and floating-point implementation for Multi-channel 5.1+4/7.1+4 for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-7 was conducted by FORCE Technology (language: DAN).

The aggregated results per condition are listed in Table 9.6-11 and are shown in Figure 9.6-8. The results aggregated over FX/FL are shown in Figure 9.6-9. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-12. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.6-11: results per condition for experiment P800-7

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.60	0.56	0.08
c02	MNRU Q = 32 dB	180	3.61	1.07	0.16
c03	MNRU Q = 27 dB	180	2.66	1.13	0.17
c04	MNRU Q = 22 dB	180	1.90	0.95	0.14
c05	MNRU Q = 17 dB	180	1.35	0.63	0.09
c06	ESDRU α = 0.8	180	4.45	0.69	0.10
c07	ESDRU α = 0.6	180	3.43	1.02	0.15
c08	ESDRU α = 0.4	180	2.80	1.04	0.15
c09	ESDRU α = 0.2	180	2.32	1.02	0.15
c10	IVAS FL enc / FX dec, 32 kbps, DTX off	180	3.67	0.97	0.14
c11	IVAS FX enc / FL dec, 48 kbps, DTX off	180	3.92	0.93	0.14
c12	IVAS FL enc / FX dec, 64 kbps, DTX off	180	4.16	0.87	0.13
c13	IVAS FX enc / FL dec, 96 kbps, DTX off	180	4.33	0.80	0.12
c14	IVAS FL enc / FX dec, 128 kbps, DTX off	180	4.49	0.69	0.10
c15	IVAS FX enc / FL dec, 160 kbps, DTX off	180	4.48	0.66	0.10
c16	IVAS FL enc / FX dec, 192 kbps, DTX off	180	4.47	0.74	0.11
c17	IVAS FX enc / FL dec, 256 kbps, DTX off	180	4.54	0.77	0.11
c18	IVAS FL enc / FX dec, 384 kbps, DTX off	180	4.63	0.59	0.09
c19	IVAS FL, 32 kbps, DTX off	180	3.69	1.04	0.15

c20	IVAS FL, 48 kbps, DTX off	180	4.00	0.88	0.13
c21	IVAS FL, 64 kbps, DTX off	180	4.11	0.82	0.12
c22	IVAS FL, 96 kbps, DTX off	180	4.33	0.78	0.11
c23	IVAS FL, 128 kbps, DTX off	180	4.47	0.77	0.11
c24	IVAS FL, 160 kbps, DTX off	180	4.55	0.70	0.10
c25	IVAS FL, 192 kbps, DTX off	180	4.42	0.69	0.10
c26	IVAS FL, 256 kbps, DTX off	180	4.52	0.72	0.11
c27	IVAS FL, 384 kbps, DTX off	180	4.66	0.63	0.09
c28	IVAS FX, 32 kbps, DTX off	180	3.51	0.97	0.14
c29	IVAS FX, 48 kbps, DTX off	180	4.09	0.85	0.13
c30	IVAS FX, 64 kbps, DTX off	180	4.02	0.86	0.13
c31	IVAS FX, 96 kbps, DTX off	180	4.33	0.75	0.11
c32	IVAS FX, 128 kbps, DTX off	180	4.42	0.77	0.11
c33	IVAS FX, 160 kbps, DTX off	180	4.46	0.71	0.10
c34	IVAS FX, 192 kbps, DTX off	180	4.43	0.76	0.11
c35	IVAS FX, 256 kbps, DTX off	180	4.54	0.69	0.10
c36	IVAS FX, 384 kbps, DTX off	180	4.58	0.68	0.10

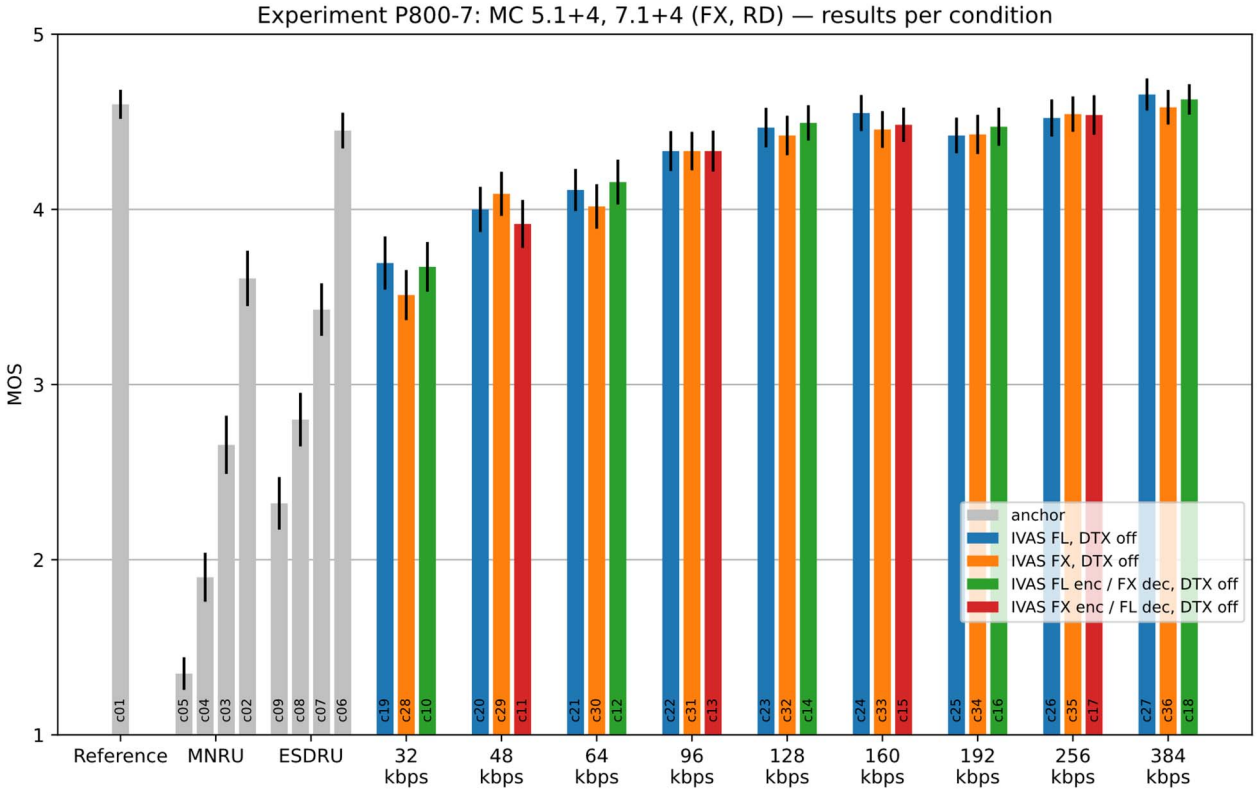


Figure 9.6-8: results per condition for experiment P800-7

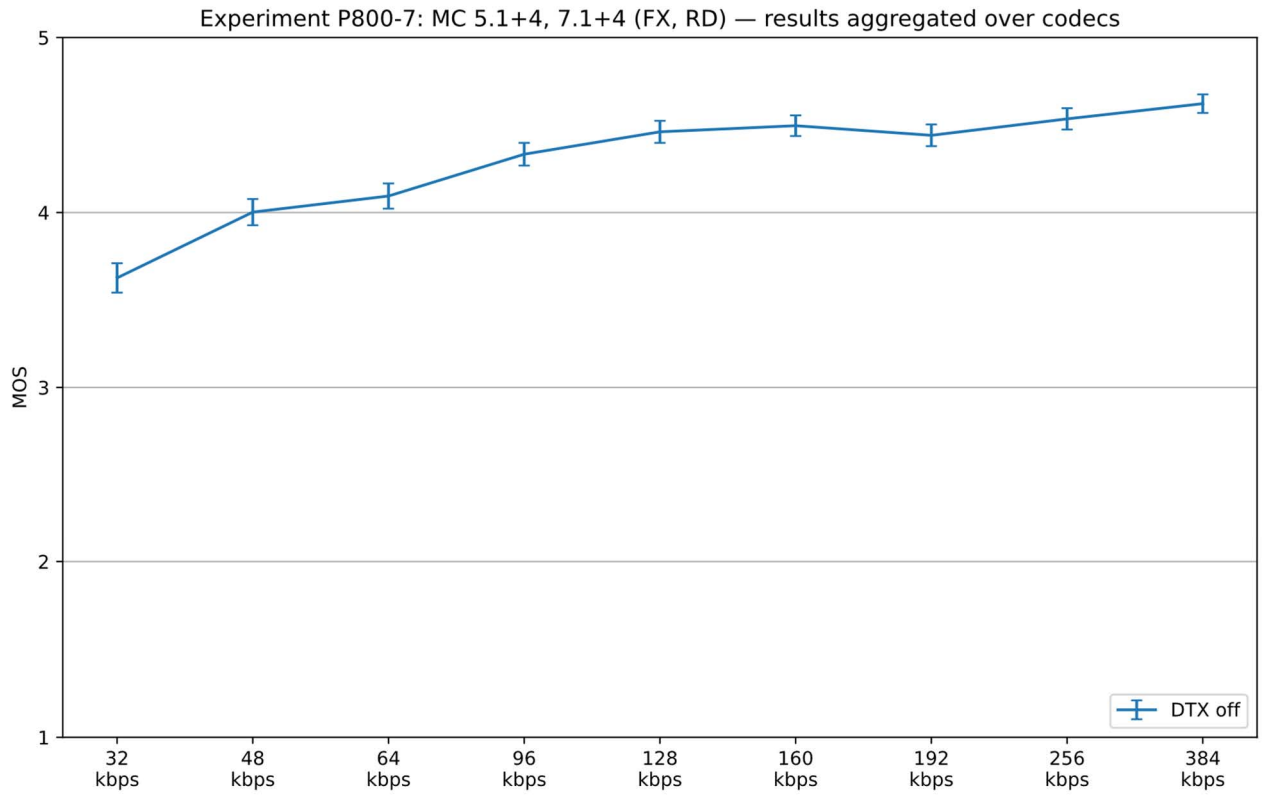


Figure 9.6-9: results aggregated over FX/FL for experiment P800-7

Table 9.6-12: statistical significance analysis for experiment P800-7

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.67	0.97	c19	3.69	1.04	-0.02	<0.001	0.582	NWT
c11	3.92	0.93	c20	4.00	0.88	-0.08	<0.001	0.807	NWT
c12	4.16	0.87	c21	4.11	0.82	0.04	0.505	0.307	NWT
c13	4.33	0.80	c22	4.33	0.78	0.00	<0.001	0.500	NWT
c14	4.49	0.69	c23	4.47	0.77	0.03	0.352	0.363	NWT
c15	4.48	0.66	c24	4.55	0.70	-0.07	<0.001	0.823	NWT
c16	4.47	0.74	c25	4.42	0.69	0.05	0.660	0.255	NWT
c17	4.54	0.77	c26	4.52	0.72	0.02	0.217	0.414	NWT
c18	4.63	0.59	c27	4.66	0.63	-0.03	<0.001	0.669	NWT
c28	3.51	0.97	c19	3.69	1.04	-0.18	<0.001	0.958	WT
c29	4.09	0.85	c20	4.00	0.88	0.09	0.975	0.165	NWT
c30	4.02	0.86	c21	4.11	0.82	-0.09	<0.001	0.855	NWT
c31	4.33	0.75	c22	4.33	0.78	0.00	<0.001	0.500	NWT
c32	4.42	0.77	c23	4.47	0.77	-0.04	<0.001	0.711	NWT
c33	4.46	0.71	c24	4.55	0.70	-0.09	<0.001	0.896	NWT
c34	4.43	0.76	c25	4.42	0.69	0.01	0.078	0.469	NWT
c35	4.54	0.69	c26	4.52	0.72	0.02	0.296	0.384	NWT
c36	4.58	0.68	c27	4.66	0.63	-0.07	<0.001	0.855	NWT

9.6.8 Characterization Experiment P800-8 (Mixed CICP, Speech and Mixed/Music, Headphone Presentation)

Characterization Experiment P800-7 evaluates the performance of the IVAS fixed-point and floating-point implementation for multi-channel mixed CICP for speech and mixed/music, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-8 was conducted by Mesaqin.com (language: CZ).

The listening lab has reported that for each panel, up to seven instead of five participants participated. The five subjects with the highest PCC were selected for result submission.

The aggregated results per condition are listed in Table 9.6-13 and are shown in Figure 9.6-10. The results aggregated over FX/FL are shown in Figure 9.6-11. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-14. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.6-13: results per condition for experiment P800-8

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.81	0.41	0.06
c02	MNRU Q = 32 dB	180	4.02	0.96	0.14
c03	MNRU Q = 27 dB	180	3.34	1.01	0.15
c04	MNRU Q = 22 dB	180	2.37	1.08	0.16
c05	MNRU Q = 17 dB	180	1.60	0.80	0.12
c06	ESDRU α = 0.8	180	4.59	0.64	0.09
c07	ESDRU α = 0.6	180	3.82	1.05	0.15
c08	ESDRU α = 0.4	180	3.17	1.26	0.19
c09	ESDRU α = 0.2	180	2.70	1.27	0.19
c10	IVAS FL enc / FX dec, FER 5%, 16.4 kbps	180	3.21	1.05	0.15
c11	IVAS FX enc / FL dec, FER 5%, 24.4 kbps	180	3.52	0.94	0.14
c12	IVAS FL enc / FX dec, FER 5%, 48 kbps	180	3.97	0.82	0.12
c13	IVAS FX enc / FL dec, FER 5%, 64 kbps	180	4.08	0.76	0.11
c14	IVAS FL enc / FX dec, FER 5%, 80 kbps	180	4.12	0.85	0.13
c15	IVAS FX enc / FL dec, FER 5%, 128 kbps	180	4.13	0.90	0.13
c16	IVAS FL enc / FX dec, FER 5%, 160 kbps	180	4.23	0.78	0.11
c17	IVAS FX enc / FL dec, FER 5%, 256 kbps	180	4.29	0.75	0.11
c18	IVAS FL enc / FX dec, FER 5%, 384 kbps	180	4.28	0.73	0.11
c19	IVAS FL, FER 5%, 16.4 kbps	180	3.23	1.12	0.17
c20	IVAS FL, FER 5%, 24.4 kbps	180	3.61	1.02	0.15
c21	IVAS FL, FER 5%, 48 kbps	180	3.88	0.95	0.14
c22	IVAS FL, FER 5%, 64 kbps	180	4.06	0.88	0.13
c23	IVAS FL, FER 5%, 80 kbps	180	4.13	0.85	0.12
c24	IVAS FL, FER 5%, 128 kbps	180	4.21	0.74	0.11
c25	IVAS FL, FER 5%, 160 kbps	180	4.17	0.75	0.11
c26	IVAS FL, FER 5%, 256 kbps	180	4.27	0.74	0.11
c27	IVAS FL, FER 5%, 384 kbps	180	4.30	0.75	0.11
c28	IVAS FX, FER 5%, 16.4 kbps	180	3.32	1.04	0.15
c29	IVAS FX, FER 5%, 24.4 kbps	180	3.65	0.99	0.15
c30	IVAS FX, FER 5%, 48 kbps	180	3.87	0.96	0.14
c31	IVAS FX, FER 5%, 64 kbps	180	3.99	0.88	0.13
c32	IVAS FX, FER 5%, 80 kbps	180	4.02	0.90	0.13
c33	IVAS FX, FER 5%, 128 kbps	180	4.24	0.76	0.11
c34	IVAS FX, FER 5%, 160 kbps	180	4.20	0.79	0.12
c35	IVAS FX, FER 5%, 256 kbps	180	4.27	0.71	0.10
c36	IVAS FX, FER 5%, 384 kbps	180	4.33	0.76	0.11

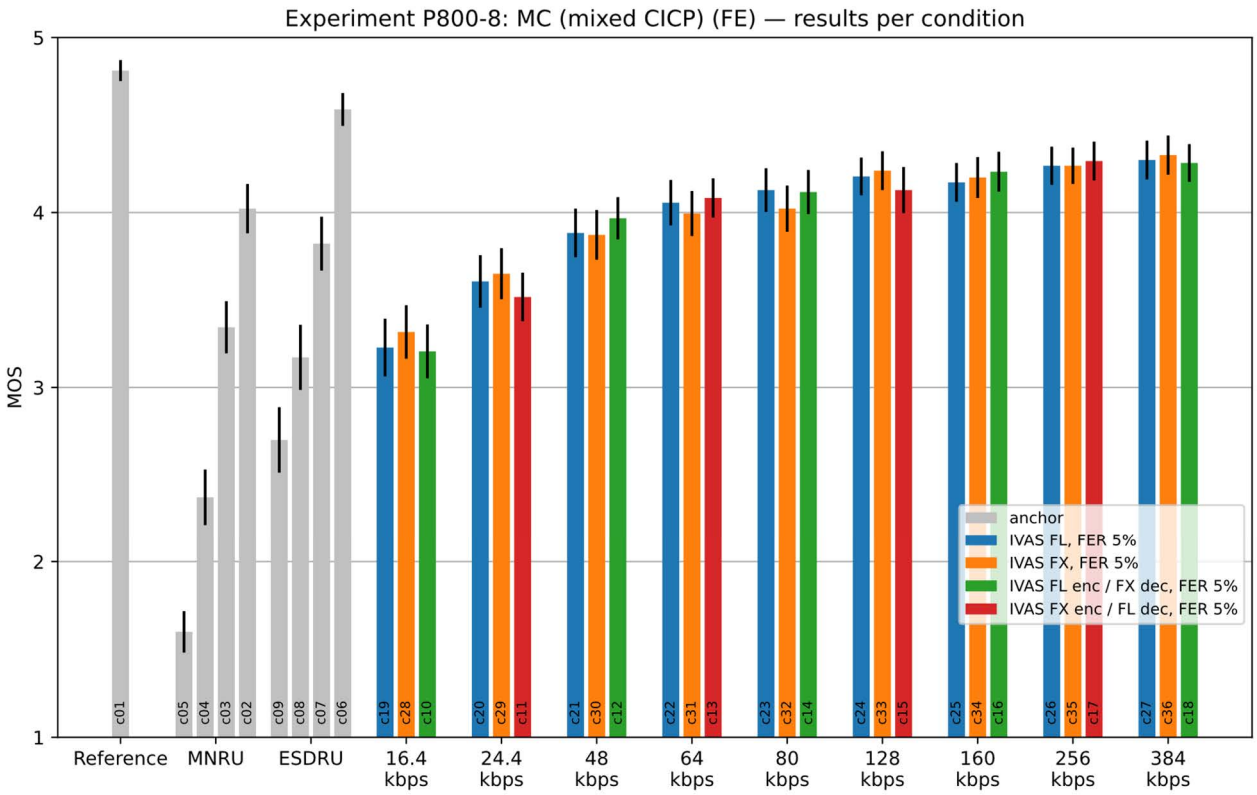


Figure 9.6-10: results per condition for experiment P800-8

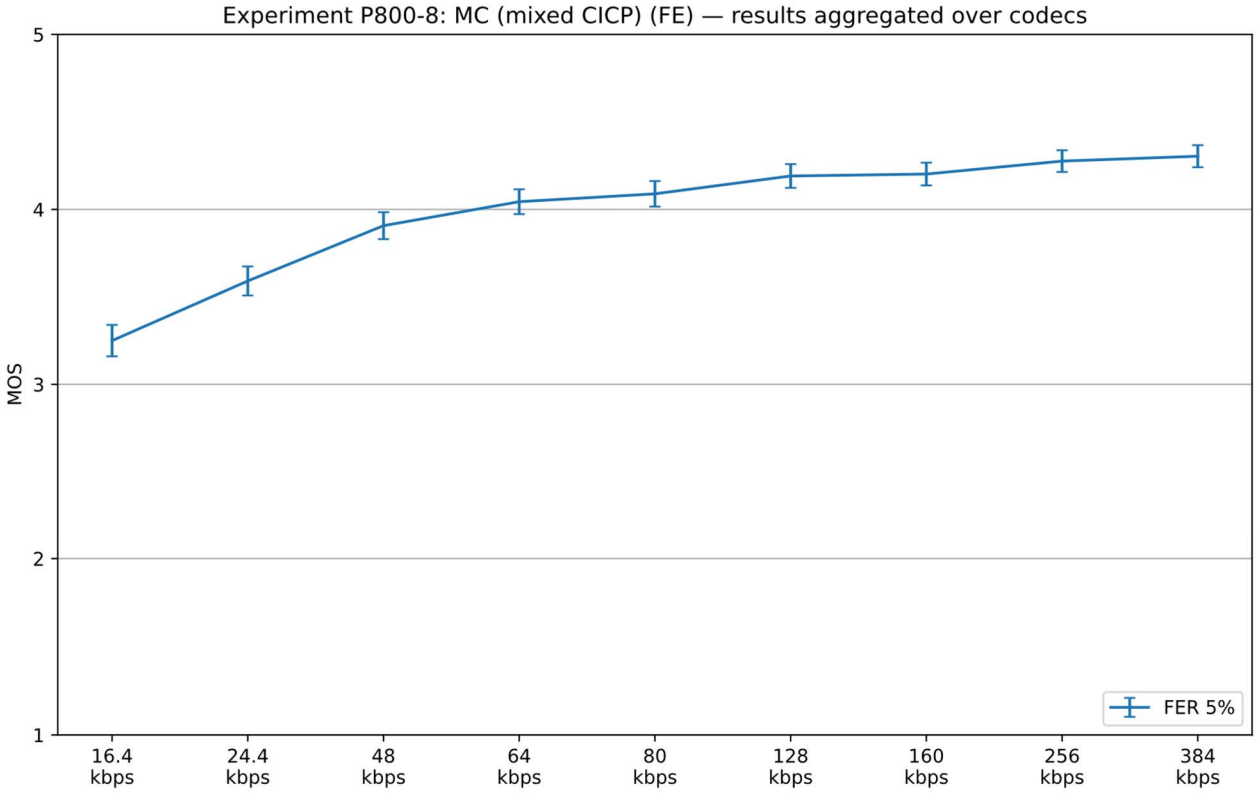


Figure 9.6-11: results aggregated over FX/FL for experiment P800-8

Table 9.6-14: statistical significance analysis for experiment P800-8

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.21	1.05	c19	3.23	1.12	-0.02	<0.001	0.576	NWT
c11	3.52	0.94	c20	3.61	1.02	-0.09	<0.001	0.805	NWT
c12	3.97	0.82	c21	3.88	0.95	0.08	0.897	0.185	NWT
c13	4.08	0.76	c22	4.06	0.88	0.03	0.311	0.378	NWT
c14	4.12	0.85	c23	4.13	0.85	-0.01	<0.001	0.549	NWT
c15	4.13	0.90	c24	4.21	0.74	-0.08	<0.001	0.816	NWT
c16	4.23	0.78	c25	4.17	0.75	0.06	0.756	0.225	NWT
c17	4.29	0.75	c26	4.27	0.74	0.03	0.342	0.366	NWT
c18	4.28	0.73	c27	4.30	0.75	-0.02	<0.001	0.586	NWT
c28	3.32	1.04	c19	3.23	1.12	0.09	0.781	0.218	NWT
c29	3.65	0.99	c20	3.61	1.02	0.04	0.414	0.340	NWT
c30	3.87	0.96	c21	3.88	0.95	-0.01	<0.001	0.543	NWT
c31	3.99	0.88	c22	4.06	0.88	-0.06	<0.001	0.748	NWT
c32	4.02	0.90	c23	4.13	0.85	-0.11	<0.001	0.874	NWT
c33	4.24	0.76	c24	4.21	0.74	0.03	0.418	0.338	NWT
c34	4.20	0.79	c25	4.17	0.75	0.03	0.343	0.366	NWT
c35	4.27	0.71	c26	4.27	0.74	0.00	<0.001	0.500	NWT
c36	4.33	0.76	c27	4.30	0.75	0.03	0.351	0.363	NWT

9.6.9 Characterization Experiment BS1534-6 (MC 5.1, Generic Audio, 16.4 – 48 kbps, Loudspeaker Presentation)

Characterization Experiment BS1534-6 evaluates the performance of the IVAS fixed-point implementation for multi-channel 5.1 for generic audio content with loudspeaker presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-6 was conducted by Ericsson LM.

The aggregated results per condition are listed in Table 9.6-15 and are shown in Figure 9.6-12. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-16. Scores with result WT are highlighted in red.

Table 9.6-15: results per condition for experiment BS1534-6

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.56	2.54	0.39
c02	LP 3.5 kHz	168	28.34	17.20	2.62
c03	Mono EVS, 16.4 kbps	168	46.65	19.08	2.91
c04	Mono EVS, 32 kbps	168	50.80	17.76	2.71
c05	IVAS, 16.4 kbps	168	52.53	23.89	3.64
c06	IVAS, 24.4 kbps	168	65.76	18.12	2.76
c07	IVAS, 32 kbps	168	67.57	16.85	2.57
c08	IVAS, 48 kbps	168	84.62	15.46	2.35

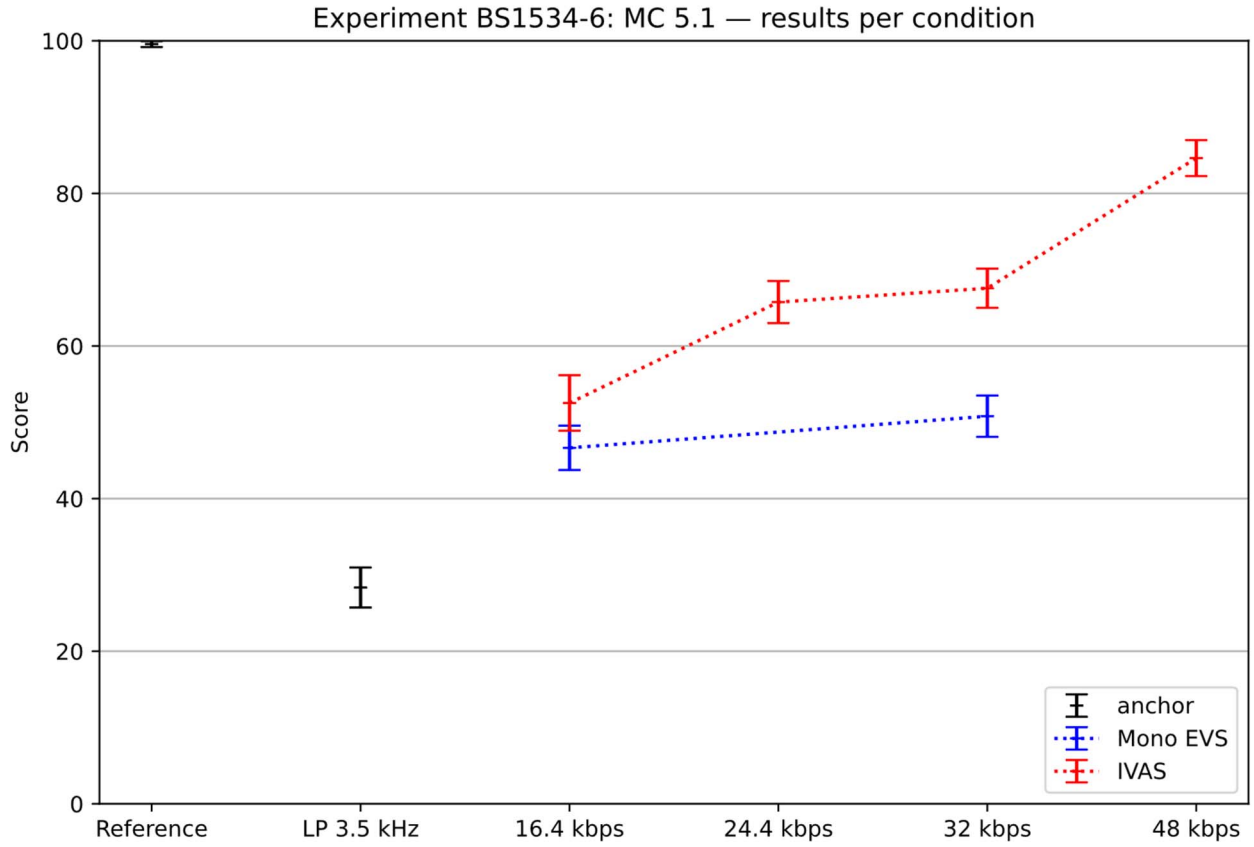


Figure 9.6-12: results per condition for experiment BS1534-6

Table 9.6-16: statistical significance analysis for experiment BS1534-6

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	52.53	23.89	c03	46.65	19.08	5.88	2.988	0.003	BT
c07	67.57	16.85	c04	50.80	17.76	16.77	9.400	<0.001	BT

9.6.10 Characterization Experiment BS1534-7 (MC 5.1/7.1, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-7 evaluates the performance of the IVAS fixed-point implementation for multi-channel 5.1/7.1 for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-7 was conducted by Dolby Laboratories, Inc.

The aggregated results per condition are listed in Table 9.6-17 and are shown in Figure 9.6-13. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-18. Scores with result WT are highlighted in red.

Table 9.6-17: results per condition for experiment BS1534-7

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.62	1.60	0.24
c02	LP 3.5 kHz	168	36.35	9.88	1.50
c03	Mono EVS, 64 kbps	168	59.30	11.26	1.72
c04	Mono EVS, 128 kbps	168	59.41	11.19	1.70
c05	IVAS, 64 kbps	168	71.76	9.49	1.45
c06	IVAS, 96 kbps	168	76.34	8.41	1.28
c07	IVAS, 128 kbps	168	77.06	8.25	1.26
c08	IVAS, 256 kbps	168	84.17	8.80	1.34

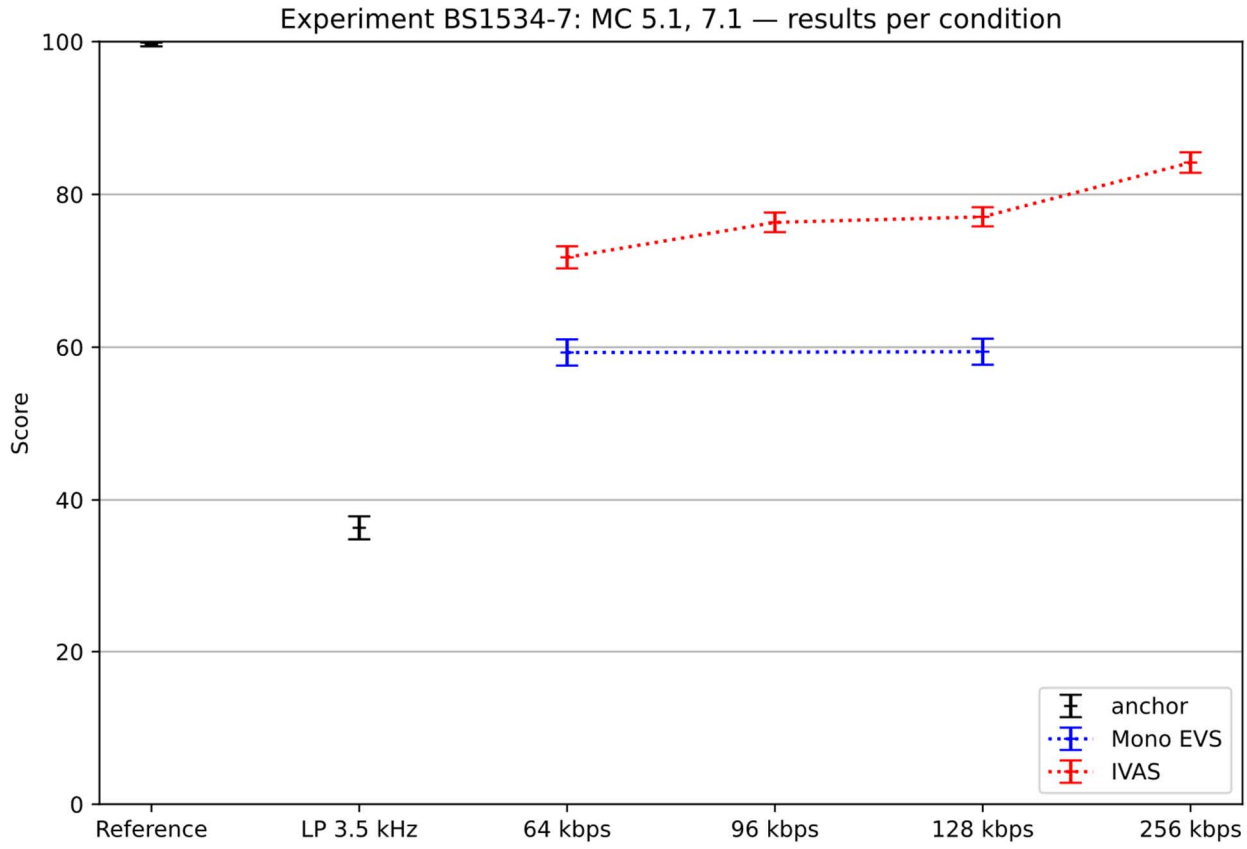


Figure 9.6-13: results per condition for experiment BS1534-7

Table 9.6-18: statistical significance analysis for experiment BS1534-7

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	71.76	9.49	c03	59.30	11.26	12.46	17.378	<0.001	BT
c07	77.06	8.25	c04	59.41	11.19	17.66	22.739	<0.001	BT

9.6.11 Characterization Experiment BS1534-8 (MC 5.1+2/5.1+4, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-8 evaluates the performance of the IVAS fixed-point implementation for multi-channel 5.1+2/5.1+4 for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-8 was conducted by Nokia.

The aggregated results per condition are listed in Table 9.6-19 and are shown in Figure 9.6-14. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-20. Scores with result WT are highlighted in red.

Table 9.6-19: results per condition for experiment BS1534-8

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	96.57	8.35	1.27
c02	LP 3.5 kHz	168	10.40	10.74	1.64
c03	Mono EVS, 64 kbps	168	35.50	16.55	2.52
c04	Mono EVS, 128 kbps	168	36.59	16.57	2.52
c05	IVAS, 64 kbps	168	55.45	19.78	3.01
c06	IVAS, 96 kbps	168	75.67	15.20	2.31
c07	IVAS, 128 kbps	168	78.81	16.41	2.50
c08	IVAS, 256 kbps	168	88.88	12.15	1.85

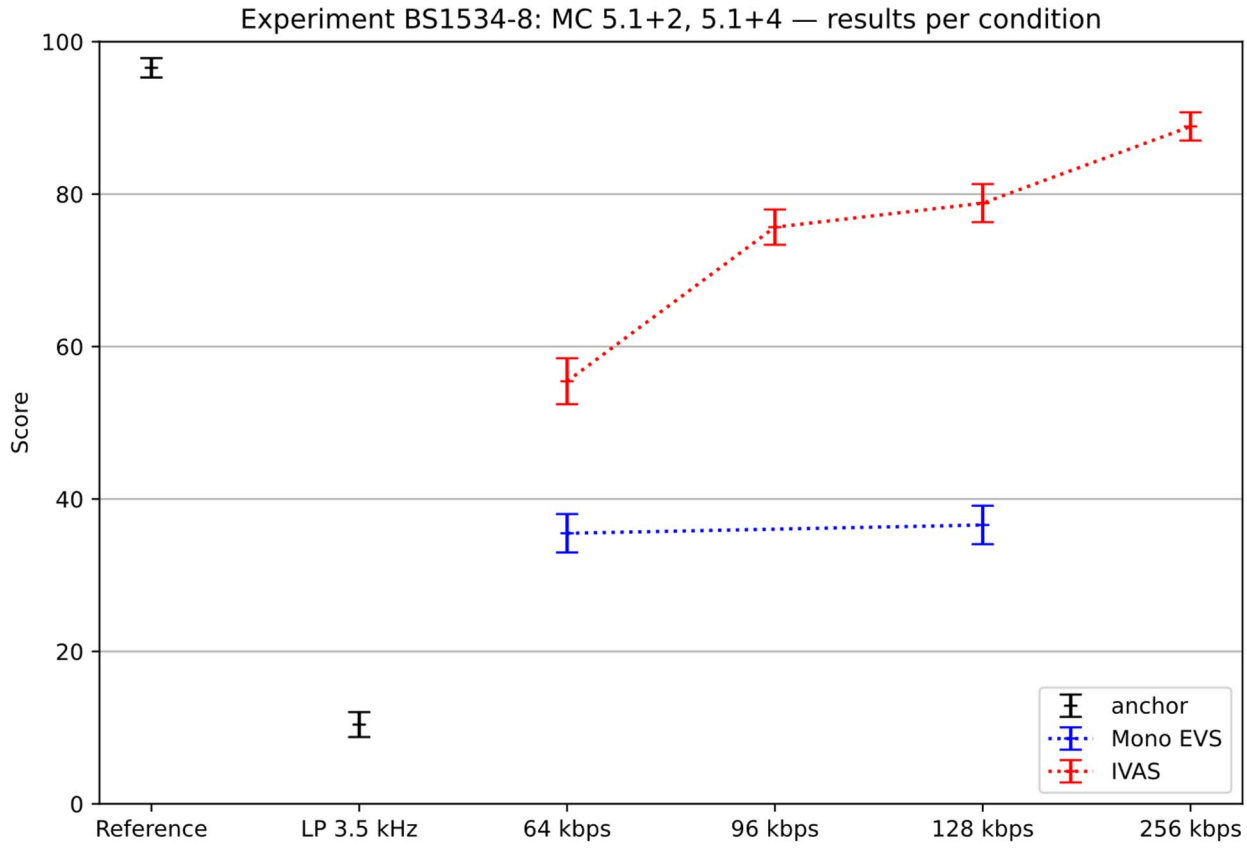


Figure 9.6-14: results per condition for experiment BS1534-8

Table 9.6-20: statistical significance analysis for experiment BS1534-8

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	55.45	19.78	c03	35.50	16.55	19.95	12.595	<0.001	BT
c07	78.81	16.41	c04	36.59	16.57	42.23	28.310	<0.001	BT

9.6.12 Characterization Experiment BS1534-9 (MC 7.1+4, Generic Audio, 64 – 256 kbps, Loudspeaker Presentation)

Characterization Experiment BS1534-9 evaluates the performance of the IVAS fixed-point implementation for multi-channel 7.1+4 for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-9 was conducted by Fraunhofer IIS.

The aggregated results per condition are listed in Table 9.6-21 and are shown in Figure 9.6-15. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.6-22. Scores with result WT are highlighted in red.

Table 9.6-21: results per condition for experiment BS1534-9

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.95	0.49	0.07
c02	LP 3.5 kHz	168	22.51	6.89	1.05
c03	Mono EVS, 64 kbps	168	33.32	10.93	1.67
c04	Mono EVS, 128 kbps	168	33.59	11.18	1.70
c05	IVAS, 64 kbps	168	60.83	17.66	2.69
c06	IVAS, 96 kbps	168	64.09	16.71	2.54
c07	IVAS, 128 kbps	168	75.30	15.21	2.32
c08	IVAS, 256 kbps	168	87.23	12.10	1.84

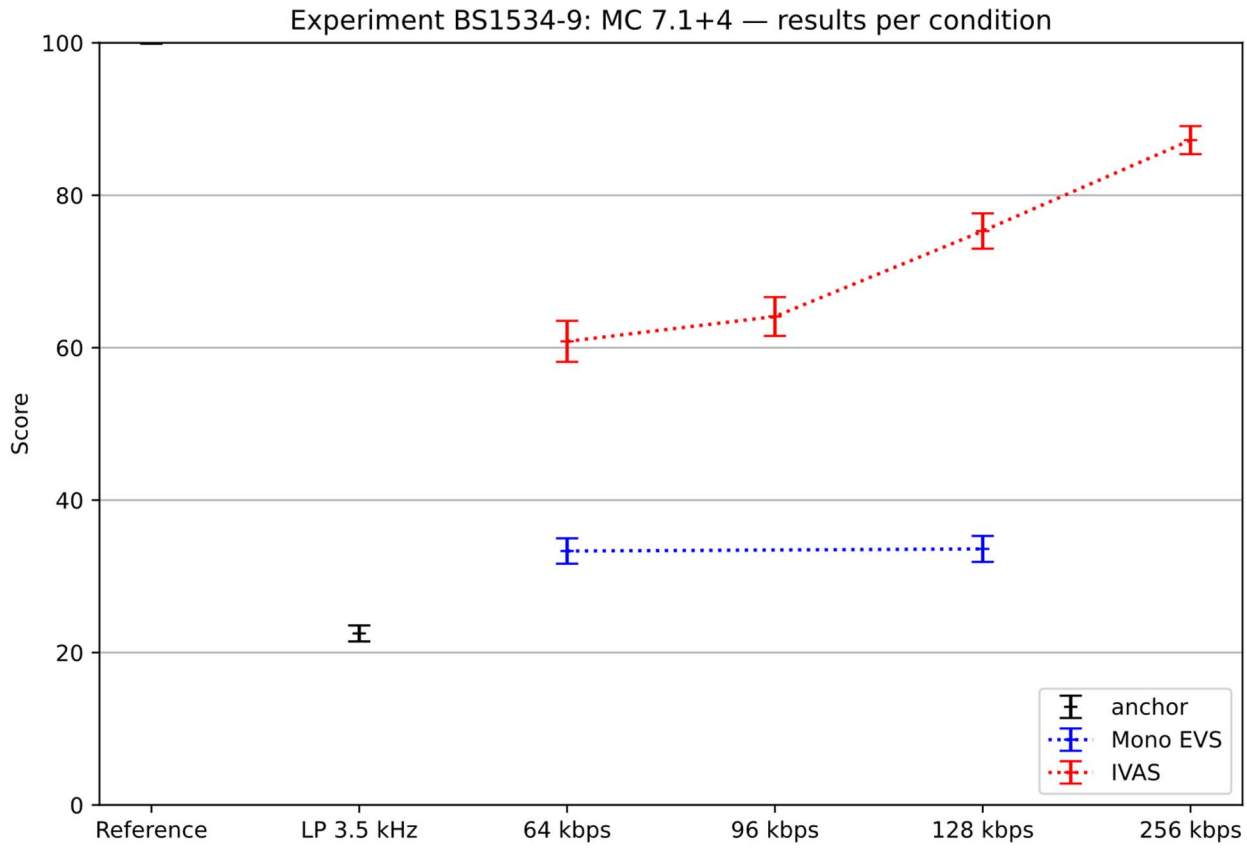


Figure 9.6-15: results per condition for experiment BS1534-9

Table 9.6-22: statistical significance analysis for experiment BS1534-9

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	60.83	17.66	c03	33.32	10.93	27.51	19.608	<0.001	BT
c07	75.30	15.21	c04	33.59	11.18	41.71	30.271	<0.001	BT

9.6.13 Conclusions on IVAS Performance for Multi-channel (MC)

Taking the results of the previous clauses in 9.6 into account, the following conclusions can be drawn for IVAS multi-channel audio modes:

During selection phase, IVAS multi-channel coding was extensively compared against EVS multi-mono. In a series of BS.1534 tests, 5.1 and 7.1+4 multi-channel coding was evaluated against EVS multi-mono. For multi-channel 5.1, IVAS outperforms EVS multi-mono for bitrates of 64, 96 and 128 kbps and is equivalent for 160 kbps. For multi-channel 7.1+4, IVAS outperforms EVS multi-mono for 128 and 160 kbps and is equivalent for 384 and 512, while these bitrates reach transparency.

During characterization phase, IVAS multi-channel was evaluated against EVS mono in a series of BS.1534 tests for various multi-channel layouts (5.1, 7.1, 5.1+2, 5.1+4, 7.1+4). Across all tested bitrates, IVAS was significantly outperforming EVS mono at the same bitrate.

Additionally, the performance of IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for multi-channel. Out of 54 tested operating points, 53 operating points were statistically equivalent.

Overall, the IVAS multi-channel coding for maintains a “good” MOS rating (>= 4.0) from bitrates starting from 32 kbps and above for 5.1/7.1 content. A level 1 IVAS encoder alreads reaches an “excellent” MOS rating of approx. 4.5

for this configuration. For multi-channel configurations including height-channels (5.1+4/7.1+4), IVAS multi-channel coding for maintains a “good” MOS rating (>= 4.0) from bitrates starting from 48 kbps and above.

9.7 Combined Objects and SBA (OSBA)

9.7.1 Overview

Provided as optional additional information for codec selection, two experiments have been conducted to evaluate the performance of the IVAS codec with the combination of combined objects and SBA content. All experiments were conducted as BS.1534 tests using headphone presentation.

- Optional additional information experiment 1: Generic audio, combined objects and SBA, 256 kbps, headphone presentation
- Optional additional information experiment 2: Generic audio, combined objects and SBA, 512 kbps, headphone presentation

In Characterization phase, five additional experiments have been conducted. While the experiments P800-15, P800-16, P800-17 were conducted as P.800 DCR tests, the experiments BS1534-16 and BS1534-17 were conducted as BS.1534 tests.

- Characterization Experiment P800-15: Speech and mixed/music + background, 1-2 objects combined with SBA, under clean channel conditions, headphone presentation
- Characterization Experiment P800-16: Speech and mixed/music + background, 3-4 objects combined with SBA, under clean channel conditions, headphone presentation
- Characterization Experiment P800-17: Speech and mixed/music + background, 1-4 objects combined with SBA, under impaired channel conditions, headphone presentation
- Characterization Experiment BS1534-16: Generic audio, 1-4 objects combined with SBA, 16.4 – 48 kbps, headphone presentation
- Characterization Experiment BS1534-17: Generic audio, 1-4 objects combined with SBA, 64 – 256 kbps, headphone presentation

9.7.2 Optional additional information provided for codec selection

The following quality assessments has been provided as optional additional information for codec selection:

IVAS provides a combined format mode to code ISMs and an SBA scene in a single instance of the codec. At low bitrates, the objects are pre-rendered into the SBA scene, which is then coded and reproduced as usual. At high bitrates, a high-quality mode exists, which allocates extra downmix channels for the objects and codes their metadata separately.

This high-quality mode can be shown to exhibit a strong quality benefit over both the separate coding of the two formats by separate instances of IVAS and pre-rendering. This is demonstrated by the results of a MUSHRA listening test plotted in Figure 9.7-1. The high-quality mode (osba_256kpbs) scores significantly higher than the separate-coding (ism_128kbps_sba_128kbps) and the pre-rendering (prerendered_256kbps).

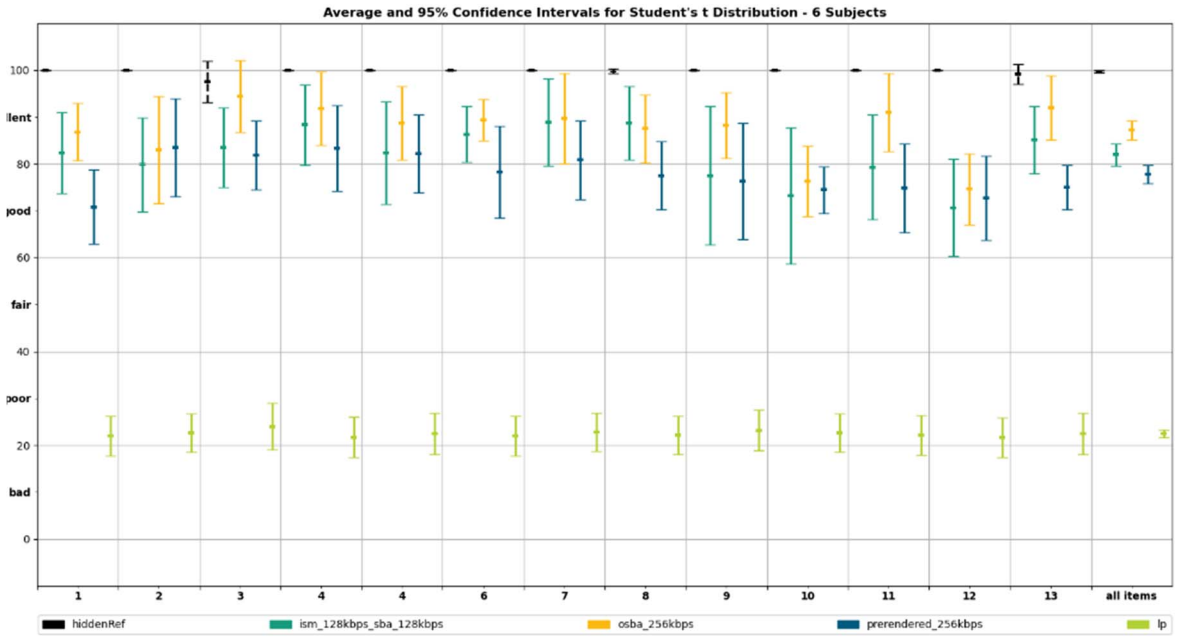


Figure 9.7-1: BS.1534 comparison between combined vs. Separate operation of ISM + SBA at 256 kbps

Further benefit can be observed at 512 kbps by the results of a separate MUSHRA listening test (6 expert listeners), shown in Figure 9.7-2. The high-quality mode (osba_512) scores significantly higher than the separate-coding (base_2x256kbps) and the pre-rendering (prerend_512), bringing the average quality into the excellent range.

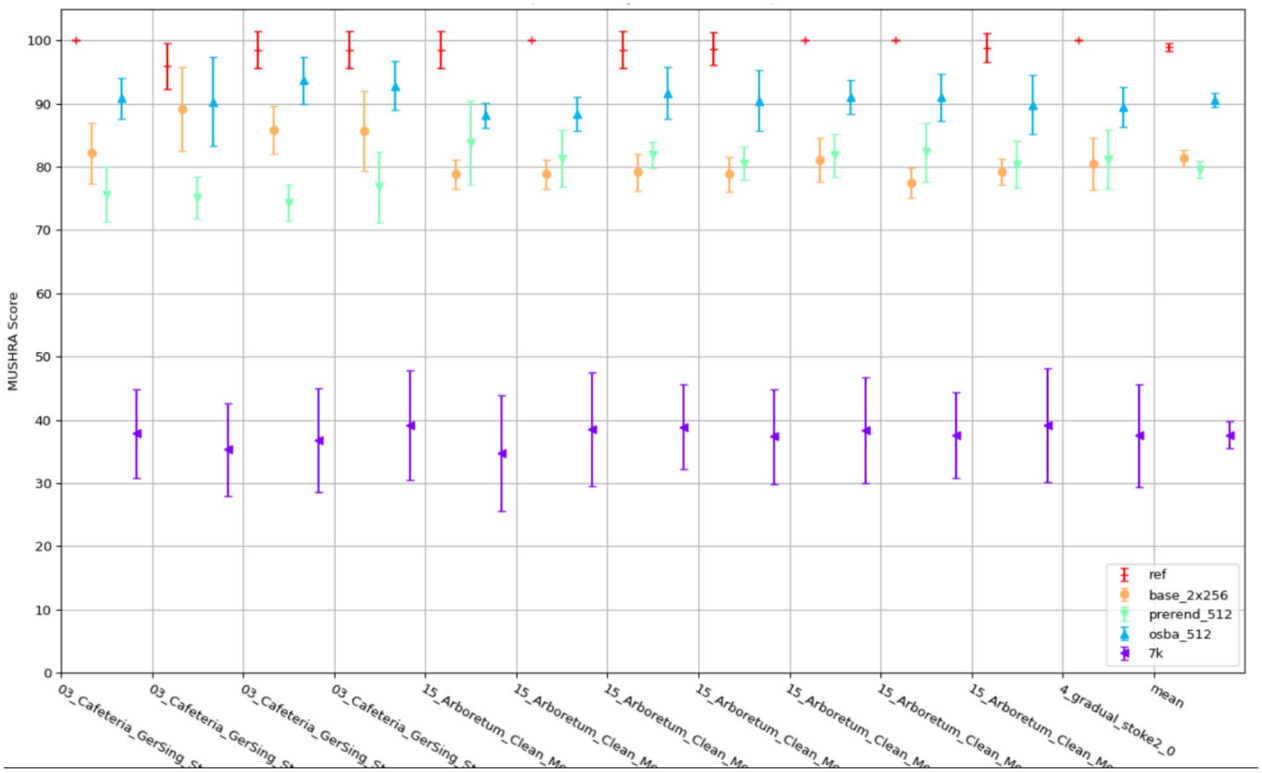


Figure 9.7-2: BS.1534 comparison between combined vs. separate operation of ISM + SBA at 512 kbps

9.7.3 Characterization Experiment P800-15 (OSBA 1-2 Objects, Speech and Mixed/Music + Background, Headphone Presentation)

Characterization Experiment P800-15 evaluates the performance of the IVAS fixed-point and floating-point implementation for OSBA with 1-2 objects for speech and mixed/music + background, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-15 was conducted by Dolby Laboratories, Inc. (language: ENG).

The aggregated results per condition are listed in Table 9.7-1 and are shown in Figure 9.7-3. The results aggregated over FX/FL are shown in Figure 9.7-4. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.7-2. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.7-1: results per condition for experiment P800-15

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.79	0.43	0.06
c02	MNRU Q = 32 dB	180	4.32	0.93	0.14
c03	MNRU Q = 27 dB	180	3.78	1.22	0.18
c04	MNRU Q = 22 dB	180	2.78	1.34	0.20
c05	MNRU Q = 17 dB	180	1.91	1.22	0.18
c06	ESDRU α = 0.8	180	4.48	0.70	0.10
c07	ESDRU α = 0.6	180	3.57	1.15	0.17
c08	ESDRU α = 0.4	180	2.71	1.19	0.17
c09	ESDRU α = 0.2	180	2.09	1.10	0.16
c10	IVAS FL enc / FX dec, 32 kbps, DTX Off	180	3.33	1.22	0.18
c11	IVAS FX enc / FL dec, 48 kbps, DTX Off	180	3.82	1.07	0.16
c12	IVAS FL enc / FX dec, 64 kbps, DTX Off	180	4.03	0.98	0.14
c13	IVAS FX enc / FL dec, 96 kbps, DTX Off	180	4.32	0.73	0.11
c14	IVAS FL enc / FX dec, 128 kbps, DTX Off	180	4.42	0.74	0.11
c15	IVAS FX enc / FL dec, 192 kbps, DTX Off	180	4.44	0.72	0.11
c16	IVAS FL enc / FX dec, 256 kbps, DTX Off	180	4.45	0.77	0.11
c17	IVAS FX enc / FL dec, 384 kbps, DTX Off	180	4.38	0.74	0.11
c18	IVAS FL enc / FX dec, 512 kbps, DTX Off	180	4.35	0.77	0.11
c19	IVAS FL, 32 kbps, DTX Off	180	3.41	1.19	0.17
c20	IVAS FL, 48 kbps, DTX Off	180	3.82	1.04	0.15
c21	IVAS FL, 64 kbps, DTX Off	180	3.99	0.96	0.14
c22	IVAS FL, 96 kbps, DTX Off	180	4.36	0.70	0.10
c23	IVAS FL, 128 kbps, DTX Off	180	4.37	0.78	0.11
c24	IVAS FL, 192 kbps, DTX Off	180	4.39	0.80	0.12
c25	IVAS FL, 256 kbps, DTX Off	180	4.44	0.68	0.10
c26	IVAS FL, 384 kbps, DTX Off	180	4.46	0.72	0.11
c27	IVAS FL, 512 kbps, DTX Off	180	4.44	0.72	0.11
c28	IVAS FX, 32 kbps, DTX Off	180	3.31	1.21	0.18
c29	IVAS FX, 48 kbps, DTX Off	180	3.76	1.09	0.16
c30	IVAS FX, 64 kbps, DTX Off	180	3.99	1.08	0.16
c31	IVAS FX, 96 kbps, DTX Off	180	4.33	0.84	0.12
c32	IVAS FX, 128 kbps, DTX Off	180	4.41	0.72	0.11
c33	IVAS FX, 192 kbps, DTX Off	180	4.39	0.77	0.11
c34	IVAS FX, 256 kbps, DTX Off	180	4.38	0.73	0.11
c35	IVAS FX, 384 kbps, DTX Off	180	4.44	0.65	0.10
c36	IVAS FX, 512 kbps, DTX Off	180	4.49	0.72	0.11

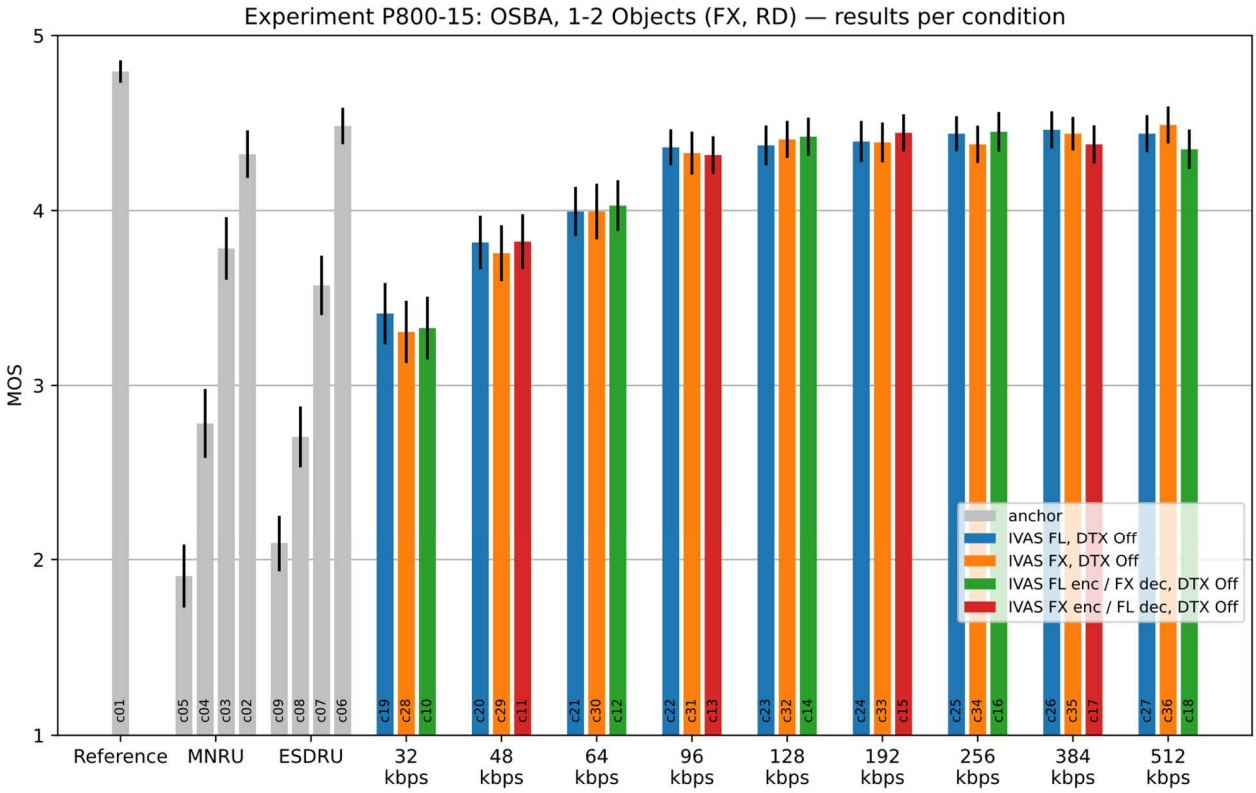


Figure 9.7-3: results per condition for experiment P800-15

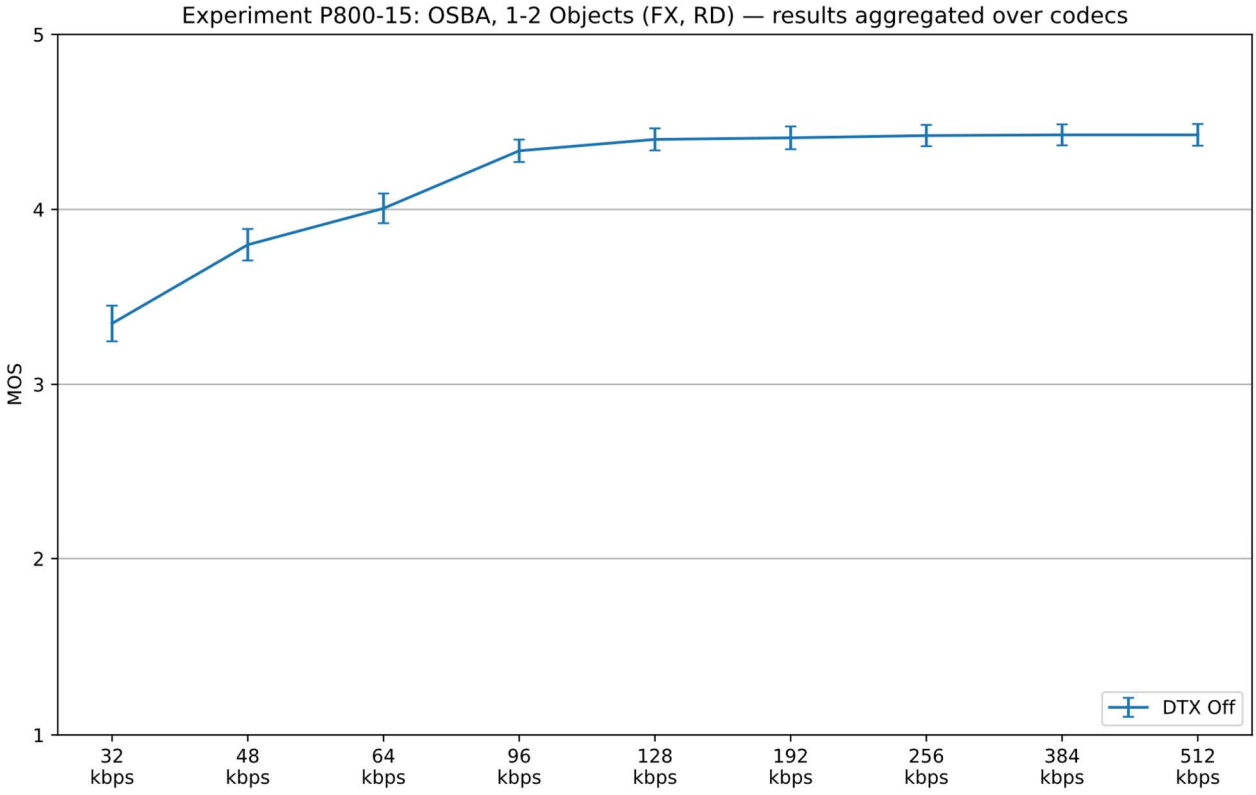


Figure 9.7-4: results aggregated over FX/FL for experiment P800-15

Table 9.7-2: statistical significance analysis for experiment P800-15

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.33	1.22	c19	3.41	1.19	-0.08	<0.001	0.743	NWT
c11	3.82	1.07	c20	3.82	1.04	0.00	0.045	0.482	NWT
c12	4.03	0.98	c21	3.99	0.96	0.03	0.332	0.370	NWT
c13	4.32	0.73	c22	4.36	0.70	-0.04	<0.001	0.721	NWT
c14	4.42	0.74	c23	4.37	0.78	0.05	0.625	0.266	NWT
c15	4.44	0.72	c24	4.39	0.80	0.05	0.623	0.267	NWT
c16	4.45	0.77	c25	4.44	0.68	0.01	0.144	0.443	NWT
c17	4.38	0.74	c26	4.46	0.72	-0.08	<0.001	0.859	NWT
c18	4.35	0.77	c27	4.44	0.72	-0.09	<0.001	0.872	NWT
c28	3.31	1.21	c19	3.41	1.19	-0.10	<0.001	0.796	NWT
c29	3.76	1.09	c20	3.82	1.04	-0.06	<0.001	0.707	NWT
c30	3.99	1.08	c21	3.99	0.96	0.00	<0.001	0.500	NWT
c31	4.33	0.84	c22	4.36	0.70	-0.03	<0.001	0.657	NWT
c32	4.41	0.72	c23	4.37	0.78	0.03	0.430	0.334	NWT
c33	4.39	0.77	c24	4.39	0.80	-0.00	<0.001	0.524	NWT
c34	4.38	0.73	c25	4.44	0.68	-0.06	<0.001	0.795	NWT
c35	4.44	0.65	c26	4.46	0.72	-0.02	<0.001	0.619	NWT
c36	4.49	0.72	c27	4.44	0.72	0.05	0.659	0.255	NWT

9.7.4 Characterization Experiment P800-16 (OSBA 3-4 Objects, Speech and Mixed/Music + Background, Headphone Presentation)

Characterization Experiment P800-16 evaluates the performance of the IVAS fixed-point and floating-point implementation for OSBA with 3-4 objects for speech and mixed/music + background, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-16 was conducted by Ericsson LM (language: SWE).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.7-3 and are shown in Figure 9.7-5. The results aggregated over FX/FL are shown in Figure 9.7-6. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.7-4. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.7-3: results per condition for experiment P800-16

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	179	4.60	0.59	0.09
c02	MNRU Q = 32 dB	180	4.39	0.73	0.11
c03	MNRU Q = 27 dB	180	3.59	1.30	0.19
c04	MNRU Q = 22 dB	179	2.42	1.22	0.18
c05	MNRU Q = 17 dB	179	1.52	0.91	0.13
c06	ESDRU α = 0.8	179	4.43	0.73	0.11
c07	ESDRU α = 0.6	180	3.84	1.03	0.15
c08	ESDRU α = 0.4	179	3.17	1.14	0.17
c09	ESDRU α = 0.2	179	2.71	1.22	0.18
c10	IVAS FL enc / FX dec, 32 kbps, DTX Off	180	3.42	1.13	0.17
c11	IVAS FX enc / FL dec, 48 kbps, DTX Off	180	3.82	0.97	0.14
c12	IVAS FL enc / FX dec, 64 kbps, DTX Off	179	4.23	0.76	0.11
c13	IVAS FX enc / FL dec, 96 kbps, DTX Off	179	4.35	0.74	0.11
c14	IVAS FL enc / FX dec, 128 kbps, DTX Off	180	4.59	0.65	0.10
c15	IVAS FX enc / FL dec, 192 kbps, DTX Off	180	4.55	0.68	0.10
c16	IVAS FL enc / FX dec, 256 kbps, DTX Off	179	4.62	0.66	0.10

c17	IVAS FX enc / FL dec, 384 kbps, DTX Off	180	4.63	0.63	0.09
c18	IVAS FL enc / FX dec, 512 kbps, DTX Off	180	4.63	0.58	0.09
c19	IVAS FL, 32 kbps, DTX Off	180	3.35	1.16	0.17
c20	IVAS FL, 48 kbps, DTX Off	180	3.92	0.92	0.14
c21	IVAS FL, 64 kbps, DTX Off	180	4.04	0.85	0.12
c22	IVAS FL, 96 kbps, DTX Off	178	4.45	0.70	0.10
c23	IVAS FL, 128 kbps, DTX Off	180	4.49	0.66	0.10
c24	IVAS FL, 192 kbps, DTX Off	179	4.60	0.56	0.08
c25	IVAS FL, 256 kbps, DTX Off	180	4.59	0.70	0.10
c26	IVAS FL, 384 kbps, DTX Off	180	4.58	0.62	0.09
c27	IVAS FL, 512 kbps, DTX Off	180	4.56	0.65	0.10
c28	IVAS FX, 32 kbps, DTX Off	180	3.38	1.07	0.16
c29	IVAS FX, 48 kbps, DTX Off	179	3.92	1.00	0.15
c30	IVAS FX, 64 kbps, DTX Off	178	4.18	0.83	0.12
c31	IVAS FX, 96 kbps, DTX Off	180	4.46	0.72	0.11
c32	IVAS FX, 128 kbps, DTX Off	180	4.58	0.60	0.09
c33	IVAS FX, 192 kbps, DTX Off	180	4.60	0.61	0.09
c34	IVAS FX, 256 kbps, DTX Off	180	4.68	0.54	0.08
c35	IVAS FX, 384 kbps, DTX Off	180	4.61	0.62	0.09
c36	IVAS FX, 512 kbps, DTX Off	179	4.60	0.69	0.10

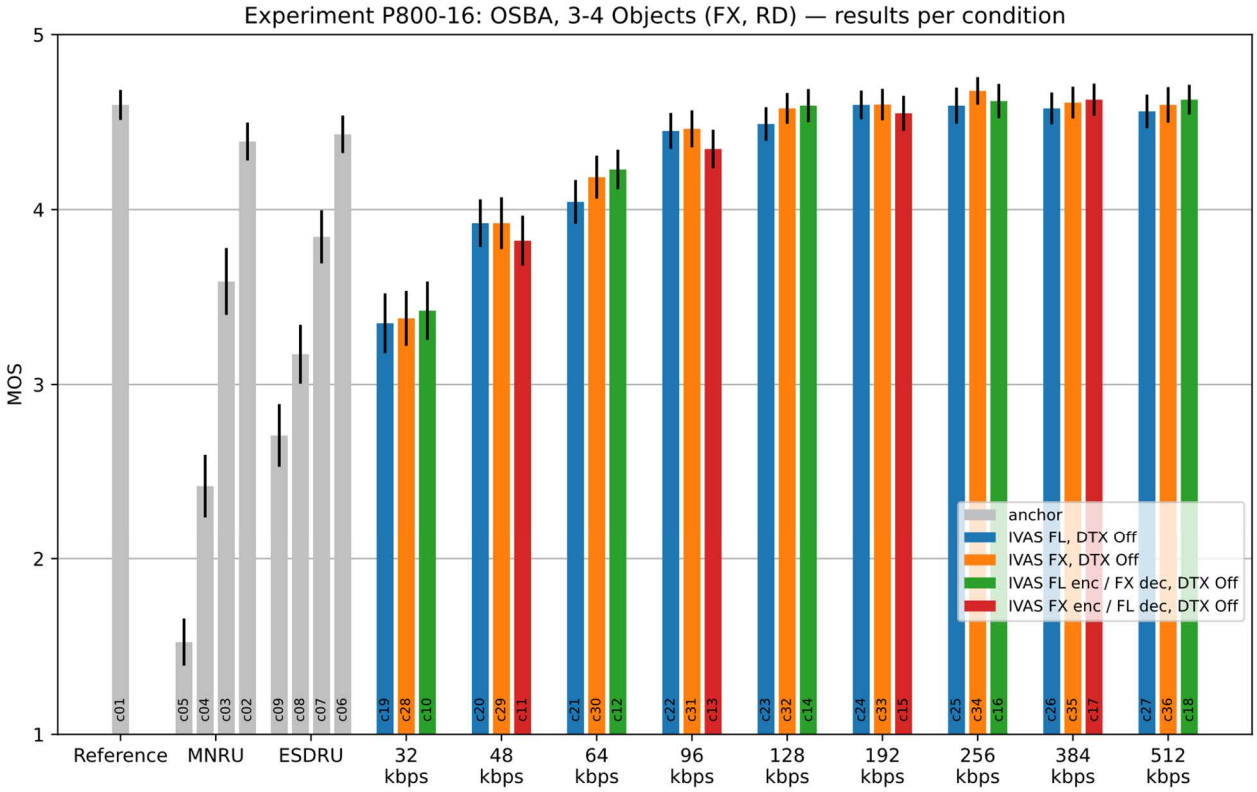


Figure 9.7-5: results per condition for experiment P800-16

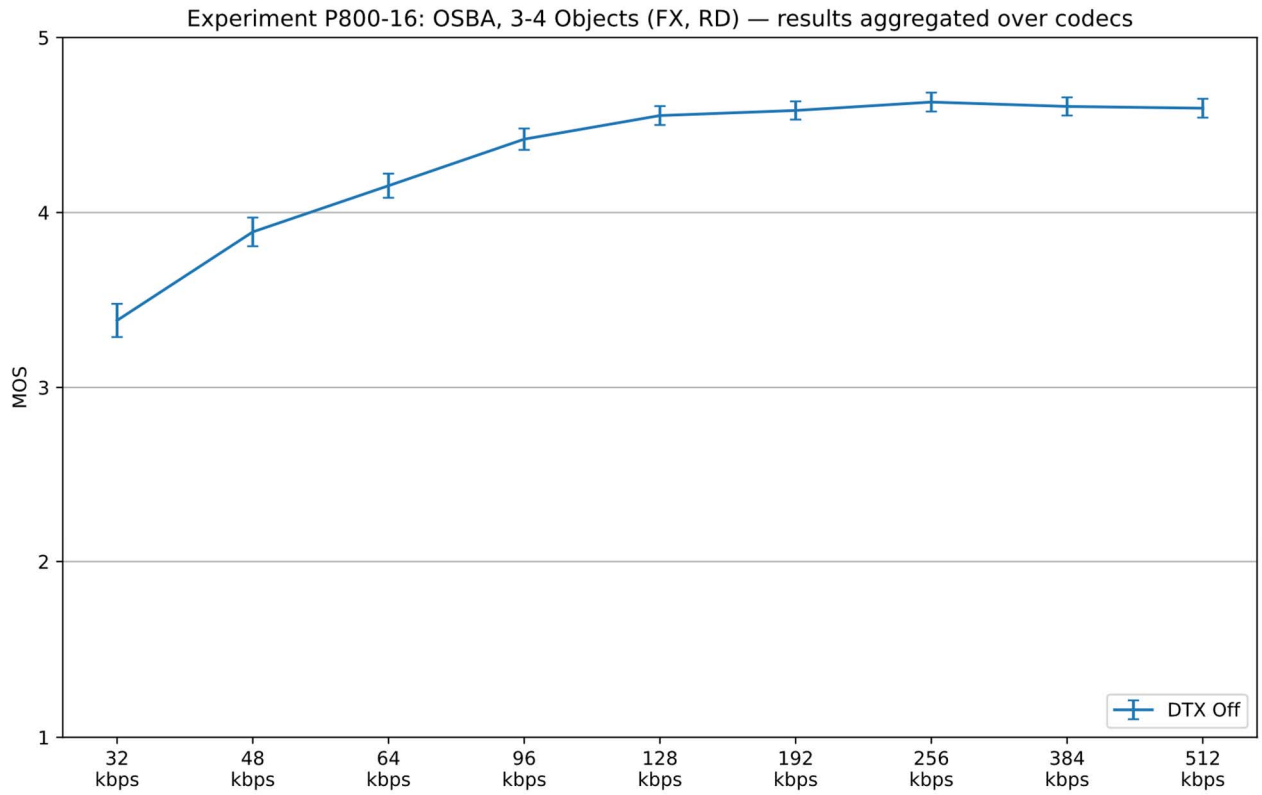


Figure 9.7-6: results aggregated over FX/FL for experiment P800-16

Table 9.7-4: statistical significance analysis for experiment P800-16

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	3.42	1.13	c19	3.35	1.16	0.07	0.596	0.276	NWT
c11	3.82	0.97	c20	3.92	0.92	-0.10	<0.001	0.841	NWT
c12	4.23	0.76	c21	4.04	0.85	0.19	2.169	0.015	BT
c13	4.35	0.74	c22	4.45	0.70	-0.10	<0.001	0.911	NWT
c14	4.59	0.65	c23	4.49	0.66	0.11	1.527	0.064	NWT
c15	4.55	0.68	c24	4.60	0.56	-0.05	<0.001	0.768	NWT
c16	4.62	0.66	c25	4.59	0.70	0.03	0.362	0.359	NWT
c17	4.63	0.63	c26	4.58	0.62	0.05	0.764	0.223	NWT
c18	4.63	0.58	c27	4.56	0.65	0.07	1.030	0.152	NWT
c28	3.38	1.07	c19	3.35	1.16	0.03	0.238	0.406	NWT
c29	3.92	1.00	c20	3.92	0.92	0.00	<0.001	0.500	NWT
c30	4.18	0.83	c21	4.04	0.85	0.14	1.584	0.057	NWT
c31	4.46	0.72	c22	4.45	0.70	0.01	0.160	0.436	NWT
c32	4.58	0.60	c23	4.49	0.66	0.09	1.345	0.090	NWT
c33	4.60	0.61	c24	4.60	0.56	0.00	0.032	0.487	NWT
c34	4.68	0.54	c25	4.59	0.70	0.08	1.280	0.101	NWT
c35	4.61	0.62	c26	4.58	0.62	0.03	0.507	0.306	NWT
c36	4.60	0.69	c27	4.56	0.65	0.04	0.521	0.301	NWT

9.7.5 Characterization Experiment P800-17 (OSBA 1-4 Objects, Speech and Mixed/Music + Background, Impaired Channel, Headphone Presentation)

Characterization Experiment P800-16 evaluates the performance of the IVAS fixed-point and floating-point implementation for OSBA with 1-4 objects for speech and mixed/music + background, under impaired channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-17 was conducted by Fraunhofer IIS (language: GER).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.7-5 and are shown in Figure 9.7-7. The results aggregated over FX/FL are shown in Figure 9.7-8. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.7-6. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.7-5: results per condition for experiment P800-17

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	150	4.77	0.47	0.07
c02	MNRU Q = 34 dB	150	4.69	0.57	0.09
c03	MNRU Q = 30 dB	150	4.73	0.54	0.09
c04	MNRU Q = 26 dB	150	4.43	0.70	0.11
c05	MNRU Q = 22 dB	150	3.62	1.15	0.19
c06	ESDRU α = 0.8	150	4.56	0.61	0.10
c07	ESDRU α = 0.6	150	3.37	1.11	0.18
c08	ESDRU α = 0.4	150	2.54	1.11	0.18
c09	ESDRU α = 0.2	150	2.16	1.10	0.18
c10	IVAS FL, FER 0%, 32 kbps, OSBA, DTX off	150	3.40	0.95	0.15
c11	IVAS FL, FER 0%, 64 kbps, OSBA, DTX off	150	4.13	0.84	0.14
c12	IVAS FL, FER 0%, 128 kbps, OSBA, DTX off	150	4.70	0.53	0.09
c13	IVAS FL, FER 0%, 256 kbps, OSBA, DTX off	150	4.67	0.55	0.09
c14	IVAS FX, FER 5%, 32 kbps, OSBA, DTX off	150	2.80	1.07	0.17
c15	IVAS FX, FER 5%, 64 kbps, OSBA, DTX off	150	3.39	1.01	0.16
c16	IVAS FX, FER 5%, 256 kbps, OSBA, DTX off	150	4.21	0.81	0.13
c17	IVAS FX, FER 0%, 13.2 kbps, OSBA, DTX off	150	1.83	1.00	0.16
c18	IVAS FX, FER 0%, 16.4 kbps, OSBA, DTX off	150	2.53	1.07	0.17
c19	IVAS FX, FER 0%, 24.4 kbps, OSBA, DTX off	150	3.23	1.02	0.16
c20	IVAS FX, FER 0%, 32 kbps, OSBA, DTX off	150	3.37	1.04	0.17
c21	IVAS FX, FER 0%, 48 kbps, OSBA, DTX off	150	3.85	0.92	0.15
c22	IVAS FX, FER 0%, 64 kbps, OSBA, DTX off	150	4.14	0.84	0.14
c23	IVAS FX, FER 0%, 80 kbps, OSBA, DTX off	150	4.10	0.86	0.14
c24	IVAS FX, FER 0%, 96 kbps, OSBA, DTX off	150	4.57	0.56	0.09
c25	IVAS FX, FER 0%, 128 kbps, OSBA, DTX off	150	4.62	0.55	0.09
c26	IVAS FX, FER 0%, 160 kbps, OSBA, DTX off	150	4.61	0.63	0.10
c27	IVAS FX, FER 0%, 192 kbps, OSBA, DTX off	150	4.65	0.67	0.11
c28	IVAS FX, FER 0%, 256 kbps, OSBA, DTX off	150	4.65	0.59	0.10
c29	IVAS FX, FER 0%, 384 kbps, OSBA, DTX off	150	4.77	0.50	0.08
c30	IVAS FX, FER 0%, 512 kbps, OSBA, DTX off	150	4.75	0.48	0.08
c31	IVAS FX, FER 0%, 48.8 kbps, ISM + HOA3, DTX off	150	4.03	0.82	0.13
c32	IVAS FX, FER 0%, 64 kbps, ISM + HOA3, DTX off	150	4.20	0.74	0.12
c33	IVAS FX, FER 0%, 96 kbps, ISM + HOA3, DTX off	150	4.37	0.70	0.11
c34	IVAS FX, FER 0%, 128 kbps, ISM + HOA3, DTX off	150	4.45	0.77	0.12
c35	IVAS FX, FER 0%, 192 kbps, ISM + HOA3, DTX off	150	4.55	0.64	0.10
c36	IVAS FX, FER 0%, 256 kbps, ISM + HOA3, DTX off	150	4.70	0.61	0.10

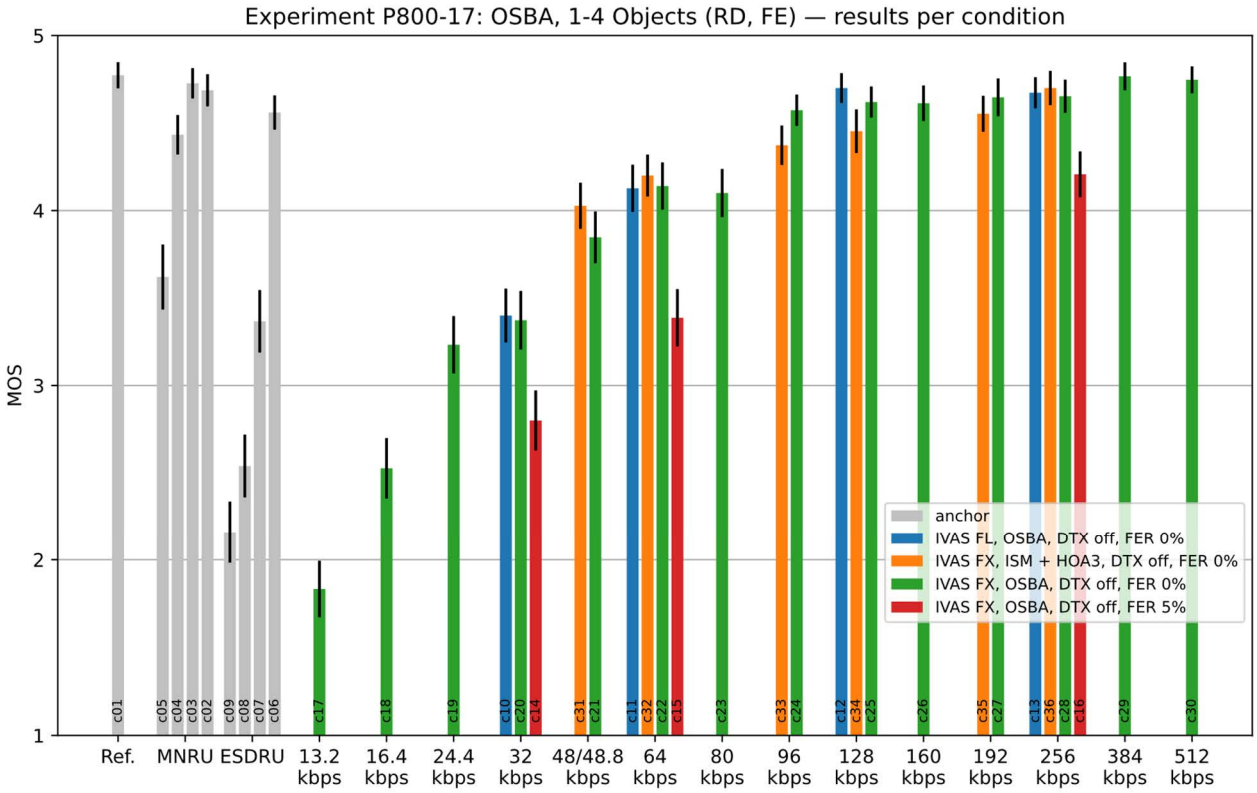


Figure 9.7-7: results per condition for experiment P800-17

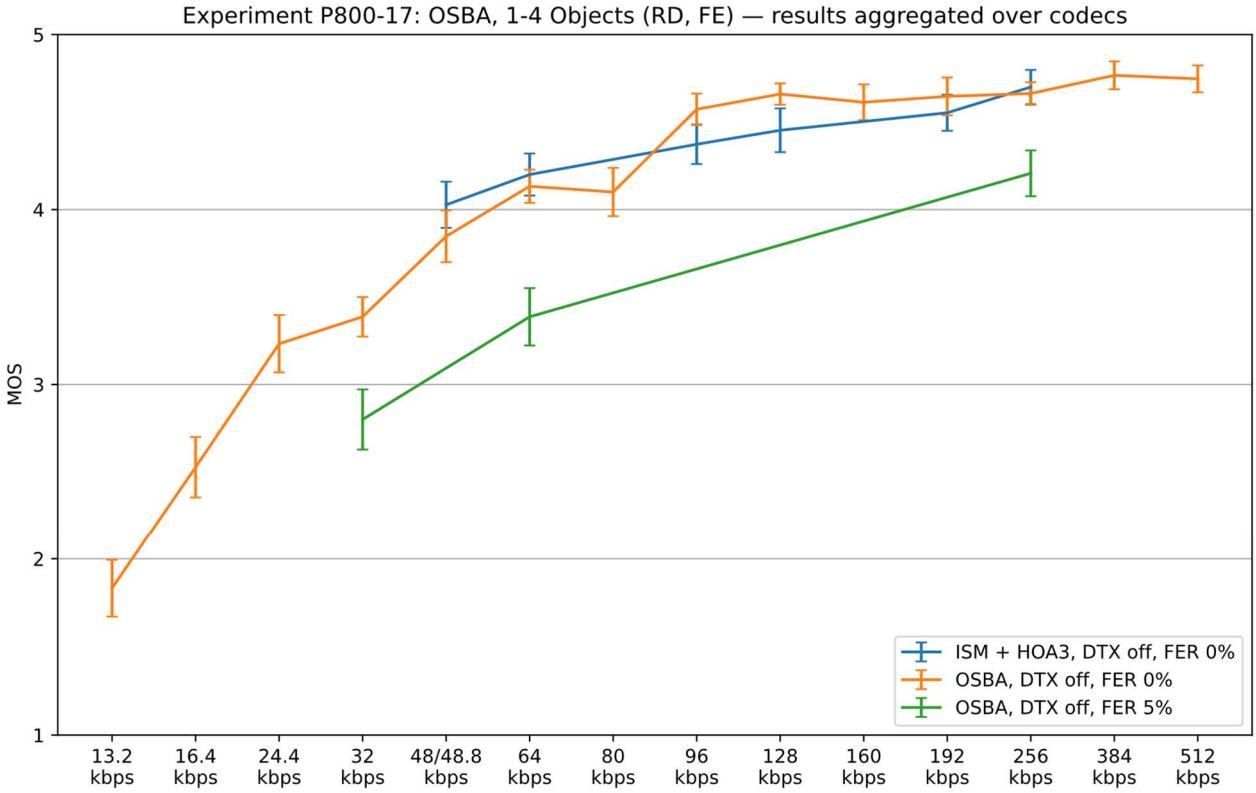


Figure 9.7-8: results aggregated over FX/FL for experiment P800-17

Table 9.7-6: statistical significance analysis for experiment P800-17

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c20	3.37	1.04	c10	3.40	0.95	-0.03	<0.001	0.593	NWT
c21	3.85	0.92	c31	4.03	0.82	-0.18	<0.001	0.963	WT
c22	4.14	0.84	c11	4.13	0.84	0.01	0.135	0.447	NWT
c22	4.14	0.84	c32	4.20	0.74	-0.06	<0.001	0.744	NWT
c24	4.57	0.56	c33	4.37	0.70	0.20	2.732	0.003	BT
c25	4.62	0.55	c12	4.70	0.53	-0.08	<0.001	0.900	NWT
c25	4.62	0.55	c34	4.45	0.77	0.17	2.151	0.016	BT
c27	4.65	0.67	c35	4.55	0.64	0.09	1.245	0.107	NWT
c28	4.65	0.59	c13	4.67	0.55	-0.02	<0.001	0.619	NWT
c28	4.65	0.59	c36	4.70	0.61	-0.05	<0.001	0.751	NWT

9.7.6 Characterization Experiment BS1534-16 (OSBA 1-4 Objects, Generic Audio, 16.4 – 48 kbps, Headphone Presentation)

Characterization Experiment BS1534-16 evaluates the performance of the IVAS fixed-point implementation for OSBA with 1-4. objects for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-16 was conducted by Fraunhofer IIS.

The aggregated results per condition are listed in Table 9.7-7 and are shown in Figure 9.7-9. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.7-8. Scores with result WT are highlighted in red.

Table 9.7-7: results per condition for experiment BS1534-16

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	100.00	0.00	0.00
c02	LP 3.5 kHz	168	23.49	7.91	1.20
c03	Mono EVS, 16.4 kbps	168	39.31	14.83	2.26
c04	Mono EVS, 32 kbps	168	43.12	14.71	2.24
c05	IVAS, 16.4 kbps	168	43.53	15.76	2.40
c06	IVAS, 24.4 kbps	168	53.04	17.02	2.59
c07	IVAS, 32 kbps	168	58.44	15.89	2.42
c08	IVAS, 48 kbps	168	63.59	14.47	2.20

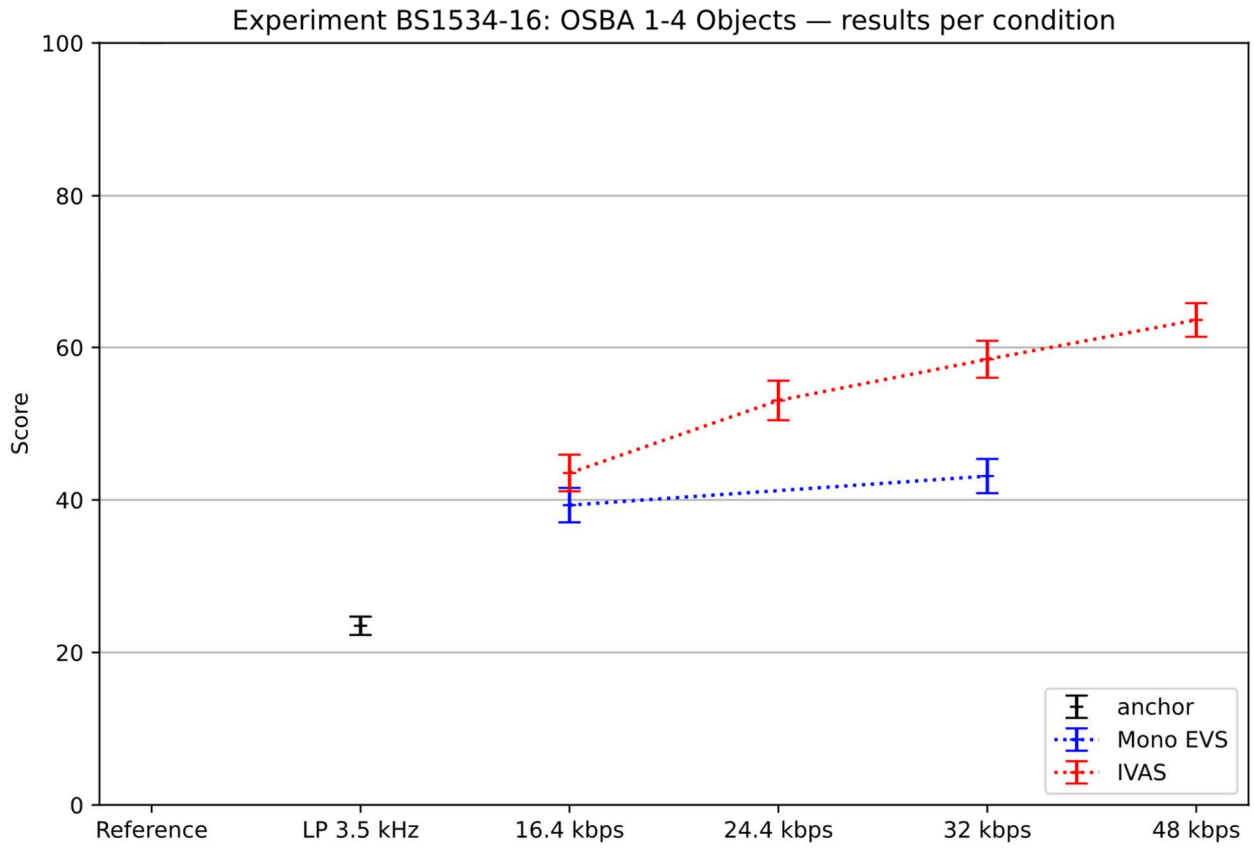


Figure 9.7-9: results per condition for experiment BS1534-16

Table 9.7-8: statistical significance analysis for experiment BS1534-16

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	43.53	15.76	c03	39.31	14.83	4.21	4.075	<0.001	BT
c07	58.44	15.89	c04	43.12	14.71	15.32	13.407	<0.001	BT

9.7.7 Characterization Experiment BS1534-17 (OSBA 1-4 Objects, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-17 evaluates the performance of the IVAS fixed-point implementation for OSBA with 1-4. objects for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-17 was conducted by Dolby Laboratories, Inc.

The aggregated results per condition are listed in Table 9.7-9 and are shown in Figure 9.7-10. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.7-10. Scores with result WT are highlighted in red.

Table 9.7-9: results per condition for experiment BS1534-17

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.80	1.18	0.18
c02	LP 3.5 kHz	168	34.17	8.16	1.24
c03	Mono EVS, 64 kbps	168	60.43	9.41	1.43
c04	Mono EVS, 128 kbps	168	60.95	9.28	1.41
c05	IVAS, 64 kbps	168	68.75	9.14	1.39
c06	IVAS, 96 kbps	168	72.02	9.20	1.40
c07	IVAS, 128 kbps	168	75.72	9.14	1.39
c08	IVAS, 256 kbps	168	80.41	10.22	1.56

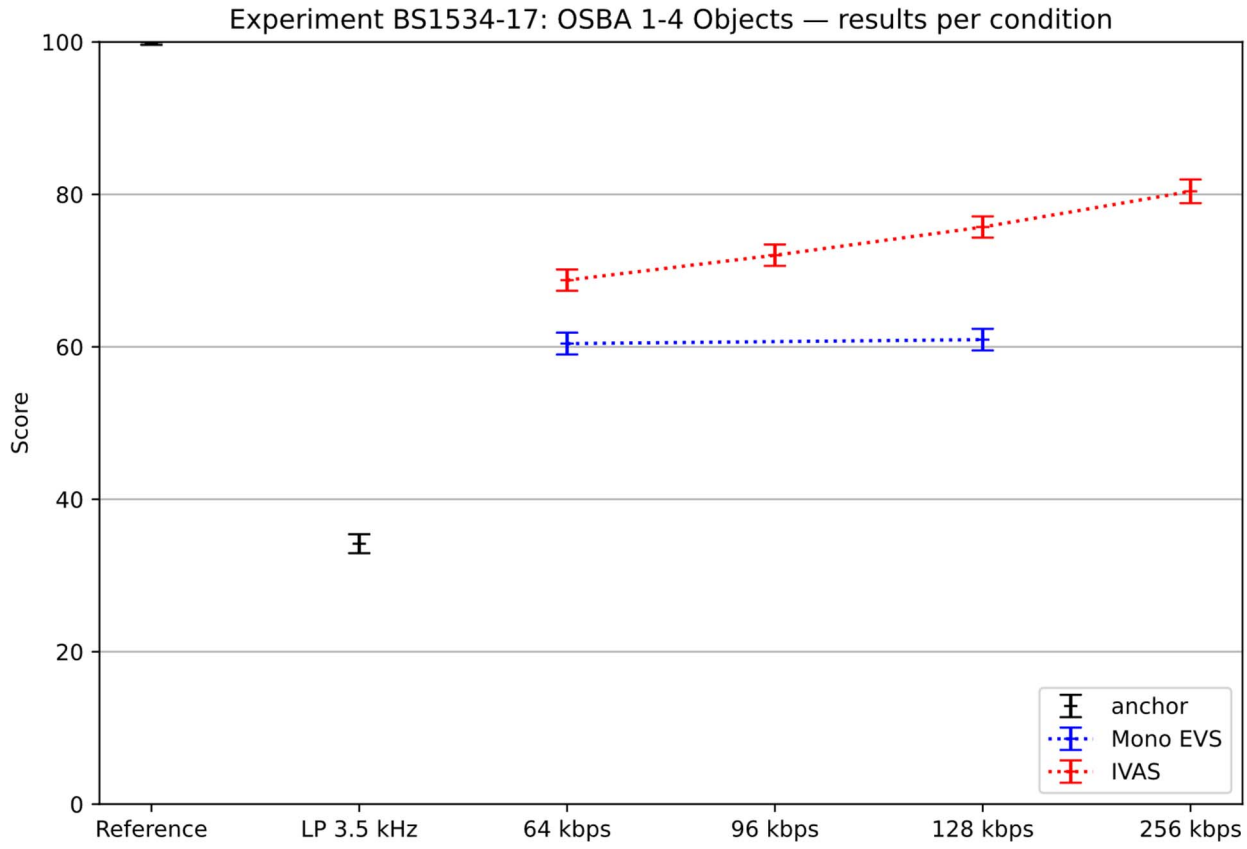


Figure 9.7-10: results per condition for experiment BS1534-17

Table 9.7-10: statistical significance analysis for experiment BS1534-17

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	68.75	9.14	c03	60.43	9.41	8.32	14.550	<0.001	BT
c07	75.72	9.14	c04	60.95	9.28	14.77	22.462	<0.001	BT

9.7.8 Conclusions on IVAS Performance for Combined Objects and SBA (OSBA)

Taking the results of the previous clauses in 9.7 into account, the following conclusions can be drawn for IVAS Combined Objects and SBA (OSBA) modes:

During standardization of the IVAS codec, additional information on the performance of OSBA for 256 and 512 kbps has been provided. At these bitrates, the combined OSBA modes outperform separate coding of SBA and audio objects significantly, with IVAS showing “excellent” quality at both 256 kbps and 512 kbps, and a MUSHRA score above 90 for 512kbps. Additional bitrates have been evaluated during characterization testing.

During characterization phase, IVAS OSBA was evaluated against EVS mono in a series of BS.1534 tests. Across all tested bitrates, IVAS was significantly outperforming EVS mono at the same bitrate, with an improvement in mean score across items and listeners of up to 15 MUSHRA points (32 kbps). IVAS OSBA maintained “good” MUSHRA quality starting from 48 kbps and reached “excellent” quality at 256 kbps.

Additionally, the performance of IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for OSBA in a series of P.800 tests. After adjustment of the significance level to account for multiple comparisons (Bonferroni method), all of the 40 tested operating points were statistically equivalent.

Overall, the IVAS OSBA coding for maintains a “good” MOS rating (>= 4.0) from bitrates starting from 64 kbps and above.

9.8 Combined Objects and MASA (OMASA)

9.8.1 Overview

Provided as optional additional information provided for codec selection, one experiments has been conducted to evaluate the performance of the IVAS codec with the combination of combined objects and MASA content. The experiment was conducted as BS.1534 test using headphone presentation.

- Optional additional information experiment 3: Generic audio, combined objects and MASA, 48, 64, 80 and 128 kbps, headphone presentation

In Characterization phase, four additional experiments have been conducted. While the experiments P800-18, P800-19 and P800-20 were conducted as P.800 DCR tests, the experiment BS1534-18 was conducted as a BS.1534 test.

- Characterization Experiment P800-18: Speech and mixed/music + background, 1-2 objects combined with MASA, under clean channel conditions, headphone presentation
- Characterization Experiment P800-19: Speech and mixed/music + background, 3-4 objects combined with MASA, under clean channel conditions, headphone presentation
- Characterization Experiment P800-20: Speech and mixed/music + background, 1-4 objects combined with MASA, under impaired channel conditions, headphone presentation
- Characterization Experiment BS1534-18: Generic audio, 1-4 objects combined with MASA, 64 – 256 kbps, headphone presentation

9.8.2 Optional additional information provided for codec selection

The following quality assessment has been provided as optional additional information for codec selection:

Combined ISM (objects) and spatial capture with MASA provide an efficient way to, e.g., capture and transmit the environment and talker(s) at the same time. Combined coding of ISMs and MASA provides both increased coding efficiency and reduced complexity and memory usage compared to separate codec instances. Combined ISM and MASA operation is possible with 1 to 4 ISMs at all bitrates from 13.2 to 512 kbps. At the lowest bitrates ISMs and MASA are pre-renderer into MASA format and with increasing bitrate objects are parametrically presented, most important object separately coded, and finally all ISMs separately coded with the accompanying metadata in addition to the spatial MASA signal. The combined format coding automatically distributes allowed bit budget optimally between ISMs, MASA, and spatial metadata.

When comparing to separate coding of ISMs and MASA, combined coding (labeled ‘oMASA’ in Figure 9.8-1) provides significant quality improvement especially at lower bitrates.

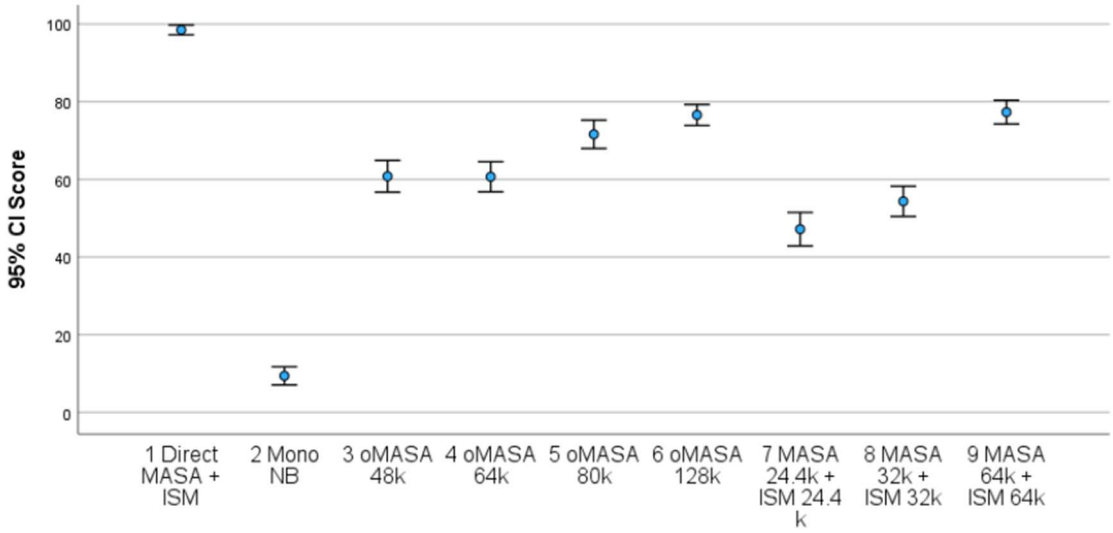


Figure 9.8-1: BS.1534 comparison between combined vs. separate operation of ISM and MASA at various bitrates

9.8.3 Characterization Experiment P800-18 (OMASA 1-2 Objects, Speech and Mixed/Music + Background, Headphone Presentation)

Characterization Experiment P800-18 evaluates the performance of the IVAS fixed-point and floating-point implementation for OMASA with 1-2 objects for speech and mixed/music + background, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-18 was conducted by Ericsson LM (language: SWE).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.8-1 and are shown in Figure 9.8-2. The results aggregated over FX/FL are shown in Figure 9.8-3. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.8-2. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.8-1: results per condition for experiment P800-18

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	179	4.60	0.58	0.09
c02	MNRU Q = 32 dB	180	4.11	0.98	0.14
c03	MNRU Q = 27 dB	180	2.96	1.20	0.18
c04	MNRU Q = 22 dB	179	1.84	1.01	0.15
c05	MNRU Q = 17 dB	180	1.22	0.50	0.07
c06	ESDRU α = 0.8	179	4.52	0.65	0.10
c07	ESDRU α = 0.6	180	3.63	0.94	0.14
c08	ESDRU α = 0.4	180	2.73	1.04	0.15
c09	ESDRU α = 0.2	180	2.21	1.02	0.15
c10	IVAS FL enc / FX dec, 13.2 kbps, DTX off	180	2.82	0.95	0.14
c11	IVAS FX enc / FL dec, 16.4 kbps, DTX off	180	3.20	1.00	0.15
c12	IVAS FL enc / FX dec, 24.4 kbps, DTX off	180	3.53	0.90	0.13
c13	IVAS FX enc / FL dec, 32 kbps, DTX off	180	3.91	0.88	0.13
c14	IVAS FL enc / FX dec, 48 kbps, DTX off	179	4.43	0.64	0.10
c15	IVAS FX enc / FL dec, 64 kbps, DTX off	180	4.48	0.66	0.10
c16	IVAS FL enc / FX dec, 80 kbps, DTX off	179	4.50	0.62	0.09
c17	IVAS FX enc / FL dec, 128 kbps, DTX off	179	4.56	0.59	0.09
c18	IVAS FL enc / FX dec, 256 kbps, DTX off	180	4.63	0.57	0.08
c19	IVAS FL, 13.2 kbps, DTX off	180	2.92	0.99	0.14

c20	IVAS FL, 16.4 kbps, DTX off	180	3.21	1.01	0.15
c21	IVAS FL, 24.4 kbps, DTX off	179	3.56	1.03	0.15
c22	IVAS FL, 32 kbps, DTX off	180	3.92	0.89	0.13
c23	IVAS FL, 48 kbps, DTX off	179	4.41	0.72	0.11
c24	IVAS FL, 64 kbps, DTX off	180	4.53	0.65	0.10
c25	IVAS FL, 80 kbps, DTX off	179	4.52	0.63	0.09
c26	IVAS FL, 128 kbps, DTX off	179	4.60	0.59	0.09
c27	IVAS FL, 256 kbps, DTX off	180	4.64	0.57	0.08
c28	IVAS FX, 13.2 kbps, DTX off	179	2.86	0.96	0.14
c29	IVAS FX, 16.4 kbps, DTX off	180	3.19	1.02	0.15
c30	IVAS FX, 24.4 kbps, DTX off	180	3.52	1.02	0.15
c31	IVAS FX, 32 kbps, DTX off	180	3.86	0.91	0.13
c32	IVAS FX, 48 kbps, DTX off	179	4.40	0.68	0.10
c33	IVAS FX, 64 kbps, DTX off	180	4.56	0.61	0.09
c34	IVAS FX, 80 kbps, DTX off	180	4.53	0.63	0.09
c35	IVAS FX, 128 kbps, DTX off	179	4.65	0.55	0.08
c36	IVAS FX, 256 kbps, DTX off	180	4.64	0.60	0.09

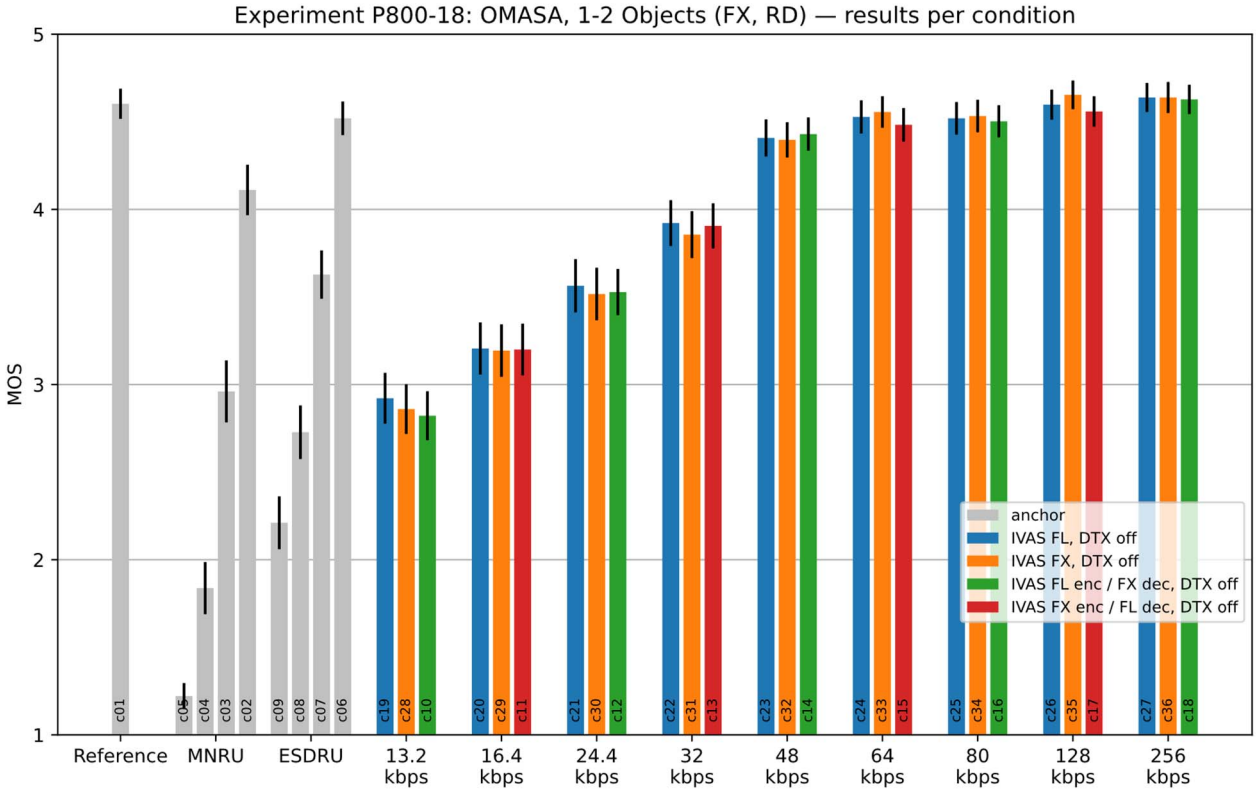


Figure 9.8-2: results per condition for experiment P800-18

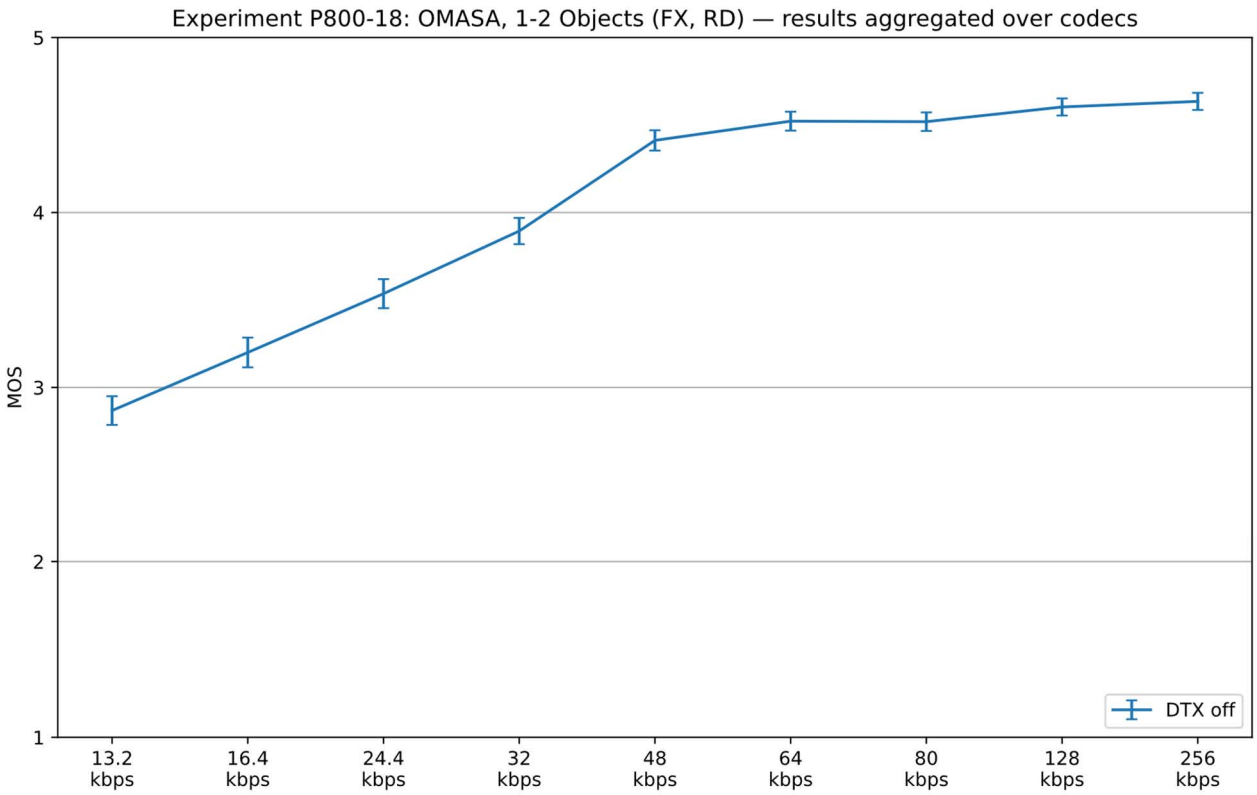


Figure 9.8-3: results aggregated over FX/FL for experiment P800-18

Table 9.8-2: statistical significance analysis for experiment P800-18

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	2.82	0.95	c19	2.92	0.99	-0.10	<0.001	0.835	NWT
c11	3.20	1.00	c20	3.21	1.01	-0.01	<0.001	0.522	NWT
c12	3.53	0.90	c21	3.56	1.03	-0.04	<0.001	0.638	NWT
c13	3.91	0.88	c22	3.92	0.89	-0.02	<0.001	0.568	NWT
c14	4.43	0.64	c23	4.41	0.72	0.02	0.306	0.380	NWT
c15	4.48	0.66	c24	4.53	0.65	-0.04	<0.001	0.744	NWT
c16	4.50	0.62	c25	4.52	0.63	-0.02	<0.001	0.601	NWT
c17	4.56	0.59	c26	4.60	0.59	-0.04	<0.001	0.734	NWT
c18	4.63	0.57	c27	4.64	0.57	-0.01	<0.001	0.573	NWT
c28	2.86	0.96	c19	2.92	0.99	-0.06	<0.001	0.726	NWT
c29	3.19	1.02	c20	3.21	1.01	-0.01	<0.001	0.545	NWT
c30	3.52	1.02	c21	3.56	1.03	-0.05	<0.001	0.668	NWT
c31	3.86	0.91	c22	3.92	0.89	-0.07	<0.001	0.756	NWT
c32	4.40	0.68	c23	4.41	0.72	-0.01	<0.001	0.559	NWT
c33	4.56	0.61	c24	4.53	0.65	0.03	0.423	0.336	NWT
c34	4.53	0.63	c25	4.52	0.63	0.01	0.196	0.423	NWT
c35	4.65	0.55	c26	4.60	0.59	0.06	0.930	0.177	NWT
c36	4.64	0.60	c27	4.64	0.57	0.00	<0.001	0.500	NWT

9.8.4 Characterization Experiment P800-19 (OMASA 3-4 Objects, Speech and Mixed/Music + Background, Headphone Presentation)

Characterization Experiment P800-19 evaluates the performance of the IVAS fixed-point and floating-point implementation for OMASA with 3-4 objects for speech and mixed/music + background, under clean channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-19 was conducted by Fraunhofer IIS (language: GER).

The listening lab has reported that some votes are missing. This was taken into account for result aggregation and statistical calculations.

The aggregated results per condition are listed in Table 9.8-3 and are shown in Figure 9.8-4. The results aggregated over FX/FL are shown in Figure 9.8-5. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.8-4. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.8-3: results per condition for experiment P800-19

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	150	4.68	0.58	0.09
c02	MNRU Q = 32 dB	150	4.59	0.61	0.10
c03	MNRU Q = 27 dB	150	4.21	0.86	0.14
c04	MNRU Q = 22 dB	150	3.15	1.21	0.20
c05	MNRU Q = 17 dB	150	1.83	0.90	0.14
c06	ESDRU α = 0.8	150	4.54	0.59	0.10
c07	ESDRU α = 0.6	150	3.55	0.95	0.15
c08	ESDRU α = 0.4	150	2.69	1.09	0.18
c09	ESDRU α = 0.2	150	2.10	1.02	0.17
c10	IVAS FL enc / FX dec, 16.4 kbps, DTX off	150	2.71	1.11	0.18
c11	IVAS FX enc / FL dec, 24.4 kbps, DTX off	150	2.69	1.04	0.17
c12	IVAS FL enc / FX dec, 48 kbps, DTX off	150	4.15	0.90	0.14
c13	IVAS FX enc / FL dec, 64 kbps, DTX off	150	4.28	0.80	0.13
c14	IVAS FL enc / FX dec, 80 kbps, DTX off	150	4.31	0.75	0.12
c15	IVAS FX enc / FL dec, 96 kbps, DTX off	150	4.33	0.76	0.12
c16	IVAS FL enc / FX dec, 128 kbps, DTX off	150	4.51	0.73	0.12
c17	IVAS FX enc / FL dec, 256 kbps, DTX off	150	4.53	0.62	0.10
c18	IVAS FL enc / FX dec, 512 kbps, DTX off	150	4.58	0.64	0.10
c19	IVAS FL, 16.4 kbps, DTX off	150	2.76	1.15	0.18
c20	IVAS FL, 24.4 kbps, DTX off	150	2.90	1.04	0.17
c21	IVAS FL, 48 kbps, DTX off	150	4.10	0.87	0.14
c22	IVAS FL, 64 kbps, DTX off	150	4.12	0.82	0.13
c23	IVAS FL, 80 kbps, DTX off	150	4.31	0.74	0.12
c24	IVAS FL, 96 kbps, DTX off	150	4.43	0.70	0.11
c25	IVAS FL, 128 kbps, DTX off	150	4.40	0.71	0.12
c26	IVAS FL, 256 kbps, DTX off	150	4.56	0.62	0.10
c27	IVAS FL, 512 kbps, DTX off	150	4.66	0.59	0.10
c28	IVAS FX, 16.4 kbps, DTX off	150	2.66	1.14	0.18
c29	IVAS FX, 24.4 kbps, DTX off	150	2.78	1.07	0.17
c30	IVAS FX, 48 kbps, DTX off	150	4.14	0.77	0.12
c31	IVAS FX, 64 kbps, DTX off	150	4.21	0.73	0.12
c32	IVAS FX, 80 kbps, DTX off	150	4.34	0.72	0.12
c33	IVAS FX, 96 kbps, DTX off	150	4.31	0.74	0.12
c34	IVAS FX, 128 kbps, DTX off	150	4.43	0.68	0.11
c35	IVAS FX, 256 kbps, DTX off	150	4.53	0.65	0.10
c36	IVAS FX, 512 kbps, DTX off	150	4.61	0.61	0.10

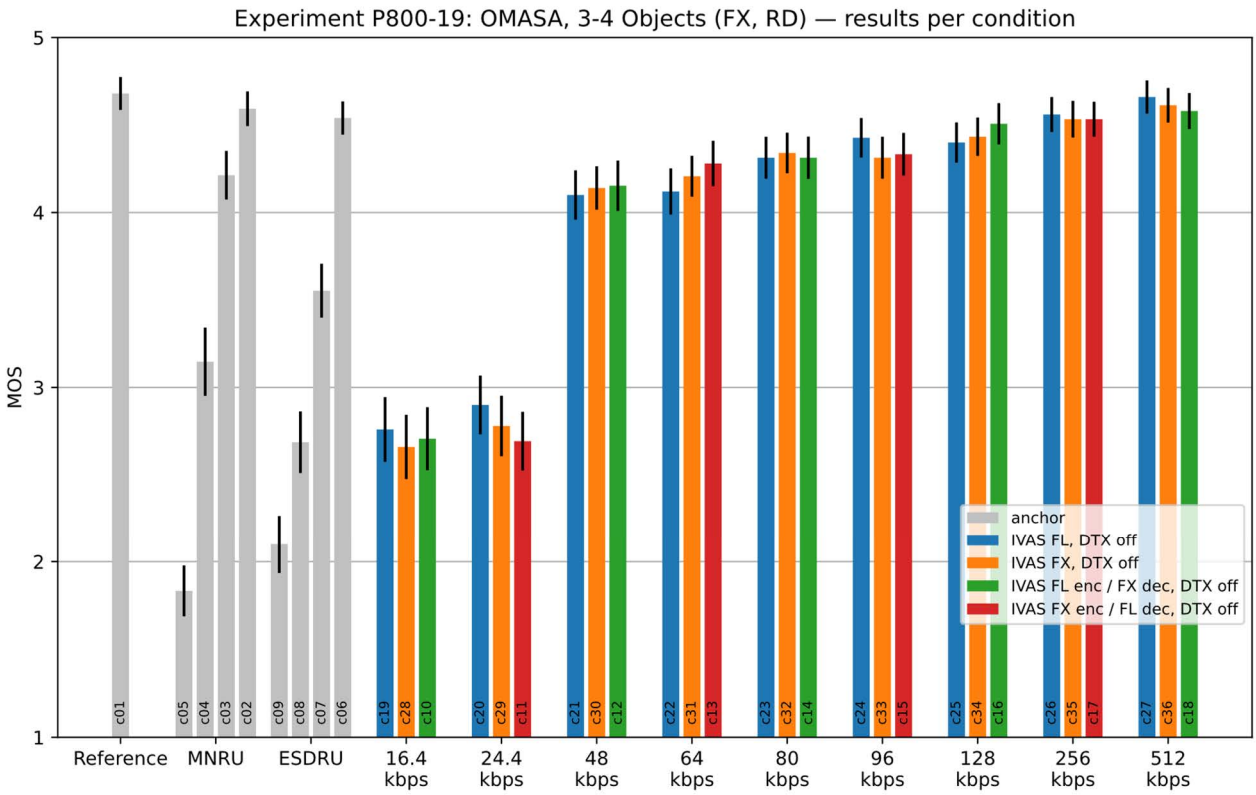


Figure 9.8-4: results per condition for experiment P800-19

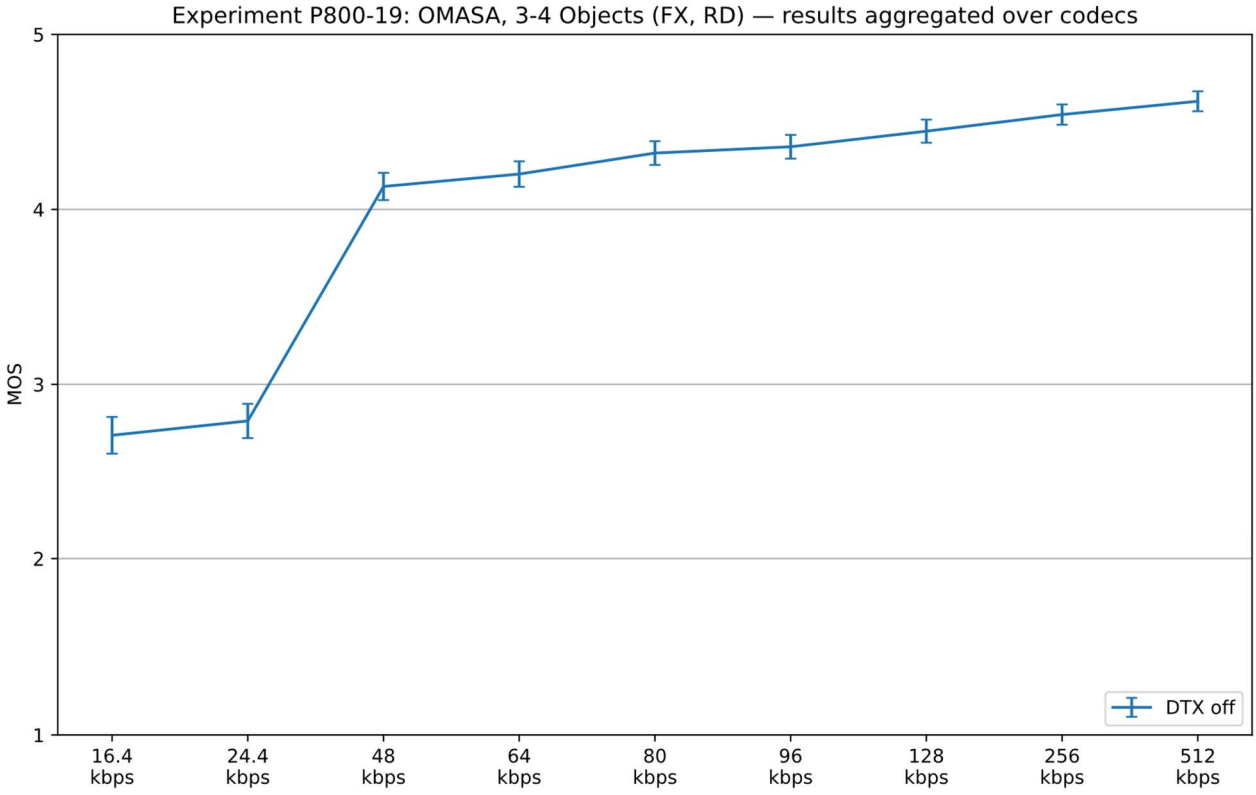


Figure 9.8-5: results aggregated over FX/FL for experiment P800-19

Table 9.8-4: statistical significance analysis for experiment P800-19

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c10	2.71	1.11	c19	2.76	1.15	-0.05	<0.001	0.658	NWT
c11	2.69	1.04	c20	2.90	1.04	-0.21	<0.001	0.957	WT
c12	4.15	0.90	c21	4.10	0.87	0.05	0.519	0.302	NWT
c13	4.28	0.80	c22	4.12	0.82	0.16	1.707	0.044	BT
c14	4.31	0.75	c23	4.31	0.74	0.00	<0.001	0.500	NWT
c15	4.33	0.76	c24	4.43	0.70	-0.09	<0.001	0.868	NWT
c16	4.51	0.73	c25	4.40	0.71	0.11	1.283	0.100	NWT
c17	4.53	0.62	c26	4.56	0.62	-0.03	<0.001	0.647	NWT
c18	4.58	0.64	c27	4.66	0.59	-0.08	<0.001	0.870	NWT
c28	2.66	1.14	c19	2.76	1.15	-0.10	<0.001	0.775	NWT
c29	2.78	1.07	c20	2.90	1.04	-0.12	<0.001	0.837	NWT
c30	4.14	0.77	c21	4.10	0.87	0.04	0.421	0.337	NWT
c31	4.21	0.73	c22	4.12	0.82	0.09	0.974	0.166	NWT
c32	4.34	0.72	c23	4.31	0.74	0.03	0.319	0.375	NWT
c33	4.31	0.74	c24	4.43	0.70	-0.11	<0.001	0.914	NWT
c34	4.43	0.68	c25	4.40	0.71	0.03	0.410	0.341	NWT
c35	4.53	0.65	c26	4.56	0.62	-0.03	<0.001	0.643	NWT
c36	4.61	0.61	c27	4.66	0.59	-0.05	<0.001	0.751	NWT

9.8.5 Characterization Experiment P800-20 (OMASA 1-4 Objects, Speech and Mixed/Music + Background, Impaired Channel, Headphone Presentation)

Characterization Experiment P800-20 evaluates the performance of the IVAS fixed-point and floating-point implementation for OMASA with 1-4 objects for speech and mixed/music + background, under impaired channel conditions with headphone presentation. See Annex E.9 for details.

Experiment P800-20 was conducted by Nokia (language: FIN).

The listening lab has reported that the presentation order of the samples to the listeners was not according to the original randomization tables due to a configuration issue in the listening test software.

The aggregated results per condition are listed in Table 9.8-5 and are shown in Figure 9.8-6. The results aggregated over FX/FL are shown in Figure 9.8-7. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.8-6. Significant different MOS values (BT or WT) are highlighted in red.

Table 9.8-5: results per condition for experiment P800-20

Designator	Condition	COUNT	MOS	STD	CI95
c01	Reference	180	4.59	0.59	0.09
c02	MNRU Q = 34 dB	180	4.58	0.62	0.09
c03	MNRU Q = 30 dB	180	4.47	0.68	0.10
c04	MNRU Q = 26 dB	180	3.81	1.08	0.16
c05	MNRU Q = 22 dB	180	2.44	1.27	0.19
c06	ESDRU α = 0.8	180	4.34	0.73	0.11
c07	ESDRU α = 0.6	180	3.04	1.03	0.15
c08	ESDRU α = 0.4	180	2.11	0.97	0.14
c09	ESDRU α = 0.2	180	1.67	0.91	0.13
c10	IVAS FL, FER 0%, 32 kbps, OMASA 2TC, DTX off	180	2.52	1.23	0.18
c11	IVAS FL, FER 0%, 64 kbps, OMASA 2TC, DTX off	180	3.79	0.97	0.14
c12	IVAS FL, FER 0%, 128 kbps, OMASA 2TC, DTX off	180	4.24	0.83	0.12
c13	IVAS FL, FER 0%, 256 kbps, OMASA 2TC, DTX off	180	4.49	0.65	0.10

c14	IVAS FX, FER 5%, 32 kbps, OMASA 2TC, DTX off	180	2.02	0.91	0.13
c15	IVAS FX, FER 5%, 64 kbps, OMASA 2TC, DTX off	180	3.12	1.04	0.15
c16	IVAS FX, FER 5%, 256 kbps, OMASA 2TC, DTX off	180	3.59	0.96	0.14
c17	IVAS FX, FER 0%, 13.2 kbps, OMASA 2TC, DTX off	180	1.51	0.69	0.10
c18	IVAS FX, FER 0%, 16.4 kbps, OMASA 2TC, DTX off	180	1.92	0.91	0.13
c19	IVAS FX, FER 0%, 24.4 kbps, OMASA 2TC, DTX off	180	2.14	0.89	0.13
c20	IVAS FX, FER 0%, 32 kbps, OMASA 2TC, DTX off	180	2.54	1.19	0.17
c21	IVAS FX, FER 0%, 48 kbps, OMASA 2TC, DTX off	180	3.73	0.97	0.14
c22	IVAS FX, FER 0%, 64 kbps, OMASA 2TC, DTX off	180	3.82	0.99	0.15
c23	IVAS FX, FER 0%, 80 kbps, OMASA 2TC, DTX off	180	3.88	0.90	0.13
c24	IVAS FX, FER 0%, 96 kbps, OMASA 2TC, DTX off	180	4.08	0.83	0.12
c25	IVAS FX, FER 0%, 128 kbps, OMASA 2TC, DTX off	180	4.31	0.76	0.11
c26	IVAS FX, FER 0%, 160 kbps, OMASA 2TC, DTX off	180	4.26	0.84	0.12
c27	IVAS FX, FER 0%, 192 kbps, OMASA 2TC, DTX off	180	4.36	0.69	0.10
c28	IVAS FX, FER 0%, 256 kbps, OMASA 2TC, DTX off	180	4.51	0.61	0.09
c29	IVAS FX, FER 0%, 384 kbps, OMASA 2TC, DTX off	180	4.50	0.66	0.10
c30	IVAS FX, FER 0%, 512 kbps, OMASA 2TC, DTX off	180	4.52	0.66	0.10
c31	IVAS FX, FER 0%, 48.8 kbps, ISM + MASA 2TC, DTX off	180	3.64	0.96	0.14
c32	IVAS FX, FER 0%, 64 kbps, ISM + MASA 2TC, DTX off	180	3.76	0.93	0.14
c33	IVAS FX, FER 0%, 96 kbps, ISM + MASA 2TC, DTX off	180	4.06	0.77	0.11
c34	IVAS FX, FER 0%, 128 kbps, ISM + MASA 2TC, DTX off	180	4.25	0.72	0.11
c35	IVAS FX, FER 0%, 192 kbps, ISM + MASA 2TC, DTX off	180	4.38	0.67	0.10
c36	IVAS FX, FER 0%, 256 kbps, ISM + MASA 2TC, DTX off	180	4.54	0.60	0.09

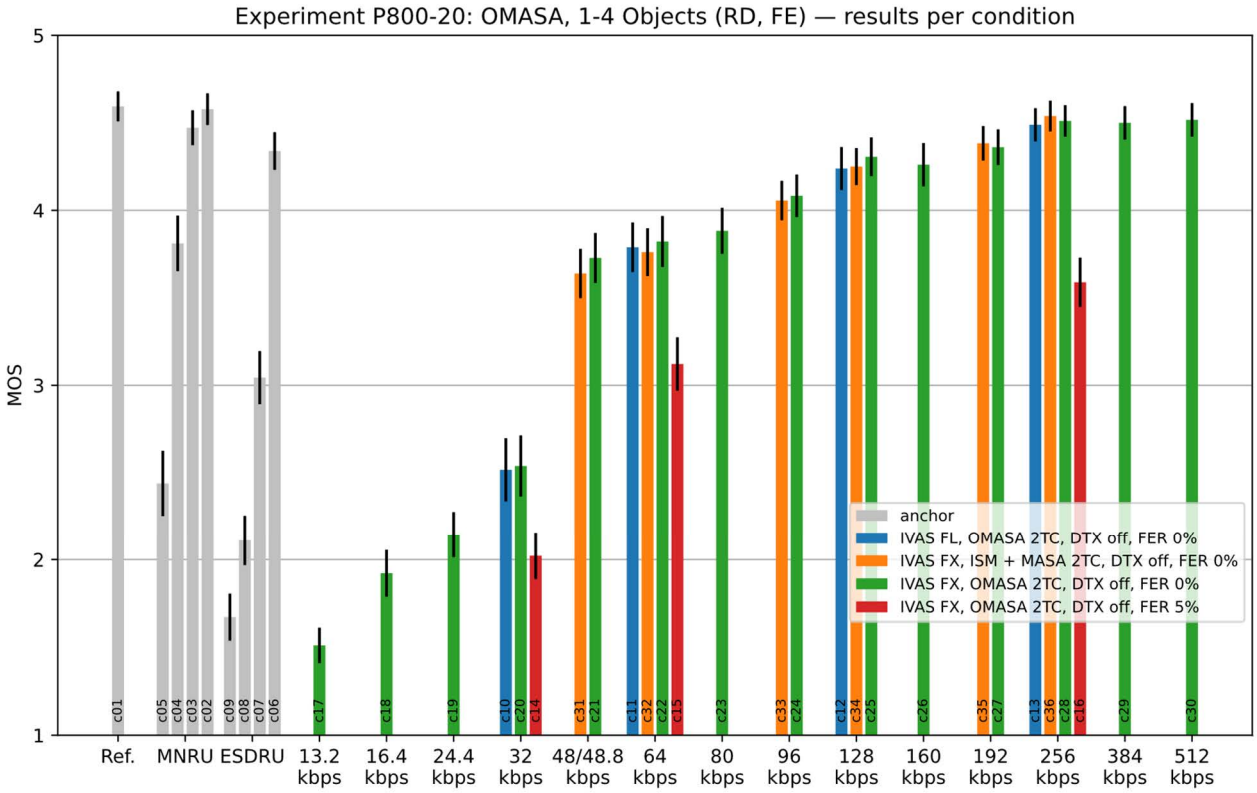


Figure 9.8-6: results per condition for experiment P800-20

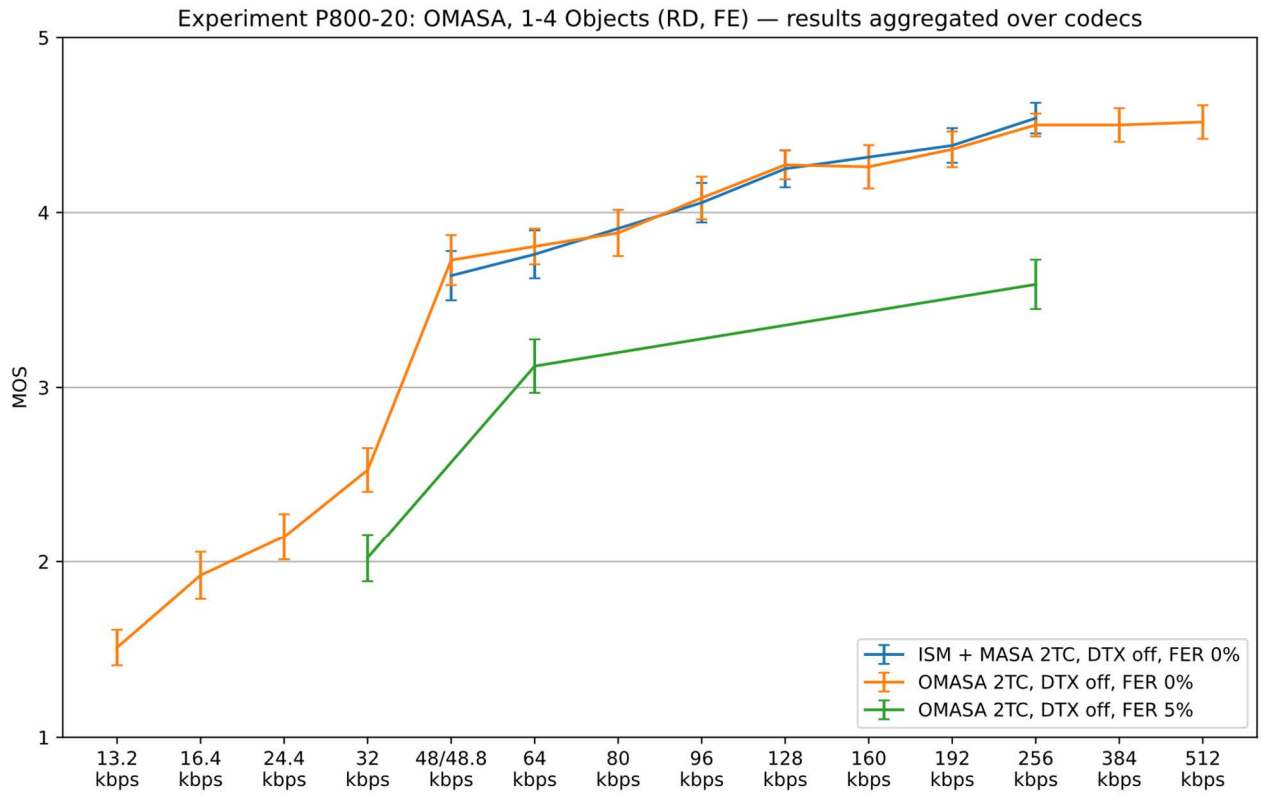


Figure 9.8-7: results aggregated over FX/FL for experiment P800-20

Table 9.8-6: statistical significance analysis for experiment P800-20

CuT			Reference			Evaluation			
Condition	MOS	STD	Condition	MOS	STD	ΔMOS	T-Stat	p-value	Result
c20	2.54	1.19	c10	2.52	1.23	0.02	0.172	0.432	NWT
c21	3.73	0.97	c31	3.64	0.96	0.09	0.872	0.192	NWT
c22	3.82	0.99	c11	3.79	0.97	0.03	0.319	0.375	NWT
c22	3.82	0.99	c32	3.76	0.93	0.06	0.602	0.274	NWT
c24	4.08	0.83	c33	4.06	0.77	0.03	0.320	0.375	NWT
c25	4.31	0.76	c12	4.24	0.83	0.07	0.798	0.213	NWT
c25	4.31	0.76	c34	4.25	0.72	0.06	0.718	0.237	NWT
c27	4.36	0.69	c35	4.38	0.67	-0.02	<0.001	0.620	NWT
c28	4.51	0.61	c13	4.49	0.65	0.02	0.331	0.370	NWT
c28	4.51	0.61	c36	4.54	0.60	-0.03	<0.001	0.669	NWT

9.8.6 Characterization Experiment BS1534-18 (OMASA 1-4 Objects, Generic Audio, 64 – 256 kbps, Headphone Presentation)

Characterization Experiment BS1534-18 evaluates the performance of the IVAS fixed-point implementation for OMASA with 1-4. objects for generic audio content with headphone presentation. This performance is evaluated against EVS (mono) at the same bitrate. See Annex E.9 for details.

Experiment BS1534-18 was conducted by Nokia.

The aggregated results per condition are listed in Table 9.8-7 and are shown in Figure 9.8-8. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.8-8. Scores with result WT are highlighted in red.

Table 9.8-7: results per condition for experiment BS1534-18

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	96.42	8.97	1.37
c02	LP 3.5 kHz	168	14.42	11.44	1.74
c03	Mono EVS, 64 kbps	168	45.88	19.69	3.00
c04	Mono EVS, 128 kbps	168	47.82	21.70	3.31
c05	IVAS, 64 kbps	168	62.60	22.20	3.38
c06	IVAS, 96 kbps	168	75.94	16.48	2.51
c07	IVAS, 128 kbps	168	80.87	16.74	2.55
c08	IVAS, 256 kbps	168	89.73	12.01	1.83

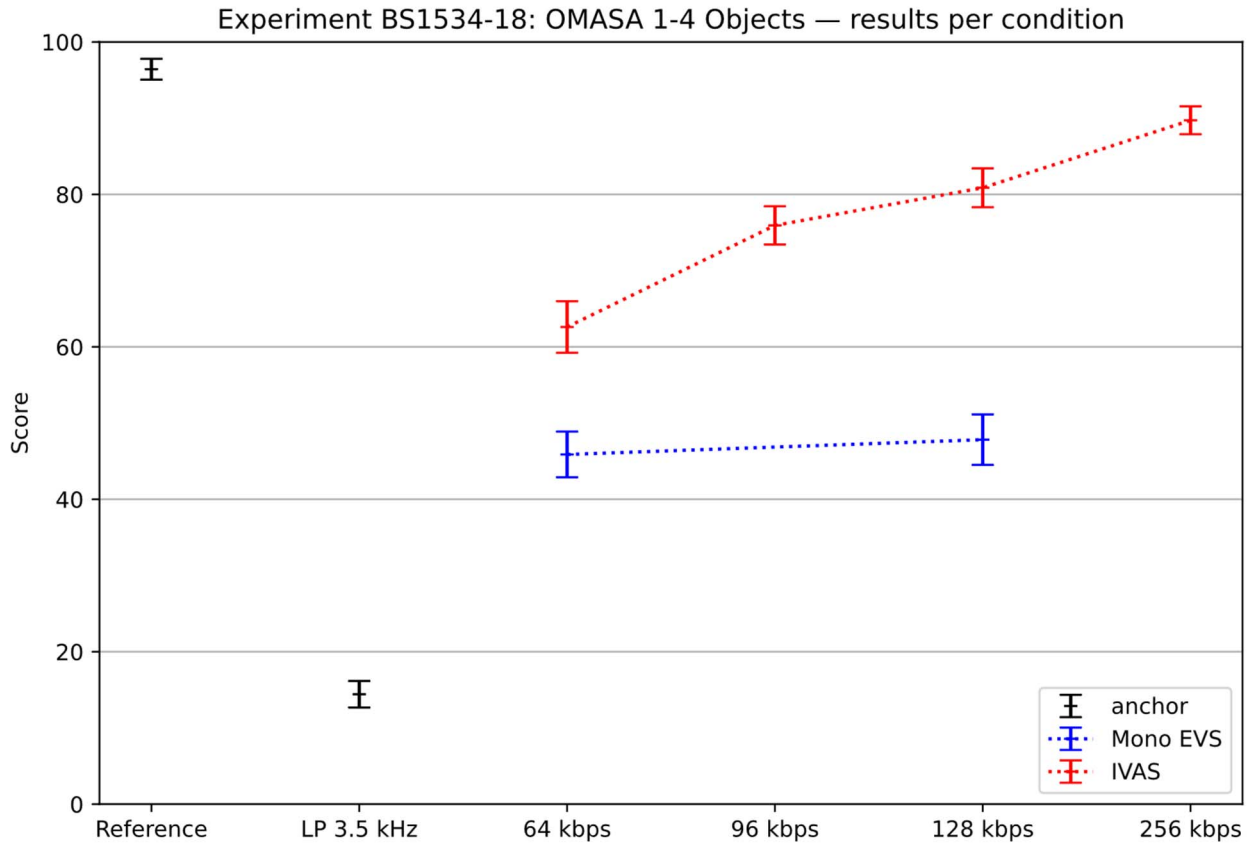


Figure 9.8-8: results per condition for experiment BS1534-18

Table 9.8-8: statistical significance analysis for experiment BS1534-18

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	62.60	22.20	c03	45.88	19.69	16.72	10.353	<0.001	BT
c07	80.87	16.74	c04	47.82	21.70	33.05	16.421	<0.001	BT

9.8.7 Conclusions on IVAS Performance for Combined Objects and MASA (OMASA)

Taking the results of the previous clauses in 9.7 into account, the following conclusions can be drawn for IVAS Combined Objects and MASA (OMASA) modes:

During standardization of the IVAS codec, additional information on the performance of OMASA for 48 - 128 kbps has been provided. For bitrates of 48 and 64 kbps, the combined OMASA modes outperform separate coding of MASA and audio objects significantly. Additional bitrates have been evaluated during characterization testing.

During characterization phase, IVAS OMASA was evaluated against EVS mono in an BS.1534 tests. Across the tested bitrates (64 kbps, 128 kbps), IVAS OMASA coding was significantly outperforming EVS mono at the same bitrate.

Additionally, the performance of IVAS floating-point and fixed-point code incl. cross-operation (FX encoder + FL decoder and vice versa) was evaluated during characterization phase for OMASA. Out of 40 tested operating points, 38 operating points were statistically equivalent.

Overall, the IVAS OMASA coding for maintains a “good” MOS rating (>= 4.0) from bitrates starting from 48 kbps and above. A level 1 IVAS encoder alreads reaches an “excellent” MOS rating of approx. 4.5 for 1-2 objects.

9.9 Stereo operation with EVS compatible mono downmix stream

9.9.1 Overview

Provided as optional additional information provided for codec selection, one experiments has been conducted to evaluate the performance of the IVAS codec for downmixing stereo content to an EVS compatible mono downmix. The experiment was conducted as P.800 DCR test using headphone presentation.

- Optional additional information experiment 4: Speech and music, stereo input, mono output, 13.2 and 24.4 kbps, headphone presentation

In Characterization phase, one additional experiment has been conducted. The experiment BS1534-19 was conducted as a BS.1534 test.

- Characterization Experiment BS1534-19: Generic Audio, stereo input, mono output, 13.2 and 24.4 kbps, headphone presentation

9.9.2 Optional additional information provided for codec selection

The following quality assessment has been provided as optional additional information for codec selection:

For dynamic downmix tool for interoperable EVS the following in house listening test has been conducted. The results are shown in Figure 9.9-1.

- P.800 DCR 48 kHz stereo input
  - Direct, MNRU, SDRU
  - CuT1: average (static downmix) + EVS
  - CuT2: trunk (dynamic downmix) + EVS
  - CuT3: tandem: mono output of IVAS stereo + EVS
- Bit rate
  - 13.2 kbps, 24.4 kbps, downmix only
- Input items:
  - Japanese speech with some music 60 items.
  - Including non-overlap, overlap talkers
  - Including 6 types of speaker and microphone position
- Listeners: 24 (naïve)

According to the results, dynamic downmix could provide better or not worse than static downmix and tandem coding. Note that tandem coding consists of EVS encoding of the mono output of IVAS stereo coding with additional delay.

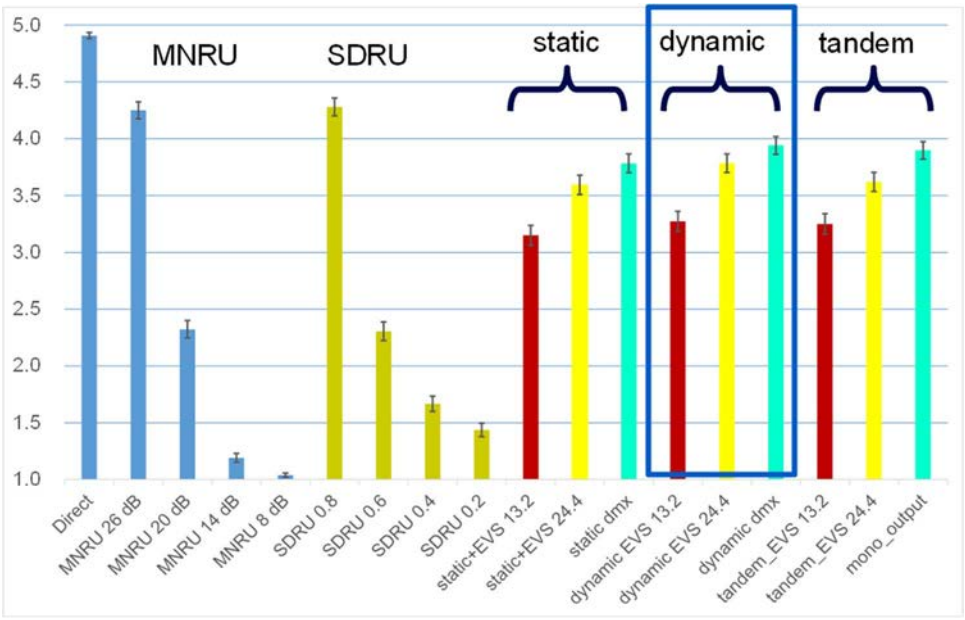


Figure 9.9-1: MOS scores of P.800 DCR test for downmix tools and tandem coding

9.9.3 Characterization Experiment BS1534-19 (Stereo Downmix, Generic Audio, 13.2 and 24.4 kbps, Headphone Presentation)

Characterization Experiment BS1534-19 evaluates the performance of the IVAS fixed-point implementation for EVS compatibe Stereo downmix for generic audio content with headphone presentation. This performance is evaluated against a static downmix followed by EVS (mono) and against 2 x EVS mono followed by a static downmix at the same bitrate. See Annex E.9 for details.

Experiment BS1534-19 was conducted by Orange.

The aggregated results per condition are listed in Table 9.9-1 and are shown in Figure 9.9-2. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.9-2. Scores with result WT are highlighted in red.

Table 9.9-1: results per condition for experiment BS1534-19

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.94	0.85	0.13
c02	LP 3.5 kHz	168	19.21	16.62	2.53
c03	Dual EVS + static dmx, 14.4 kbps	168	25.22	18.74	2.85
c04	Dual EVS + static dmx, 26.4 kbps	168	51.17	20.62	3.14
c05	Static dmx +EVS, 13.2 kbps	168	45.77	22.23	3.38
c06	Static dmx +EVS, 24.4 kbps	168	57.13	20.44	3.11
c07	CuT (dynamic dmx +EVS), 13.2 kbps	168	46.39	22.84	3.48
c08	CuT (dynamic dmx +EVS), 24.4 kbps	168	59.24	20.35	3.10

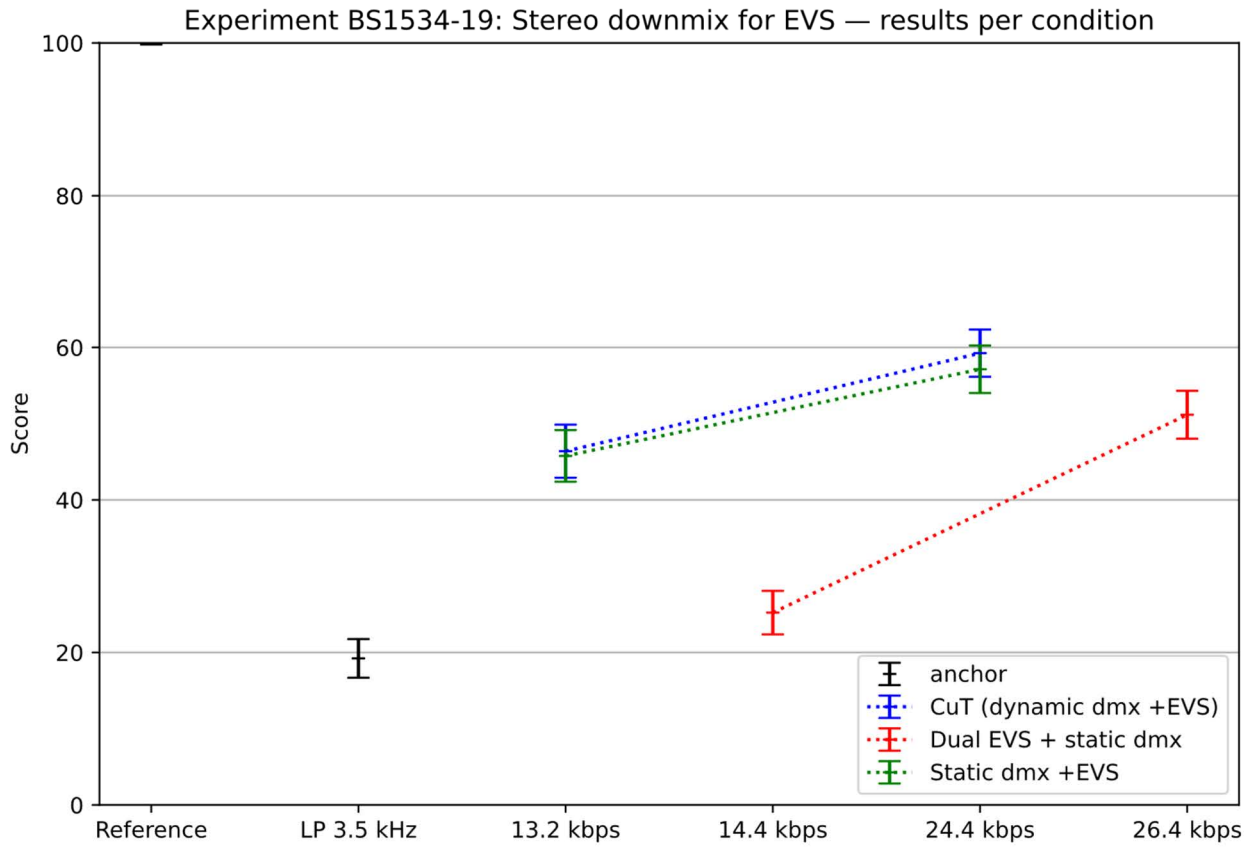


Figure 9.9-2: results per condition for experiment BS1534-19

Table 9.9-2: statistical significance analysis for experiment BS1534-19

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c07	46.39	22.84	c03	25.22	18.74	21.17	16.108	<0.001	BT
c07	46.39	22.84	c05	45.77	22.23	0.62	0.509	0.611	NWT
c08	59.24	20.35	c04	51.17	20.62	8.08	7.448	<0.001	BT
c08	59.24	20.35	c06	57.13	20.44	2.11	2.668	0.008	BT

9.9.4 Conclusions on IVAS Performance for Stereo operation with EVS compatible mono downmix stream

Taking the results of the previous clauses in 9.7 into account, the following conclusions can be drawn for stereo operation with EVS compatible mono downmix stream:

During standardization of the IVAS codec, additional information on the performance for the IVAS stereo operation with EVS compatible mono downmix stream was provided. At the tested bitrates of 13.2 and 24.4 kbps, dynamic downmix could provide better or not worse than static downmix and tandem coding.

During characterization phase, an additional test has been performed, comparing the performance for the IVAS stereo operation with EVS compatible mono downmix stream against a) a static downmix before coding with a subsequent EVS mono coding and b) a dual-mono EVS coding with a subsequent static downmix. It was found that the EVS compatible mono downmix stream outperforms a) at 24.4 kbps and b) at both tested bitrates, 13.2 and 24.4 kbps.

9.10 Rendering

9.10.1 Overview

In Characterization phase, five experiments have been conducted to evaluate the performance of the rendering functionality of the IVAS codec for various content. Experimemnt BS1534-20 was conducted using BS.1534 test methodology, whereas the experiments ROOM-1 – ROOM-4 were conducted using a dedicated preference test methodology for room acoustics.

- Characterization Experiment BS1534-20: Generic Audio, 4 objects, 64 and 512 kbps, evaluating 6-DoF and directivity rendering, headphone presentation
- Characteriaztion Experiment ROOM-1: Generic Audio, FOA, 80 kbps, evaluating binaural rendering with and without room acoustics, headphone presentation
- Characteriaztion Experiment ROOM-2: Generic Audio, FOA, 96 kbps, evaluating binaural rendering with and without room acoustics, headphone presentation
- Characteriaztion Experiment ROOM-3: Generic Audio, multi-channel 5.1, 96 kbps, evaluating binaural rendering with and without room acoustics, headphone presentation
- Characteriaztion Experiment ROOM-4: Generic Audio, multi-channel 5.1+2, 128 kbps, evaluating binaural rendering with and without room acoustics, headphone presentation

9.10.2 Characterization Experiment BS1534-20 (4 Objects, Generic Audio, 64 and 512 kbps, 6-DoF and Directivity Rendering, Headphone Presentation)

Characterization Experiment BS1534-20 evaluates the performance of the IVAS fixed-point implementation for 4 objects of generic audio content, testing extended object metadata which allows for 6-DoF and directivity rendering with headphone presentation. This performance is evaluated against IVAS 4 objects without extended object metadata at the same bitrate. See Annex E.9 for details.

Experiment BS1534-20 was conducted by Ericsson LM.

The aggregated results per condition are listed in Table 9.10-1 and are shown in Figure 9.10-1. The plot shows means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The statistical significance analysis is shown in Table 9.10-2. Scores with result WT are highlighted in red.

Table 9.10-1: results per condition for experiment BS1534-20

Designator	Condition	COUNT	Score	STD	CI95
c01	Reference	168	99.64	1.46	0.22
c02	LP 3.5 kHz	168	25.48	13.37	2.04
c03	IVAS, no ext. metadata, 64 kbps	168	54.60	20.27	3.09
c04	IVAS, no ext. metadata, 512 kbps	168	56.83	18.42	2.81
c05	IVAS, extended metadata, 64 kbps	168	91.61	11.19	1.71
c06	IVAS, extended metadata, 512 kbps	168	99.13	3.51	0.53

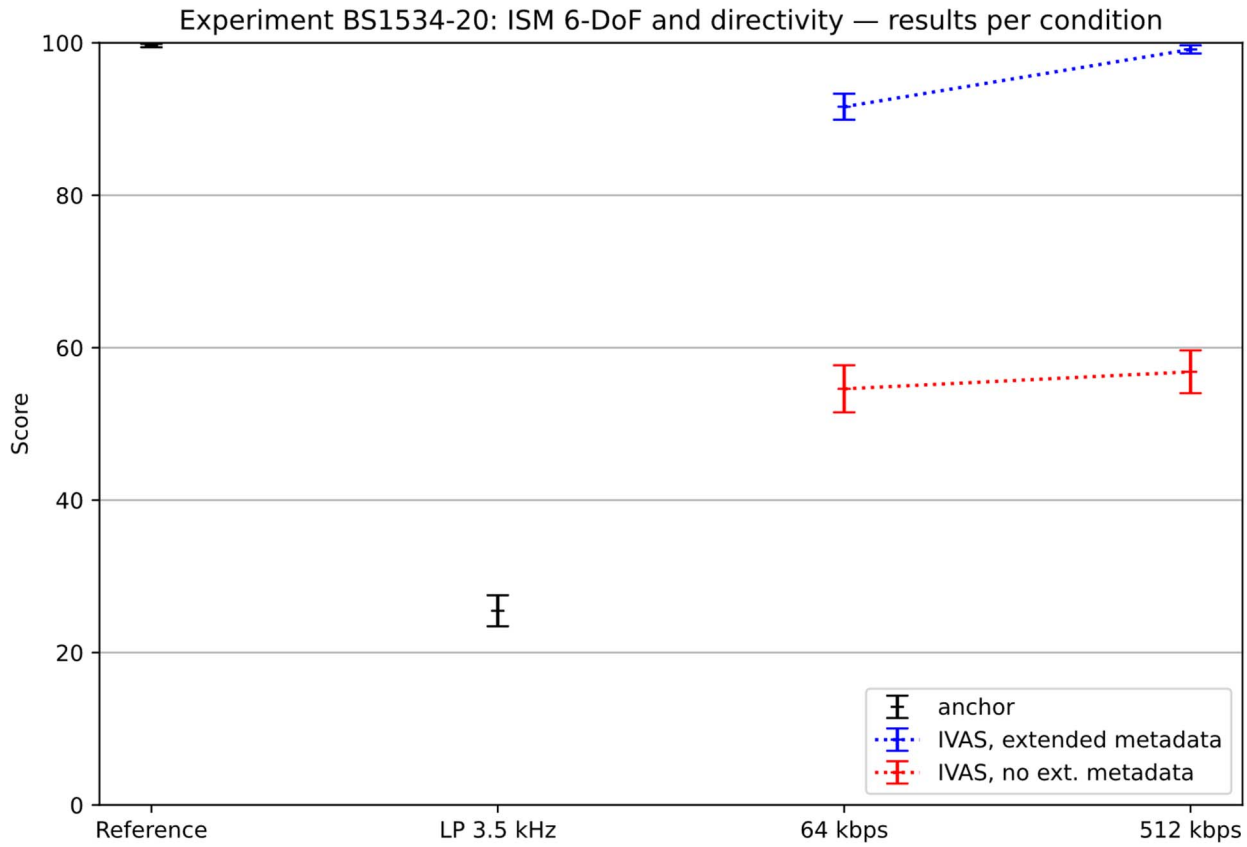


Figure 9.10-1: results per condition for experiment BS1534-20

Table 9.10-2: statistical significance analysis for experiment BS1534-20

CuT			Reference			Evaluation			
Condition	Score	STD	Condition	Score	STD	ΔScore	T-Stat	p-value	Result
c05	91.61	11.19	c03	54.60	20.27	37.01	23.490	<0.001	BT
c06	99.13	3.51	c04	56.83	18.42	42.30	28.701	<0.001	BT

9.10.3 Characterization Experiment ROOM-1 (FOA, Generic Audio, 80 kbps, Room Acoustic Rendering, Headphone Presentation)

Characterization Experiment ROOM-1 evaluates the performance of the three different IVAS binaural rendering modes againt a Python reference renderer. More precisely, the following conditions are evaluated against each other: IVAS rendering with (anechoic) HRTF (reference), IVAS rendering with BRIR, IVAS rendering with synthetic reverb, rendering with BRIR (Python implementation). Experiment ROOM-1 performs this evaluation for the Parametric binaural renderer. See Annex E.9 for details.

Experiment ROOM-1 was conducted by Philips International B.V.

The aggregated results per condition are listed in Table 9.10-3 and shown in Figure 9.10-2. The results separated by category and separated by sample are shown in Figure 9.10-3 left and right, respectively. All boxplots illustrate the distribution of votes by means of average ("x" in black), median (horizontal lines in red), inter-quartile-range (IQR, lower/upper box limit) and outliers (indicated by whiskers at ±1.5x IQR).

The ANOVA and Tukey's HSD analyses are provided in Table 9.10-4 and Table 9.10-5, respectively.

Table 9.10-3: results per condition for experiment ROOM-1

Designator	Condition	COUNT	SCORE	STD	CI95
c01	IVAS Binaural (default)	192	27.00	21.81	3.10
c02	Binaural Reference Renderer (RoomIR)	192	81.02	17.47	2.49
c03	IVAS Binaural (RoomIR)	192	77.81	18.22	2.59
c04	IVAS Binaural (RoomReverb)	192	85.84	16.32	2.32

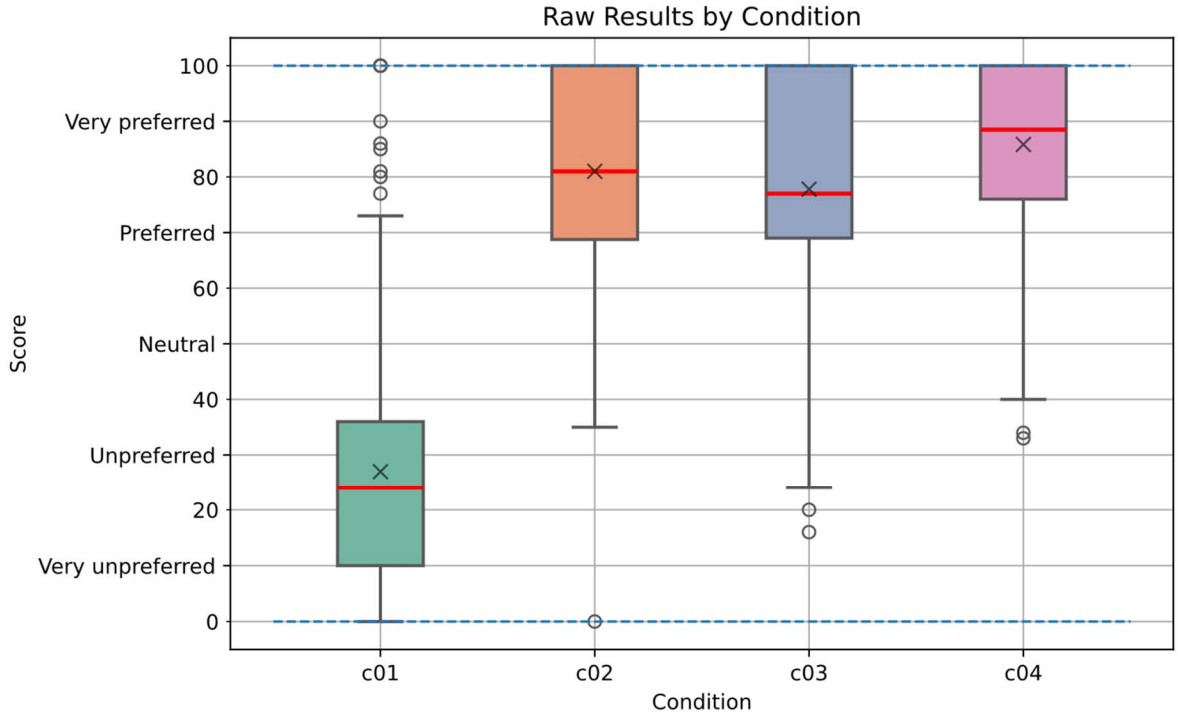


Figure 9.10-2: results per condition for experiment ROOM-1

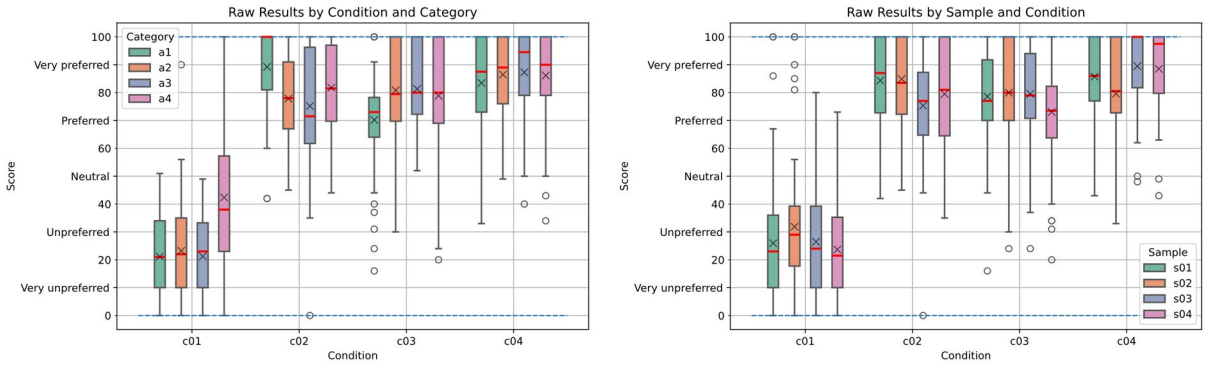


Figure 9.10-3: results per condition versus category (left) and sample (right) for experiment ROOM-1

Table 9.10-4: ANOVA results for experiment ROOM-1

Category Variables	Sum of Squares	Degrees of Freedom	F-Value	P-Value
C(Sample)	941.0	3	1.0100	0.3869
C(Condition)	434974.2	3	467.6100	<0.0001
C(Category)	5045.3	3	5.4200	0.0011
C(Category):C(Condition)	19928.6	9	7.1400	<0.0001

C(Sample):C(Condition)	8038.8	9	2.8800	0.0024
Residual	229452.8	740		

Table 9.10-5: Tukey's HSD results for experiment ROOM-1

Group 1	Group 2	Mean Difference	P-Value	Lower CI	Upper CI	Significant
c01	c02	54.02	<0.0001	49.14	58.90	True
c01	c03	50.82	<0.0001	45.94	55.70	True
c01	c04	58.85	<0.0001	53.97	63.73	True
c02	c03	-3.20	0.3300	-8.08	1.68	False
c02	c04	4.83	0.0540	-0.05	9.71	False
c03	c04	8.03	<0.0001	3.15	12.91	True

9.10.4 Characterization Experiment ROOM-2 (FOA, Generic Audio, 96 kbps, Room Acoustic Rendering, Headphone Presentation)

Characterization Experiment ROOM-2 evaluates the performance of the three different IVAS binaural rendering modes againt a Python reference renderer. More precisely, the following conditions are evaluated against each other: IVAS rendering with (anechoic) HRTF (reference), IVAS rendering with BRIR, IVAS rendering with synthetic reverb, rendering with BRIR (Python implementation). Experiment ROOM-2 performs this evaluation for the FastConv binaural renderer. See Annex E.9 for details.

Experiment ROOM-2 was conducted by Philips International B.V.

The aggregated results per condition are listed in Table 9.10-6 and shown in Figure 9.10-4. The results separated by category and separated by sample are shown in Figure 9.10-5 left andright, respectively. All boxplots illustrate the distribution of votes by means of average ("x" in black), median (horizontal lines in red), inter-quartile-range (IQR, lower/upper box limit) and outliers (indicated by whiskers at ±1.5x IQR).

The ANOVA and Tukey's HSD analyses are provided in Table 9.10-7 and Table 9.10-8, respectively.

Table 9.10-6: results per condition for experiment ROOM-2

Designator	Condition	COUNT	SCORE	STD	CI95
c01	IVAS Binaural (default)	192	25.86	22.39	3.19
c02	Binaural Reference Renderer (RoomIR)	192	81.28	19.52	2.78
c03	IVAS Binaural (RoomIR)	192	80.08	17.45	2.48
c04	IVAS Binaural (RoomReverb)	192	83.92	21.96	3.13

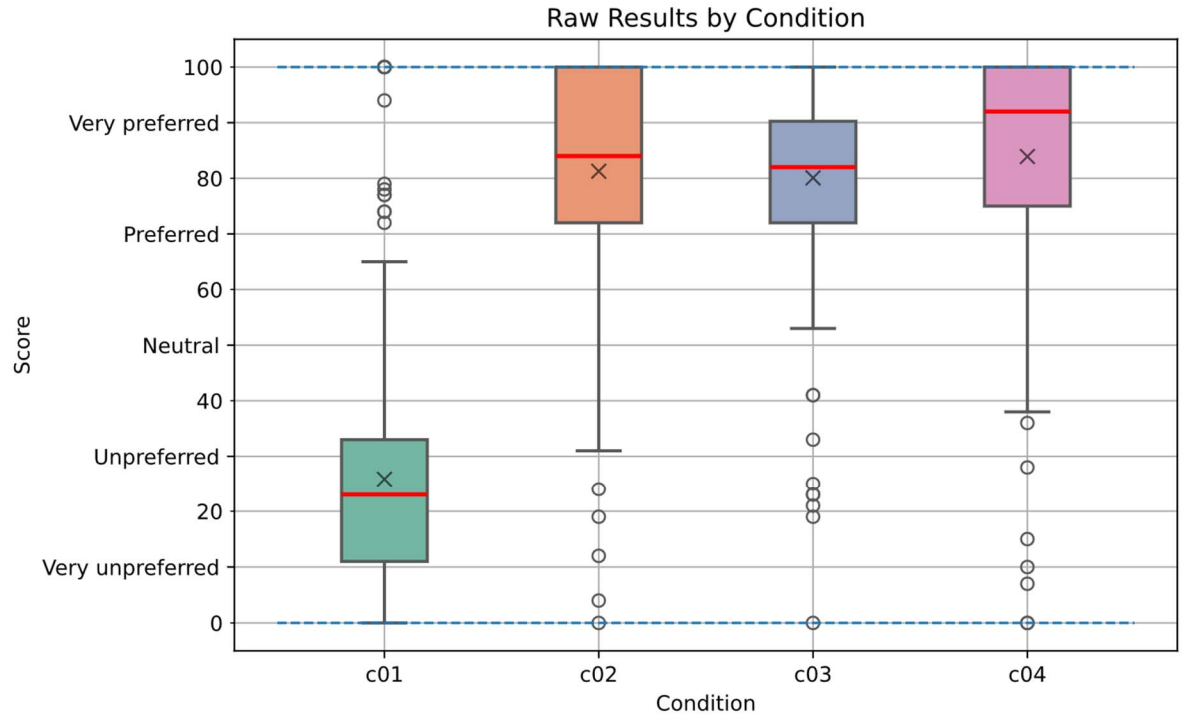


Figure 9.10-4: results per condition for experiment ROOM-2

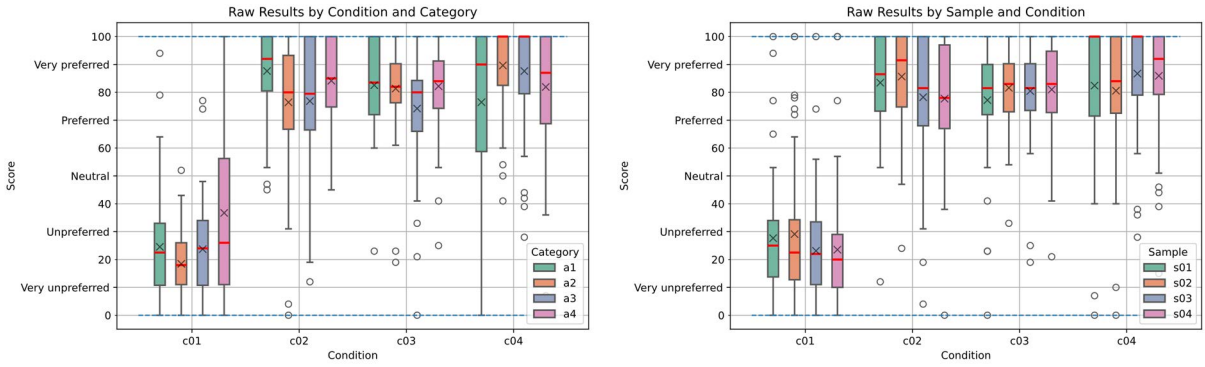


Figure 9.10-5: results per condition versus category (left) and sample (right) for experiment ROOM-2

Table 9.10-7: ANOVA results for experiment ROOM-2

Category Variables	Sum of Squares	Degrees of Freedom	F-Value	P-Value
C(Sample)	594.2	3	0.5000	0.6824
C(Condition)	451330.8	3	379.7900	<0.0001
C(Category)	3564.9	3	3.0000	0.0300
C(Category):C(Condition)	16800.1	9	4.7100	<0.0001
C(Sample):C(Condition)	4671.8	9	1.3100	0.2274
Residual	293130.3	740		

Table 9.10-8: Tukey's HSD results for experiment ROOM-2

Group 1	Group 2	Mean Difference	P-Value	Lower CI	Upper CI	Significant
c01	c02	55.41	<0.0001	50.04	60.78	True
c01	c03	54.21	<0.0001	48.85	59.58	True

c01	c04	58.05	<0.0001	52.68	63.42	True
c02	c03	-1.20	0.9400	-6.57	4.17	False
c02	c04	2.64	0.5840	-2.73	8.01	False
c03	c04	3.84	0.2550	-1.53	9.21	False

9.10.5 Characterization Experiment ROOM-3 (MC 5.1, Generic Audio, 96 kbps, Room Acoustic Rendering, Headphone Presentation)

Characterization Experiment ROOM-3 evaluates the performance of the three different IVAS binaural rendering modes againt a Python reference renderer. More precisely, the following conditions are evaluated against each other: IVAS rendering with (anechoic) HRTF (reference), IVAS rendering with BRIR, IVAS rendering with synthetic reverb, rendering with BRIR (Python implementation). Experiment ROOM-3 performs this evaluation for the Time-domain object renderer / Crend binaural renderer. See Annex E.9 for details.

Experiment ROOM-3 was conducted by Philips International B.V.

The aggregated results per condition are listed in Table 9.10-9 and shown in Figure 9.10-6. The results separated by category and separated by sample are shown in Figure 9.10-7 left and right, respectively. All boxplots illustrate the distribution of votes by means of average ("x" in black), median (horizontal lines in red), inter-quartile-range (IQR, lower/upper box limit) and outliers (indicated by whiskers at ±1.5x IQR).

The ANOVA and Tukey's HSD analyses are provided in Table 9.10-10 and Table 9.10-11, respectively.

Table 9.10-9: results per condition for experiment ROOM-3

Designator	Condition	COUNT	SCORE	STD	CI95
c01	IVAS Binaural (default)	208	26.43	19.96	2.73
c02	Binaural Reference Renderer (RoomIR)	208	85.01	13.48	1.84
c03	IVAS Binaural (RoomIR)	208	84.71	14.11	1.93
c04	IVAS Binaural (RoomReverb)	208	87.62	17.19	2.35

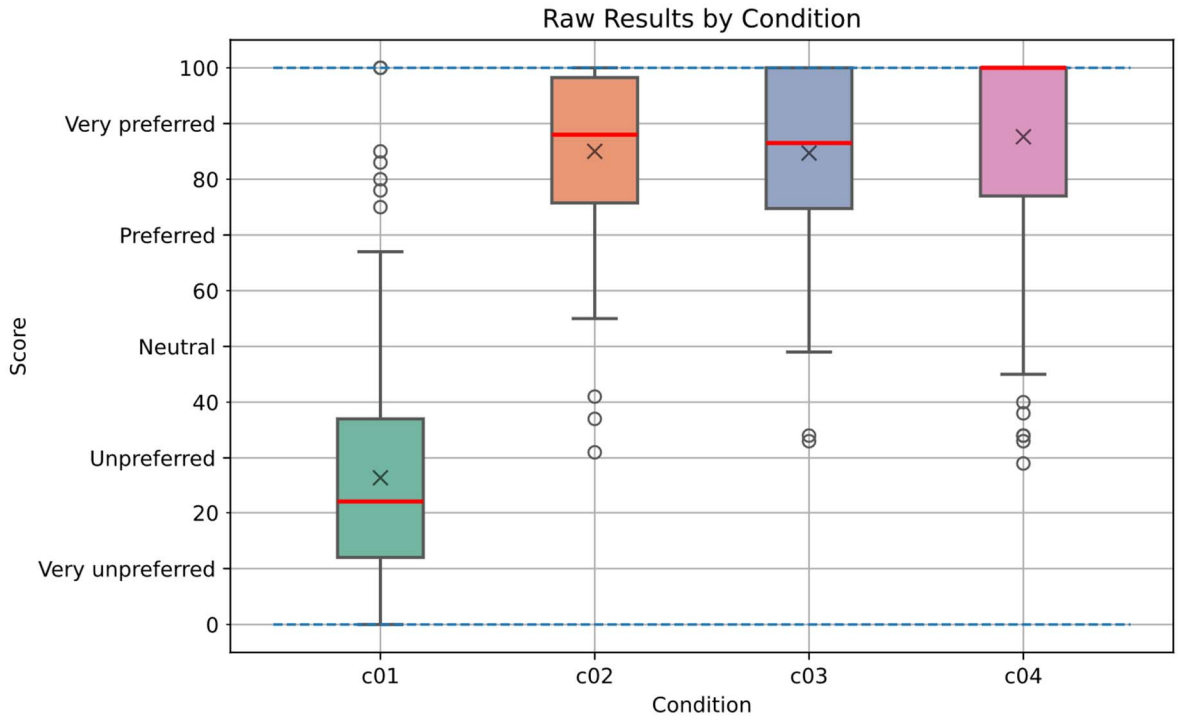


Figure 9.10-6: results per condition for experiment ROOM-3

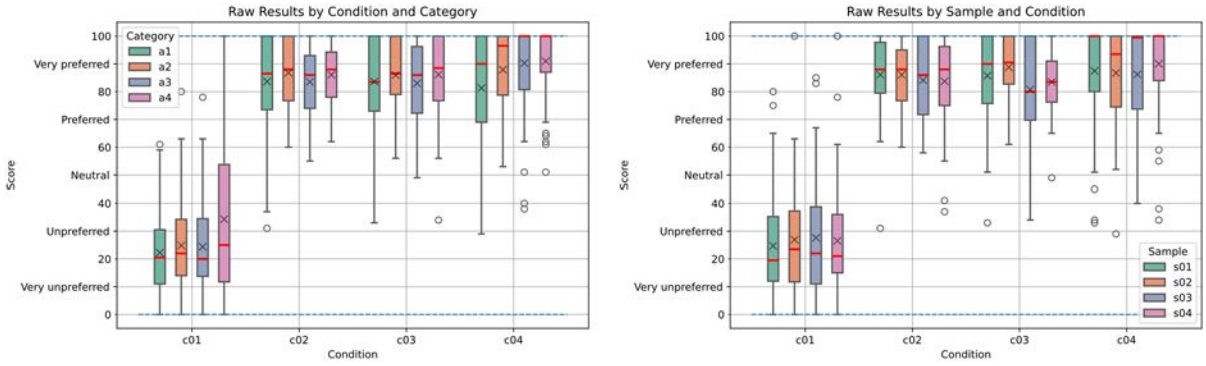


Figure 9.10-7: results per condition versus category (left) and sample (right) for experiment ROOM-3

Table 9.10-10: ANOVA results for experiment ROOM-3

Category Variables	Sum of Squares	Degrees of Freedom	F-Value	P-Value
C(Sample)	627.0	3	0.7900	0.4969
C(Condition)	550555.6	3	697.9900	<0.0001
C(Category)	4762.4	3	6.0400	0.0005
C(Category):C(Condition)	3602.0	9	1.5200	0.1356
C(Sample):C(Condition)	2125.4	9	0.9000	0.5263
Residual	211391.0	804		

Table 9.10-11: Tukey's HSD results for experiment ROOM-3

Group 1	Group 2	Mean Difference	P-Value	Lower CI	Upper CI	Significant
c01	c02	58.58	<0.0001	54.44	62.72	True
c01	c03	58.27	<0.0001	54.14	62.41	True
c01	c04	61.19	<0.0001	57.05	65.33	True
c02	c03	-0.31	0.9980	-4.45	3.83	False
c02	c04	2.61	0.3660	-1.53	6.75	False
c03	c04	2.92	0.2670	-1.22	7.06	False

9.10.6 Characterization Experiment ROOM-4 (MC 5.1+2, Generic Audio, 128 kbps, Room Acoustic Rendering, Headphone Presentation)

Characterization Experiment ROOM-4 evaluates the performance of the three different IVAS binaural rendering modes againt a Python reference renderer. More precisely, the following conditions are evaluated against each other: IVAS rendering with (anechoic) HRTF (reference), IVAS rendering with BRIR, IVAS rendering with synthetic reverb, rendering with BRIR (Python implementation). Experiment ROOM-4 performs this evaluation for the Crend binaural renderer. See Annex E.9 for details.

Experiment ROOM-4 was conducted by Philips International B.V.

The aggregated results per condition are listed in Table 9.10-12 and shown in Figure 9.10-8. The results separated by category and separated by sample are shown in Figure 9.10-9 left and right, respectively. All boxplots illustrate the distribution of votes by means of average ("x" in black), median (horizontal lines in red), inter-quartile-range (IQR, lower/upper box limit) and outliers (indicated by whiskers at ±1.5x IQR).

The ANOVA and Tukey's HSD analyses are provided in Table 9.10-13 and Table 9.10-14, respectively.

Table 9.10-12: results per condition for experiment ROOM-4

Designator	Condition	COUNT	SCORE	STD	CI95
c01	IVAS Binaural (default)	192	24.86	22.54	3.21
c02	Binaural Reference Renderer (RoomIR)	192	83.49	13.34	1.90
c03	IVAS Binaural (RoomIR)	192	82.25	15.83	2.25
c04	IVAS Binaural (RoomReverb)	192	87.61	14.77	2.10

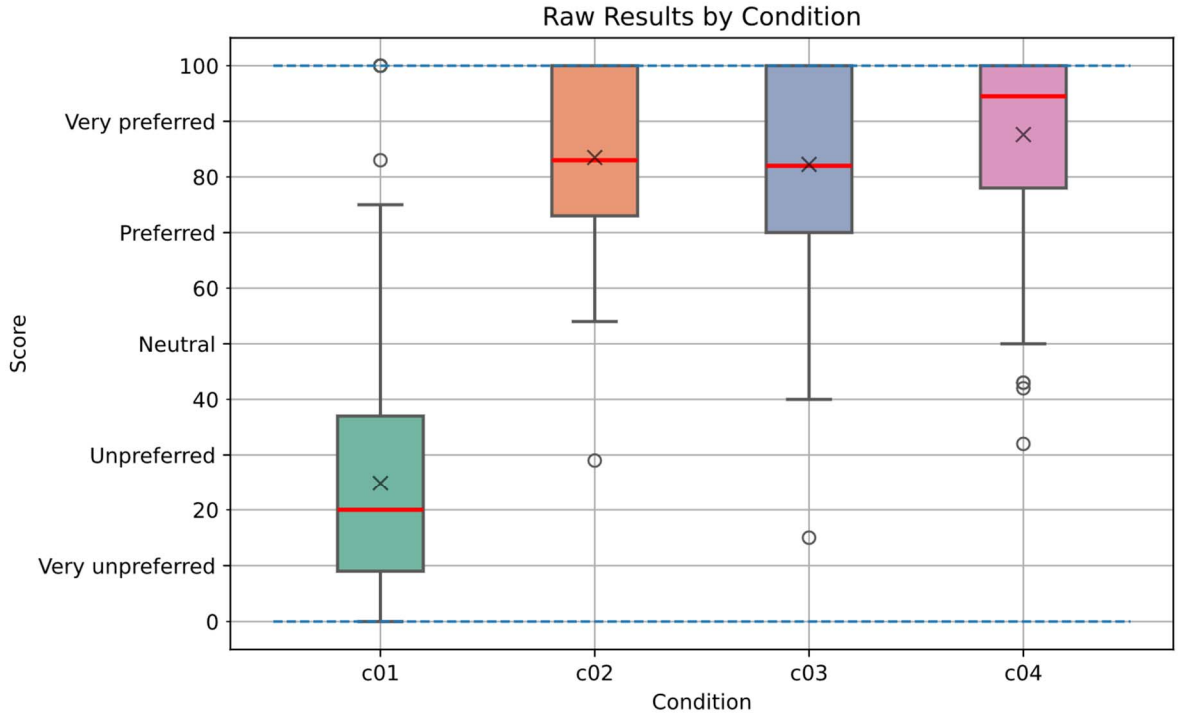


Figure 9.10-8: results per condition for experiment ROOM-4

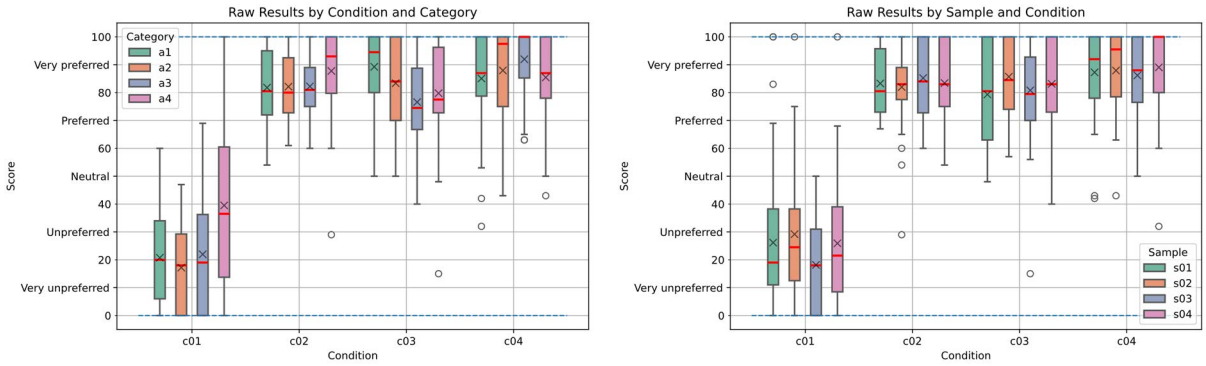


Figure 9.10-9: results per condition versus category (left) and sample (right) for experiment ROOM-4

Table 9.10-13: ANOVA results for experiment ROOM-4

Category Variables	Sum of Squares	Degrees of Freedom	F-Value	P-Value
C(Sample)	1452.7	3	1.8400	0.1381
C(Condition)	514428.5	3	652.3300	<0.0001
C(Category)	3513.4	3	4.4600	0.0041
C(Category):C(Condition)	17723.9	9	7.4900	<0.0001

C(Sample):C(Condition)	3322.1	9	1.4000	0.1820
Residual	194523.1	740		

Table 9.10-14: Tukey's HSD results for experiment ROOM-4

Group 1	Group 2	Mean Difference	P-Value	Lower CI	Upper CI	Significant
c01	c02	58.63	<0.0001	54.17	63.09	True
c01	c03	57.40	<0.0001	52.93	61.86	True
c01	c04	62.76	<0.0001	58.29	67.22	True
c02	c03	-1.23	0.8920	-5.70	3.23	False
c02	c04	4.12	0.0820	-0.34	8.59	False
c03	c04	5.36	0.0110	0.90	9.82	True

9.10.7 Conclusions on IVAS Performance for Rendering

Taking the results of the previous clauses in 9.10 into account, the following conclusions can be drawn for the IVAS rendering operations:

During characterization, the performance of the 6-DoF and directivity for object-based audio (ISMs) was evaluated at 64 and 256 kbps in a BS.1534 test. At both tested bitrates, a significant improvement over 3-DoF was found. However, it should be noted that the reference was 6-DoF, meaning that the result should not be read as if the 3-DoF rendering is inferior to 6-DoF rendering in general.

Furthermore, rendering using room acoustics was evaluated against an anechoic reference in a series of preference tests. The subjects preferred room acoustics over the anechoic reference, and there was no significant difference found between the IVAS room acoustic implementations (BRIR, synthetic room reverb) and a Python-based reference implementation. Users/implementers may select a room acoustics rendering mode without compromising on rendering quality based on their preferences, availability of room acoustics data in each format, availability of CPU resources, etc.

Additionally, for the selection experiments, pass-through operation was tested, with a Python based external reference rendering, used to present conditions to binaural or loudspeaker playback configurations. For characterization, the native IVAS rendering was used. The fact that the overall performance at all operating points meets the expected quality, indicates that the native IVAS rendering is equivalent in quality to the python external rendering.

9.11 Split Rendering

9.11.1 General

An IVAS specific split rendering solution was added to IVAS as part of the Workitem “Immersive Audio for Split Rendering Scenarios” (ISAR). A detailed performance characterization of ISAR can be found in TR 26.996 (“Immersive Audio for Split Rendering Scenarios; Performance characterization”) [20]. In order for this document to contain a full characterization of all IVAS immersive audio features, the test results of the IVAS specific solutionare replicated here. For details on test setup, please see TR 26.996.

9.11.2 Overview

In the ISAR Selection phase, four experiments have been conducted to evaluate the performance of the IVAS specific ISAR solution. Table 9.11-1shows a high-level overview of the experiments. Each experiment was carried out twice (in experiments a and b), once by the solution proponent and once by a cross-checker (XC).

Table 9.11-1: High-level overview of ISAR selection experiments

Exp	Input format	Source material	Listening environment	Bitrates (kbps)	Listening Lab
ISAR_BS1534-1a ISAR_BS1534-1b	SBA (HOA3)	Generic Audio	Headphones	IVAS: 512, CuT: 768	Dolby Qualcomm (XC)
ISAR_BS1534-2a ISAR_BS1534-2b	Multi-channel 7.1+4	Generic Audio	Headphones	IVAS: 512, CuT: 768	Fraunhofer Ittiam (XC)
ISAR_BS1534-3a ISAR_BS1534-3b	Objects (ISM-4)	Generic Audio	Headphones	IVAS: 512, CuT: 768	Fraunhofer Nokia (XC)
ISAR_BS1534-4a ISAR_BS1534-4b	MASA (2 TC)	Generic Audio	Headphones	IVAS: 512, CuT: 768	Dolby Bytedance (XC)

In addition to the selection tests evaluating the performance of the single ISAR candidate solution for IVAS described above, analogous experiments were carried out with the selected ISAR solution operated at 384 and 512 kbps. These experiments were provided by a company contribution and are summarized in Table 9.11-1a.

Table 9.11-1a: High-level overview of ISAR low-rate experiments

Exp	Input format	Source material	Listening environment	Bitrates kbps	Listening Lab
ISAR_BS1534-1-LR	SBA (HOA3)	Generic Audio	Headphones	IVAS: 512, CuT: 384 IVAS: 512, CuT: 512	Dolby
ISAR_BS1534-2-LR	Multi-channel 7.1+4	Generic Audio	Headphones	IVAS: 512, CuT: 384 IVAS: 512, CuT: 512	Dolby
ISAR_BS1534-3-LR	Objects (ISM-4)	Generic Audio	Headphones	IVAS: 512, CuT: 384 IVAS: 512, CuT: 512	Dolby
ISAR_BS1534-4-LR	MASA (2 TC)	Generic Audio	Headphones	IVAS: 512, CuT: 384 IVAS: 512, CuT: 512	Dolby

9.11.3 ISAR Selection Phase Requirements and Objectives

All experiments check the same requirements defined in TR 26.865 [2], namely that the QoE of the ISAR split rendering system is no worse than the 0-DOF native transcoding reference system using the same operation point of the native coding system (IVAS coding at 512 kbps) and best possible operation point for transcoding (IVAS stereo at 256 kbps). The 4 experiments evaluate the requirement for the 4 different main head-trackable IVAS coding formats, i.e., SBA (HOA3), MC 7.1.4, ISM-4 and MASA.

The objective is that QoE provided by split rendering solution should be as close as possible to quality of native coding reference system using same operation point. There is no statistical test to verify if this objective is met. However, a statement will be made based on the observed test scores how close the quality of the tested ISAR split rendering solution for the given immersive audio input format is to the quality of the native coding reference system.

Conclusion of all 8 experiments testing the requirement that the ISAR split rendering solution for IVAS shall be no worse than the 0-DOF transcoding reference system is that this requirement is met across all tested immersive input audio formats. It can generally be observed that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range close to the quality of the native IVAS coding reference system is achieved whereas the 0-DOF transcoding alternative offers substantially lower quality.

9.11.4 ISAR Selection Experiment ISAR_BS1534-1 (HOA3, Generic Audio, 768 kbps Split Rendering Link, Headphone Presentation)

9.11.4.1 Overview

Selection Experiment ISAR_BS1534-1 evaluates Split Rendering for IVAS HOA3 audio at 768 kbps. See Annex E.10 for details.

9.11.4.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (Ref):** Native coding reference system operated at 512 kbps with IVAS decoding and head-tracked rendering at end-device
- **c02 (LP7 anchor):** 7 kHz low-pass filtered native coding reference system
- **c03 (0-DOF):** 0-DOF native transcoding reference system with 512 kbps coded IVAS content, decoded and binaurally rendered to (outdated) pose available at pre-renderer, subsequently transcoded to IVAS stereo at 256 kbps
- **c04 (CuT):** Codec under Test, i.e., ISAR split rendering system operated at the required bit rate of 768 kbps for the intermediate immersive audio representation

9.11.4.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the two ISAR_BS1534-1 experiments.

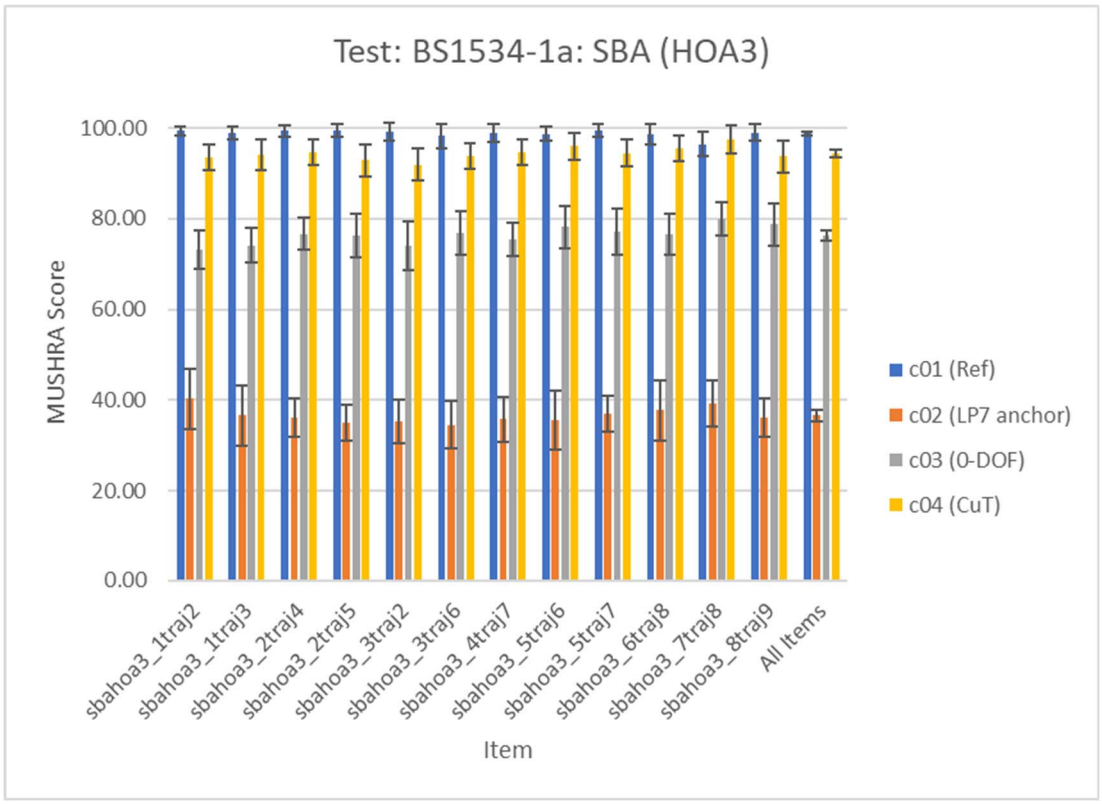


Figure 9.11-1: Results of ISAR_BS1534-1a test for SBA input audio

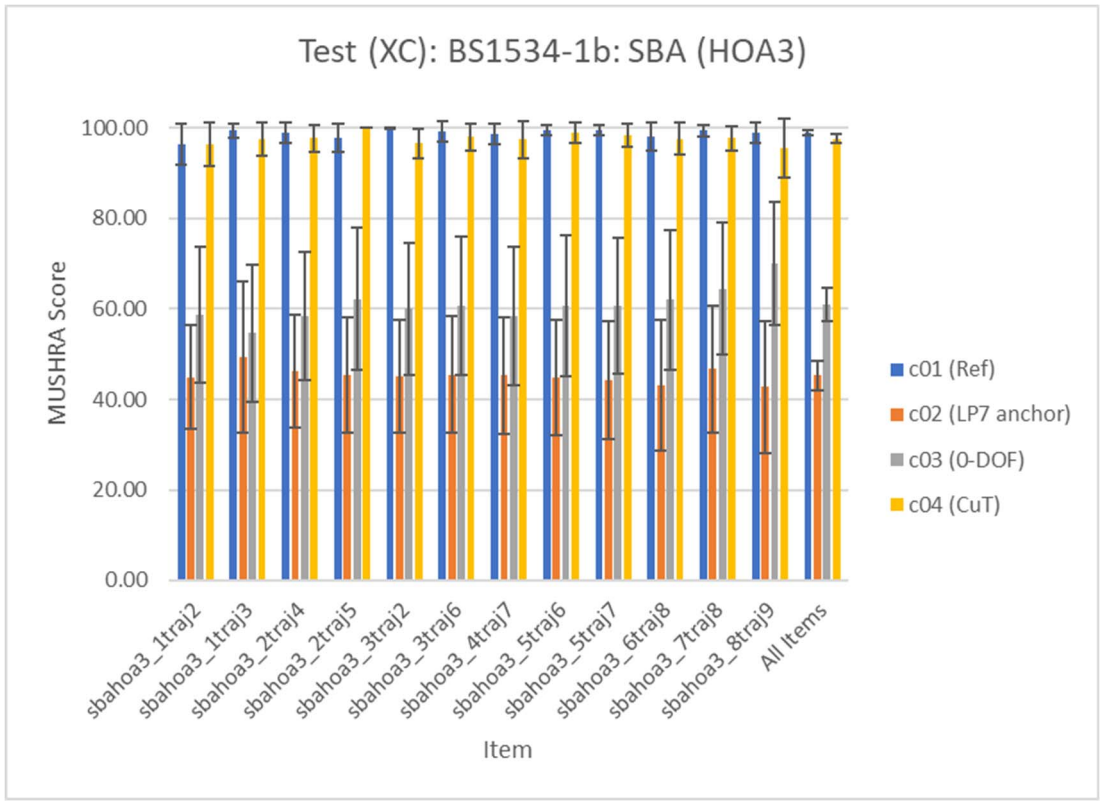


Figure 9.11-2: Results of ISAR_BS1534-1b test for SBA input audio

9.11.4.4 Statistical analysis

Table 9.11-2: Result of statistical analysis of ISAR_BS1534-1a test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-17.9917	5.7941	0.5289	-34.0156	1.0000	Pass

Table 9.11-3: Result of statistical analysis of ISAR_BS1534-1b test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-36.6583	20.3336	1.8562	-19.7492	1.0000	Pass

9.11.4.5 Experimental conclusions

Conclusion of both experiments is that the ISAR split rendering solution for SBA input meets the requirement to be no worse than the 0-DOF transcoding reference system. The experiments indicate that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range is achieved compared to the 0-DOF transcoding reference providing quality in the ‘good’ range. The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range which indicates only very minor audible differences.

9.11.5 ISAR Selection Experiment ISAR_BS1534-2 (MC 7.1+4, Generic Audio, 768 kbps Split Rendering Link, Headphone Presentation)

9.11.5.1 Overview

Selection Experiment ISAR_BS1534-2 evaluates Split Rendering for IVAS 7.1+4 audio at 768 kbps. See Annex E.10 for details.

9.11.5.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (Ref):** Native coding reference system operated at 512 kbps with IVAS decoding and head-tracked rendering at end-device
- **c02 (LP7 anchor):** 7 kHz low-pass filtered native coding reference system
- **c03 (0-DOF):** 0-DOF native transcoding reference system with 512 kbps coded IVAS content, decoded and binaurally rendered to (outdated) pose available at pre-renderer, subsequently transcoded to IVAS stereo at 256 kbps
- **c04 (CuT):** Codec under Test, i.e., ISAR split rendering system operated at the required bit rate of 768 kbps for the intermediate immersive audio representation

9.11.5.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the two ISAR_BS1534-2 experiments.

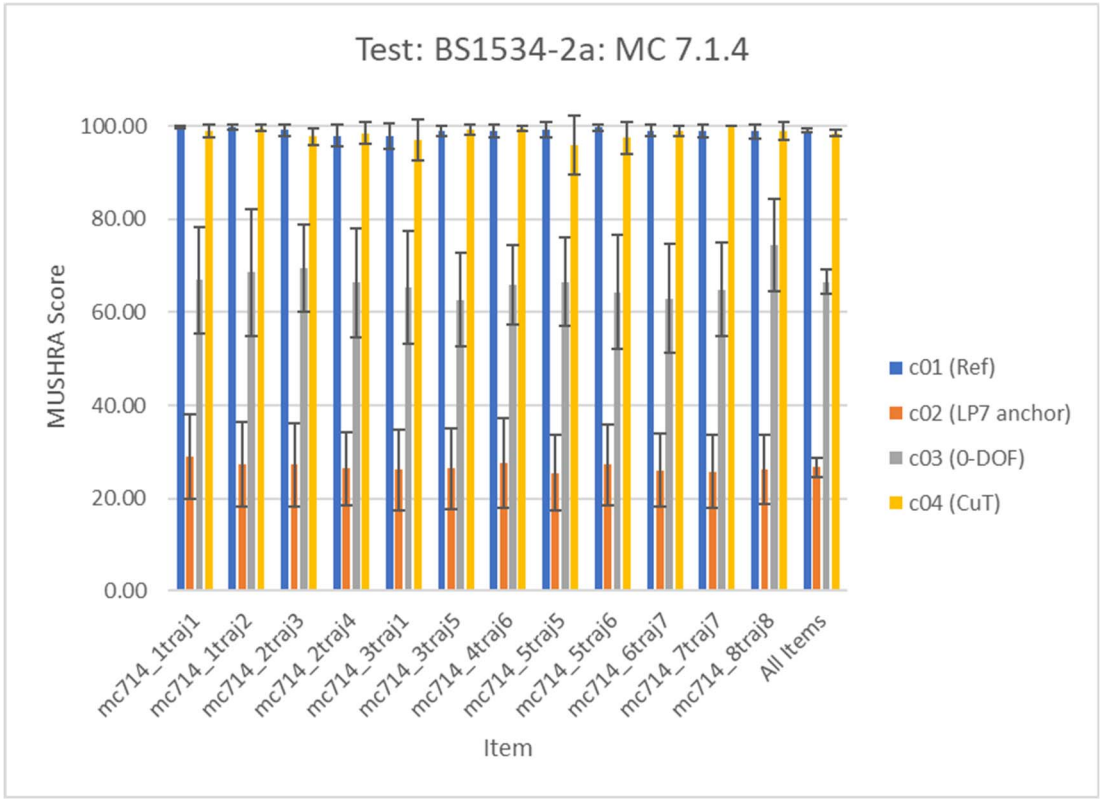


Figure 9.11-3: Results of ISAR_BS1534-2a test for MC 7.1.4 input audio

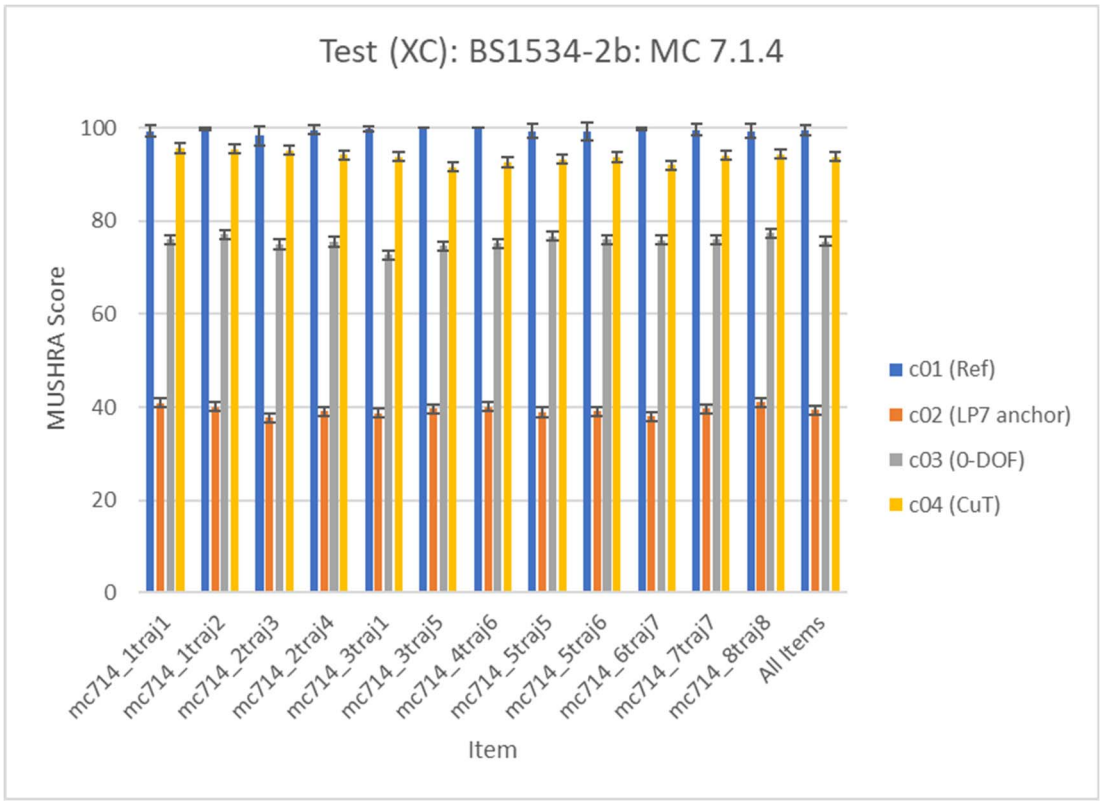


Figure 9.11-4: Results of ISAR_BS1534-2b test for MC 7.1.4 input audio

9.11.5.4 Statistical analysis

Table 9.11-4: Result of statistical analysis of ISAR_BS1534-2a test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-31.9583	15.1321	1.3814	-23.1353	1.0000	Pass

Table 9.11-5: Result of statistical analysis of ISAR_BS1534-2b test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-18.1667	6.0881	0.5558	-32.6879	1.0000	Pass

9.11.5.5 Experimental conclusions

Conclusion of both experiments is that the ISAR split rendering solution for Multi-Channel 7.1+4 input meets the requirement to be no worse than the 0-DOF transcoding reference system. The experiments indicate that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range is achieved compared to the 0-DOF transcoding reference providing quality in the ‘good’ range. The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

9.11.6 ISAR Selection Experiment ISAR_BS1534-3 (4 Objects, Generic Audio, 768 kbps Split Rendering Link, Headphone Presentation)

9.11.6.1 Overview

Selection Experiment ISAR_BS1534-3 evaluates Split Rendering for 4 object audio (ISM4) at 768 kbps. See Annex E.10 for details.

9.11.6.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (Ref):** Native coding reference system operated at 512 kbps with IVAS decoding and head-tracked rendering at end-device
- **c02 (LP7 anchor):** 7 kHz low-pass filtered native coding reference system
- **c03 (0-DOF):** 0-DOF native transcoding reference system with 512 kbps coded IVAS content, decoded and binaurally rendered to (outdated) pose available at pre-renderer, subsequently transcoded to IVAS stereo at 256 kbps
- **c04 (CuT):** Codec under Test, i.e., ISAR split rendering system operated at the required bit rate of 768 kbps for the intermediate immersive audio representation

9.11.6.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the two ISAR_BS1534-3 experiments.

Figure 9.11-5: Results of ISAR_BS1534-3a test for ISM-4 input audio

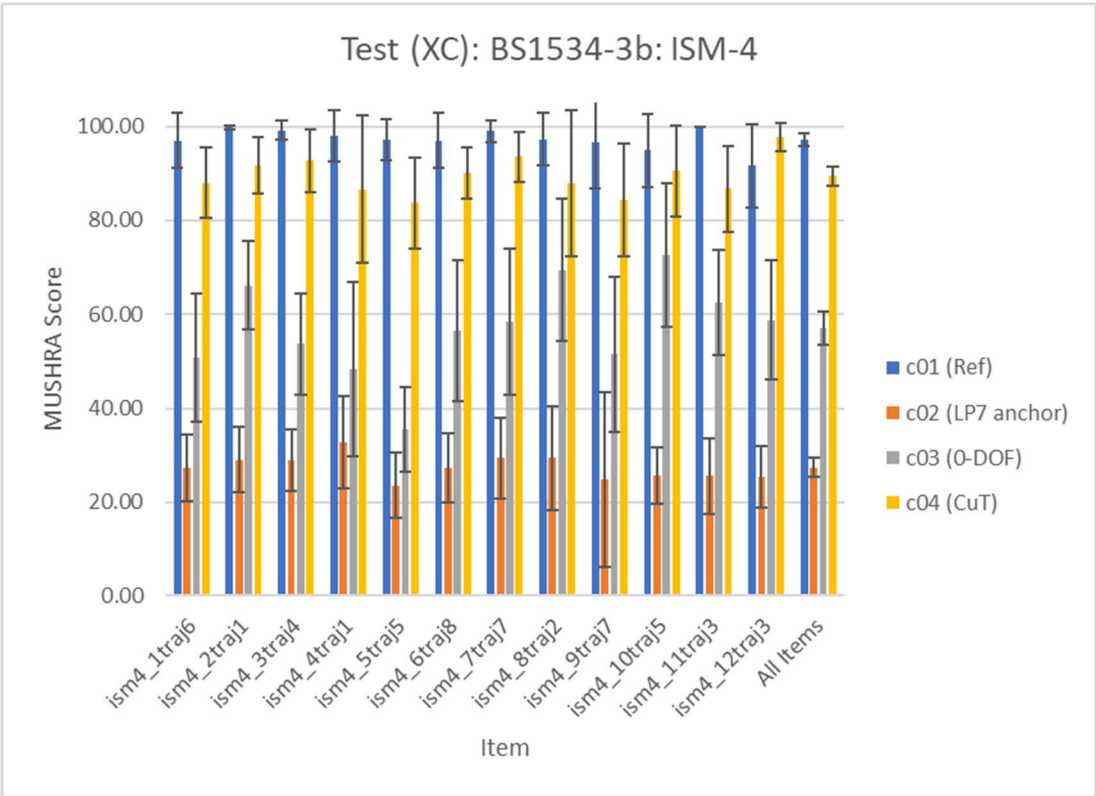


Figure 9.11-6: Results of ISAR_BS1534-3b test for ISM-4 input audio

9.11.6.4 Statistical analysis

Table 9.11-6: Result of statistical analysis of ISAR_BS1534-3a test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-26.6500	14.7310	1.3448	-19.8178	1.0000	Pass

Table 9.11-7: Result of statistical analysis of ISAR_BS1534-3b test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-32.4083	22.0761	2.0153	-16.0814	1.0000	Pass

9.11.6.5 Experimental conclusions

Conclusion of both experiments is that the ISAR split rendering solution for ISM-4 input meets the requirement to be no worse than the 0-DOF transcoding reference system. The experiments indicate that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range is achieved compared to the 0-DOF transcoding reference providing quality in the ‘good’ range. The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

9.11.7 ISAR Selection Experiment ISAR_BS1534-4 (MASA 2TC, Generic Audio, 768 kbps Split Rendering Link, Headphone Presentation)

9.11.7.1 Overview

Selection Experiment ISAR_BS1534-4 evaluates Split Rendering for IVAS MASA audio using 2 transport channels (TC) at 768 kbps. See Annex E.10 for details.

9.11.7.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (Ref):** Native coding reference system operated at 512 kbps with IVAS decoding and head-tracked rendering at end-device
- **c02 (LP7 anchor):** 7 kHz low-pass filtered native coding reference system
- **c03 (0-DOF):** 0-DOF native transcoding reference system with 512 kbps coded IVAS content, decoded and binaurally rendered to (outdated) pose available at pre-renderer, subsequently transcoded to IVAS stereo at 256 kbps
- **c04 (CuT):** Codec under Test, i.e., ISAR split rendering system operated at the required bit rate of 768 kbps for the intermediate immersive audio representation

9.11.7.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the two ISAR_BS1534-4 experiments.

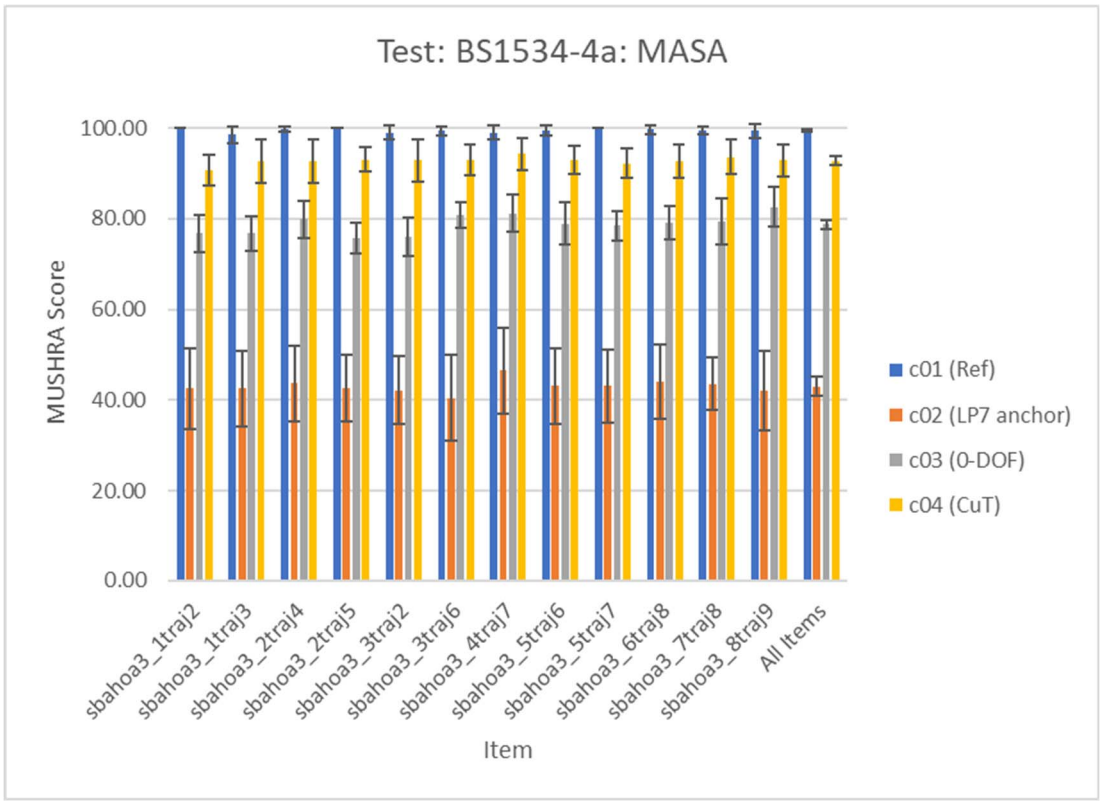


Figure 9.11-7: Results of ISAR_BS1534-4a test for MASA input audio

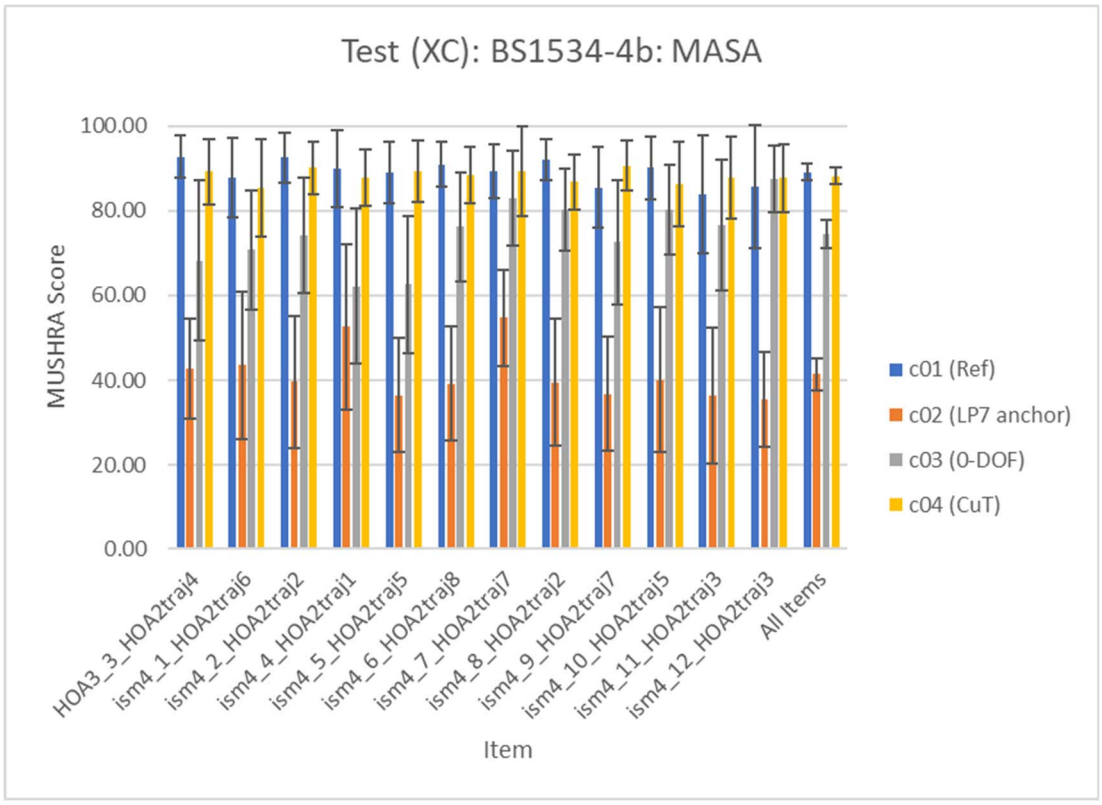


Figure 9.11-8: Results of ISAR_BS1534-4b test for MASA input audio

9.11.7.4 Statistical analysis

Table 9.11-8: Result of statistical analysis of ISAR_BS1534-4a test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-14.0167	5.0493	0.4609	-30.4091	1.0000	Pass

Table 9.11-9: Result of statistical analysis of ISAR_BS1534-4b test checking CuT NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-13.7000	22.1500	2.0220	-6.7754	1.0000	Pass

9.11.7.5 Experimental conclusions

Conclusion of both experiments is that the ISAR split rendering solution for MASA input meets the requirement to be no worse than the 0-DOF transcoding reference system. The experiments indicate that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range is achieved compared to the 0-DOF transcoding reference providing quality in the ‘good’ range. The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

9.11.8 ISAR Experiment ISAR_BS1534-1-LR (HOA3, Generic Audio, 384 and 512 kbps Split Rendering Link, Headphone Presentation)

9.11.8.1 Overview

The experiment ISAR_BS1534-1 evaluates Split Rendering for IVAS HOA3 audio at 384 and 512 kbps.

9.11.8.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (REF):** Hidden reference: Native coding system (IVAS at 512kbps rendered to post renderer pose)
- **c02 (LP7):** LP7 anchor: Hidden reference, 7 kHz LP filtered
- **c03 (0-DOF):** 0-DOF native transcoding reference (IVAS at 512kbps binaurally rendered to pre-renderer pose, IVAS stereo coded@256kbps)
- **c04 (CuT1):** 3-DOF system 1 under test (IVAS at512 kbps split-rendered with ISAR operating at 512 kbps)
- **c05 (CuT2):** 3-DOF system 1 under test (IVAS at 512 kbps split-rendered with ISAR operating at 384 kbps)

9.11.8.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the ISAR_BS1534-1-LR experiment.

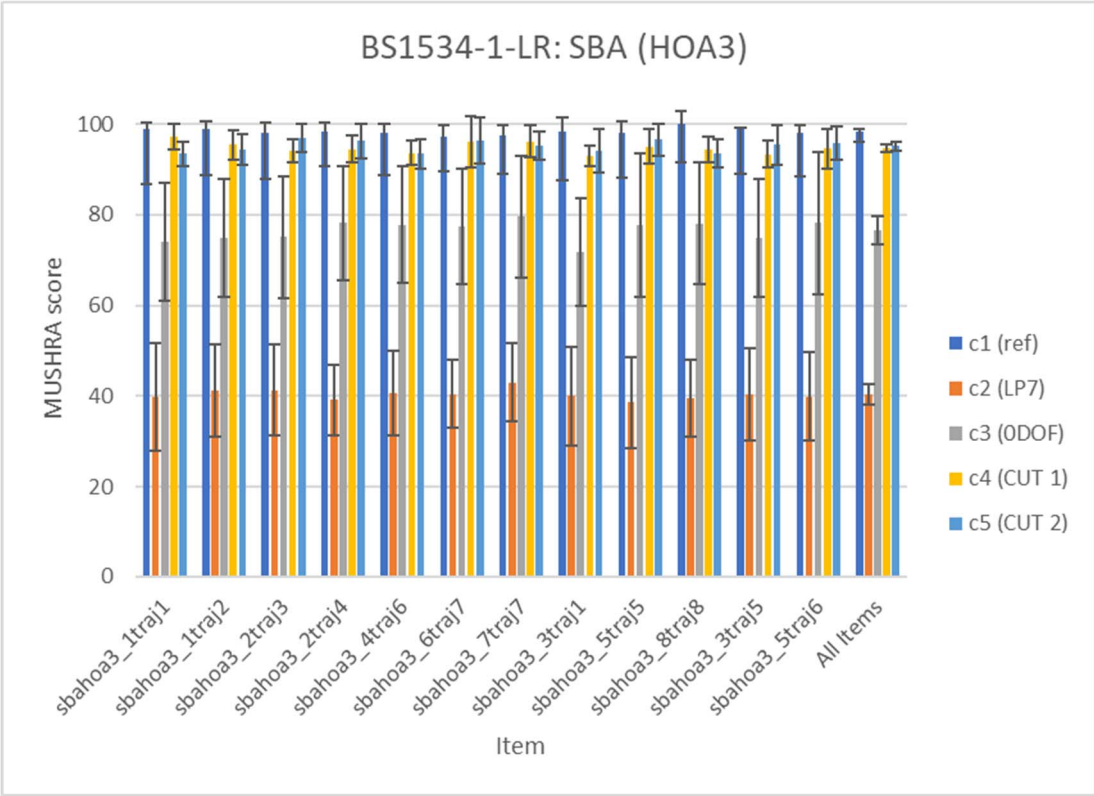


Figure 9.11-9: Results of ISAR_BS1534-1-LR test for SBA input audio

9.11.8.4 Statistical analysis

Table 9.11-10: Result of statistical analysis of BS1534-1-LR test checking CuT1 NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-18.3021	14.2564	1.4550	-12.5784	1.0000	Pass

Table 9.11-11: Result of statistical analysis of BS1534-1-LR test checking CuT2 NWT 0-DOF Reference

Mean Diff. (c03 - c05)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-18.6354	14.3455	1.4641	-12.7280	1.0000	Pass

9.11.8.5 Experimental conclusions

Conclusion of the low-rate experiment ISAR_BS1534-1-LR is that the ISAR solution for IVAS, when operated at 384 and 512 kbps, respectively, still clearly meets the requirement that it shall be no worse than the 0-DOF transcoding reference system. It can be observed that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range close to the quality of the native IVAS coding reference system is achieved. In contrast, the 0-DOF transcoding alternative offers substantially lower quality.

The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

Despite the lower bit rates, this conclusion is the same as for the highest tested bit rate of 784 kbps.

9.11.9 ISAR Experiment ISAR_BS1534-2-LR (MC 7.1+4, Generic Audio, 384 and 512 kbps Split Rendering Link, Headphone Presentation)

9.11.9.1 Overview

The experiment ISAR_BS1534-2-LR evaluates Split Rendering for IVAS 7.1+4 audio at 384 and 512 kbps.

9.11.9.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (REF):** Hidden reference: Native coding system (IVAS at 512kbps rendered to post renderer pose)
- **c02 (LP7):** LP7 anchor: Hidden reference, 7 kHz LP filtered
- **c03 (0-DOF):** 0-DOF native transcoding reference (IVAS at 512kbps binaurally rendered to pre-renderer pose, IVAS stereo coded@256kbps)
- **c04 (CuT1):** 3-DOF system 1 under test (IVAS at512 kbps split-rendered with ISAR operating at 512 kbps)
- **c05 (CuT2):** 3-DOF system 1 under test (IVAS at 512 kbps split-rendered with ISAR operating at 384 kbps)

9.11.9.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the ISAR_BS1534-2 experiment.

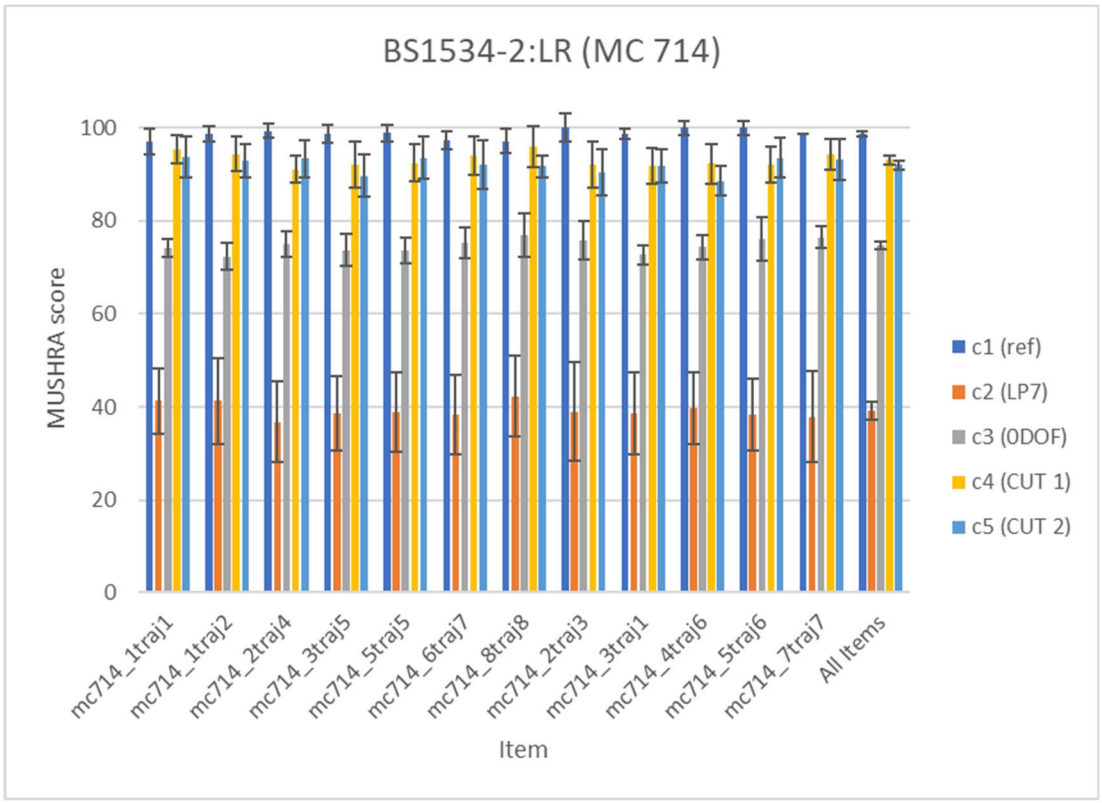


Figure 9.11-10: Results of ISAR_BS1534-2-LR test for MC 7.1.4 input audio

9.11.9.4 Statistical analysis

Table 9.11-12: Result of statistical analysis of BS1534-2-LR test checking CuT1 NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-18.4479	4.4506	0.4542	-40.6130	1.0000	Pass

Table 9.11-13: Result of statistical analysis of BS1534-2-LR test checking CuT2 NWT 0-DOF Reference

Mean Diff. (c03 - c05)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-17.3542	5.0304	0.5134	-33.8016	1.0000	Pass

9.11.9.5 Experimental conclusions

Conclusion of the low-rate experiment ISAR_BS1534-2-LR is that the ISAR solution for IVAS, when operated at 384 and 512 kbps, respectively, still clearly meets the requirement that it shall be no worse than the 0-DOF transcoding reference system. It can be observed that the achievable quality is even clearly better whereby a quality level in the 'excellent' range close to the quality of the native IVAS coding reference system is achieved. In contrast, the 0-DOF transcoding alternative offers substantially lower quality.

The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high 'excellent' range, which indicates only very minor audible differences.

Despite the lower bit rates, this conclusion is the same as for the highest tested bit rate of 784 kbps.

9.11.10 ISAR Experiment ISAR_BS1534-3-LR (4 Objects, Generic Audio, 384 and 512 kbps Split Rendering Link, Headphone Presentation)

9.11.10.1 Overview

The experiment ISAR_BS1534-3-LR evaluates Split Rendering for 4 object audio (ISM4) at 384 and 512 kbps.

9.11.10.2 Test conditions

The test conditions in all experiments are as follows:

The test conditions in all experiments are as follows:

- **c01 (REF):** Hidden reference: Native coding system (IVAS at 512kbps rendered to post renderer pose)
- **c02 (LP7):** LP7 anchor: Hidden reference, 7 kHz LP filtered
- **c03 (0-DOF):** 0-DOF native transcoding reference (IVAS at 512kbps binaurally rendered to pre-renderer pose, IVAS stereo coded@256kbps)
- **c04 (CuT1):** 3-DOF system 1 under test (IVAS at512 kbps split-rendered with ISAR operating at 512 kbps)
- **c05 (CuT2):** 3-DOF system 1 under test (IVAS at 512 kbps split-rendered with ISAR operating at 384 kbps)

9.11.10.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the ISAR_BS1534-3-LR experiment.

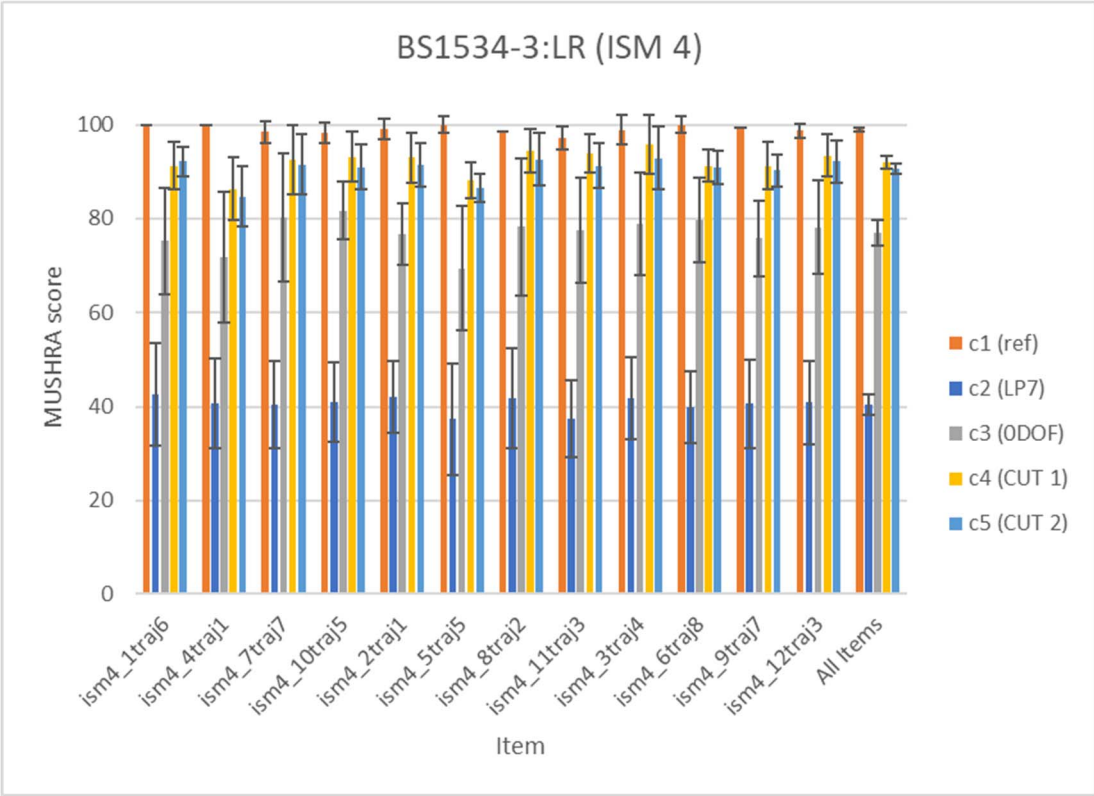


Figure 9.11-11: Results of ISAR_BS1534-3a test for ISM-4 input audio

9.11.6.4 Statistical analysis

Table 9.11-14: Result of statistical analysis of BS1534-4-LR test checking CuT1 NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-15.0938	10.4419	1.0657	-14.1629	1.0000	Pass

Table 9.11-15: Result of statistical analysis of BS1534-3-LR test checking CuT2 NWT 0-DOF Reference

Mean Diff. (c03 - c05)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-13.6771	9.9202	1.0125	-13.5085	1.0000	Pass

9.11.8.5 Experimental conclusions

Conclusion of the low-rate experiment ISAR_BS1534-3-LR is that the ISAR solution for IVAS, when operated at 384 and 512 kbps, respectively, still clearly meets the requirement that it shall be no worse than the 0-DOF transcoding reference system. It can be observed that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range close to the quality of the native IVAS coding reference system is achieved. In contrast, the 0-DOF transcoding alternative offers substantially lower quality.

The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

Despite the lower bit rates, this conclusion is the same as for the highest tested bit rate of 784 kbps.

9.11.11 ISAR Experiment ISAR_BS1534-4-LR (MASA 2TC, Generic Audio, 384 and 512 kbps Split Rendering Link, Headphone Presentation)

9.11.11.1 Overview

The experiment ISAR_BS1534-4-LR evaluates Split Rendering for IVAS MASA audio using 2 transport channels (TC) at 384 and 512 kbps.

9.11.11.2 Test conditions

The test conditions in all experiments are as follows:

- **c01 (REF):** Hidden reference: Native coding system (IVAS at 512kbps rendered to post renderer pose)
- **c02 (LP7):** LP7 anchor: Hidden reference, 7 kHz LP filtered
- **c03 (0-DOF):** 0-DOF native transcoding reference (IVAS at 512kbps binaurally rendered to pre-renderer pose, IVAS stereo coded@256kbps)
- **c04 (CuT1):** 3-DOF system 1 under test (IVAS at512 kbps split-rendered with ISAR operating at 512 kbps)
- **c05 (CuT2):** 3-DOF system 1 under test (IVAS at 512 kbps split-rendered with ISAR operating at 384 kbps)

9.11.11.3 Result plots

Provided below are the result plots and tables with statistical analysis result for the ISAR_BS1534-4-LR experiment.

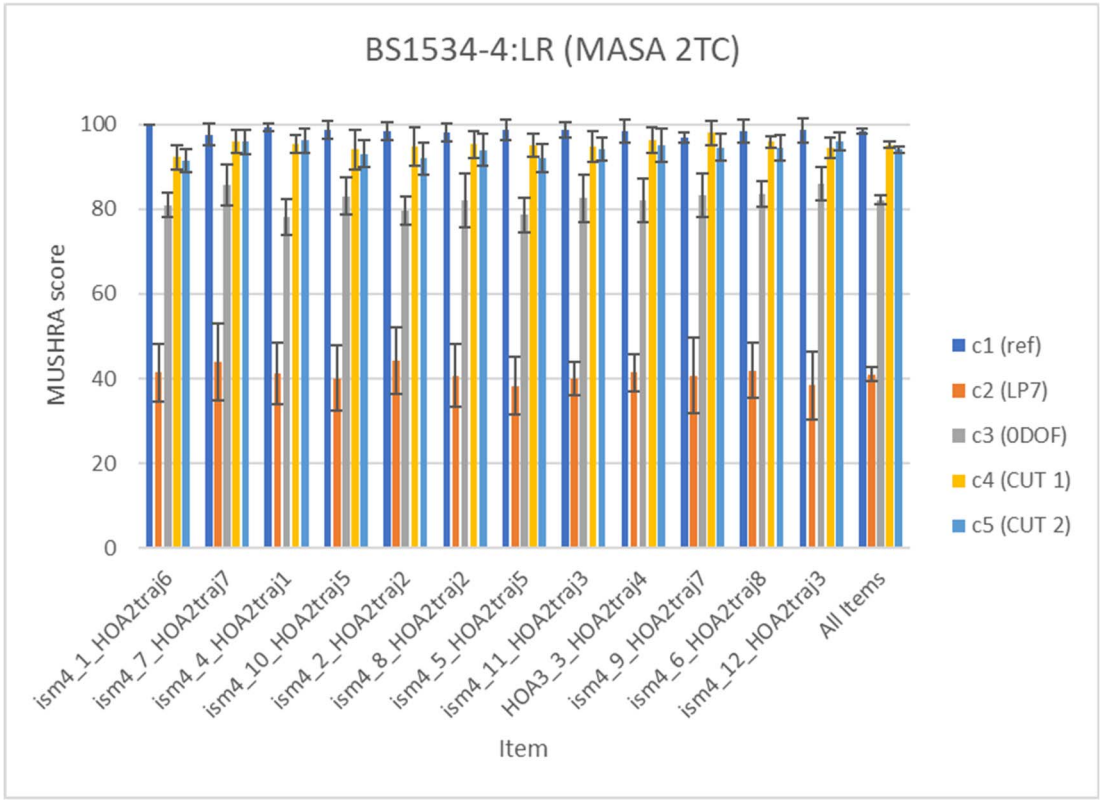


Figure 9.11-12: Results of ISAR_BS1534-4a test for MASA input audio

9.11.11.4 Statistical analysis

Table 9.11-16: Result of statistical analysis of BS1534-4-LR test checking CuT1 NWT 0-DOF Reference

Mean Diff. (c03 - c04)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-13.0104	5.7647	0.5884	-22.1132	1.0000	Pass

Table 9.11-17: Result of statistical analysis of BS1534-4-LR test checking CuT2 NWT 0-DOF Reference

Mean Diff. (c03 - c05)	Stdev Diff.	SE_{MD}	t	Prob.	ToR
-11.8958	5.6819	0.5799	-20.5134	1.0000	Pass

9.11.11.5 Experimental conclusions

Conclusion of the low-rate experiment ISAR_BS1534-4-LR is that the ISAR solution for IVAS, when operated at 384 and 512 kbps, respectively, still clearly meets the requirement that it shall be no worse than the 0-DOF transcoding reference system. It can be observed that the achievable quality is even clearly better whereby a quality level in the ‘excellent’ range close to the quality of the native IVAS coding reference system is achieved. In contrast, the 0-DOF transcoding alternative offers substantially lower quality.

The objective to provide a quality level as close as possible to the native coding reference system is met in the sense that the quality score of the split rendering system is in the high ‘excellent’ range, which indicates only very minor audible differences.

Despite the lower bit rates, this conclusion is the same as for the highest tested bit rate of 784 kbps.

9.12 Multi-Format

9.12.1 Overview

In Characterization phase, one experiment has been conducted to evaluate the performance of the IVAS codec for multiple formats. Experimemnt ACR-1 was conducted using P.800 ACR test methodology.

- Characterization Experiment ACR-1: Speech and mixed/music, Stereo/FOA/HOA3/MASA 1TC/MASA 2TC/2 objects, under clean and impaired channel conditions, headphone presentation

9.12.2 Characterization Experiment ACR-1 (Multi-Format, Speech and Mixed/Music, Clean and Impaired Channel, Headphone Presentation)

Characterization Experiment ACR-1 evaluates the performance of the IVAS fixed-point implementation for some IVAS spatial formats (Stereo, FOA, HOA3, MASA 1TC, MASA 2TC, ISM2) for speech and mixed/music + background for different bandwidths. The test was conducted using headphone presentation. See Annex E.9 for details.

Experiment ACR-1 was conducted by Nokia (language: FIN).

The listening lab has reported that the presentation order of the samples to the listeners was not according to the original randomization tables due to a configuration issue in the listening test software.

The aggregated results for all condition are listed in Table 9.12-1.

The experiment contains multiple references at different bandwidths from narrowband to fullband. The scores of the references are shown in Figure 9.12-1. The results for the test conditions are shown in the different subfigures of Figure 9.12-2 separated by bandwidth (WB, SWB, and FB). In addition, the spatial reference for the respective

bandwidth is shown as a horizontal, dashed line. All plots show means and 95% confidence intervals calculated according to ITU-T Recommendation P.1401 [23].

The additional plot in Figure 9.12-3 presents the data for FER 0% and DTX off in a different way by showing the results for all bandwidths WB, SWB, and FB in a single plot.

Table 9.12-1: results per condition for experiment ACR-1

Designator	Condition	COUNT	MOS	STD	CI95
c01	Spatial FB	180	4.37	0.70	0.10
c02	Spatial SWB (16 kHz)	180	4.31	0.76	0.11
c03	Spatial (14 kHz)	180	4.33	0.71	0.10
c04	Spatial SSWB (12 kHz)	180	4.31	0.84	0.12
c05	Spatial (10 kHz)	180	4.14	0.83	0.12
c06	Spatial WB (8 kHz)	180	3.89	0.85	0.12
c07	Spatial MB (6 kHz)	180	3.32	0.96	0.14
c08	Spatial NB (4 kHz)	180	2.66	1.06	0.16
c09	MNRU Q = 30 dB	180	2.92	1.28	0.19
c10	MNRU Q = 26 dB	180	2.02	1.17	0.17
c11	MNRU Q = 22 dB	180	1.59	0.98	0.14
c12	MNRU Q = 18 dB	180	1.24	0.68	0.10
c13	ESDRU α = 0.8	180	4.26	0.78	0.11
c14	ESDRU α = 0.6	180	3.47	0.99	0.15
c15	ESDRU α = 0.4	180	2.92	0.98	0.14
c16	ESDRU α = 0.2	180	2.53	1.05	0.15
c17	IVAS-WB, 13.2 kbps, FER 0%, DTX off	180	2.81	0.97	0.14
c18	IVAS-WB, 16.4 kbps, FER 0%, DTX off	180	3.16	0.91	0.13
c19	IVAS-WB, 24.4 kbps, FER 0%, DTX off	180	3.33	0.94	0.14
c20	IVAS-WB, 32 kbps, FER 0%, DTX off	180	3.36	0.92	0.14
c21	IVAS-WB, 48 kbps, FER 0%, DTX off	180	3.57	0.87	0.13
c22	IVAS-WB, 64 kbps, FER 0%, DTX off	180	3.70	0.84	0.12
c23	IVAS-WB, 80 kbps, FER 0%, DTX off	180	3.77	0.83	0.12
c24	IVAS-WB, 96 kbps, FER 0%, DTX off	180	3.79	0.83	0.12
c25	IVAS-SWB, 13.2 kbps, FER 0%, DTX off	180	2.66	1.03	0.15
c26	IVAS-SWB, 16.4 kbps, FER 0%, DTX off	180	3.10	0.95	0.14
c27	IVAS-SWB, 24.4 kbps, FER 0%, DTX off	180	3.58	0.95	0.14
c28	IVAS-SWB, 32 kbps, FER 0%, DTX off	180	3.63	0.87	0.13
c29	IVAS-SWB, 48 kbps, FER 0%, DTX off	180	3.81	0.94	0.14
c30	IVAS-SWB, 64 kbps, FER 0%, DTX off	180	3.98	0.81	0.12
c31	IVAS-SWB, 80 kbps, FER 0%, DTX off	180	4.03	0.82	0.12
c32	IVAS-SWB, 96 kbps, FER 0%, DTX off	180	4.16	0.82	0.12
c33	IVAS-FB, 24.4 kbps, FER 0%, DTX off	180	3.56	0.95	0.14
c34	IVAS-FB, 32 kbps, FER 0%, DTX off	180	3.68	0.96	0.14
c35	IVAS-FB, 48 kbps, FER 0%, DTX off	180	3.72	0.88	0.13
c36	IVAS-FB, 64 kbps, FER 0%, DTX off	180	3.97	0.81	0.12
c37	IVAS-FB, 80 kbps, FER 0%, DTX off	180	4.11	0.85	0.12
c38	IVAS-FB, 96 kbps, FER 0%, DTX off	180	4.13	0.79	0.12
c39	IVAS-FB, 128 kbps, FER 0%, DTX off	180	4.26	0.83	0.12
c40	IVAS-FB, 160 kbps, FER 0%, DTX off	180	4.28	0.75	0.11
c41	IVAS-FB, 192 kbps, FER 0%, DTX off	180	4.14	0.81	0.12
c42	IVAS-FB, 256 kbps, FER 0%, DTX off	180	4.28	0.76	0.11
c43	IVAS-WB, 13.2 kbps, FER 0%, DTX on	180	2.81	1.01	0.15
c44	IVAS-WB, 16.4 kbps, FER 0%, DTX on	180	3.09	0.91	0.13
c45	IVAS-SWB, 16.4 kbps, FER 0%, DTX on	180	3.10	0.84	0.12
c46	IVAS-SWB, 24.4 kbps, FER 0%, DTX on	180	3.42	0.94	0.14
c47	IVAS-FB, 24.4 kbps, FER 0%, DTX on	180	3.42	0.85	0.12
c48	IVAS-FB, 32 kbps, FER 0%, DTX on	180	3.69	0.91	0.13
c49	IVAS-WB, 24.4 kbps, FER 5%, DTX off	180	2.63	0.86	0.13
c50	IVAS-WB, 32 kbps, FER 5%, DTX off	180	2.89	0.96	0.14
c51	IVAS-SWB, 32 kbps, FER 5%, DTX off	180	2.91	0.94	0.14
c52	IVAS-SWB, 48 kbps, FER 5%, DTX off	180	2.85	0.94	0.14
c53	IVAS-FB, 64 kbps, FER 5%, DTX off	180	3.09	0.89	0.13
c54	IVAS-FB, 80 kbps, FER 5%, DTX off	180	3.14	0.93	0.14

c55	IVAS-WB, 16.4 kbps, FER 5%, DTX on	180	2.53	0.83	0.12
c56	IVAS-WB, 24.4 kbps, FER 5%, DTX on	180	2.63	0.91	0.13
c57	IVAS-SWB, 24.4 kbps, FER 5%, DTX on	180	2.58	0.84	0.12
c58	IVAS-SWB, 32 kbps, FER 5%, DTX on	180	2.86	0.79	0.12
c59	IVAS-FB, 48 kbps, FER 5%, DTX on	180	2.87	0.90	0.13
c60	IVAS-FB, 64 kbps, FER 5%, DTX on	180	3.01	0.95	0.14

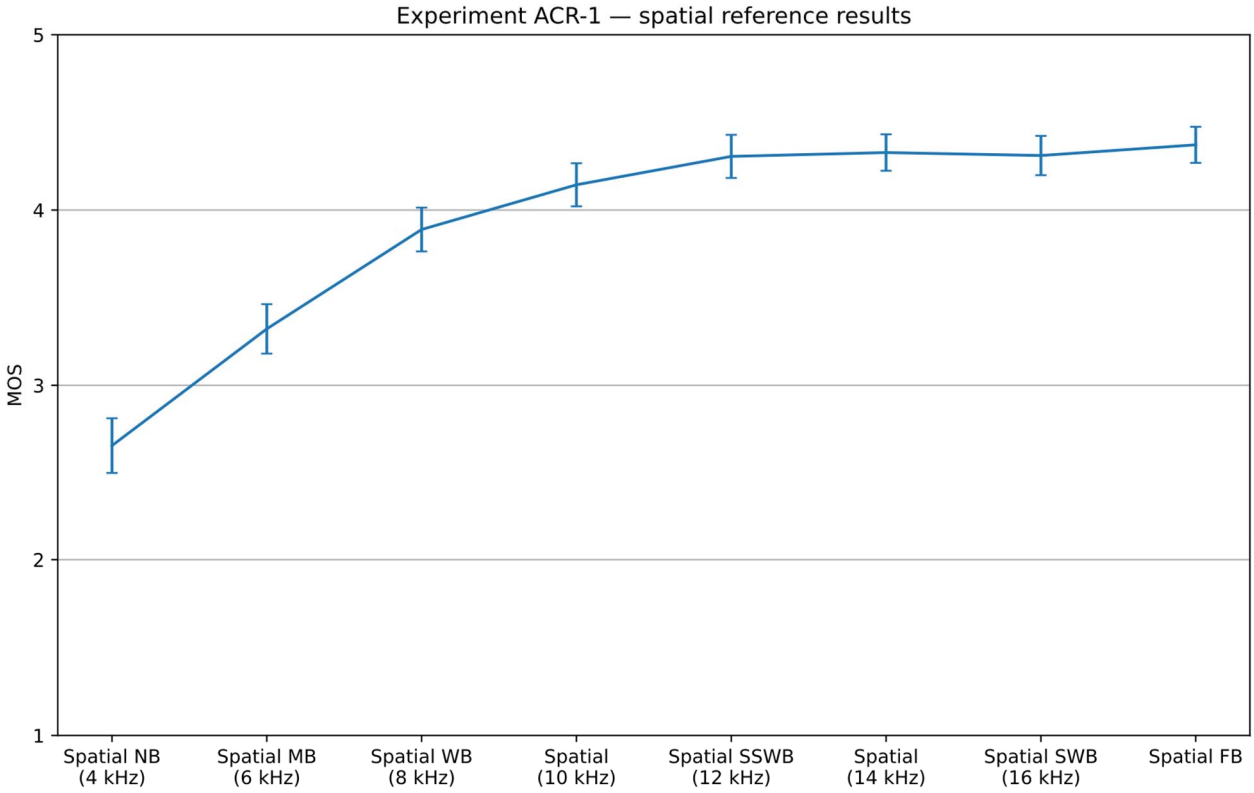
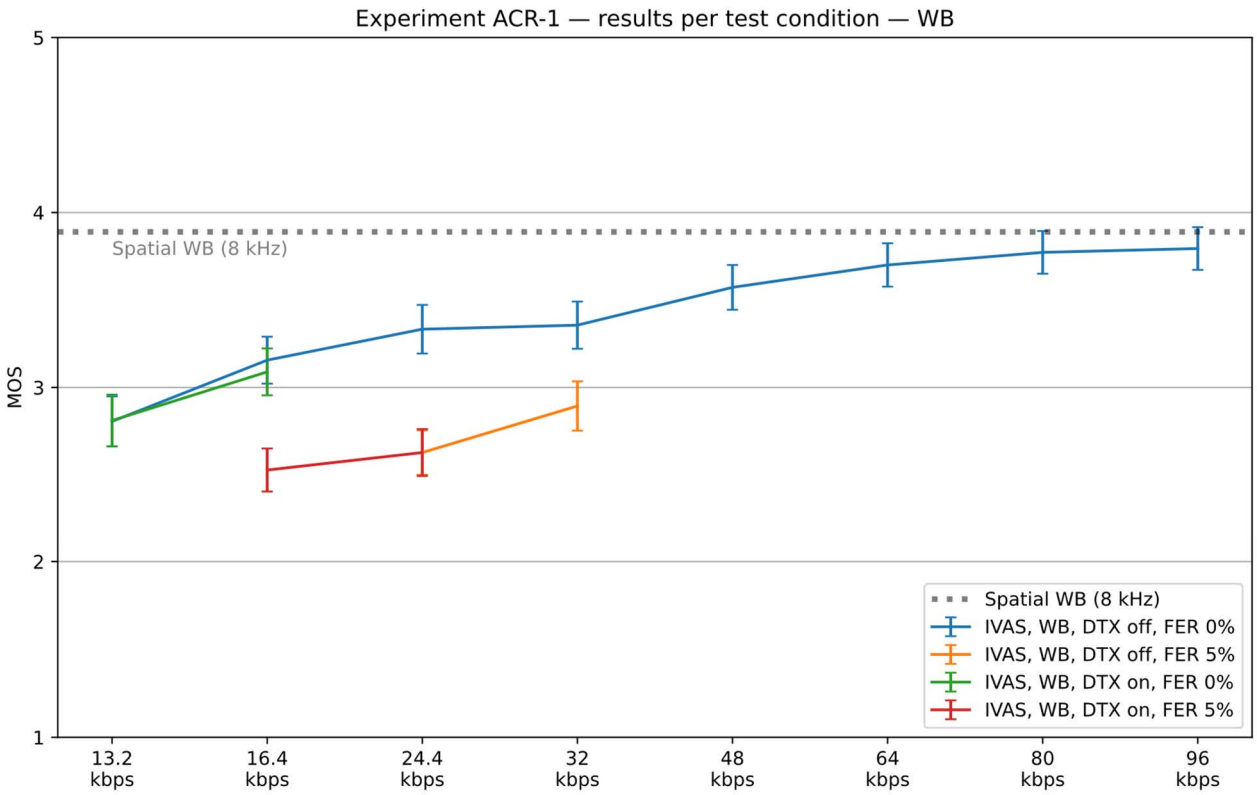
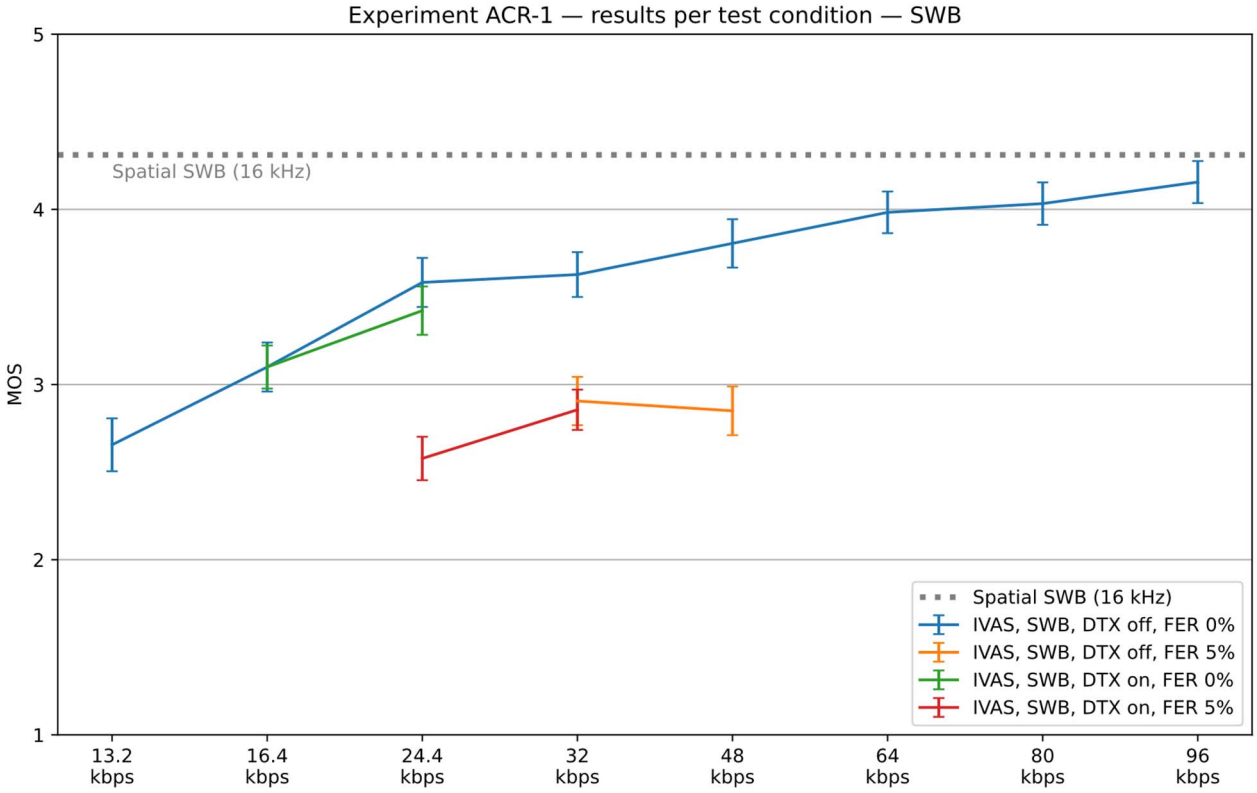


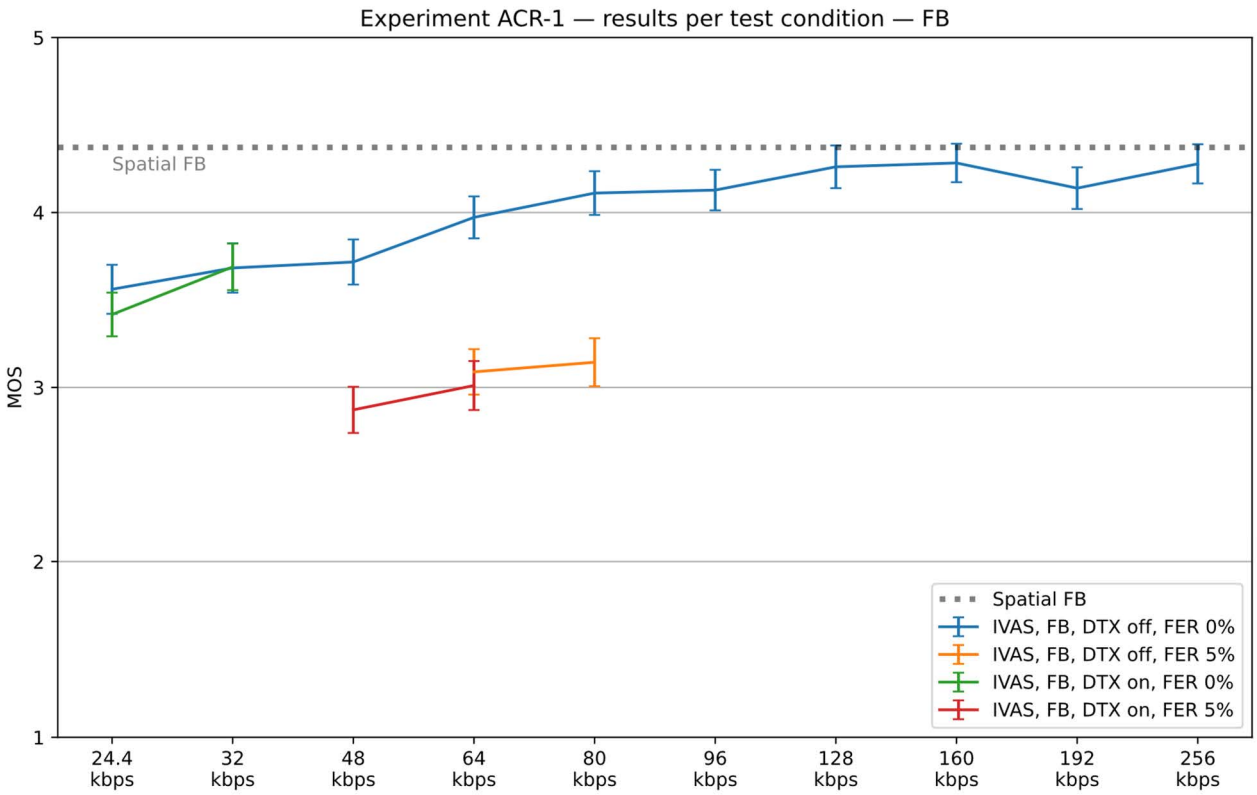
Figure 9.12-1: spatial reference conditions for experiment ACR-1


Experiment ACR-1 — results per test condition — WB
MOS
5
4
3
2
1
Spatial WB (8 kHz)
13.2 kbps
16.4 kbps
24.4 kbps
32 kbps
48 kbps
64 kbps
80 kbps
96 kbps
Spatial WB (8 kHz)
IVAS, WB, DTX off, FER 0%
IVAS, WB, DTX off, FER 5%
IVAS, WB, DTX on, FER 0%
IVAS, WB, DTX on, FER 5%

(a)


Experiment ACR-1 — results per test condition — SWB
MOS
5
4
3
2
1
Spatial SWB (16 kHz)
13.2 kbps
16.4 kbps
24.4 kbps
32 kbps
48 kbps
64 kbps
80 kbps
96 kbps
Spatial SWB (16 kHz)
IVAS, SWB, DTX off, FER 0%
IVAS, SWB, DTX off, FER 5%
IVAS, SWB, DTX on, FER 0%
IVAS, SWB, DTX on, FER 5%

(b)



(c)

Figure 9.12-2: results per condition for experiment ACR-1

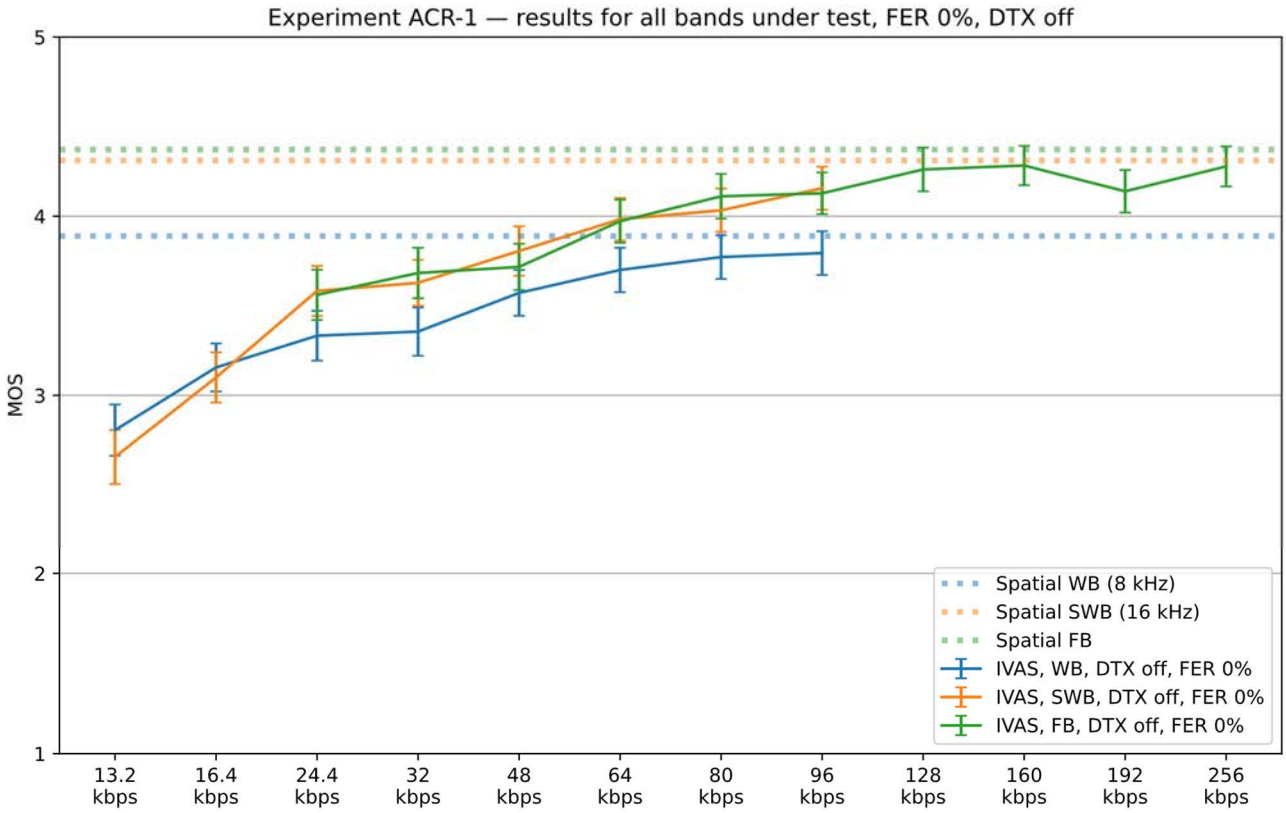


Figure 9.12-3: test conditions for all bands under test, FER 0%, DTX off for experiment ACR-1

9.12.3 Conclusions on IVAS Performance for Multi-Format

Taking the results of the previous clauses in 9.10 into account, the following conclusions can be drawn:

During characterization phase, the performance of the IVAS codec for different audio bandwidths was evaluated. The ACR test demonstrates the suitability of the IVAS codec for all supported bandwidths (WB, SWB, FB). For all bandwidths, a monotonously increasing quality was shown for bitrates up to 160 kbps, approaching transparency.

10 Objective Evaluations

10.1 Complexity and Delay Analysis

10.1.1 Complexity

A complexity evaluation is expected to be included.

10.1.2 Algorithmic Delay

The input signals (audio, or audio and metadata) are processed using 20 ms frames. The codec algorithmic delay depends on the input/output audio formats as described in Table 10.1-1.

Table 10.1-1: IVAS algorithmic delay for different input/output format combinations (rounded to integer milliseconds; in case multiple values are provided they depend on the bitrate)

		Decoder output format						
		Mono	**Stereo**	**Multi-Channel**	**Binaural audio**	**Scene-based audio**	**Object-based audio**	**Metadata-assisted spatial audio**
Encoder input format	**Mono**	32						
	Stereo	32	32	32				
	Binaural	32			32			
	Multi-channel	32	32	32 / 37	32 / 37	32 / 37		
	Scene-based audio	33	33	38	38	38		
	Object-based audio	32	32	32 / 37	32 / 37	32 / 37	32 / 37	
	Metadata-assisted spatial audio	32 / 37	37	37	37	37		32 (NOTE
	OSBA	33	33	38	38	38		
	OMASA	32 / 37	37	37	37	37		

NOTE: Metadata-assisted spatial audio (MASA) decoder output allows also for mono or stereo decoder output at 32 ms algorithmic delay by stripping the metadata file

The algorithmic delay related to the core-coder coding in IVAS is 32 ms similarly as in EVS though its splitting between the encoder and the decoder is slightly different. It consists of 8.75 ms for the encoder look-ahead and 3.25 ms for the decoder delay related to the time-domain BWE and resampling in the DFT domain.

Further, the IVAS delay consists of 5 ms delay related to the rendering to the related output configuration, thus making the overall delay of 32 ms in some set-ups and 37 ms in other set-ups.

Finally, in SBA format, an additional encoder delay of 1 ms is present and it is related to the filter-bank analyses prior to the encoding.

It is also noted that the delay figures exclude any HRIR/BRIR induced delay.

The codec delay for mono (EVS) operation and stereo downmix operation in EVS compatible operation is 32 ms as described in clause 4.3 in [19].

Annex A:
ToR Tests in Selection Phase

A.1 ToR Tests for Requirements

Table A.1-1 summarizes the results for the Requirements ToR tests over the 23 selection experiments as analyzed in the GAL report (Annex F.1). Each row of the table shows results of ToR tests for a single Experiment - results for Test 1 on the left and for Test 2 on the right. For each Experiment, the table shows the test name, #Requirement ToR's followed by the Test 1 listening lab (LL), #ToRs passed, #ToRs failed, Test 2 LL, #ToRs passed, #ToRs failed and finally the Percent of ToRs passed across both Tests within the Experiment.

On the far right side of the table, "Percent ToRs Passed" values are shown for each of the four Groups of Experiments: 100% for Stereo, 100% for Multichannel, 98.6% for SBA, 93.8% for Objects, and 100% for MASA. Finally, at the bottom of the table, "Percent ToRs Passed" for the entire Selection Phase is reported - 5 of 288 ToRs were failed for a Percent Passed value of 98.3%.

Table A.1-1: Summary ToR Test Results for Requirements

Test	#Req	Test 1			Test 2			Passed	Group
		LL	#Pass	#Fail	LL	#Pass	#Fail		
P800-1	12	a	12	0	d	12	0	100%	Stereo 100%
P800-2	11	b	11	0	d	11	0	100%	
P800-3	13	a	13	0	d	13	0	100%	
BS1534-1a	2	a	2	0	d	2	0	100%	
BS1534-1b	2	b	2	0	d	2	0	100%	
BS1534-2a	2	a	2	0	b	2	0	100%	Multichan. 100%
BS2534-2b	2	a	2	0	b	2	0	100%	
BS1534-3a	2	a	2	0	b	2	0	100%	
BS1534-3b	2	a	2	0	b	2	0	100%	
P800-4	13	a	13	0	c	13	0	100%	SBA 98.6%
P800-5	13	a	13	0	b	13	0	100%	
BS1534-4a	3	a	3	0	d	3	0	100%	
BS1534-4b	2	b	2	0	d	2	0	100%	
BS1534-5a	2	a	2	0	d	1	1	75.0%	
BS1534-5b	2	a	2	0	b	2	0	100%	
P800-6	13	a	10	3	c	12	1	84.6%	Objects 93.8%
P800-7	13	a	13	0	d	13	0	100%	
BS1534-6a	3	b	3	0	d	3	0	100%	
BS1534-6b	3	b	3	0	d	3	0	100%	
P800-8	12	a	12	0	b	12	0	100%	MASA 100%
P800-9	13	a	13	0	d	13	0	100%	
BS1534-7a	2	b	2	0	d	2	0	100%	
BS1534-7b	2	b	2	0	d	2	0	100%	
Total	144		141	3		142	2	98.3%	

The tests were carried out by the listening labs according to Table A.1-2:

Table A.1-2: List of listening labs

LL identifier	Listening Lab Company	P800 tests	BS1534 tests (headphone)	BS1534 tests (loudspeaker)
a	FORCE Technology	8	3	5
b	HEAD acoustics GmbH / IKS (RWTH Aachen)	3	6	5
c	Macquarie University	2	0	0
d	Mesaqin.com	5	9	0

Across all P.800 experiments, there are 226 requirement ToR tests (i.e. 113 requirements per labs). The CuT failed four of them, it passed 94 times and exceeded the requirement 128 times. The four failed tests are all in P800-6.

Table A.1-3: Overview of all failed tests in P.800 experiments

Exp.	Lab	Cond.	Bitrate	DTX	FER	ToR	Status
P800-6	a	c29	64	Off		NWT C16	FAIL
		c31	16.4	Off	5%	NWT C18	FAIL
		c36	24.4	On		NWT C23	FAIL
	c	c28	48	Off		NWT C15	FAIL

Across all BS.1534 experiments, there are 93 requirement ToR tests (i.e. 31 requirements per lab + a combined analysis for both labs). It should be noted that the GAL report includes for BS.1534 experiments also an analysis of the combination of the data sets from both labs. The CuT failed one of them, it passed 48 times and exceeded the requirement 44 times. The failed test is in BS1534-5A.

Table A.1-3: Overview of all failed tests in BS.1534 experiments

Exp.	Lab	Cond.	Bitrate	ToR	Status
BS1534-5A	d	c07	256	NWT c05 OR BT c04	FAIL

In total, the CuT failed on four out of 226 requirement ToR tests for the P.800 experiments, it passed 94 times and exceeded the requirement 128 times. In the BS.1534 tests, the CuT failed on one of the 93 requirement ToR tests, it passed 48 times and exceeded the requirement 44 times. There is no instance in any experiment where the CuT failed a requirement ToR test for the data sets of both listening laboratories. Additionally, the CuT did not fail any requirement ToR test for the joint analyses of the BS.1534 data sets.

Annex B:
Overall Characterization of the IVAS Codec

IVAS is the next generation codec in 3GPP. It is an extension of the 3GPP Enhanced Voice Services (EVS) codec offering:

- Complete bit-exact EVS codec functionality for mono speech/audio signal input
- Support of stereo and binaural audio
- Support of audio formats beyond stereo which include multi-channel audio (5.1, 5.1.2, 5.1.4, 7.1, 7.1.4), scene-based audio (Ambisonics up to 3rd order), metadata-assisted spatial audio (MASA), and object-based audio.
- Support of combined immersive audio formats: object-based audio with scene-based audio (OSBA) and object-based audio with metadata-assisted spatial audio (OMASA)
- VAD/DTX/CNG for rate efficient stereo and immersive conversational voice transmissions
- Error concealment mechanisms to combat the effects of transmission errors and lost packets
- Jitter buffer management
- Binaural rendering functionality for headphone playback including head-tracking and scene orientation control, and loudspeaker rendering functionality for loudspeaker playback
- Binaural split rendering functionality with pre-rendering and transcoding to a head-trackable intermediate representation that can be transmitted to a post-rendering end-device for headphone playback.

The codec is optimized for services over 5G mobile networks and implementations on 5G devices with:

- Operation on 20 ms audio frames
- Multi-rate/multi-mode operation at the following discrete bit rates [kbps]: 13.2, 16.4, 24.4, 32, 48, 64, 80, 96, 128, 160, 192, 256, 384, and 512
- Ability to switch bitrate upon command
- Support of sampling frequencies of 8 kHz (only EVS interoperable coding), 16 kHz, 32 kHz and 48 kHz (fullband audio content)
- Low algorithmic delay ($\leq$38 ms)
- Complexity and memory footprint within design constraint limits defining three levels, suitable for different device types and application scenarios

IVAS is the first 3GPP communication codec which natively supports stereo and immersive audio. This enables completely new use-cases beyond traditional telephony: In a stereo or immersive telephony use-case, for example, a participant can capture and convey an immersive scene to a remote participant, e.g., to share the full immersive experience of an event. For spatial conferencing applications, the flexibility of the IVAS codec will provide multiple options for:

- Ad-hoc conferencing calls with the transmission of the physical immersive scene picked up by a UE, e.g., placed on a table. Rendering of the immersive scene makes it easier to distinguish the talkers' voice, clearly separated from ambient sounds, leading to more natural and effort-less conferencing.
- More complex scenarios with multiple participants, transmitted as individual streams and spatially rendered on the receiving UE to match the video scene, for example.
- Scenarios where an intermediate call server combines multiple participants into an immersive scene.

Moreover, the flexibility of the IVAS codec with respect to supported audio formats, bitrates, rendering options and suitability for different device types might enable completely new usage scenarios for communication. In addition, the IVAS codec will support content distribution use-cases including streaming of stereo/immersive content and advanced VR/AR applications.

3GPP's rigorous and transparent standardization process involved the definition of demanding terms of reference (ToRs). During the selection phase, the IVAS codec was tested against in total 144 ToRs in 46 subjective tests performed in six languages by four independents test labs. During these tests, the IVAS codec has met the requirements in 98.3% of the cases with no systematic fails in both labs.

Annex C : IVAS Selection Phase Testplan

C.1 Experiment P800-1: Stereo/Binaural Clean Speech Test

C.1.1 Experiment setup

Tables C.1-1 to C.1-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.1-1: Conditions for Experiment P800-1

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48
DTX	DTX on and off
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Codec references	
Codec references	EVS dual mono
Bitrates	2x7.2, 2x8.0, 2x9.6, 2x13.2, 2x16.4, 2x24.4, 2x32 kbps
Input level	-26 LKFS [22]
DTX	DTX on and off
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q=16, 20, 24, 28 dB
ESDRU	α = 0.1, 0.4, 0.7
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based relying on convolution of raw mono clean speech sentences with Room Impulse Responses respective to various talker positions relative to a capture point as described in the ITU-T Reverberation Tool [21]
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to SWB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Japanese, French

Table C.1-2: Preliminaries for Experiment P800-1

Trial #	Label	Sample	Condition	Bitrate	FER/Profile
1	c19		EVS	2x13.2	5%
2	c02		MNRU Q=28 dB	-	-
3	c06		ESDRU α = 0.7	-	-
4	c13		EVS	2x16.4	No errors
5	c08		ESDRU α = 0.1	-	-
6	c16		EVS	2x7.2	5%
7	c04		MNRU Q=20 dB	-	-
8	c01		Reference	-	-
9	c15		EVS	2x32	No errors
10	c07		ESDRU α = 0.4	-	-
11	c05		MNRU Q=16 dB	-	-
12	c21		EVS	2x24.4	5%

Table C.1-3: Test conditions for Experiment P800-1, clean speech under clean and impaired channel conditions

Label	Condition	Bitrate [kbps]	DTX	FER/Profile	ToR
c01	Reference	-	-	-	
c02	MNRU Q=28 dB	-	-	-	
c03	MNRU Q=24 dB	-	-	-	
c04	MNRU Q=20 dB	-	-	-	
c05	MNRU Q=16 dB	-	-	-	
c06	ESDRU α = 0.7	-	-	-	
c07	ESDRU α = 0.4	-	-	-	
c08	ESDRU α = 0.1	-	-	-	
c09	EVS	2x7.2	off	No errors	
c10	EVS	2x8	off	No errors	
c11	EVS	2x9.6	off	No errors	
c12	EVS	2x13.2	off	No errors	
c13	EVS	2x16.4	off	No errors	
c14	EVS	2x24.4	off	No errors	
c15	EVS	2x32	off	No errors	
c16	EVS	2x7.2	off	5%	
c17	EVS	2x8	off	5%	
c18	EVS	2x9.6	off	5%	
c19	EVS	2x13.2	off	5%	
c20	EVS	2x16.4	off	5%	
c21	EVS	2x24.4	off	5%	
c22	EVS	2x32	off	5%	
c23	EVS	2x13.2	on	No errors	
c24	EVS	2x8	on	5%	
c25	CuT	13.2	off	No errors	NWT c10 OR BT c09
c26	CuT	16.4	off	No errors	NWT c11 OR BT c10
c27	CuT	24.4	off	No errors	NWT c12 OR BT c11
c28	CuT	32	off	No errors	NWT c13 OR BT c12
c29	CuT	48	off	No errors	NWT c15 OR BT c14
c30	CuT	13.2	off	5%	NWT c17 OR BT c16
c31	CuT	16.4	off	5%	NWT c18 OR BT c17
c32	CuT	24.4	off	5%	NWT c19 OR BT c18
c33	CuT	32	off	5%	NWT c20 OR BT c19
c34	CuT	48	off	5%	NWT c22 OR BT c21
c35	CuT	24.4	on	No errors	NWT c23
c36	CuT	13.2	on	5%	NWT c24

C.1.2 Content type categories and scene definitions

Table C.1-4: Content type categories and scene definitions

Category	Room	Reverb	Microphone Setup	***Background***	SNR [dB]	Overtalk [s] See NOTE 1	Bandwidth	Talker positions See NOTE 2	Talker selection by panel
cat 1	small	anechoic	A-B (100cm)	Low level idle noise	45	1	Max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	large	anechoic	A-B (150cm)	Low level idle noise	45	-1	max available up to SWB	5-11 1-6 3-7 5-8 9-7 10-9	P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	small	anechoic	M-S	Low level idle noise	45	-1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	small	echoic	A-B (100cm)	Low level idle noise	45	1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	large	echoic	A-B (150cm)	Low level idle noise	45	1	max available up to SWB	2-8 9-4 6-10 11-8	P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2

								10-12 12-1	
cat 6	small	echoic	Binaural	Low level idle noise	45	-1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1
NOTE 1: Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences. NOTE 2: The talker positions are part of the scene definition of the different categories. They correspond to the talker positions as depicted in Figures 14.3 and 14.5 of [21] for the large and the small room, respectively.									

C.2 Experiment P800-2: Stereo Speech and Background Test

C.2.1 Experiment setup

Tables C.2-1 to C.2-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.2-1: Conditions for Experiment P800-2

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48, 64
DTX	DTX on and off
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	15dB
Error Conditions	No errors
Codec references	
Codec references	EVS dual mono
Bitrates	2x7.2, 2x8.0, 2x9.6, 2x13.2, 2x16.4, 2x24.4, 2x32, 2x448 kbps
Input level	-26 LKFS [22]
DTX	DTX off, DTX on
Input frequency mask	HP50
Noise	15dB
Error Conditions	No errors
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q=12, 17, 22, 27, 32 dB
ESDRU	α = 0.1, 0.3, 0.5, 0.7
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based relying on convolution of raw mono clean speech sentences with Room Impulse Responses respective to various talker positions relative to a capture point as described in the ITU-T Reverberation Tool [21] and impulse responses provided by MC.
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to SWB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments with different background types (car, office, street) and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	German, Mandarin

Table C.2-2: Preliminaries for Experiment P800-2

Trial #	Label	Sample	Condition	Bitrate	DTX
1	c22		EVS	2x13.2	on
2	c02		MNRU Q=32 dB	-	-
3	c07		ESDRU α = 0.7	-	-
4	c15		EVS	2x16.4	off
5	c10		ESDRU α = 0.1	-	-
6	c19		EVS	2x7.2	on
7	c04		MNRU Q=22 dB	-	-
8	c01		Reference	-	-
9	c17		EVS	2x32	off
10	c08		ESDRU α = 0.5	-	-
11	c05		MNRU Q=17 dB	-	-
12	c24		EVS	2x24.4	on

Table C.2-3: Test conditions for Experiment P800-2, clean speech with background noise under clean channel conditions

Label	Condition	Bitrate [kbps]	DTX	ToR
c01	Reference	-	-	
c02	MNRU Q=32 dB	-	-	
c03	MNRU Q=27 dB	-	-	
c04	MNRU Q=22 dB	-	-	
c05	MNRU Q=17 dB	-	-	
c06	MNRU Q=12 dB			
c07	ESDRU α = 0.7	-	-	
c08	ESDRU α = 0.5	-	-	
c09	ESDRU α = 0.3			
c10	ESDRU α = 0.1	-	-	
c11	EVS	2x7.2	off	
c12	EVS	2x8	off	
c13	EVS	2x9.6	off	
c14	EVS	2x13.2	off	
c15	EVS	2x16.4	off	
c16	EVS	2x24.4	off	
c17	EVS	2x32	off	
c18	EVS	2x48	off	
c19	EVS	2x7.2	on	
c20	EVS	2x8	on	
c21	EVS	2x9.6	on	
c22	EVS	2x13.2	on	
c23	EVS	2x16.4	on	
c24	EVS	2x24.4	on	
c25	EVS	2x32	on	
c26	CuT	13.2	off	NWT c12 OR BT c11
c27	CuT	16.4	off	NWT c13 OR BT c12
c28	CuT	24.4	off	NWT c14 OR BT c13
c29	CuT	32	off	NWT c15 OR BT c14
c30	CuT	48	off	NWT c17 OR BT c16
c31	CuT	64	off	NWT c18 OR BT c17
c32	CuT	13.2	on	NWT c20 OR BT c19
c33	CuT	16.4	on	NWT c21 OR BT c20
c34	CuT	24.4	on	NWT c22 OR BT c21
c35	CuT	32	on	NWT c23 OR BT c22
c36	CuT	48	on	NWT c25 OR BT c24

C.2.2 Content type categories and scene definitions

Table C.2-4: Content type categories and scene definitions

Category	Room	Reverb	Microphone Setup	Background	SNR In dB	Overtalk [s] see NOTE 1	Bandwidth	Talker positions see NOTE 2	Talker selection by panel
cat 1	car	car	A-B Cardioid pair20 cm	car	15	-1	Max available up to SWB	Driver-Passenger BackRight-Driver Driver-BackCenter BackLeft-Driver BackRight-BackLeft BackCenter-BackRight	P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	-	-	A-B (20cm)	street	15	-1	max available up to SWB	Regularly spaced around the microphone	P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	small	echoic	M-S	office	15	-1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	small	echoic	A-B (100cm)	office	15	-1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	large	echoic	A-B (150cm)	office	15	-1	max available up to SWB	2-8 9-4 6-10 11-8 10-12 12-1	P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2

cat 6	small	echoic	Binaural	office	15	-1	max available up to SWB	1-7 5-3 2-6 4-1 3-4 7-2	P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1
NOTE 1: Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences. NOTE 2: The talker positions are part of the scene definition of the different categories. For office noise, they correspond to the talker positions as depicted in Figures 14.3 and 14.5 of [21] for the large and the small room.									

C.3 Experiment P800-3: Stereo Mixed and Music Test

C.3.1 Experiment setup

Tables C.3-1 to C.3-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.3-1: Conditions for Experiment P800-3

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48, 64
DTX	DTX on and off
Input level	-26 LKFS [22]
Input frequency mask	20KBP
Noise	No noise
Error Conditions	0%, 5%
Codec references	
Codec references	EVS dual mono
Bitrates	2x7.2, 2x8.0, 2x9.6, 2x13.2, 2x16.4, 2x24.4, 2x32, 2x48 kbps
Input level	-26 LKFS [22]
DTX	DTX off
Input frequency mask	20KBP
Noise	No noise
Error Conditions	0%, 5%
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q = 12, 17, 22, 27 dB
ESDRU	α = 0.1, 0.4, 0.7
Input frequency mask	20KBP
Common Conditions	
Categories	Following clause 7.1.3.4
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Number of categories	3 mixed content and 3 music
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Danish, Mandarin

Table C.3-2: Preliminaries for Experiment P800-3

Trial #	Label	Sample	Condition	Bitrate	FER/Profile
1	C20		EVS	2x13.2	5%
2	c02		MNRU Q=28 dB	-	-
3	c06		ESDRU α = 0.7	-	-
4	c13		EVS	2x16.4	No errors
5	c08		ESDRU α = 0.1	-	-
6	c17		EVS	2x7.2	5%
7	c04		MNRU Q=23 dB	-	-
8	c01		Reference	-	-
9	c15		EVS	2x32	No errors
10	c07		ESDRU α = 0.4	-	-
11	c05		MNRU Q=12 dB	-	-
12	c22		EVS	2x24.4	5%

Table C.3-3: Test conditions for Experiment P800-3, mixed contents and music under clean and impaired channel conditions

Label	Condition	Bitrate [kbps]	DTX	FER/Profile	ToR
c01	Reference	-	-	-	
c02	MNRU Q=27 dB	-	-	-	
c03	MNRU Q=22 dB	-	-	-	
c04	MNRU Q=17 dB	-	-	-	
c05	MNRU Q=12 dB	-	-	-	
c06	ESDRU $\alpha = 0.7$	-	-	-	
c07	ESDRU $\alpha = 0.4$	-	-	-	
c08	ESDRU $\alpha = 0.1$	-	-	-	
c09	EVS	2x7.2	off	No errors	
c10	EVS	2x8	off	No errors	
c11	EVS	2x9.6	off	No errors	
c12	EVS	2x13.2	off	No errors	
c13	EVS	2x16.4	off	No errors	
c14	EVS	2x24.4	off	No errors	
c15	EVS	2x32	off	No errors	
c16	EVS	2x48	off	No errors	
c17	EVS	2x7.2	off	5%	
c18	EVS	2x8	off	5%	
c19	EVS	2x9.6	off	5%	
c20	EVS	2x13.2	off	5%	
c21	EVS	2x16.4	off	5%	
c22	EVS	2x24.4	off	5%	
c23	EVS	2x32	off	5%	
c24	CuT	13.2	off	No errors	NWT c10 OR BT c09
c25	CuT	16.4	off	No errors	NWT c11 OR BT c10
c26	CuT	24.4	off	No errors	NWT c12 OR BT c11
c27	CuT	32	off	No errors	NWT c13 OR BT c12
c28	CuT	48	off	No errors	NWT c15 OR BT c14
c29	CuT	64	off	No errors	NWT c16 OR BT c15
c30	CuT	13.2	off	5%	NWT c18 OR BT c17
c31	CuT	16.4	off	5%	NWT c19 OR BT c18
c32	CuT	24.4	off	5%	NWT c20 OR BT c19
c33	CuT	32	off	5%	NWT c21 OR BT c20
c34	CuT	48	off	5%	NWT c23 OR BT c22
c35	CuT	24.4	on	No errors	NWT c12 OR BT c11[(1]
c36	CuT	13.2	on	5%	NWT c18 OR BT c17[(1]

[(1]These references are not requirements, they are just for information.

C.4 Experiment P800-4: FOA Clean Speech Test

C.4.1 Experiment setup

Tables C.4-1 to C.4-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.4-1: Conditions for Experiment P800-4 (Clean Speech)

Main Codec Conditions	
Candidate	CuT
Bitrates	IVAS candidate operated at 16.4, 24.4, 32, 48, 64, 80 and 96 kbps
DTX	DTX off
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Codec references	
Codec references	Multi-mono EVS
Bitrates	3x7.2, 4x7.2, 4x8, 4x9.6, 4x13.2, 4x16.4, 4x24.4, 4x32 kbps
Input level	-26 LKFS [22]
DTX	DTX off
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q= 20, 24, 28, 32 dB
ESDRU	α = 0.4, 0.6, 0.8
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based relying on convolution of raw mono clean speech sentences convolved with (FOA) Spatial Room Impulse Responses respective to various talker positions relative to a capture point and spatial (FOA) ambient noise mixing.
Binaural renderer	FOA to binaural (external) rendering
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Japanese, English

Table C.4-2: Preliminaries for Experiment P800-4

Trial #	Label	Sample	Condition	Bitrate	FER/Profile
1	c20		EVS	4x13.2	5%
2	c02		MNRU Q=32 dB	-	-
3	c06		ESDRU α = 0.8	-	-
4	c14		EVS	4x16.4	No errors
5	c08		ESDRU α = 0.55	-	-
6	c17		EVS	4x7.2	5%
7	c04		MNRU Q=24 dB	-	-
8	c01		Reference	-	-
9	c16		EVS	4x32	No errors
10	c07		ESDRU α = 0.675	-	-
11	c05		MNRU Q=20 dB	-	-
12	c22		EVS	4x24.4	5%

Table C.4-3: Test conditions for Experiment P800-4, clean speech under clean and impaired channel conditions

Label	Condition	Bitrate [kbps]	FER/Profile	ToR
c01	Reference	-	-	
c02	MNRU Q=32 dB	-	-	
c03	MNRU Q=28 dB	-	-	
c04	MNRU Q=24 dB	-	-	
c05	MNRU Q=20 dB	-	-	
c06	ESDRU $\alpha = 0.8$	-	-	
c07	ESDRU $\alpha = 0.675$	-	-	
c08	ESDRU $\alpha = 0.55$	-	-	
c09	EVS	3x7.2	No errors	
c10	EVS	4x7.2	No errors	
c11	EVS	4x8	No errors	
c12	EVS	4x9.6	No errors	
c13	EVS	4x13.2	No errors	
c14	EVS	4x16.4	No errors	
c15	EVS	4x24.4	No errors	
c16	EVS	4x32	No errors	
c17	EVS	4x7.2	5%	
c18	EVS	4x8	5%	
c19	EVS	4x9.6	5%	
c20	EVS	4x13.2	5%	
c21	EVS	4x16.4	5%	
c22	EVS	4x24.4	5%	
c23	EVS	4x32	5%	
c24	CuT	16.4	No errors	NWT c09
c25	CuT	24.4	No errors	NWT c10
c26	CuT	32	No errors	NWT c12 or BT c11
c27	CuT	48	No errors	NWT c14 or BT c13
c28	CuT	64	No errors	NWT c15 or BT c14
c29	CuT	80	No errors	NWT c15 or BT c14
c30	CuT	96	No errors	NWT c16 or BT c15
c31	CuT	24.4	5%	NWT c17
c32	CuT	32	5%	NWT c19 or BT c18
c33	CuT	48	5%	NWT c21 or BT c20
c34	CuT	64	5%	NWT c22 or BT c21
c35	CuT	80	5%	NWT c22 or BT c21
c36	CuT	96	5%	NWT c23 or BT c22

C.4.2 Content type categories and scene definitions

Table C.4-4: Content type categories and scene definitions

Category	Environment	Background	SNR [dB]	Overtalk [s] see NOTE 1	Bandwidth	Talker positions	Talker selection by panel see NOTE 2
cat 1	room_1_FOA	room_1_cleanbg_FOA	45	1	Max		P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	room_4_FOA	room_[1/4]_cleanbg_FOA	45	-1	Max		P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	room_2_FOA	room_2_cleanbg_FOA	45	1	Max		P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	room_5_FOA	room_[2/5]_cleanbg_FOA	45	-1	Max		P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	room_3_FOA	room_3_cleanbg_FOA	45	1	Max		P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2
cat 6	room_6_FOA	room_[3/6]_cleanbg_FOA	45	-1	Max		P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1

NOTE 1 Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences.

NOTE 2 All sentences by the 6 talkers shall be unique.

C.5 Experiment P800-5: FOA Speech+background Test

C.5.1 Experiment setup

Tables C.5-1 to C.5-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.5-1: Conditions for Experiment P800-5

Main Codec Conditions	
Candidate	CuT
bitrates	16.4, 24.4, 32, 48, 64, 80, 96 kbps
DTX	DTX off, DTX on
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	10, 15 dB
Error Conditions	No errors
Codec references	
Codec references	Multi-mono EVS
Bitrates	3x7.2, 4x7.2, 4x8, 4x9.6, 4x13.2, 4x16.4, 4x24.4, 4x32 kbps
Input level	-26 LKFS [22]
DTX	DTX off, DTX on
Input frequency mask	HP50
Noise	10, 15dB
Error Conditions	No errors
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q= 17, 21, 25, 29 dB
ESDRU	α = 0.4, 0.6, 0.8
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based relying on convolution of raw mono clean speech sentences convolved with (FOA) Spatial Room Impulse Responses respective to various talker positions relative to a capture point and spatial (FOA) ambient noise mixing
Binaural renderer	FOA to binaural (external) rendering
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments with different background types (car, office, street) and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Danish, German

Table C.5-2: Preliminaries for Experiment P800-5

Trial #	Label	Sample	Condition	Bitrate	DTX
1	c21		EVS	4x13.2	on
2	c02		MNRU Q=29 dB	-	-
3	c06		ESDRU α = 0.8	-	-
4	c14		EVS	4x16.4	off
5	c08		ESDRU α = 0.4	-	-
6	c18		EVS	4x7.2	on
7	c04		MNRU Q=21 dB	-	-
8	c01		Reference	-	-
9	c16		EVS	4x32	off
10	c07		ESDRU α = 0.6	-	-
11	c05		MNRU Q=17 dB	-	-
12	c23		EVS	4x24.4	on

Table C.5-3: Test conditions for Experiment P800-5, clean speech with background noise under clean channel conditions

Label	Condition	Bitrate [kbps]	DTX	ToR
c01	Reference	-	-	
c02	MNRU Q=29 dB	-	-	
c03	MNRU Q=25 dB	-	-	
c04	MNRU Q=21 dB	-	-	
c05	MNRU Q=17 dB	-	-	
c06	ESDRU α = 0.8	-	-	
c07	ESDRU α = 0.6	-	-	
c08	ESDRU α = 0.4	-	-	
c09	EVS	3x7.2	off	
c10	EVS	4x7.2	off	
c11	EVS	4x8	off	
c12	EVS	4x9.6	off	
c13	EVS	4x13.2	off	
c14	EVS	4x16.4	off	
c15	EVS	4x24.4	off	
c16	EVS	4x32	off	
c17	EVS	3x7.2	on	
c18	EVS	4x7.2	on	
c19	EVS	4x8	on	
c20	EVS	4x9.6	on	
c21	EVS	4x13.2	on	
c22	EVS	4x16.4	on	
c23	EVS	4x24.4	on	
c24	CuT	16.4	off	NWT c09
c25	CuT	24.4	off	NWT c10
c26	CuT	32	off	NWT c12 or BT c11
c27	CuT	48	off	NWT c14 or BT c13
c28	CuT	64	off	NWT c15 or BT c14
c29	CuT	80	off	NWT c15 or BT c14
c30	CuT	96	off	NWT c16 or BT c15
c31	CuT	16.4	on	NWT c17
c32	CuT	24.4	on	NWT c18
c33	CuT	32	On	NWT c20 or BT c19
c34	CuT	48	On	NWT c22 or BT c21
c35	CuT	64	On	NWT c23 or BT c22
c36	CuT	80	On	NWT c23 or BT c22

C.5.2 Content type categories and scene definitions

Table C.5-4: Content type categories and scene definitions

Category	Environment	Background	SNR[dB]	Overtalk [s] see NOTE 1	Bandwidth	Talker position	Talker selection by panel see NOTE 2

cat 1	car_1_FOA	car_1_bg_FOA	10	-1	Max		P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	car_[1/2]_FOA	car_[1/2]_bg_FOA	10	-1	Max		P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	out_1_FOA	street_1_bg_FOA	15	-1	Max		P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	out_[1/2]_FOA	[park_1_bg_FOA / nature_1_bg_FOA / event_1_bg_FOA / street_[1/2]_bg_FOA]	15	-1	Max		P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	room_[X]_FOA	office_1_bg_FOA	15	-1	Max		P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2
cat 6	room_[X]_FOA	[cafeteria_1_bg_FOA / mall_1_bg_FOA/ office[1/2]_bg_FOA]	15	-1	Max		P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1
NOTE 1	Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences.						
NOTE 2	All sentences by the 6 talkers shall be unique.						

C.6 Experiment P800-6: 1-Object Clean Speech Test

C.6.1 Experiment setup

Tables C.6-1 to C.6-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

The test Categories correspond to different talkers.

Table E.6.1: Conditions for Experiment P800-6

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48, 64 kbps
DTX	DTX ON and OFF
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5% FERs
Codec references	
Codec references	EVS
Bitrates	13.2, 16.4, 24.4, 32, 48, 64 kbps
DTX	DTX ON and OFF
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5% FERs
Other references	
Direct	-26 LKFS [22]
MNRU	Q= 15, 23, 31, 39, 47 dB
ESDRU	0.1, 0.3, 0.5, 0.7
Input frequency mask	HP50
Common Conditions	
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Number of talkers (categories)	3 male and 3 female
Number of speech samples	6 + 1 (preliminaries) sentence pairs per talker
Listening Level	73 dB SPL
Listeners	Naïve Listeners
Randomizations (panels)	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Languages	Japanese, English

Table C.6-2: Preliminaries for Experiment P800-6

Trial	**Label**	**Sample**	**Condition**	**Bitrate [kb/s]**	**DTX**	**FER [%]**
1	C19		EVS	24.4	Off	5
2	C04		MNRU 31	-	-	-
3	C21		EVS	13.2	On	0
4	C09		ESDRU 0.5	-	-	-
5	C14		EVS	32	Off	0
6	C02		MNRU 15	-	-	-
7	C12		EVS	16.4	Off	0

8	C01		Direct	-	-	-
9	C17		EVS	13.2	Off	5
10	C06		MNRU 47	-	-	-
11	C23		EVS	24.4	On	0
12	C07		ESDRU 0.1	-	-	-

Table C.6-3 Test conditions for Experiment P800-6

Label	**Condition**	**Bitrate [kb/s]**	**DTX**	**FER [%]**	**Ref condition**	**ToR**
C01	Direct		-	-	-	-
C02	MNRU 15		-	-	-	-
C03	MNRU 23		-	-	-	-
C04	MNRU 31		-	-	-	-
C05	MNRU 39		-	-	-	-
C06	MNRU 47		-	-	-	-
C07	ESDRU 0.1		-	-	-	-
C08	ESDRU 0.3		-	-	-	-
C09	ESDRU 0.5		-	-	-	-
C10	ESDRU 0.7		-	-	-	-
C11	EVS	13.2	Off	0	-	-
C12	EVS	16.4	Off	0	-	-
C13	EVS	24.4	Off	0	-	-
C14	EVS	32	Off	0	-	-
C15	EVS	48	Off	0	-	-
C16	EVS	64	Off	0		
C17	EVS	13.2	Off	5	-	-
C18	EVS	16.4	Off	5	-	-
C19	EVS	24.4	Off	5	-	-
C20	EVS	32	Off	5		
C21	EVS	13.2	On	0	-	-
C22	EVS	16.4	On	0	-	-
C23	EVS	24.4	On	0	-	-
C24	CuT	13.2	Off	0	C11	NWT
C25	CuT	16.4	Off	0	C12	NWT
C26	CuT	24.4	Off	0	C13	NWT
C27	CuT	32	Off	0	C14	NWT

C28	CuT	48	Off	0	C15	NWT
C29	CuT	64	Off	0	C16	NWT
C30	CuT	13.2	Off	5	C17	NWT
C31	CuT	16.4	Off	5	C18	NWT
C32	CuT	24.4	Off	5	C19	NWT
C33	CuT	32	Off	5	C20	NWT
C34	CuT	13.2	On	0	C21	NWT
C35	CuT	16.4	On	0	C22	NWT
C36	CuT	24.4	On	0	C23	NWT

C.6.2 Content type categories and scene definitions

A leading and trailing silence is present for each sample, in accordance with IVAS-7a. The metadata corresponds to the whole duration of the samples. This means that for moving objects, only a part of the trajectory corresponds to active speech. The following scenes are used:

a) Talker sitting at a table (elevation 0°), at different azimuths.

b) Standing talker (elevation 35°), at different azimuths.

c) Smaller talker (child) walking around a table in the positive sense (counterclockwise), elevation 0°. Azimuth varies continuously for the sentence pair.

d) Adult talker walking around a table in the negative sense (clockwise), elevation 35°. Azimuth varies continuously for the sentence pair.

e) Elevation displacement: Elevation varies continuously for the sentence pair. Azimuth is constant for a sentence pair, but different for each sentence pair.

f) Azimuth and elevation displacement: Azimuth and elevation vary continuously.

Each of the sentences uttered by a certain talker is encoded using different scene. To balance the test, in addition to listeners of each panel listening to all talkers (Categories), all scenes are also covered in each panel. Allocation of scenes to each panel is given in the Table below:

Table C.6-4: Allocation of scenes for each to talker (category cat 1-cat 6) and listening panel (P1-P6)

Category	**Talker initial elevation**	**Elevation change** see NOTE 1	**Talker initial azimuth**	**Azimuth change** see NOTE 2	**Panel**
	0°	static	0°	static	P1
cat 1:	35°	static	180°	static	P2
M1	0°	static	120°	1°/ frame	P3
	35°	static	180°	-1°/ frame	P4
	-90°	0.3°/ frame	120°	static	P5
	35°	-0.2°/ frame	0°	0.5°/ frame	P6
	35°	static	120°	static	P1
cat 2:	0°	static	60°	1°/ frame	P2
F1	35°	static	120°	-1°/ frame	P3

	-90°	0.3°/ frame	60°	static	P4
	35°	-0.2°/ frame	300°	0.5°/ frame	P5
	0°	static	300°	static	P6
	0°	static	0°	1°/ frame	P1
cat 3:	35°	static	60°	-1°/ frame	P2
M2	-90°	0.3°/ frame	0°	static	P3
	35°	-0.2°/ frame	240°	0.5°/ frame	P4
	0°	static	240°	static	P5
	35°	static	60°	static	P6
	35°	static	0°	-1°/ frame	P1
cat 4:	-90°	0.3°/ frame	300°	static	P2
F2	35°	-0.2°/ frame	180°	0.5°/ frame	P3
	0°	static	180°	static	P4
	35°	static	0°	static	P5
	0°	static	300°	1°/ frame	P6
	-90°	0.3°/ frame	240°	static	P1
cat 5:	35°	-0.2°/ frame	120°	0.5°/ frame	P2
M3	0°	static	120°	static	P3
	35°	static	300°	static	P4
	0°	static	240°	1°/ frame	P5
	35°	static	300°	-1°/ frame	P6
	35°	-0.2°/ frame	60°	0.5°/ frame	P1
cat 6:	0°	static	60°	static	P2
F3	35°	static	240°	static	P3
	0°	static	180°	1°/ frame	P4
	35°	static	240°	-1°/ frame	P5
	-90°	0.3°/ frame	180°	static	P6
NOTE 1 The positive sense for elevation is from bottom up NOTE 2 The positive sense for azimuth is counterclockwise					

C.7 Experiment P800-7: 2-Objects Clean Speech Test

C.7.1 Experiment setup

Tables C.7-1 to C.7-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

The test Categories correspond to different talker pairs.

Table C.7-1: Conditions for Experiment P800-7

Main Codec Conditions	
Candidate	CuT
Bitrates	16.4, 24.4, 32, 48, 64 kbps
DTX	DTX ON and OFF
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5% FERs
Codec references	
Codec references	EVS
Bitrates	2x8.0, 2x13.2, 2x16.4, 2x24.4, 2x32 kbps
DTX	DTX ON and OFF
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5% FERs
Other references	
Direct	-26 LKFS [22]
MNRU	Q= 15, 23, 31, 39, 47 dB
ESDRU	0.1, 0.3, 0.5, 0.7
Input frequency mask	HP50
Common Conditions	
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Number of talkers (categories)	3 male and 3 female
Number of speech samples	6 + 1 (preliminaries) sentence pairs per talker.
Listening Level	73 dB SPL
Listeners	Naïve Listeners
Randomizations (panels)	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Languages	Danish, Mandarin

Table C.7-2: Preliminaries for Experiment P800-7

Trial	**Label**	**Sample**	**Condition**	**Bitrate [kb/s]**	**DTX**	**FER [%]**
1	C18		EVS	2 x 16.4	Off	5
2	C04		MNRU 31	-	-	-
3	C20		EVS	2 x 8.0	On	0
4	C09		ESDRU 0.5	-	-	-
5	C14		EVS	2 x 24.4	Off	0
6	C02		MNRU 15	-	-	-
7	C12		EVS	2 x 13.2	Off	0
8	C01		Direct	-	-	-
9	C16		EVS	2 x 8.0	Off	5
10	C06		MNRU 47	-	-	-
11	C22		EVS	2 x 16.4	On	0
12	C07		ESDRU 0.1	-	-	-

Table C.7-3 Test conditions for Experiment P800-7:

Label	Condition	Bitrate [kb/s]	DTX	FER [%]	Ref condition	ToR
C01	Direct	-	-	-	-	-
C02	MNRU 15	-	-	-	-	-
C03	MNRU 23	-	-	-	-	-
C04	MNRU 31	-	-	-	-	-
C05	MNRU 39	-	-	-	-	-
C06	MNRU 47	-	-	-	-	-
C07	ESDRU 0.1	-	-	-	-	-
C08	ESDRU 0.3	-	-	-	-	-
C09	ESDRU 0.5	-	-	-	-	-
C10	ESDRU 0.7	-	-	-	-	-
C11	EVS	2 x 8.0	Off	0	-	-
C12	EVS	2 x 13.2	Off	0	-	-
C13	EVS	2 x 16.4	Off	0	-	-
C14	EVS	2 x 24.4	Off	0	-	-
C15	EVS	2 x 32	Off	0	-	-
C16	EVS	2 x 8.0	Off	5	-	-
C17	EVS	2 x 13.2	Off	5	-	-
C18	EVS	2 x 16.4	Off	5	-	-
C19	EVS	2 x 24.4	Off	5	-	-
C20	EVS	2 x 8.0	On	0	-	-
C21	EVS	2 x 13.2	On	0	-	-
C22	EVS	2 x 16.4	On	0	-	-
C23	EVS	2 x 24.4	On	0		
C24	CuT	16.4	Off	0	C11	NWT
C25	CuT	24.4	Off	0	C12	NWT
C26	CuT	32	Off	0	C13	NWT
C27	CuT	48	Off	0	C14	NWT
C28	CuT	64	Off	0	C15	NWT
C29	CuT	16.4	Off	5	C16	NWT
C30	CuT	24.4	Off	5	C17	NWT
C31	CuT	32	Off	5	C18	NWT
C32	CuT	48	Off	5	C19	NWT
C33	CuT	16.4	On	0	C20	NWT

C34	CuT	24.4	On	0	C21	NWT
C35	CuT	32	On	0	C22	NWT
C36	CuT	48	On	0	C23	NWT

C.7.2 Content type categories and scene definitions

The listening database consists of artificially created spatial audio samples from monophonic clean speech recordings where always 1 female and 1 male talker are combined in conversation-like scenarios following the Scene descriptions below.

A leading and trailing silence is present for each artificially created spatial audio sample, in accordance with IVAS-7a. The metadata corresponds to the whole duration of the sample. This means that for moving objects, only a part of the trajectory corresponds to active speech.

In one half of the samples, the 2nd talker's utterance follows the 1st talker's utterance simulating natural conversation. The gap between the utterances is set to 1 s. In the other half of the samples, the situation is similar, but the utterances partially overlap. The targeted overlap is also 1 s. Non-overlapping sentence pairs are used for Scenes a., c., and e. as described below. Overlapping sentence pairs are used for Scenes b., d., and f. The following scenes are used:

a) Two talkers sitting at a table (elevation 0°), at different azimuths. To increase positional variation, both the absolute azimuths and the difference of the azimuths of both talkers vary for each sentence pair.

b) Two standing talkers (elevation 35°), at different azimuths. To increase positional variation, both the absolute azimuths and the difference of the azimuths of both talkers vary for each sentence pair.

c) One talker sitting at a table (elevation 0°), second talker standing beside the table (elevation 45°). Non-overlapping utterances.

d) One talker sitting at a table (elevation 0°), second talker walking around the table (elevation 45°). The azimuth of the 2nd talker varies continually, positive sense is counter clockwise.

e) Two talkers walking side-by-side around the table (elevation 45°). The azimuth is the same for both talkers and varies continually.

f) Two talkers walking around the table in opposite directions (elevation 30°), starting at the same position. Azimuths of both talkers vary continually.

The following table lists the test Categories corresponding to different talker pairs. Each of the sentence pairs uttered by a certain talker pair is associated to different scenes. To balance the test, listeners of each panel listen to all talker pairs, and all scenes are covered in each panel.

Table C.7-4: Allocation of scenes for each talker pair (category cat 1 – cat 6) and listening panel (P1-P6)

Category	**Overtalk [s]** Note 1	**1st talker elevation**	**2nd talker elevation**	**1st talker initial azimuth**	**1st talker azimuth change** see NOTE 2	**2nd talker initial azimuth**	**2nd talker azimuth change** see NOTE 2	**Panel**
	-1	0°	0°	0°	static	50°	static	P1
cat 1:	1	35°	35°	10°	static	110°	static	P2
M1 + F1	-1	0°	45°	20°	static	170°	static	P3
	1	0°	45°	200°	static	30°	-1°/ frame	P4
	-1	45°	45°	340°	-1°/ frame	340°	-1°/ frame	P5
	1	30°	30°	120°	1°/ frame	120°	-1°/ frame	P6

	1	35°	35°	20°	static	170°	static	P1
cat 2:	-1	0°	45°	30°	static	230°	static	P2
M2 + F2	1	0°	45°	250°	static	340°	-1°/ frame	P3
	-1	45°	45°	290°	-1°/ frame	290°	-1°/ frame	P4
	1	30°	30°	180°	1°/ frame	180°	-1°/ frame	P5
	-1	0°	0°	10°	static	110°	static	P6
	-1	0°	45°	40°	static	290°	static	P1
cat 3:	1	0°	45°	300°	static	290°	-1°/ frame	P2
M3 + F3	-1	45°	45°	180°	1°/ frame	180°	1°/ frame	P3
	1	30°	30°	240°	1°/ frame	240°	1°/ frame	P4
	-1	0°	0°	20°	static	170°	static	P5
	1	35°	35°	30°	static	230°	static	P6
	1	0°	45°	50°	static	180°	1°/ frame	P1
cat 4:	-1	45°	45°	130°	1°/ frame	130°	1°/ frame	P2
F1 + M2	1	30°	30°	300°	1°/ frame	300°	-1°/ frame	P3
	-1	0°	0°	30°	static	230°	static	P4
	1	35°	35°	40°	static	290°	static	P5
	-1	0°	45°	50°	static	350°	static	P6
	-1	45°	45°	80°	1°/ frame	80°	1°/ frame	P1
cat 5:	1	30°	30°	0°	1°/ frame	0°	-1°/ frame	P2
F2 + M3	-1	0°	0°	40°	static	290°	static	P3
	1	35°	35°	50°	static	350°	static	P4
	-1	0°	45°	0°	static	50°	static	P5
	1	0°	45°	100°	static	130°	1°/ frame	P6
	1	30°	30°	60°	1°/ frame	60°	-1°/ frame	P1
cat 6:	-1	0°	0°	50°	static	350°	static	P2
F3 + M1	1	35°	35°	0°	static	50°	static	P3
	-1	0°	45°	10°	static	110°	static	P4
	1	0°	45°	150°	static	80°	1°/ frame	P5
	-1	45°	45°	30°	-1°/ frame	30°	-1°/ frame	P6
NOTE 1	Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences.							
NOTE 2	The positive sense for azimuth is counterclockwise							

C.8 Experiment P800-8: MASA Clean Speech Test

C.8.1 Experiment setup

Stereo-MASA inputs are tested in IVAS selection phase. The stereo-MASA inputs can have 1-direction spatial metadata (based, e.g., on MASA analysis of original FOA content) or 2-direction spatial metadata (based, e.g., on MASA analysis of original HOA2 content).

Tables C.8-1 to C.8-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.8-1: Conditions for Experiment P800-8 (Clean Speech)

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48, 64, 80 kbps
DTX	DTX off
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Codec references	
Codec references	Multi-mono EVS, Dual-mono EVS with unquantized metadata
Bitrates	3x7.2, 4x7.2, 4x8, 4x9.6, 4x16.4, 4x24.4 kbps 2x7.2, 2x8, 2x9.6, 2x16.4, 2x24.4 kbps
Input level	-26 LKFS [22]
DTX	DTX off
Input frequency mask	HP50
Noise	No noise
Error Conditions	0%, 5%
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q = 20, 24, 28, 32 dB
ESDRU	α = 0.4, 0.6, 0.8
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based generation according to convolution of raw mono clean speech sentences with FOA Spatial Room Impulse Responses corresponding to the talker positions relative to a capture point and spatial FOA background. MASA format generation from FOA according to MASA analysis
Binaural renderer	IVAS MASA C Reference Software binaural rendering
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Danish, German

Table C.8-2: Preliminaries for Experiment P800-8

Trial #	Label	Condition	Bitrate	FER/Profile
1	c23	EVS	4x9.6	5%
2	c02	MNRU Q=32 dB	-	-
3	c06	ESDRU α = 0.8	-	-
4	c13	EVS	4x16.4	No errors
5	c08	ESDRU α = 0.4	-	-
6	c21	EVS	4x7.2	5%
7	c04	MNRU Q=24 dB	-	-
8	c01	Reference	-	-
9	c14	EVS	4x24.4	No errors
10	c07	ESDRU α = 0.6	-	-
11	c05	MNRU Q=20 dB	-	-
12	c22	EVS	4x8	5%

Table C.8-3: Test conditions for Experiment P800-8, clean speech under clean and impaired channel conditions

Label	Condition	Bitrate [kbps]	FER/Profile	ToR
c01	Reference	-	-	
c02	MNRU Q=32 dB	-	-	
c03	MNRU Q=28 dB	-	-	
c04	MNRU Q=24 dB	-	-	
c05	MNRU Q=20 dB	-	-	
c06	ESDRU α = 0.8	-	-	
c07	ESDRU α = 0.6	-	-	
c08	ESDRU α = 0.4	-	-	
c09	EVS	3x7.2	No errors	
c10	EVS	4x7.2	No errors	
c11	EVS	4x8	No errors	
c12	EVS	4x9.6	No errors	
c13	EVS	4x16.4	No errors	
c14	EVS	4x24.4	No errors	
c15	EVS(+unquantized metadata)	2x7.2	No errors	
c16	EVS(+unquantized metadata)	2x8	No errors	
c17	EVS(+unquantized metadata)	2x9.6	No errors	
c18	EVS(+unquantized metadata)	2x16.4	No errors	
c19	EVS(+unquantized metadata)	2x24.4	No errors	
c20	EVS	3x7.2	5%	
c21	EVS	4x7.2	5%	
c22	EVS	4x8	5%	
c23	EVS	4x16.4	5%	
c24	EVS	4x24.4	5%	
c25	CuT	13.2	No errors	NWT c09 OR NWT c15
c26	CuT	16.4	No errors	NWT c10 OR NWT c15
c27	CuT	24.4	No errors	NWT c11 OR NWT c16
c28	CuT	32	No errors	NWT c12 OR NWT c17
c29	CuT	48	No errors	NWT c13 OR NWT c18
c30	CuT	64	No errors	NWT c14 OR NWT c19
c31	CuT	80	No errors	NWT c14 OR NWT c19
c32	CuT	13.2	5%	NWT c20
c33	CuT	16.4	5%	NWT c21
c34	CuT	24.4	5%	NWT c22
c35	CuT	48	5%	NWT c23
C36	CuT	64	5%	NWT c24

C.8.2 Content type categories and scene definitions (Exp P800-8: Clean speech)

Table C.8-4 Content type categories and scene definitions

Category	Environment	Background	SNR [dB]	Overtalk [s] see NOTE 1	Bandwidth	Talker positions	Talker selection by panel see NOTE 2
cat 1	room_1_MASA	room_1_cleanbg_MASA	45	1	Max		P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	room_4_MASA	room_4_cleanbg_MASA	45	-1	Max		P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	room_2_MASA	room_2_cleanbg_MASA	45	1	Max		P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	room_5_MASA	room_5_cleanbg_MASA	45	-1	Max		P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	room_3_MASA	room_3_cleanbg_MASA	45	1	Max		P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2
cat 6	room_6_MASA	room_6_cleanbg_MASA	45	-1	Max		P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1
NOTE 1	Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences.						
NOTE 2	All sentences by the 6 talkers shall be unique.						

C.9 Experiment P800-9: MASA Speech+Background Test

C.9.1 Experiment setup

Stereo-MASA inputs are tested in IVAS selection phase. The stereo-MASA inputs can have 1-direction spatial metadata (based, e.g., on MASA analysis of original FOA content) or 2-direction spatial metadata (based, e.g., on MASA analysis of original HOA2 content).

Tables C.9-1 to C.9-3 show conditions to be used for this experiment, list of preliminaries and full list of conditions, respectively.

Table C.9-1: Conditions for Experiment P800- 9 (Speech+Background)

Main Codec Conditions	
Candidate	CuT
Bitrates	13.2, 16.4, 24.4, 32, 48, 64, 80
DTX	DTX ON and OFF
Input level	-26 LKFS [22]
Input frequency mask	HP50
Noise	10, 15dB
Error Conditions	No errors
Codec references	
Codec references	Multi-mono EVS + Dual-mono with unquantized metadata
Bitrates	3x7.2, 4x7.2, 4x8, 4x9.6, 4x16.4, 4x24.4 kbps 2x7.2, 2x9.6, 2x16.4 kbps
Input level	-26 LKFS [22]
DTX	DTX On and OFF
Input frequency mask	HP50
Noise	10, 15dB
Error Conditions	No errors
Other references	
Direct	-26 LKFS [22]
P.50 MNRU	Q= 17, 21, 25, 29 dB
ESDRU	α = 0.4, 0.6, 0.8
Input frequency mask	HP50
Common Conditions	
Test item generation: pre-processing incl. spatialization	Model-based generation according to convolution of raw mono clean speech sentences with FOA Spatial Room Impulse Responses corresponding to the talker positions relative to a capture point and spatial FOA background. MASA format generation from FOA according to MASA analysis
Binaural renderer	IVAS MASA C Reference Software binaural rendering
Audio sampling frequency/bandwidth	48 kHz/maximum available audio bandwidth up to FB
Kind of samples	Sentence pair uttered by different talkers and genders (3 male and 3 female)
Number of categories	6 Different environments and talker interactions
Number of samples	6 + 1 (preliminaries) samples per category
Listening Level	73 dB SPL
Listeners	Naïve listeners
Randomizations	6 panels of 5 listeners
Rating Scale	Following clause 7.1.2.1
Listening System	Headphones, in accordance with clause 7.1.4
Listening Environment	No room noise
Languages	Japanese, French

Table C.9-2: Preliminaries for Experiment P800-9

Trial #	Label	Sample	Condition	Bitrate	DTX
1	c21		EVS	4x9.6	on
2	c02		MNRU Q=29 dB	-	-
3	c06		ESDRU α = 0.8	-	-
4	c13		EVS	4x16.4	off
5	c08		ESDRU α = 0.4	-	-
6	c19		EVS	4x7.2	on
7	c04		MNRU Q=21 dB	-	-
8	c01		Reference	-	-
9	c14		EVS	4x24.4	off
10	c07		ESDRU α = 0.6	-	-
11	c05		MNRU Q=17 dB	-	-
12	c20		EVS	4x8	on

Table C.9-3: Test conditions for Experiment P800-9, speech and background under clean channel conditions

Label	Condition	Bitrate [kbps]	DTX	ToR
c01	Reference	-	-	
c02	MNRU Q=29 dB	-	-	
c03	MNRU Q=25 dB	-	-	
c04	MNRU Q=21 dB	-	-	
c05	MNRU Q=17 dB	-	-	
c06	ESDRU α = 0.8	-	-	
c07	ESDRU α = 0.6	-	-	
c08	ESDRU α = 0.4	-	-	
c09	EVS	3x7.2	off	
c10	EVS	4x7.2	off	
c11	EVS	4x8	off	
c12	EVS	4x9.6	off	
c13	EVS	4x16.4	off	
c14	EVS	4x24.4	off	
c15	EVS(+unquantized metadata)	2x7.2	on	
c16	EVS(+unquantized metadata)	2x9.6	on	
c17	EVS(+unquantized metadata)	2x16.4	on	
c18	EVS	3x7.2	on	
c19	EVS	4x7.2	on	
c20	EVS	4x8	on	
c21	EVS	4x9.6	on	
c22	EVS	4x16.4	on	
c23	EVS	4x24.4	on	
c24	CuT	13.2	off	NWT c09
c25	CuT	16.4	off	NWT c10
c26	CuT	24.4	off	NWT c11
c27	CuT	32	off	NWT c12
c28	CuT	48	off	NWT c13
c29	CuT	64	off	NWT c14
c30	CuT	80	off	NWT c14
c31	CuT	13.2	on	NWT c18 or NWT c15
c32	CuT	16.4	on	NWT c19 or NWT c15
c33	CuT	24.4	on	NWT c20
c34	CuT	32	on	NWT c21 or NWT c16
c35	CuT	48	on	NWT c22 or NWT c17
c36	CuT	64	on	NWT c23

C.9.2 Content type categories and scene definitions (Exp P800-9: Speech+Background)

Table C.9-4: Content type categories and scene definitions

Category	Environment[1]	Background[2]	SNR [dB]	Overtalk [s][3]	Bandwidth	Talker positions[4]	Talker selection

							by panel([5]
cat 1	car_1_MASA	car_1_bg_MASA	10	-1	Max		P1: f1m1 P2: m2f2 P3: f3m3 P4: m1f1 P5: f2m2 P6: m3f3
cat 2	car_[1/2]_MASA	car_[1/2]_bg_MASA	10	-1	Max		P1: m3f3 P2: f1m1 P3: m2f2 P4: f3m3 P5: m1f1 P6: f2m2
cat 3	out_1_MASA	street_1_bg_MASA	15	-1	Max		P1: f2m2 P2: m3f3 P3: f1m1 P4: m2f2 P5: f3m3 P6: m1f1
cat 4	out_[1/2]_MASA	[park_1_bg_MASA / nature_1_bg_MASA / event_1_bg_MASA / street_[1/2]_bg_MASA]	15	-1	Max		P1: m1f1 P2: f2m2 P3: m3f3 P4: f1m1 P5: m2f2 P6: f3m3
cat 5	room_[X]_MASA	office_1_bg_MASA	15	-1	Max		P1: f3m3 P2: m1f1 P3: f2m2 P4: m3f3 P5: f1m1 P6: m2f2
cat 6	room_[X]_MASA	[cafeteria_1_bg_MASA / mall_1_bg_MASA/ office[1/2]_bg_MASA]	15	-1	Max		P1: m2f2 P2: f3m3 P3: m1f1 P4: f2m2 P5: m3f3 P6: f1m1
NOTE 1	Overtalk [s] means the duration in seconds by which the two sentences in the sound item uttered by different talkers are overlapping. A negative number means that there is a corresponding pause between the two sentences.						
NOTE 2	All sentences by the 6 talkers shall be unique.						

C.10 Experiment BS1534-1a: Stereo

Table C.10-1: Conditions (BS1534-1a Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with stereo audio input at 48 and 64 kbps DTX off at 0% FER
Codec references	
Codec references	Dual-mono EVS 2*24.4 kbps, 2*32 kbps, 2*48 kbps DTX off at 0% FER
Other references	
Reference	Direct signal, Nominal input level
Hidden Reference	Direct signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered signal, nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality headphone for diotic presentation, in accordance with clause 7.1.4
Listening Environment	No room noise

Table C.10-2: Test conditions for Experiment BS1534-1a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	2x24.4	-
c04	EVS	2x32	-
c05	EVS	2x48	-
c06	CuT	48	NWT c04 OR BT c03
c07	CuT	64	NWT c05 OR BT c04

C.11 Experiment BS1534-1b: Stereo

Table C.11-1: Conditions (BS1534-1b Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with stereo audio input at 96 and 128 kbps DTX off at 0% FER
Codec references	
Codec references	Dual-mono EVS 2*48 kbps, 2*64 kbps, 2*96 kbps DTX off at 0% FER
Other references	
Reference	Direct signal, Nominal input level
Hidden Reference	Direct signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered signal, nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener

Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality headphone for diotic presentation, in accordance with clause 7.1.4
Listening Environment	No room noise

Table C.11-2: Test conditions for Experiment BS1534-1b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	2x48	-
c04	EVS	2x64	-
c05	EVS	2x96	-
c06	CuT	96	NWT c04 OR BT c03
c07	CuT	128	NWT c05 OR BT c04

C.12 Experiment BS1534-2a: Multi-channel 5.1

Table C.12-1: Conditions (BS1534-2a Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 64, 96 kbps at 0% FER
Codec references	
Codec references	Multi-mono EVS operated at 5*13.2, 5*16.4, 5*24.4 kbps with the LFE channel processed using EVS operated at 9.6 kbps NB (for all Codec references) at 0% FER
Other references	
Reference / Direct	Direct 5.1 signal, nominal level
Hidden Reference	Direct 5.1 signal, nominal level
LP7 anchor	7 kHz lowpass filtered direct 5.1 signal, nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection *BS.1534* tests.
Loudspeaker Rendering	5.1 channels direct playback
Audio sampling frequency / bandwidth	48 kHz / maximum available audio bandwidth (SWB, FB)
Input frequency mask	*20KBP*

Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listener
Rating Scale	Continuous BS.1534 scale: 0-100
Listening System	5.1 high-quality loudspeaker setup following IVAS-7a
Listening Environment	No room noise

Table C.12-2: Test conditions for Experiment BS1534-2a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	5x13.2	-
c04	EVS	5x16.4	-
c05	EVS	5x24.4	-
c06	CuT	64	NWT c04 OR BT c03
c07	CuT	96	NWT c05 OR BT c04

C.13 Experiment BS1534-2b: Multi-channel 5.1

Table C.13-1: Conditions (BS1534-2b Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 128, 160 kbps at 0% FER
Codec references	
Codec references	Multi-mono EVS operated at 5*24.4, 5*32, 5*48 kbps with the LFE channel processed using EVS operated at 9.6 kbps NB (for all Codec references) at 0% FER
Other references	
Reference	Direct 5.1 signal, nominal input level
Hidden Reference	Direct 5.1 signal, nominal input level
LP7 anchor	7 kHz lowpass filtered direct 5.1 signal, nominal level
Common Conditions	*(see BS1534-2a for full list)*
Audio sampling frequency / bandwidth	48 kHz / FB

Table C.13-2: Test conditions for Experiment BS1534-2b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	5x24.4	-
c04	EVS	5x32	-
c05	EVS	5x48	-
c06	CuT	128	NWT c04 OR BT c03
c07	CuT	160	NWT c05 OR BT c04

C.14 Experiment BS1534-3a: Multi-channel 7.1+4

Table C.14-1: Conditions (BS1534-3a Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with multi-channel 7.1+4 input at 128 and 160 kbps DTX off at 0% FER
Codec references	
Codec references	Multi-mono EVS coding multi-channel 7.1+4 input at 11*9.6 kbps, 11*13.2 kbps, 11*16.4 kbps DTX off at 0% FER LFE coded with 9.6kbps NB (IVAS-3)
Other references	
Reference	Direct signal, Nominal input level
Hidden Reference	Direct signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered signal, nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality loudspeaker: 7.1+4 overhead speaker setup with the configuration following IVAS-7a

Listening Environment	No room noise

Table C.14-2: Test conditions for Experiment BS1534-3a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	11x9.6	-
c04	EVS	11x13.2	-
c05	EVS	11x16.4	-
c06	CuT	128	NWT c04 OR BT c03
c07	CuT	160	NWT c05 OR BT c04

C.15 Experiment BS1534-3b: Multi-channel 7.1+4

Table C.15-1: Conditions (BS1534-3b Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with multi-channel 7.1+4 audio input at 384 and 512 kbps DTX off at 0% FER
Codec references	
Codec references	Multi-mono EVS coding multi-channel 7.1+4 input at 11*32 kbps, 11*48 kbps, 11*64 kbps DTX off at 0% FER LFE coded with 9.6kbps NB (IVAS-3)
Other references	
Reference	Direct signal, Nominal input level
Hidden Reference	Direct signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered signal, nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality loudspeaker: 7.1+4 overhead speaker setup with the configuration following IVAS-7a

Listening Environment	No room noise

Table C.15-2: Test conditions for Experiment BS1534-3b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	11x32	-
c04	EVS	11x48	-
c05	EVS	11x64	-
c06	CuT	384	NWT c04 OR BT c03
c07	CuT	512	NWT c05 OR BT c04

C.16 Experiment BS1534-4a: FOA

Table C.16-1: Conditions (BS1534-4a, Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with audio input truncated to FOA, HOA3 output, at 96 kbps, 128 kbps, 160 kbps
Codec references	
Codec references	Multi-mono EVS operated with audio input truncated to FOA at 4*24.4 kbps, 4*32 kbps, 4*48 kbps
Other references	
Reference	Direct rendering of HOA3 signal, Nominal input level
Hidden Reference	Direct rendering of HOA3 signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered direct rendered HOA3 signal: nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Binaural renderer	Ambisonics to binaural (external) rendering
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality headphone for diotic presentation, in accordance with 7.1.4

Listening Environment	No room noise

Table C.16-2: Test conditions for Experiment BS1534-4a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x24.4	-
c04	EVS	4x32	-
c05	EVS	4x48	-
c06	CuT	96	NWT c04 OR BT c03
c07	CuT	128	NWT c05 OR BT c04
c08	CuT	160	NWT c05 OR BT c04

C.17 Experiment BS1534-4b: HOA2

Table C.17-1: Conditions (BS1534-4b, Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with audio input truncated to HOA2, HOA3 output, at 160, 192 kbps kbps.
Codec references	
Codec references	Multi-mono EVS operated with audio input truncated to FOA at 4*32 kbps, 4*48 kbps, 4*64 kbps.
Other references	
Reference	Direct rendering of HOA3 signal, Nominal input level
Hidden Reference	Direct rendering of HOA3 signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered direct rendered HOA3 signal: nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Binaural renderer	Ambisonics to binaural (external) rendering
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	High-quality headphone for diotic presentation, in accordance with clause 7.1.4

Listening Environment	No room noise

Table C.17-2: Test conditions for Experiment BS1534-4b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x32	-
c04	EVS	4x48	-
c05	EVS	4x64	-
c06	CuT	160	NWT c04 OR BT c03
c07	CuT	192	NWT c05 OR BT c04

C.18 Experiment BS1534-5a: HOA3

Table C.18-1: Conditions (BS1534-5a, Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with HOA3 input at 192 kbps, 256 kbps.
Codec references	
Codec references	Multi-mono EVS operated with audio input truncated to FOA at 4*48 kbps, 4*64, 4*96 kbps.
Other references	
Reference	Direct rendering of HOA3 signal, Nominal input level
Hidden Reference	Direct rendering of HOA3 signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered direct rendered HOA3 signal: nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Binaural renderer	Ambisonics to binaural (external) rendering
Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100

Listening System	High-quality headphone for diotic presentation, in accordance with clause 7.1.4.
Listening Environment	No room noise

Table C.18-2: Test conditions for Experiment BS1534-5a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x48	-
c04	EVS	4x64	-
c05	EVS	4x96	-
c06	CuT	192	NWT c04 OR BT c03
c07	CuT	256	NWT c05 OR BT c04

C.19 Experiment BS1534-5b: HOA3

Table C.19-1: Conditions (BS1534-5b, Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated with HOA3 input at 384 kbps, 512 kbps.
Codec references	
Codec references	Multi-mono EVS operated with audio input truncated to FOA at 4*96 kbps, 4*128 kbps.
Other references	
Reference	Direct rendering of HOA3 signal, Nominal input level
Hidden Reference	Direct rendering of HOA3 signal, Nominal input level
LP7 anchor	7 kHz lowpass filtered direct rendered HOA3 signal: nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection BS.1534 tests.
Renderer	Ambisonics to loudspeaker (external) rendering

Audio sampling frequency/bandwidth	48 kHz/FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listeners
Rating Scale	Continuous BS.1534 scale from 0-100
Listening System	Calibrated and conformant 7.1.4 listening room with single listener in center.
Listening Environment	No room noise

Table C.19-2: Test conditions for Experiment BS1534-5b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x96	-
c04	EVS	4x128	-
c05	CuT	384	NWT c04 OR BT c03
c06	CuT	512	NWT c04

C.20 Experiment BS1534-6a: Objects

Test with 3 Objects in BS1534-6a at mid bitrates.

Table C.20-1: Conditions (BS1534-6a Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 48, 64, 96 kbps
Codec references	
Codec references	Multi-mono EVS operated at 3*16.4, 3*24.4, 3*32 kbps
Other references	
Reference / Direct	Direct Object signal(s), nominal input level

Hidden Reference	Direct Object signal(s), nominal input level
LP7 anchor	7 kHz lowpass filtered direct Object signal(s), nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection *BS.1534* tests.
Binaural rendering	Objects to binaural (external) rendering
Audio sampling frequency / bandwidth	48 kHz / FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listener
Rating Scale	Continuous BS.1534 scale: 0-100
Listening System	High-quality headphones, diotic presentation, in accordance with clause 7.1.4
Listening Environment	No noise

Table C.20-2: Test conditions for Experiment BS1534-6a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	3x16.4	-
c04	EVS	3x24.4	-
c05	EVS	3x32	-
c06	CuT	48	NWT c03
c07	CuT	64	NWT c04
c08	CuT	96	NWT c05

C.21 Experiment BS1534-6b: Objects

Test with 4 Objects in BS1534-6b at mid/high bitrates.

Table C.21-1: Conditions (BS1534-6b Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 96, 128, 256 kbps with DTX off at 0% FER
Codec references	
Codec references	Multi-mono EVS operated at 4*24.4, 4*32, 4*64 kbps with DTX off at 0% FER
Other references	
Reference / Direct	Direct Object signal(s), nominal input level
Hidden Reference	Direct Object signal(s), nominal input level
LP7 anchor	7 kHz lowpass filtered direct Object signal(s), nominal level
Common Conditions	*(see BS1534-6a for full list)*

Table C.21-2: Test conditions for Experiment BS1534-6b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x24.4	-
c04	EVS	4x32	-
c05	EVS	4x64	-
c06	CuT	96	NWT c03
c07	CuT	128	NWT c04
c08	CuT	256	NWT c05

C.22 Experiment BS1534-7a: MASA

Table C.22-1: Conditions (BS1534-7a Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 96, 128 kbps
Codec references	
Codec references	Multi-mono EVS operated at 4*32, 4*48 kbps with DTX off at 0% FER
	Multi-mono EVS with unquantized metadata operated at 2*32, 2*48 kbps with DTX off at 0% FER
Other references	
Reference / Direct	Direct MASA signal(s), nominal input level
Hidden Reference	Direct MASA signal(s), nominal input level
LP7 anchor	7 kHz lowpass filtered direct MASA signal(s), nominal level
Common Conditions	
Test item generation	According to material collection procedure for IVAS selection *BS.1534* tests
Binaural rendering	MASA to binaural (external) rendering
Audio sampling frequency / bandwidth	48 kHz / FB
Input frequency mask	20KBP
Nominal output loudness	-26 LKFS [22]
Listening Level	Adjusted by listener
Listeners	Experienced Listeners
Randomizations	Individual per listener
Rating Scale	Continuous BS.1534 scale: 0-100

Listening System	High-quality headphones, diotic presentation, in accordance with clause 7.1.4
Listening Environment	No noise

Table C.22-2: Test conditions for Experiment BS1534-7a

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x32	-
c04	EVS	4x48	-
c05	EVS	2x32, unquantized MD	-
c06	EVS	2x48, unquantized MD	-
c07	CuT	96	NWT c03 OR NWT c05
c08	CuT	128	NWT c04 OR NWT c06

C.23 Experiment BS1534-7b: MASA

Table C.23-1: Conditions (BS1534-7b Generic Audio)

Main Codec Conditions	
Codec under Test (CuT)	IVAS candidate operated at 192, 256 kbps with DTX off at 0% FER
Codec references	
Codec references	Multi-mono EVS operated at 4*64, 4*96 kbps with DTX off at 0% FER
	Multi-mono EVS with unquantized metadata operated at 2*64, 2*96 kbps with DTX off at 0% FER
Other references	
Reference / Direct	Direct MASA signal(s), nominal input level
Hidden Reference	Direct MASA signal(s), nominal input level
LP7 anchor	7 kHz lowpass filtered direct MASA signal(s), nominal level
Common Conditions	*(see BS1534-7a for full list)*

Table C.23-2: Test conditions for Experiment BS1534-7b

Label	Condition	Bitrate [kbps]	ToR
c01	Reference	-	-
c02	LP7 anchor	-	-
c03	EVS	4x64	-
c04	EVS	4x96	-
c05	EVS	2x64, unquantized MD	-
c06	EVS	2x96, unquantized MD	-
c07	CuT	192	NWT c03 OR NWT c05
c08	CuT	256	NWT c04 OR NWT c06

Annex D:
IVAS Characterization Phase Testplan

NOTE: The IVAS Characterization Phase Testplan will be added after completion of the characterization phase.

Annex E:
Relevant Permanent Documents

E.1 Overview

The standardization of the IVAS codec is described in a series of permanent project documents. They contain the most important guidelines, rules and decisions. The following clauses contain the relevant permanent documemnts as electronic attachments.

The latest version of these documents can be found in the following link:
https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/IVAS_Permanent_Documents

E.2 IVAS-3: IVAS Performance Requirements

The IVAS-3 Permanent Document: "IVAS Performance Requirements" can be found in the electronic attachment to Annex E.2.

E.3 IVAS-4: IVAS Design Constraints

The IVAS-4 Permanent Document: "IVAS Design Constraints" can be found in the electronic attachment to Annex E.3.

E.4 IVAS-5: IVAS Selection Rules

The IVAS-5 Permanent Document: "IVAS Selection Rules" can be found in the electronic attachment to Annex E.4.

E.5 IVAS-6: IVAS Selection Deliverables

The IVAS-6 Permanent Document: "IVAS Selection Deliverables" can be found in the electronic attachment to Annex E.5.

E.6 IVAS-7a: Processing Plan for Selection Phase

The IVAS-7a Permanent Document: "Processing Plan for Selection Phase" can be found in the electronic attachment to Annex E.6.

E.7 IVAS-7b: Processing Plan for Characterization Phase

The IVAS-7b Permanent Document: "Processing Plan for Characterization Phase" can be found in the electronic attachment to Annex E.7.

E.8 IVAS-8a: Test Plan for Selection Phase

The IVAS-8a Permanent Document: "Test Plan for Selection Phase" can be found in the electronic attachment to Annex E.8.

E.9 IVAS-8b: Test Plan for Characterization Phase

The IVAS-8b Permanent Document: "Test Plan for Selection Phase" can be found in the electronic attachment to Annex E.9.

E.10 ISAR Permanent Document: Testing Aspects for Phase/Track 2/a

The ISAR Permanent Document “Testing Aspects for Phase/Track 2/”a can be found in the electronic attachment to Annex E.10.

Annex F:
Subjective Test Raw Data and Analyses

F.1 IVAS Selection Phase: Global Analysis Laboratory report

The Global Analysis Laboratory (GAL) report for the IVAS selection phase including raw subjective tests data can be found in the electronic attachment to Annex F.1.

F.2 IVAS Characterization Phase: Global Analysis Laboratory report

The Global Analysis Laboratory (GAL) report for the IVAS characterization phase including raw subjective tests data can be found in the electronic attachment to Annex F.2

Annex G (informative): Change history

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2024-04	SA4#127-bis-e	S4-240626				Initial Version	0.0.1
2024-04	SA4#127-bis-e	S4-240792				Endorsed by SA4	0.1.0
2024-05	S4#128	S4-240982				New input by rapporteurs, lots of added text	0.1.1
2024-05	S4#128	S4-241245				Submitted to SA4 for agreement	0.2.0
2024-05	S4#128	S4-241342				Submitted to SA4 for agreement	0.3.0
2024-06						Version 1.0.0 created by MCC	1.0.0
2024-06						Version 18.0.0 created by MCC upon TSG approval	18.0.0
2024-09	SA#105	SP-241114	001	2	F	Corrections to TR 26.997	18.1.0
2025-12	SA#110	SP-251435	0002	1	B	Documentation of IVAS Characterization Test Results	19.0.0

History

Version	Date	Status
V19.0.0	February 2026	Publication