|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A logo of a flag  Description automatically generated | World Telecommunication Standardization Assembly (WTSA-24) New Delhi, 15–24 October 2024 | | | A blue logo with a globe and lightning  Description automatically generated |
|  | | | | |
|  | |  | | |
| PLENARY MEETING | | Document | 12-E | |
|  | | October 2024 | | |
|  | | Original: English | | |
|  | | | | |
| ITU‑T Study Group 12 Performance, quality of service (QoS) and quality of experience (QoE) | | | | |
| Report of ITU-T SG12 to the World Telecommunication Standardization Assembly (WTSA-24), Part II: Questions proposed for study during the next study period (2025-2028) | | | | |
|  | | | | |
|  | | | | |

|  |  |  |
| --- | --- | --- |
| **Abstract:** | This contribution contains the text of the Study Group 12 Questions proposed for approval by the Assembly for the 2025-2028 study period. | |
| **Contact:** | Ms Tania Villa Chair ITU-T SG12 Mexico | Tel: +52 55 50154146 E-mail: [tania.villa@ift.org.mx](mailto:tania.villa@ift.org.mx) |

Note by the TSB:

The report of Study Group 12 to the WTSA-24 is presented in the following documents:

Part I: **Document** [**11**](https://www.itu.int/md/T22-WTSA.24-C-0011/en) – General

Part II: **Document** [**12**](https://www.itu.int/md/T22-WTSA.24-C-0012/en) – Questions proposed for study during the study period 2025-2028

[1 List of Questions proposed by Study Group 12 5](#_Toc170206080)

[2 Wording of Questions 6](#_Toc170206081)

[DRAFT QUESTION A/12 SG12 work programme and quality of service/quality of experience (QoS/QoE) coordination in ITU-T 6](#_Toc170206082)

[A.1 Motivation 6](#_Toc170206083)

[A.2 Question 6](#_Toc170206084)

[A.3 Tasks 6](#_Toc170206085)

[A.4 Relationships 7](#_Toc170206086)

[DRAFT QUESTION B/12 Definitions, guides and frameworks related to quality of service/quality of experience (QoS/QoE) 8](#_Toc170206087)

[B.1 Motivation 8](#_Toc170206088)

[B.2 Question 8](#_Toc170206089)

[B.3 Tasks 8](#_Toc170206090)

[B.4 Relationships 8](#_Toc170206091)

[DRAFT QUESTION C/12 Objective methods for speech and audio evaluation in vehicles 10](#_Toc170206092)

[C.1 Motivation 10](#_Toc170206093)

[C.2 Question 10](#_Toc170206094)

[C.3 Tasks 11](#_Toc170206095)

[C.4 Relationships 12](#_Toc170206096)

[DRAFT QUESTION D/12 Telephonometric methodologies for handset and headset terminals 13](#_Toc170206097)

[D.1 Motivation 13](#_Toc170206098)

[D.2 Question 13](#_Toc170206099)

[D.3 Tasks 13](#_Toc170206100)

[D.4 Relationships 14](#_Toc170206101)

[DRAFT QUESTION E/12 Analysis methods for speech and audio using complex measurement signals 15](#_Toc170206102)

[E.1 Motivation 15](#_Toc170206103)

[E.2 Question 15](#_Toc170206104)

[E.3 Tasks 16](#_Toc170206105)

[E.4 Relationships 16](#_Toc170206106)

[DRAFT QUESTION F/12 Methodologies, tools and test plans for the subjective assessment of speech, audio and audiovisual quality interactions 18](#_Toc170206107)

[F.1 Motivation 18](#_Toc170206108)

[F.2 Question 18](#_Toc170206109)

[F.3 Tasks 19](#_Toc170206110)

[F.4 Relationships 19](#_Toc170206111)

[DRAFT QUESTION G/12 Perceptual-based objective methods and corresponding evaluation guidelines for voice and audio quality measurements in telecommunication services 20](#_Toc170206112)

[G.1 Motivation 20](#_Toc170206113)

[G.2 Question 20](#_Toc170206114)

[G.3 Tasks 21](#_Toc170206115)

[G.4 Relationships 21](#_Toc170206116)

[DRAFT QUESTION H/12 Conferencing and telemeeting assessment 22](#_Toc170206117)

[H.1 Motivation 22](#_Toc170206118)

[H.2 Question 22](#_Toc170206119)

[H.3 Tasks 23](#_Toc170206120)

[H.4 Relationships 24](#_Toc170206121)

[DRAFT QUESTION I/12 Operational aspects of telecommunication network service quality and end-to-end performance considerations 25](#_Toc170206122)

[I.1 Motivation 25](#_Toc170206123)

[I.2 Question 25](#_Toc170206124)

[I.3 Tasks 26](#_Toc170206125)

[I.4 Relationships 26](#_Toc170206126)

[DRAFT QUESTION J/12 Quality of experience (QoE), quality of service (QoS) and performance requirements and assessment methods for multimedia applications 27](#_Toc170206127)

[J.1 Motivation 27](#_Toc170206128)

[J.2 Question 27](#_Toc170206129)

[J.3 Tasks 27](#_Toc170206130)

[J.4 Relationships 28](#_Toc170206131)

[DRAFT QUESTION K/12 Development of models and tools for multimedia quality assessment of packet-based video services 29](#_Toc170206132)

[K.1 Motivation 29](#_Toc170206133)

[K.2 Question 30](#_Toc170206134)

[K.3 Tasks 31](#_Toc170206135)

[K.4 Relationships 32](#_Toc170206136)

[DRAFT QUESTION L/12 Parametric and E-model-based planning, prediction and monitoring of conversational speech and audio-visual quality 33](#_Toc170206137)

[L.1 Motivation 33](#_Toc170206138)

[L.2 Question 33](#_Toc170206139)

[L.3 Tasks 34](#_Toc170206140)

[L.4 Relationships 35](#_Toc170206141)

[DRAFT QUESTION M/12 Performance of packet-based networks and other networking technologies 36](#_Toc170206142)

[M.1 Motivation 36](#_Toc170206143)

[M.2 Question 36](#_Toc170206144)

[M.3 Tasks 38](#_Toc170206145)

[M.4 Relationships 38](#_Toc170206146)

[DRAFT QUESTION N/12 Objective and subjective methods for evaluating perceptual audiovisual quality in multimedia and television services 40](#_Toc170206147)

[N.1 Motivation 40](#_Toc170206148)

[N.2 Question 41](#_Toc170206149)

[N.3 Tasks 43](#_Toc170206150)

[N.4 Relationships 43](#_Toc170206151)

[DRAFT QUESTION O/12 Perceptual and field assessment principles for quality of service (QoS) and quality of experience (QoE) of digital financial services (DFS) 44](#_Toc170206152)

[O.1 Motivation 44](#_Toc170206153)

[O.2 Question 44](#_Toc170206154)

[O.3 Tasks 45](#_Toc170206155)

[O.4 Relationships 45](#_Toc170206156)

# 1 List of Questions proposed by Study Group 12

| Question number | Question title | Status |
| --- | --- | --- |
| A/12 | SG12 work programme and quality of service/quality of experience (QoS/QoE) coordination in ITU-T | Continuation of Question 1/12 |
| B/12 | Definitions, guides and frameworks related to quality of service/quality of experience (QoS/QoE) | Continuation of Question 2/12 |
| C/12 | Objective methods for speech and audio evaluation in vehicles | Continuation of Question 4/12 |
| D/12 | Telephonometric methodologies for handset and headset terminals | Continuation of Question 5/12 |
| E/12 | Analysis methods for speech and audio using complex measurement signals | Continuation of Question 6/12 |
| F/12 | Methodologies, tools and test plans for the subjective assessment of speech, audio and audiovisual quality interactions | Continuation of Question 7/12 |
| G/12 | Perceptual-based objective methods and corresponding evaluation guidelines for voice and audio quality measurements in telecommunication services | Continuation of Question 9/12 |
| H/12 | Conferencing and telemeeting assessment | Continuation of Question 10/12 |
| I/12 | Operational aspects of telecommunication network service quality and end-to-end performance considerations | Continuation of Question 12/12 |
| J/12 | Quality of experience (QoE), quality of service (QoS) and performance requirements and assessment methods for multimedia applications | Continuation of Question 13/12 |
| K/12 | Development of models and tools for multimedia quality assessment of packet-based video services | Continuation of Question 14/12 |
| L/12 | Parametric and E-model-based planning, prediction and monitoring of conversational speech and audio-visual quality | Continuation of Question 15/12 |
| M/12 | Performance of packet-based networks and other networking technologies | Continuation of Question 17/12 |
| N/12 | Objective and subjective methods for evaluating perceptual audiovisual quality in multimedia and television services | Continuation of Question 19/12 |
| O/12 | Perceptual and field assessment principles for quality of service (QoS) and quality of experience (QoE) of digital financial services (DFS) | Continuation of Question 20/12 |

# 2 Wording of Questions

## DRAFT QUESTION A/12 SG12 work programme and quality of service/quality of experience (QoS/QoE) coordination in ITU-T

(Continuation of Question 1/12)

### A.1 Motivation

A Study Group should identify new or revised Questions to enable its programme of work to evolve. But for new work proposals, a home is needed when they are not directly related to existing Questions. This Question provides that home. Additionally, this Question can address actions requested of the Study Group that have no associated Question or Rapporteur.

SG12 is the Lead Study Group on QoS/QoE, and this Question is where SG12 can provide cross-ITU SG coordination for the many aspects of QoS in order to foster consistency within the ITU, and with related external organizations (e.g., 3GPP, IETF).

SG12 works proactively to help bridge the standardization gap in the area of QoS/QoE. The Regional Groups for Africa and the Americas were created by SG12 in support of the needs of two of the world's regions, and any issues related to SG12 being its parent group are addressed in this Question.

Consistent with the above, this Question itself does not usually produce any Recommendations.

### A.2 Question

This Question asks, but is not limited to, the following:

– What new/revised Questions are needed to evolve the work programme of SG12?

– When contributions or liaisons are addressed to SG12, on topics not covered by any existing Questions, what is the SG12 view—and any recommended action?

– What are the results of TSB initiatives, or actions of other SGs or SDOs, that need to be considered under the Study Group work programme?

– What ITU-T coordination is needed for the studies carried out on QoS/QoE?

– Is harmonization needed among ITU-T Recommendations on QoS/QoE?

– What collaboration is needed on QoS issues with other bodies in the industry?

– What are the needs and issues expressed by developing countries on QoS and QoE, and how can SG12 provide support in the course of its work?

– What contributions from groups for which SG12 is the parent, such as the Regional Group for Africa, can be implemented in Recommendations, Guides or Handbooks?

### A.3 Tasks

Tasks include, but are not limited to:

– identify new/revised Questions needed in the SG12 work programme to address QoS/QoE issues in the rapidly changing ICT marketplace;

– coordinate QoS/QoE-related activities in the ITU-T (ongoing);

– collaborate on QoS/QoE with other standards bodies (ongoing);

– provide leadership on QoS/QoE related issues to TSAG and the TSB, as needed;

– create other SG12 Regional Groups, as needed;

– respond to actions required in liaisons addressed to SG12 on issues for which no other Question is responsible.

An up-to-date status of work under this Question is contained in the SG12 work programme at <https://itu.int/ITU-T/workprog/wp_search.aspx?sg=12>.

### A.4 Relationships

Recommendations:

– All Recommendations under the responsibility of SG12

Questions:

– All/12

Study Groups:

– All ITU-T Study Groups with activities related to QoS/QoE

Other bodies:

– All standards-related organizations working on QoS/QoE, such as ETSI, IETF, ATIS, TIA, IEEE, 3GPP, MEF, BBF, etc.

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION B/12 Definitions, guides and frameworks related to quality of service/quality of experience (QoS/QoE)

(Continuation of Question 2/12)

### B.1 Motivation

This Question is the focal point for terms and definitions needed for supporting new or revised Recommendations developed in the other Study Group 12 Questions.

Additionally, this Question addresses the need for new participants in the ITU-T to understand the concepts and the Recommendations on QoS, QoE, transmission quality, etc. Tutorials and guides can be developed to serve this purpose. To help all members and inform them on the work done in the Study Group, it is useful to create tutorials, frameworks, FAQ, reference implementations, etc., and post them on the Study Group website.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility:

– Recommendations ITU-T P.10/G.100, G.100.1, G.191, G.192, P.800.1, P.800.2, G.1000;

– Handbook on QoS; Handbook on Network Planning; Handbook on Practical Subjective Testing Procedures; Handbook on Telephonometry.

### B.2 Question

Study items to be considered include, but are not limited to:

– What new or revised definitions need to be included in Recommendation P.10/G.100?

– What are the new sections to be written to update the guides or tutorials? How could we ensure greater visibility and better use of these materials?

– What kind of materials (FAQ, reference implementations, tutorials, etc.) could be made available on the Study Group website?

– What guides would be needed to help the users to implement the new Recommendations?

### B.3 Tasks

Tasks include, but are not limited to:

– drive actions to update existing Recommendations, or to create new Recommendations on definitions;

– update or produce guides or tutorials for the benefit of the users of the Recommendations;

– create tools that could help non-experts to understand and implement the new recommendations. Some of these tools should be implemented on the Study Group website.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### B.4 Relationships

Recommendations:

– All Recommendations under the responsibility of SG12

Questions:

– All/12

Study Groups:

– ITU-T, ITU-R and ITU-D Study Groups with activities related to QoS/QoE

Other bodies:

– ETSI

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION C/12 Objective methods for speech and audio evaluation in vehicles

(Continuation of Question 4/12)

### C.1 Motivation

Car infotainment systems, telematic services and all types of mobile communication services are used increasingly in vehicles; an increasing number of modern cars are equipped with integrated infotainment, communication systems and connection possibilities to personal devices such as smartphones. In order to provide a good user experience, low driver distraction, satisfying communication quality and optimum dialog quality for all speech-based services, a variety of user interfaces and technologies have to seamlessly interact and to be optimized for the car environment under all driving conditions. All services and technologies deployed in the car should not distract the driver from his main task. Advanced hands-free devices are required which require sophisticated signal processing adapted to the individual car, to provide superior speech quality for the driver as well as for far end conversational partner. The special needs for emergency calls need to be addressed, too. Sophisticated speech recognition and dialog systems are needed to use speech-based services in the car. In-Car Communications (ICC) systems need to be optimized to provide a mostly natural speech enhancement for all types of in-car communications. Zoning concepts allowing the use of different audio-/speech-based services in different zones within vehicles need to be considered.

The use of hands-free devices is becoming mandated in an increasing number of countries and states throughout the world. A large percentage of the target market for these vehicles will own headsets prior to purchasing a vehicle equipped with infotainment systems. They will expect to continue to use them in the vehicle, and thus will expect the vehicle to exploit the headset. The introduction of wireless headsets (e.g., Bluetooth, 802.11, DECT) requires the definition of standard behaviour and interactions with the vehicle.

So far Recommendations were developed describing the transmission requirements and test methods for narrowband, wideband, super-wideband speakerphones, for subsystems in cars, for emergency call communication and in-car communication.

The study within the Question is based on the existing Recommendations P.340, P.313, P.501, P.502, P.583, P.1100, P.1110, P.1120, P.1130, P.1140, P.1150. The main focus of the Question will be updated tests and requirements for hands-free systems, including emergency call systems, subsystem requirements in cars, in-car communication systems, speech recognition and speech dialog systems. A special focus needs to be given on the requirements for autonomous driving in the context of speech and audio in cars.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: P.1100, P.1110, P.1120, P.1130, P.1140, P.1150.

### C.2 Question

The following items are to be considered within the study of the Question:

– How can the driving situation be simulated while covering the most relevant parameters influencing driver distraction and the speech quality within a laboratory environment?

– Are there conversational speech quality parameters in the driving situation not yet covered by the existing Recommendations?

– What additional aspects need to be taken into account in emergency call communications?

– Which additional parameters determine the quality of in-car communication systems and how can they be assessed?

– What are the most influential parameters for speech recognition systems in the driving situation?

– How can we assess and quantify the dialog quality of human-machine interfaces in cars?

– Which of the newly developed methodologies known in ITU can be used and/or adapted to the car hands-free situation?

– Do different mobile networks and network configurations or OTT solutions require individual setups for testing?

– Which quality parameters have to be assessed when connecting accessory devices to the vehicle head unit (like e.g., smartphones, headsets or hearing aids)?

– What enhancements of the Recommendations P.1100, P.1110, P.1120, P.1130, P.1140 and P.1150 are needed to be developed ensuring seamless support for users of hands-free devices and ICC systems?

– Which applications and requirements in the context of speech and audio needs to be addressed for autonomous driving?

### C.3 Tasks

Tasks include, but are not limited to:

– define the typical operating conditions to be simulated covering the most relevant parameters influencing the speech quality within a laboratory environment;

– define the typical operating conditions to be simulated covering the most relevant parameters influencing the quality of in-car communication systems within a laboratory environment;

– define the typical operating conditions to be simulated covering the most relevant parameters influencing automated speech recognition performance within a laboratory environment;

– define the typical operating conditions to be simulated covering the most relevant parameters influencing dialog systems performance within a laboratory environment;

– definition of the environmental conditions for testing the car hands-free terminal and verifying its acoustical performance characteristics under typical operating conditions;

– definition of the environmental conditions for testing the car hands-free subsystems and verifying their performance characteristics under typical operating conditions including the definition of QoS classes for such (sub-)systems;

– specification of all relevant transmission characteristics;

– definition of test signals and testing techniques for emergency call systems with special focus on speech intelligibility/listening effort;

– definition of test procedures for evaluating automated speech recognition;

– definition of test procedures for dialog systems in cars;

– define requirements for ICT systems that interact with drivers of vehicles;

– identify the needs in the area of speech and audio for autonomous driving and derive relevant test scenarios and requirements.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### C.4 Relationships

Recommendations:

– P.340, P.313, P.381, P.382, P.383, P.501, P.502, P.570, P.581, P.582, P.700

Questions:

– D/12, E/12, G/12

Study Groups:

– ITU-T SG16

Other bodies:

– 3GPP SA4, ETSI TC STQ, ETSI TC ITS, Bluetooth SIG, ISO TC22, ISO TC204, CEN/TC 278, IEEE CEA, UNECE WP.29

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION D/12 Telephonometric methodologies for handset and headset terminals

(Continuation of Question 5/12)

### D.1 Motivation

Multimedia evolution leads to an increase of the audio signal bandwidth as well as spatial audio in the new generation networks. Beside the existing narrowband and wideband, super-wideband as well as fullband is being developed for the next years. Also, telecommunication is moving from monaural towards binaural as well as immersive audio.

This situation brings new challenges in terms of standardization which need to be covered. The emerging of immersive audio leads to a need of harmonization of the algorithms aiming to calculate loudness ratings and loudness.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: P.16, P.32, P.48, P.51, P.52, P.53, P.54, P.55, P.57, P.58, P.61, P.64, P.75, P.76, P.78, P.79, P.350, P.360, P.370, P.380, P.570, P.581, P.700, P-series Suppls 10, 16, 20.

### D.2 Question

Study items to be considered include, but are not limited to:

– What enhancements in existing Recommendations need to be defined to accommodate for the evolution in immersive audio?

– What new Recommendations are required to address changes in user behaviour or user interaction with regards to new measurement methodologies?

– What new measurement setups are required to support eXtended reality applications (Augmented Reality, Virtual Reality, Mixed Reality) in related devices?

### D.3 Tasks

Tasks include, but are not limited to:

– Examine if the “non-standardise” handset positions used during a conversation could form the basis for a study which potentially could accommodate an extended range of new test position complementary to those specified in P.64.

– Extending the Type 4 Ear Simulator family with Type 4.1 and Type 4.2 Ear Simulator types.

– Improvement of the specifications for acoustic frontends, mainly artificial ears, to better match the measurement needs for immersive audio.

– Investigate different geometries of HATS as alternatives to the current geometry of HATS specified by P.58. The geometries should be considered for certain application. Specifically different geometries of the head and pinnae should be investigated.

– Investigate how to aggregate measurements from multiple test positions into an overall measure of transmission performance. This is intended to address the situation that users are holding and positioning communication devices in many ways.

– Investigate measurement setups for devices that make use of bone conducting technology.

– Investigate measurement setups for wearable devices e.g., smart watches.

– Investigate possible simulation of the human behaviour with regards to movements when using communication devices in real life situations. This is relevant for all types of communication devices e.g., handset, headset and handsfree.

– Investigate the directivity as well as extended frequency range of the artificial mouth, aiming to revise Recommendation P.58 as well as P.51. The investigation should also include performance behind the lip-plane of humans.

– Maintenance of Recommendations that fall under the responsibility of the Question.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### D.4 Relationships

Recommendations:

– P.1100 series

Questions:

– C/12, E/12, F/12, H/12

Study Groups:

– ITU-T SG16

Other bodies:

– 3GPP SA4, CENELEC, ETSI TC STQ, IEC TC 29, IEEE CEA

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION E/12 Analysis methods for speech and audio using complex measurement signals

(Continuation of Question 6/12)

### E.1 Motivation

Terminal and network equipment increasingly includes complex signal processing techniques, including artificial intelligence (AI). In addition, super-wideband and fullband systems are more and more established in the market. Most devices cannot be regarded as linear, time-invariant systems anymore. The subjectively relevant transmission characteristics of such equipment need to be correctly determined using adequate measurement methods. There is a need of having reproducible and well-defined measurement methods, which can be used for certification labs as well as for developers, and which ideally should be combined to one quality value.

Test signals and analysis techniques for use in telephonometry have been continuously collected in previous study periods. New test signals allow evaluating many different parameters more realistically and are no longer limited to narrowband and wideband signals. However, there is still lack of analysis methods for mixed content such as speech and music. Modern speech codecs allow the transmission of signals of any kind. Existing test methods and to some extent also signals need to be adapted, since they may no longer be appropriate for latest signal processing methods. In addition, the interaction of consecutive signal processing blocks in an end-to-end connection that may degrade overall quality need to be investigated more in detail.

The evaluation methodologies for speech and audio processing are still incomplete and need further improvement. New technologies in hands-free, conference systems, and speech processing require the adaptation of existing testing methodologies and the study of new procedures. There is a need to produce new or update existing product-/application-oriented Recommendations for future hands-free communication terminals that support e.g., conferencing, audio-visual aspects, or immersive audio services.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: P.50, P.59, P.300, P.310, P.311, P.313, P.330, P.340, P.341, P.342, P.381, P.382, P.383, P.501, P.502, P.505.

### E.2 Question

The following items are to be considered within the study of the Question, special consideration should be given to super-wideband/fullband systems, signal processing of circuit-switched and packet-based terminals, and maintenance of existing Recommendations:

– What kind of new complex signal processing used in terminals, systems and networks may influence speech and audio transmission quality and what objective testing methodology can be used?

– What kind of techniques can be used to simulate time-variant use and time-variant behaviour of telecommunication equipment?

– What additional type of test signals and testing techniques are needed for wideband, super-wideband and fullband transmission systems?

– Which type of test signals and analysis procedures can be used for spatial audio?

– Which type of test signals can be used for AI-based signal processing in terminals?

– Which test signals other than speech and noise are needed and how can they be defined?

– Which test signals can be used for the simulation of noisy environments?

– What methods are suitable for the objective assessment of background noise transmission and to what extent can the background noise transmission be assessed without the knowledge of the near-end background noise signal?

– What testing methods/signals can be used to optimize background noise transmission in combination with VAD and comfort noise insertion techniques?

– What testing methods are needed for speech and audio enhancement devices and what are the limits for the different quality determining parameters identified?

– What are the consequences on the speech quality of speech processing implemented in hands-free terminals and new types of conferencing devices for, e.g., Smart Home? What characteristics and limits can apply?

– What characteristics and limits can apply to other speech processing techniques such as speech recognition systems?

– What are the implications of the interaction between terminal signal processing and network signal processing on speech quality?

– How can existing and/or new speech quality parameters be combined to a single speech quality representation covering all conversational aspects?

### E.3 Tasks

Tasks include, but are not limited to:

– improve/adapt existing test signals and objective speech quality testing methodologies;

– identify and study new basic objective testing methodologies in telecommunications;

– identify and study new basic objective testing methodologies for audio;

– identify and study new basic objective testing methodologies for spatial audio;

– identify and study new testing methodologies for real-time signal processing techniques used e.g., in ICC (in-car communication);

– identify and study new testing methodologies for background noise transmission quality;

– identify and study the impact of time-variant user behaviour and time-variant signal processing by defining new test methods and setups;

– improve testing methods for speech enhancement devices;

– add new and improve existing testing methodologies for modern hands-free and conference terminals;

– study applications to multichannel sound pick-up (microphone arrays) and multichannel/multi-device sound reproduction (including stereo and immersive audio).

– maintenance of Recommendations: P.50, P.59, P.300, P.310, P.311, P.313, P.330, P.340, P.341, P.342, P.381, P.382, P.383, P.501, P.502, P.505.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### E.4 Relationships

Recommendations:

– P.57, P.58, P.64, P.79, G.161, G.168, G.169, P.1100, P.1110, P.1120, P.1130, P.1140, P.1150, P.370, P.380, P.570, P.581, P.700

Questions:

– C/12, D/12, F/12, G/12, H/12

Study Groups:

– ITU-T SG16

Other bodies:

– ETSI TC STQ, 3GPP SA4, TIA, IEEE, IEC

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION F/12 Methodologies, tools and test plans for the subjective assessment of speech, audio and audiovisual quality interactions

(Continuation of Question 7/12)

### F.1 Motivation

The work of this Question concerns new methods of assessing the subjective impact of time-varying impairments and includes the design of laboratory and remote testing of speech/noisy-speech/music/mixed content and audiovisual signals. These methods and tools apply to narrowband, wideband, superwideband and fullband audio telephony.

As done so far, considering that the need for standard subjective testing methodologies will continue to exist for the effective assessment of the transmission performance of new communication systems, like speech/music, immersive coders (for audio frequency bandwidths), or other devices and equipment designed for carrying voice and audiovisual signals, the Question will continue to provide the necessary support to produce test/processing plans to execute appropriate subjective tests. Input could also be provided due to the relevant work in other standards organizations, like ISO/MPEG or fora/consortia/partnership projects like 3GPP.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: P.85, P.800, P.804, P.805, P.806, P.807, P.808, P.809, P.810, P.811, P.812, P.830, P.835, P.840, P.851, P.852,P.880, P.918, P.1501, P-series Suppls 24, 25, 29, Handbook STP.

### F.2 Question

Study items to be considered include, but are not limited to:

– What new Recommendations need to be developed to evaluate new speech/noisy-speech/music and mixed content quality requirements?

– What new Recommendations need to be developed for multi-dimensional subjective tests in a telephone conversation or multi-party call?

– What enhancements to existing Recommendations need to be defined to improve the evaluation of degradations using immersive codec aspects?

– What enhancements to existing Recommendations need to be defined to improve the subjective evaluation of speech-based or multimodal interactive services?

– What enhanced subjective test methodology is required for evaluating the performance of gaming based applications, in terms of perceived QoS/QoE by game players?

– What new or revised subjective assessment methods are required for evaluating the effects of time-varying impairments (such as delayed packets or packet losses), and what guidance can be provided for the appropriate provision of sample/noise or music material for the testing?

– What modifications to existing or new Recommendations need to be developed to assess new speech/music/mixed content digital coding systems, e.g., narrowband/wideband/superwideband/fullband speech and/or music and/or mixed content and/or immersive codecs operated over fixed and/or 5G mobile networks (including Internet Multimedia Services)?

– What new test plans are needed to evaluate (subjectively) end-to-end communications over fixed and/or 5G mobile networks using data obtained by means of "crowdsourcing"?

– What guidance can be provided for collection and post-screening of subjective test results, and global analysis of results from internationally coordinated exercises in general?

– What are the relationships between various subjective test measures, for example in the auditory modality, between intelligibility, listening effort and QoS/QoE measures?

– What guidance can be provided for collection and evaluation of cultural/language/ nationality dependence of subjective quality?

– What guidance can be provided for collection and evaluation of physiological measures as an additional test method for speech quality assessment?

– Which Questions within SG12, and other standardization activities within ITU, require support for subjective testing?

### F.3 Tasks

Tasks include, but are not limited to:

– maintenance and enhancement of Recommendations in the P-series with regards to subjective testing methods and with regards to the Handbook on Subjective Testing Practical Procedures;

– revise existing Recommendations (e.g., crowdsourcing, gaming, etc.) and draft new ones, e.g., P.PHYSIO and all new Recommendations originated from new work items.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### F.4 Relationships

Recommendations:

– P-series, G.700-series

Questions:

– E/12, G/12, H/12, J/12, L/12, N/12

Study Groups:

– ITU-T SG9, SG16

– ITU-R WP5C, WP6C

Other bodies:

– ISO-MPEG, 3GPP, IETF, ETSI, ANSI

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION G/12 Perceptual-based objective methods and corresponding evaluation guidelines for voice and audio quality measurements in telecommunication services

(Continuation of Question 9/12)

### G.1 Motivation

The work of this Question will focus on objective, perceptual and mainly signal-based methods for evaluating quality parameters in telecommunication scenarios. Primarily, the methods under study should concentrate on user-perceived quality characteristics. Consequently, these methods and algorithms include perceptual approaches. They model results and procedures, which are applicable in subjective tests. So that subjective procedures will get an objective counterpart by using the same scaling and basic procedures.

An example for that is the successful standardization of Recommendations P.863 up to fullband audio, a perceptual based method which models objectively Listening Only Tests with Absolute Category Rating for the evaluating of the Listening Speech Quality according to Recommendation P.800. Recommendation P.863 is also extended by Recommendation P.863.2 that provides quality predictions for individual quality dimensions. This Question will extend the objective evaluation of Listening Quality – the by far most often measured speech quality parameter up to now – to other quality aspects of voice telephony like talking quality and quality dimensions in no-reference and full-reference setups, including perceptual, signal-based models for objective rating of multi-channel and spatial audio in telecommunication services. Under consideration of new generation telecommunication services, also other media than speech like music should be taken into account.

Furthermore, the evaluation of transmitted noise – especially after processing by noise suppression systems – should be covered by the work of this Question, the same as objective prediction of speech intelligibility.

This Question also considers new speech and audio processing techniques based on Machine Learning and Artificial Intelligence and the objective, perceptual evaluation of those.

This Question analyses and recommends also methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models and gives guidance for developing quality prediction models in general (Recommendation P.1401) and especially by means of machine learning and artificial intelligence (Recommendation P.1402).

This Question will also continue and finalize the ongoing work on (P.SAMD).

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.1029, P.563, P.863, P.863.1, P.863.2, P.1401, P.1402.

### G.2 Question

Study items to be considered include, but are not limited to:

– An already defined work item is the objective assessment of talking quality. Therefore at first a reliable subjective test method has to be established. In a second step, an objective model can be developed.

– In addition to the existing objective models like P.863 or P.563 that are producing single numbers describing the overall quality; a need for additional information about possible quality degradations and quality dimensions are requested by the market. This is studied under P.863.2 (full-reference) and P.SAMD (no-reference).

– Furthermore, the objective assessment of audio signals such as music transmitted over telecommunication links like WCDMA, LTE and 5G with modern codecs and terminals should be investigated.

– Perceptual, signal-based models for objective rating of multi-channel and spatial audio in telecommunication services are interesting under the scope of this Question.

– The determination of the quality of synthesized speech in an instrumental way, e.g., using the objective perceptual methods, is an interesting topic in this Question as well as methods for objective prediction of speech intelligibility.

– Perceptual quality prediction of speech processed by audio processing techniques based on Machine Learning and Artificial Intelligence. Examples are ML-based coding, speech enhancements, talker anonymization or deep fake speech.

– This Question analyses and recommends methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models. These statistics can be applied to objective prediction models which can be translated to an estimated subjective judgment of a dedicated subjective test procedure. This Question discusses frameworks, metrics and example procedures for those statistical analyses and reporting. Furthermore, this question gives guidance to develop quality prediction models in general and specifically by means of machine learning and artificial intelligence as in Recommendation P.1402.

### G.3 Tasks

Tasks include, but are not limited to:

– maintenance and enhancement of P-series Recommendations with regards to objective quality testing methods and perceptual models as P.863, P.863.1 and P.863.2;

– completion of Recommendations on objective estimation of individual quality dimensions as no-reference approach P.SAMD;

– development of a Recommendation for objective, perceptual quality prediction of non-speech signals (e.g., music) in telecommunication services;

– development of a Recommendation for perceptual, signal-based models for objective qualitative rating the perception of multi-channel and spatial audio in telecommunication services;

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### G.4 Relationships

Recommendations:

– P-series, G.100- and G.1000-series

Questions:

– C/12, E/12, F/12, K/12, L/12, N/12

Study Groups:

– ITU-T SG16

Other bodies:

– ETSI TC STQ, 3GPP

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION H/12 Conferencing and telemeeting assessment

(Continuation of Question 10/12)

### H.1 Motivation

In today's society, audio and audio-visual telemeetings and audio- and video-conferences are gaining in importance. The term telemeeting is used here to cover with one term all means of audio or audio-visual communication between distant locations.

If the perceived quality is good enough, telemeetings can be used instead of face-to-face meetings, which will reduce the needs for travelling and in turn reduce the negative effects on our climate. The travel time and cost can also be reduced.

To achieve this goal there is a need to develop an agreed way of quantifying the quality of experience of multi-party services that are conversational and interactive.

A telemeeting is often a multipoint communication, where the participants can use different types of equipment to connect to the (virtual or real) meeting space, e.g., by fixed phone, mobile phone, PC, videoconferencing or eXtended Reality (AR, VR, MR) equipment. To obtain a good evaluation of the telemeeting quality of experience, the quality perceived by all participants in a telemeeting needs to be assessed.

There are standardized subjective and objective test methods for several components in a telemeeting, such as speech, audio and video codecs, characterized by bit rate (fixed or variable), frame rate, resolution, noise cancellation, acoustic reproduction quality, background noise, and synchronization and transmission impairments. Some recommendations on how to assess the interaction between these factors are available, too. In a telemeeting context, however, these factors need to be assessed in the light of multiple users connected via possibly asymmetric links. The initial focus has been on subjective assessment strategies. The results from performed tests can then form a base for objective quality assessment of telemeetings and can provide insights on quality aspects for telemeeting services. So, the scope of Q10 includes multimedia subjective assessment, objective modelling as well as QoE.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.1092, P.1301, P.1302, P.1305, P.1310, P.1311, P.1312, P.1320, P-series Suppl. 26.

### H.2 Question

Study items to be considered include, but are not limited to:

– How can the quality of experience of multiparty audio, audiovisual and XR telemeetings / meetings in the metaverse be evaluated?

– What is the quality impact of the different ways of connecting to a telemeeting?

– What is the quality impact of multiple users connected to the telemeeting from one single-location, from multiple locations or via links of highly different quality?

– What aspects of communication performance need to be addressed when it comes to multiparty interaction across links with delay or limited resources for audio and video?

– How can different quality aspects related to telemeeting quality be quantified, and how can their relative importance for the whole telemeeting quality be assessed with standardized subjective and objective evaluation methods?

– How do telemeeting assessment methods scale with the number of participants?

– Which performance criteria need to be assessed, when it comes to telemeetings in a group-collaboration context?

– How can spatial sound and video be evaluated in a telemeeting (via headphone- or loudspeaker reproduction, with problems such as the microphone placement, echo-cancellation, camera adjustment, lighting conditions, etc.)?

– What are the relative roles of the transmission, the conference bridge or server, and the terminal equipment being employed on quality perception, with regard to the user experience of the service?

– What is the additional impact of data media such as presentation slides on user perception?

– Which are the new challenges when it comes to the use of XR technologies for telemeetings in the metaverse?

– Which measures beyond conventional quality scores, (e.g., communication behaviour, cognitive effort, or task completion) should be considered for a comprehensive assessment of telemeeting quality?

### H.3 Tasks

Tasks include, but are not limited to:

– maintain a Recommendation (P.1301) on how to subjectively quantify the quality of audio and audiovisual multiparty telemeetings, where the participants can have different types of connections to the meeting;

– maintain a Recommendation (P.1305) on how different delays for different participants affect the meeting quality. Suitable test tasks for evaluation methods of interactive multiparty audio and audiovisual telemeetings are needed;

– maintain a Recommendation (P.1302) on subjective and objective methods for simulated conversation tests addressing audio and audiovisual call quality;

– maintain a series of Recommendations (P.1310, P.1311, P.1312) on how to evaluate the perceived quality of telemeetings using spatial audio. The methods should be applicable to listening through both headphones and loudspeakers;

– maintain a Recommendation (P.1320) on how to assess the QoE of eXtended Reality (XR) telemeetings;

– develop a Recommendation (P.IXC) on Interactive test methods for subjective assessment of extended reality communications

– develop a Recommendation on the use of auditory and visual cues for high-quality telemeetings in different application contexts such as business and private meetings (including, for example, aspects such as eye-contact and other visual cues, e.g., in the light of technical characteristics such as screen sizes);

– develop a Recommendation on how the quality impact of separate components in a telemeeting that have been tested separately can be weighted together to give an overall telemeeting quality value;

– maintain a Recommendation (G.1092) on a taxonomy of telemeetings from a QoE perspective, listing relevant QoE aspects including system, human, and contextual aspects such as time to join, screen sharing, application feedback, etc.;

– develop a Recommendation on remote operations including communication aspects;

– develop a Recommendation on the QoE aspects of haptics in remote control and telemeetings;

– develop a Recommendation on the importance of audio-visual congruence (congruence between individual audio and video streams, placement of participants on the screen).

– develop a Recommendation (P.UXV) on the user experience related to multimedia in videotelephony and videoconferencing services, by analyzing multimedia components and multimedia functions in nowadays’ videotelephony or videoconferencing services, such as, for example, screen sharing, background blurring/replacement, AR effects, etc.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### H.4 Relationships

Recommendations:

– P-series, G-series

Questions:

– D/12, E/12, F/12, G/12, J/12, K/12, L/12, N/12

Study Groups:

– ITU-T SG5, SG9, SG16

– ITU-R WP6C

Other bodies:

– ISO-MPEG, 3GPP, IETF, ETSI, VQEG, Qualinet

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION I/12 Operational aspects of telecommunication network service quality and end-to-end performance considerations

(Continuation of Question 12/12)

### I.1 Motivation

It is essential to specify network service quality parameters to enable telecommunication services to be offered to customers/users in order to satisfy customers'/users' quality of service expectations.

These parameters relate to both implementation and ongoing use of the service. Service quality is also related to all aspects of network assessment and management.

Service quality of networks needs to be assessed as a total connection, focusing on the end-to-end network service offered at all times.

Service quality parameters are required in order to meet customers'/users' expectation of a service, and related network performance parameters should relate to service quality parameters.

Network providers must plan, dimension and operate their networks to parameters which will ensure that services offered to customers/users meet the latter's quality of service expectations. In addition, Regulators need guidance in order to ensure that customers are experiencing an acceptable level of quality of service.

In addition, there is a need to give guidance on general transmission planning and keeping it up to date with technological evolution.

With the increasing industry focus on new and future technologies, there is a need for guidance on the associated end-to-end QoS, performance and resource management issues. This includes interworking aspects between different networks (e.g., cellular, wireless, wireline and also such of different generations) and packet-based technologies.

Issues and guidelines for transmission performance necessary to ensure high end-user satisfaction must be reconsidered in light of introduction of new services and their interconnection with existing networks.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: E.420, E.421, E.422, E.423, E.424, E.425, E.426, E.427, E.428, E.431, E.432, E.433, E.434, E.436, E.437, E.438, E.439, E.440, E.450, E.451, E.452, E.453, E.454, E.455, E.456, E.457, E.458, E.459, E.460, E.470, E.475, E.801, E.802, E.803, E.804, E.804.1, E.805, E.805.1, E.806, E.807, E.810, E.811, E.812, E.813, E.820, E.830, E.840, E.845, E.846, E.847, E.850, E.855, E.800‑series Suppl. 8, Suppl. 9, Suppl. 10, G.101, G.102, G.103, G.105, G.108, G.108.1, G.108.2, G.109, G.111, G.114, G.115, G.116, G.117, G.120, G.121, G,122, G.126, G.131, G.136, G.142, G.172, G.173, G.174, G.175, G.176, G.177, G.1028, G.1028.1, G.1028.2, G-series Suppl. 61, I.352, I.354, I.358, I.359, I.371, I.378, P.11, Y.1221, Y.1222, Y.1223, Y.1530, Y.1531, Y.1542, Y.1545, Y.1545.1, ESTR-KPI-RAN.

### I.2 Question

Study items to be considered include, but are not limited to:

– How can existing Recommendations covering quality of service and network performance be interpreted to meet customers'/users' expectations of service quality under operational scenarios?

– What new or revised Recommendations are required to ensure that adequate network service quality can be provided to meet customers'/users' expectations under operational scenarios? A key focus of these new or revised recommendations relates to service providers, regulators and vendors.

– What guidance can be provided for transmission planning for voice, data and multimedia services, taking into account that end-to-end connections are established via heterogeneous and interconnected networks with different transmission technologies? This includes guidance on transmission planning for the interconnection of evolving networks.

– What guidance can be provided on reference models, performance parameters and interworking requirements in end-to-end communication paths, and how can the values of performance parameters be managed across multiple network segments? In addition, the effects of the transmission delay on services and applications including multimedia are studied.

### I.3 Tasks

Tasks include, but are not limited to:

– Maintain and revise the existing literature under responsibility of the Question, as the need arises,

– Develop new guidance on issues described above, to address the needs of service providers, regulators and vendors, with the goal of meeting customers'/users' expectations of service quality.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### I.4 Relationships

Recommendations:

– P.10/G.100, P.863, Y.1540

Questions:

– A/12, B/12, G/12, J/12, K/12, M/12

Study Groups:

– ITU-T SG2, SG3, SG13, ITU-R, ITU-D

Other bodies:

– ETSI TC STQ, 3GPP

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION J/12 Quality of experience (QoE), quality of service (QoS) and performance requirements and assessment methods for multimedia applications

(Continuation of Question 13/12)

### J.1 Motivation

A major challenge for emerging IP-based networks is to provide adequate Quality of Experience (QoE) and Quality of Service (QoS) for new multimedia services and applications. An example is Extended Reality (XR) applications, including Augmented Reality (AR), Mixed Reality (MR), and Virtual Reality (VR). In such applications, QoE is critical because bad quality may cause people using it have nauseas and sickness. Another example is new services emerging in fixed, mobile, ultra reliable low latency communications (uRLLC) and massive machine type communications (mMTC) broadband. All of these services are inherently multi-media, incorporating audio, video, environments, and interactive control functions, and the QoE is affected by many different categories of factors. The performance requirements and associated measurement methodologies for each of these aspects need to be defined.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.1010, G.1011, G.1030, G.1031, G.1032, G.1034, G.1035, G.1036, G.1040, G.1050, G.1070, G.1071, G.1072, G.1080, G.1081, G.1082, G.1091, G-series Suppls 73 and 77, P.917, P.919, P.1010, Y.1562.

### J.2 Question

Study items to be considered include, but are not limited to:

– development of new Recommendations providing guidance on QoE target evaluation and measurement;

– identify end-user performance expectations and associated metrics for audio, video, text, graphics quality and control functionality;

– define the key performance parameters and values required to satisfy end-user expectations;

– determine how these requirements can be related to the underlying network, server, and terminal;

– identify simple analysis techniques for estimating end-to-end performance for multimedia applications;

– identify QoS/QoE monitoring methodologies for multimedia services;

– identify sets of KPIs and QoS metrics for different services and investigate the relationship with QoE;

– investigate techniques and methods to perform complex data processing and to make consistent and significant decisions for quality management and assurance;

– multimedia performance considerations for IP gateways;

– QoS and QoE considerations for new services in fixed, mobile broadband, ultra reliable low latency communications (uRLLC) and massive machine type communications (mMTC).

### J.3 Tasks

Tasks include, but are not limited to:

– development of new Recommendations providing guidance on end-user performance expectations for multimedia applications, such as high quality audio and video immersive applications, and gaming;

– development of new Recommendations on planning models for estimating end-to-end multimedia performance;

– development of new Recommendations providing guidance on performance monitoring methods for multimedia applications, such as high quality audio and video immersive applications, and gaming;

– development of new Recommendations (and other documents as needed) on QoS and QoE aspects related to new services in fixed, mobile broadband, ultra reliable low latency communications (uRLLC) and massive machine type communications (mMTC).

– revision of Recommendations G.1010, G.1011, G.1030, G.1031, G.1032, G.1034, G.1035, G.1036, G.1040, G.1050, G.1070, G.1071, G.1072, G.1080, G.1081, G.1082, G.1091, Y.1562, P.917, P.919 and P.1010 as necessary.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### J.4 Relationships

Recommendations:

– G.1000-series, Y.1000-series, P.800.1, P.800.2, P.1201, P.1203, P.1204, Y.1540, Y.1541, Y.1544

Questions:

– C/12, E/12, G/12, H/12, K/12, L/12, M/12, N/12

Study Groups:

– ITU-T SG9, SG16

Other bodies:

– IETF, ETSI TC STQ, 3GPP, TIA TR-41, TIA TR30.3, ATIS IIF, MPEG

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION K/12 Development of models and tools for multimedia quality assessment of packet-based video services

(Continuation of Question 14/12)

### K.1 Motivation

A major challenge for emerging IP-based networks is to provide adequate Quality of Experience (QoE) and Quality of Service (QoS) for current and evolving multimedia services and applications such as internet media including over-the-top (OTT) video, and immersive video.

A number of Recommendations have been developed by this Question, in particular:

– In the P.1203 series of standards, an integral model for audiovisual quality-assessment of streaming using reliable transport is described. It enables integral quality estimates for videos between 1 min and 5 min duration, based on short-term audio and video quality modules (Pa/P.1203.2, Pv/P.1203.1), as well as a long-term integration module (Pq/P.1203.3).

– In the P.1204 series of standards, a set of models is described, for bitstream-, pixel-based and hybrid video quality estimation up to 4K resolution, and covering the codecs H.264, HEVC and VP9. It is the first activity of its kind to cover all aforementioned types of relevant video-quality modelling approaches, using an identical dataset for training and validation. Performance figures for the models indicate their strong prediction accuracy.

Both of these standard series can be used for monitoring HTTP adaptive streaming services (such as HLS or DASH), both for TCP or QUIC-type transport. Hence, they represent tools widely applicable in the market.

The previous developments leveraged a unique set of training and validation databases – with a significantly larger number of samples than previously published research – and a modular framework, which enables future, iterative enhancements of the individual modules that are part of the parent recommendations (P.1203 and P.1204). Part of the databases have been published in an open-source manner and have consequently been used in academic research. It will therefore be beneficial for this Question to explore possibilities to further disseminate subjective test database material (video samples, ratings), model training, and/or model reference code in a reproducible manner for the public to validate the results. This in turn should result in a quicker feedback loop in which this Question may obtain comments from different stakeholders and consider those in future updates to Recommendations.

A primary aspect of the continued work is on the inclusion of long-term integration in conjunction with the existing P.1203 and P.1204 standards series. This work has started and is continued here, aiming for a harmonized view on longer-term session quality for the case of adaptive-streaming type services. In particular, the aspect of screen types, different viewing angles and/or display resolutions shall be considered to cover the variety of real-life applications of the models.

Moreover, the inclusion of further video codecs, or different parameterizations of such codecs (e.g., in terms of encoding speed), in updates or extensions of the P.1203 and P.1204 standards will be addressed.

Since today’s over-the-top services increasingly involve encrypted transport, passive, network-centric quality monitoring becomes more and more challenging. Bitstream or media-related information may not be readily available, and respective monitoring algorithms may need to apply heuristics to infer application-level performance or quality indicators from network-level features. If network operators wish to assess the quality of the media services offered over their networks, they often need to rely on proprietary solutions. These proprietary offerings are, to the extent known, not using current, standardized approaches, and have not been independently validated. Here, it will be needed to provide the market with means to validate those proprietary tools in terms of their prediction accuracy for different Key Performance Indicators such as buffer behaviour and/or QoE/MOS predictions. Also, the ever-changing nature of OTT services may require the development of a framework to continuously adapt standardized models to the current needs of the market in a more rapid fashion. To address this aspect, the Question will work on approaches for non-intrusive assessment of encrypted video, in collaboration with QM/12.

Further work items will be addressing extensions of the P.1203 and P.1204 framework towards High Dynamic Range and Wide Color Gamut, as well as work on IP-based 360° video quality assessment.

Recent advancements in machine learning / deep learning / artificial intelligence based approaches in video processing (enhancement, coding, etc.) necessitates the development of video quality models capable of taking into account degradations or possibly improvements introduced by such approaches. To address this, the Question will work towards creation of work items that will build on existing standards such as P.1203 and P.1204 to exploit the modular nature of the model development process to consider machine-learning-related effects in terms of quality prediction.

Some more interactive services have video and audiovisual-quality related components, which could take advantage of the experience and results that have led to the existing P.1200-models developed by this Question. The P.BBQCG work item was started to handle the video quality assessment for cloud gaming services, but also to create a holistic QoE estimation model taking interactivity into account.

Along these lines, the work on eXtended Reality (XR) service QoE will be continued and developed towards a QoE monitoring work item. The exact details of the work will be specified during this Study Period.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: P.1200-series.

### K.2 Question

Study items to be considered include, but are not limited to:

– What further aspects of a continued characterization of P.1201, P.1202, P.1203 and P.1204 models should be considered?

– How do P.1201, P.1202, P.1203 and P.1204 need to be maintained, and what further application guidance towards, for example, network-centric monitoring solutions needs to be provided?

– How can existing and future models be adapted/constructed to cover or support an extended scope, e.g., in terms of viewing distance, encoder parameterization, or new error concealment, without having to build a new set of models?

– What are relevant subjective test methodologies, especially when it comes to capabilities of 4K/UHD and 8K, and respective high dynamic range, enhanced color gamut and high-framerate, and which respective new standards need to be developed (possibly in cooperation with other standardization bodies)?

– How can 8K or HDR video quality be assessed using pixel-, bitstream-based or hybrid modelling approaches?

– How can audiovisual quality be monitored for streams for these cases, and how can audio and video quality be integrated?

– How can bitstream-, signal-based and hybrid models be evaluated for these extended services in a comprehensive standardization activity on the same type of data?

– What relationship exists between the subjective responses of users at the terminals and the objective measurements made from the point at which the assessment system is connected?

– How can audiovisual synchronization be reflected in models such as P.1201, P.1202, P.1203, and P.1204?

– How can long-term integration be addressed for streaming of higher resolutions up to 4 and 8K or HDR content?

– What are the requirements on future updates of the P.1203 and P.1204 standards series for HTTP-based video quality monitoring?

– How can diagnosis assessment be done when using P.1203 and P.1204 standards?

– How can knowledge on short-term measurements and their temporal pooling for longer-term predictions be generalized to complete sessions of multimedia quality monitoring?

– How can video-quality estimation modules for conversational quality estimation models be derived from existing standards or new work within this Question?

– How can video, audiovisual quality and other effects for 360° / omnidirectional video and accompanying audio be monitored?

– How can video- and audiovisual quality prediction best benefit from different machine-learning approaches?

– How can quality of cloud gaming services be assessed?

– How can existing models be adapted/extended to assess the quality of user-generated content?

– How can different XR systems be assessed? This includes defining new test methodologies and instrumental assessment methods.

– How can existing standards (P.1204.X) be extended for quality assessment of AI-related aspects (coding, enhancements, etc.)?

– How can aspects of sustainable streaming, that is, energy or more holistically resource consumption be considered jointly with multimedia quality?

– How can existing models be adapted, or assisted, to be able to handle encrypted streams?

### K.3 Tasks

Tasks include, but are not limited to:

– maintenance of Recommendations P.1201, P.1202, P.1203 and P.1204;

– development of new Recommendation(s) on guidance for the use of P.1201, P.1202, P.1203 and P.1204 in different applications or operational contexts;

– considerations on bitstream-based audio quality evaluation;

– development of tools that are used in the course of model development;

– development of models for assessing video formats such as HDR, wide color gamut, high framerate;

– development of models for monitoring video quality in the context of conversational and conferencing services;

– development of modelling approaches for quality and QoE evaluation of 360° / omnidirectional video streaming and accompanying audio;

– development of modelling approaches for quality and QoE evaluation of network-based XR systems and services;

– development and maintenance of a new Recommendation on non-intrusive assessment of TLS-encrypted, TCP-based multimedia streaming quality (P.ENATS).

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### K.4 Relationships

Recommendations:

– P.564, G.1000-series, J series recommendations on video quality

Questions:

– J/12, M/12

Study Groups:

– ITU-T SG13, SG16

– ITU-R WP6C

Other bodies:

– 3GGP SA4, ATIS, Broadband Forum, ETSI TC STQ, HGI, IETF, MPEG, VQEG

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION L/12 Parametric and E-model-based planning, prediction and monitoring of conversational speech and audio-visual quality

(Continuation of Question 15/12)

### L.1 Motivation

The telecommunications industry is working to adopt more flexible infrastructure to control costs and facilitate the introduction of new services. Examples are 5G or generally next generation IP-networks which provide flexible transmission bandwidths and user interface connections, however at the expense or quality which varies with the transmission scenario and with time. A proper transmission planning, as well as flexible prediction and monitoring of Quality of Experience (QoE) are useful in managing the efficient operation and the effective services of such networks.

Regarding transmission planning of such scenarios, Study Group 12 has established the E-model, a computational model for use in transmission planning, see Recommendation G.107. This model is now frequently applied to plan traditional, narrow-band and handset-terminated networks, and to an increasing extent also for wideband, fullband telephonies and packet-based networks, using the extensions of the E-model described in Recommendations G.107.1 and G.107.2. While being popular, the E-model still shows a considerable number of limitations, namely when applying it in super-wideband and fullband networks, which non-handset terminal equipment, and with speech processing devices (such as echo cancellers, noise reduction, or alike) integrated in the network or in the terminal. Codecs and other speech processing devices may also be expected to rely on machine-learning methods.

Regarding the quality prediction and monitoring of such scenarios, the industry is already benefiting from ITU-T Recommendations for objective speech quality assessment. However, most of the techniques described in these recommendations are signal based and address listening only contexts. Typical communications involve interactive, two-way, conversations. IP and mobile networks can be particularly deleterious to interactive applications, including voice conversation; for example, due to increased delay, which in turn will increase the probability of double-talk and increase the perceptibility of echo. Thus, there is a need for a real-time, or near real-time, conversational speech quality assessment and monitoring.

In the end, what is needed is the integration of listening-only, talking-only and interaction quality on a common scale which could be used for planning, predicting and monitoring conversational quality in real-life networks. Such a scale would allow for an easier interpretation of the QoE provided by the different network and service scenarios, and thus make use of the flexibility offered by the respective networks in order to provide optimum services to the customer.

It is envisaged that new methods under this question would be developed collaboratively.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.107, G.107.1, G.107.2, G.113, P.56, P.561, P.562, P.564, P.565, P.565.1, P.833, P.833.1, P.833.2, P.834, P.834.1, P.834.2, P.836.

### L.2 Question

Study items to be considered include, but are not limited to:

– How can the E-model be used to facilitate transmission planning in wide-band, super-wideband, fullband, and mixed-band scenarios?

– How are the relations between degradation covered by the E-model in various audio bandwidths?

– Which quality issues have to be taken into account when extending the E-model to terminal equipment other than standard handset telephones (e.g., HFTs, headsets)? Which parameters can be used to describe such terminal equipment?

– How can the perceptual effects introduced by speech-processing devices included in the network or in the terminal equipment (e.g., (acoustic) echo cancellers, level control devices, voice activity detectors, noise suppression devices) be covered by the E- model? Which parameters need to be used for speech processing which is based on machine learning?

– Is the E-model suitable for quality monitoring? How would such a monitoring application take into account strongly time-variant channel characteristics, e.g., due to bursty frame or packet loss, or in a cellular network?

– What is the impairment effect of each new coding algorithms, especially algorithms relying on machine learning, so that it can be considered in the context of Recommendation G.113 and applied together with Recommendation G.107, G.107.1 or G.107.2?

– How can non-intrusive measurements of voice quality at the IP layers be implemented and improved, for instance by taking into account signalling protocols not yet used by existing methods (e.g., SIP SDP, RTCP XR) or network technologies not covered by existing methods (mobile VoIP, WebRTC GetStats API)?

– What relationship exists between the subjective responses of users at the terminals and the objective measurements made from the point at which the non-intrusive assessment system is connected?

– What are the critical components of conversational speech and audio-visual quality? What existing models and measures addressing these components could be used as inputs and building blocks for the development of new methods?

– What subjective test methods should validation of new objective methods for the assessment of perceived conversational quality be based on?

– How can talking quality and conversational quality be measured in a non-intrusive way?

– How can existing measurement methods for voice quality be applicable for other services than telephony, in particular for video-telephony?

### L.3 Tasks

Tasks include, but are not limited to:

– maintenance and enhancement of the E-model described in Recommendation G.107, G.107.1 and G.107.2 and input to depending Recommendations;

– frequent update of Appendices to G.113;

– maintenance of the Recommendations P.833 and P.834 and corresponding wideband and fullband Recommendations for determining equipment impairment factors;

– changes and/or improvements to existing ITU-T Recommendations P.56, P.561, P.562, P.564 and P.565 to take into account new technologies;

– development of new models (both parametric and signal-based), to combine multiple objective measurements to provide an objective assessment of the perceived conversational speech and audio-visual quality;

– changes and/or improvements to existing ITU-T Recommendation P.836 on simulation-based approaches to model conversational behaviour;

– development of new models and/or relative conformance testing methodologies to assess the perceived listening and/or conversational quality of mobile IP voice and videotelephony services.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### L.4 Relationships

Recommendations:

– E.804, G.108, G.108.1, G.108.2, G.109, G.114, G.115, G.131, G.1050, G.1070, P.11, P.340, P.800, P.800.1, P.805, P.831, P.832, P.863

Questions:

– E/12, F/12, G/12, H/12, I/12, J/12, K/12, M/12

Study Groups:

– ITU-T SG16

Other bodies:

– ETSI TC STQ, IETF (IPPM, XRBLOCK)

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION M/12 Performance of packet-based networks and other networking technologies

(Continuation of Question 17/12)

### M.1 Motivation

There is constant evolution in network technologies. However, while network performance is ever increasing, it remains important to the user's experience as demand for performance also increases. When several network operators work together to provide end-to-end communications, each needs to understand how to achieve the end-to-end performance objectives. Such objectives must be both adequate for the service being offered and feasible based on the available networking technologies.

A framework is needed to guide the development of Recommendations for performance aspects of new network capabilities, transmission facilities, and transport services (e.g., forward error correction and retransmission protocols), including those supported by the emerging and heterogeneous infrastructure. Such a framework is also essential for relating performance to the properties determining network behaviour.

There is a continuing need for packet network performance metrics, and methods of measurement and analysis, and these needs are met by contributions to, and subsequently the approved Recommendations developed by this Question. Other Questions, ITU Study Groups, and some Standardization Bodies should expect that unique needs in the area of packet network performance metrics will be satisfied by this Question’s work, so that they can continue with their unique work plans without overlap.

When new networking technologies are proposed, it is not clear whether they will become sufficiently important to warrant the development of one or more new Recommendations on performance parameters, methods of measurement, and/or numerical objectives. Some investigation of each technology is worthwhile to determine whether it is an appropriate candidate.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.1021, G.1022, G.1023, G.1051, G.1052, I.350, I.351, I.353, I.355, I.356, I.357, I.381, Y.800, Y.1540, Y.1541, Y.1543, Y.1544, Y.1546, Y.1550, Y.1560, Y.1561, Y.1563, Y.1564, Y.1565, Y.1566, Y.1567, Y-series Suppl. 60.

### M.2 Question

Study items to be considered include, but are not limited to:

– General and cross-technology performance studies

• How should the generic measurement points, reference events, communication functions, performance outcomes, and performance parameters defined in ITU-T Recommendations be supplemented to address new network capabilities (e.g., multipoint connections, multi-connection calls, and modification of connection attributes), new access arrangements (e.g., wireless, satellites, HFC, xDSL, Passive Optical Networking), and new services/applications (e.g., interactive multimedia communications, personal and terminal mobility including IMT-2020 systems, flexible routing and charging, security, IP network service access, web browsing, Network Function Virtualization, NFV, and virtual private networks)?

• How can the measurement of packet networks be improved, for example, to support more meaningful service level specifications between network operators and their customers?

• How can the measurement of packet networks be coordinated, to address the issues and complexities associated with large network scale?

• How should Recommendations on network performance address communications built on heterogeneous networking technologies, such as seamless wired-wireless communications support?

• What new metrics can be developed and specified to serve the packet network infrastructure, including the needs of measurement systems and other fundamental applications (such as timing systems)?

• How can the definition or the measurement of packet loss be improved to discriminate events that affect end systems and user applications?

• How can the definition or the measurement of latency and packet delay variation be improved to provide more information to end-system designers?

– Network performance, including new technologies and existing technologies such as virtual network overlays, IP, MPLS, and Ethernet:

• Which layer(s) or other conventions have end-to-end significance in specifying performance the new technology?

• What reference events will be available to define performance parameters for these networks?

• What performance parameters and statistics should be standardized for such networks?

• How can complex topologies be assessed, such as multipoint-to-multipoint?

• What QoS levels will be needed by the services supported on these networks?

• How will the end-to-end QoS objectives for new services be achieved when more than one network participates in the provision of communications?

• To what extent will QoS commitments depend on the existence of traffic contracts that completely specify the characteristics of the offered traffic?

• How will QoS commitments of networks be verified?  
The above technologies are being deployed in new network domains, such as wired and wireless, access and transport, and within the home and business. The scope of this Question includes all these domains.

– What QoS class descriptions can assist the interconnection of network domains?

– IP network performance

• What additional performance objectives for systems employing application-layer packet loss compensation should be specified in Recommendation Y.1541?

• How will the end-to-end QoS objectives for IP-based services be achieved when more than one IP network participates in the provision of communications?

• How will users of IP-based services communicate their need for an IP QoS commitment?

• What additional performance objectives for compressed data (e.g., MPEG video, G.72x codec signals) should be specified in Recommendation Y.1541?

• In addition to the applications and services mentioned above, will machine to machine (M2M) and camera and sensor networks influence the objectives or require new QoS classes?

– TCP, UDP, QUIC, and other transport protocol performance

• How will evolution of these protocols be reflected in new performance parameters?

• How will evolution of these protocols influence IP objectives or QoS classes?

– Modelling transmission-related components of end-systems

• What end-system components should be modelled, so that the UNI-UNI performance can be estimated in mid-path measurement deployment?

• What verification procedures are useful, when models of performance cannot be standardized, but available systems can be tested?

– Considerations for realizing virtual measurement systems

• Areas for further study are discussed in Recommendation ITU-T Y.1550 Appendix I.

– How should the study items areas be organized into tasks?

### M.3 Tasks

Tasks include, but are not limited to:

– draft new Recommendation on new technology performance parameters;

– updates and maintenance of the Recommendation QoS class mapping between domains;

– updates and maintenance of the Recommendation on various performance parameters;

– updates and maintenance on Y.1540 IP performance parameters and Y.1541 IP-based network objectives;

– update the fundamental Recommendation on general aspects of quality of service and network performance in digital networks, I.350;

– continue to develop and expand the current Recommendations on assessment (testing) of key performance parameters to serve many audiences, including diagnostic and monitoring operations;

– new or revised Recommendation on IP/packet performance parameters;

– updates and maintenance on Y.1550 on considerations for virtualized measurement systems;

– additions and updates to other existing Recommendations.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### M.4 Relationships

Recommendations:

– I.371, I.381, I.610, O.191, G.828, Y.1710, Y.1711, Y.1731

Questions:

– J/12, K/12

Study Groups:

– ITU-T SG2, SG13, SG15, SG16, SG17

– ITU-R SG5, SG6

Other bodies:

– MEF, IETF working groups on performance issues, IEEE 802 LAN/MAN Standards Committee, 3GPP, Broadband Forum, ETSI, ANSI, GSMA

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION N/12 Objective and subjective methods for evaluating perceptual audiovisual quality in multimedia and television services

(Continuation of Question 19/12)

### N.1 Motivation

In digital transmission systems, the perceptual quality of the audiovisual signal is influenced by a number of interacting factors, such as source coding and compression, bit rate (fixed or variable), delay, bandwidth, synchronization between the media, transmission impairments, and many others. New services that use IP, wireless, mobile, NGN, etc. are providing ubiquitous access for multimedia services. Audiovisual multimedia cover multichannel audio, television, and 3D video applications including interactive ones, in addition to other applications such as videoconferencing, personal computer desktop conferencing, interactive educational and training services, groupware, interactive gaming, and videotelephony. This Question focuses on perceptual impacts of compression, transmission, and decompression on audiovisual quality of these multimedia services and applications.

The effect of the source and display is particularly important and necessary for the case of 3DTV and high-dynamic range (HDR) displays, as both these technologies are not mature and still introduce quality problems. Display technologies are evolving from 2D to 3D, high-definition to ultra-high definition, low dynamic range to wide-gamut and high-dynamic range displays. In particular, HDR images are currently typically displayed on low-dynamic range (LDR) displays because of the limited availability of HDR displays. In order to visualize HDR images on LDR displays, tone mapping is necessary and this creates information loss that can deteriorate the quality and details of the HDR image. Recently, HDR displays have appeared on the market but they use internal processing that can affect the video quality. 3DTVs exhibit crosstalk to various degrees and can impact negatively the viewing experience. For these new technologies, the quality impact of the display and transmission (or camera, production and transmission) cannot always be separated. Although bandwidths available in cable transmission are well suited for ultra-high definition television (UHDTV), maintaining adequate video quality still represents a challenge. ITU‑R has recommended methods for the subjective assessment of picture quality (e.g., BT.500-13, BT.1788, BT.2021). There is a need to confirm that those subjective assessment methods and set-up requirements (including selection of the display, settings/calibration of the display, viewing distance, angle, luminance levels etc.) are equally applicable to the case of next-generation visual media, such as television transmission on digital or mixed analogue-digital chains, 3D, HDR and UHDTV images. There is also a need to study and develop evaluation methods (objective and subjective) for special tasks such as object recognition.

Concerning the measurement of the overall quality of experience (QoE), it includes not only a single impairment of each mono-media but also inter-media relation and response time of user operation. There is a need to identify the group of parameters that can provide objective measurement of the overall QoE and continuous in-service monitoring and control of it along the transmission chain.

In order to develop the two-way measurement techniques required for conversational applications, a basis in one-way audio and video quality evaluation must first be defined and validated. Considering the spread of broadband connections to business and the home, the bandwidths will support both low resolution, e.g., quarter video graphics array (QVGA), and standard, high and ultra-high definition imagery. As an example, audio multimedia applications currently range from audio for narrow-band applications, e.g., video telephony, to the enhanced audio contained in 7.1 surround sound systems for interactive gaming. In the future, HDR, 3D programmes and 3D games are expected to become more widely available. Objective and subjective methods for assessing the perceptual quality of these media services are needed, particularly those relating to transmission.

Objective methods: Current objective quality measuring techniques for audiovisual applications do not correlate to the user opinion on the perceived audiovisual quality with the desirable accuracy. It is therefore necessary to identify objective techniques for measuring the various individual and combined effects of factors such as digital compression, transmission, storage, and others on the perceived quality of audiovisual systems. It is also important to verify that these techniques are meaningful by correlating proposed objective tests with corresponding subjective test data.

Subjective methods: There is a need to continue to develop new subjective methods to address new audiovisual services. The perceived quality depends on the kind of application and on the tasks the applications are used for. For example, in a free conversation through a videophone or videoconferencing application, the perceived quality may primarily depend on delay, lip-synchronization and audio quality, while in a mainly one-way application like remote-teaching the perceived quality could be primarily related to the quality of graph and low motion picture sequences.

These studies include the maintenance of and enhancements to existing Recommendations, and the development of new Recommendations as needed.

Much of the work on this Question (and its predecessors) was and will be done in cooperation with the video quality experts group (VQEG).

### N.2 Question

Study items to be considered include, but are not limited to:

– Interaction of media: What subjective and objective measurement methods should be used to evaluate end-to-end quality of each medium (e.g., video, audio, television, 3D video) and the interactions between the media, with particular attention to the audiovisual quality assessment of systems used for videoconferencing/videotelephony and other interactive multimedia services? What are the quality levels that can be defined by objective or subjective methods in different applications (or tasks) taking into account the interactions between media?

– Transmission errors: What objective methods could be used for in-service measurement and monitoring of transmission systems for such multimedia services in the presence of transmission errors? What new subjective measurement methods should be used for the evaluation of transmission quality of real time audiovisual services by expert observers resulting in the identification of specific flaws in the transmission equipment or environment? What procedures should be used, and which dimensions, transforms, and partial or differential signals should be viewed by experts to evaluate specific impairments of real time audiovisual services? What objective and subjective methods can be used to evaluate audiovisual signals with time-varying quality?

– Impairment characterizations: Among the most significant factors (e.g., spatial resolution, temporal resolution, colour fidelity, audio and visual artefacts, media synchronization, delay, cross-talk etc.) affecting the overall quality of multimedia services, what objective and subjective methods assess the extent of or can differentiate between these factors? How can the mutual interaction between these factors be objectively and subjectively measured with respect to their influence on overall audiovisual quality? For what applications can the assessment methods be shown to be useful and robust over a range of conditions? What kind of artificial impairment generator would be useful for subjective or objective methods?

– Evaluation of specific services: What assessment methods (objective and subjective) can be used to characterize the quality effects of multipoint distribution for interactive communication and other new audiovisual services such as remote monitoring, interactive gaming, and mobile audiovisual communication?

– Test methodologies: What subjective methods and assessment tools are required to fully describe perceived visual or audiovisual impairments in terms of measurable system parameters? What kind of references should be used in subjective tests? What methods can be used to measure the video quality of 3D video? What new subjective methods are needed when analysing new applications and usage scenarios? What kind of service or application design is needed to minimize visual fatigue in 3D video applications? What methods can be used to measure the visual fatigue level introduced into a 3D video signal by the source content (e.g., amount of motion, depth of field), compression and transmission?

– Combination of test results: In some cases it may be useful to combine objective measures (e.g., video measures, audio measures, media synchronization) to provide a single figure of merit. In this regard, which objective measures and/or techniques should be combined, and in what manner, so that the figure of merit correlates satisfactorily with subjective test results?

– Test sequences: While the library of test sequences has increased greatly recently (e.g., www.cdvl.org), there is still a need for more test sequences, especially those with audio included and 3D. Which audiovisual test material (e.g., audiovisual test sequences, 3D video) can be standardized for subjective and objective evaluations? In addition to the definitions of SI and TI in P.910, which criteria (objective and/or subjective) should be used to characterize and classify multimedia test material?

– Validation and applicability of objective methods: There are three basic methodologies of objective picture quality measurement. Full-reference (FR) uses the full bandwidth video input. Reduced-reference (RR) uses lower bandwidth features extracted from the video input. No reference (NR) has no information about the video input. What objective methodology should be used for different multimedia applications? What subjective methods should be used to validate each of the three basic objective methodologies? How can hybrid perceptual/bitstream (hybrid) methodologies use information about the encoded bit‑stream to supplement FR, RR or NR methodologies?

– What enhancements to existing Recommendations are required to provide energy savings directly or indirectly in information and communication technologies (ICTs) or in other industries? What enhancements to developed or new Recommendations are required to provide such energy savings?

– What are the quality requirements for transmission of UHDTV?

– Are the current methods recommended for subjective assessment of digital picture quality also applicable to scenarios where the display is not transparent, such as in 3DTV or HDR images? Are the current quality assessment methods applicable to ultra-high definition television?

– How should the impairment introduced by the display be taken into account in evaluation of the viewing experience?

– How should the impairments introduced by the transmission chain be taken into account, such as those introduced by digital or mixed analogue-digital television transmission chains?

– How should the impairment introduced by the (stereo-) camera be taken into account in evaluation of the viewing experience?

– What objective methodology can be used to jointly analyse the perceptual quality of the entire stream, including the quality of both the camera and the display?

– How should the objective measurement of impairments introduced by digital or mixed analogue-digital transmission networks be carried out?

– Which network parameters should be used to provide objective measurement of the overall QoE and should be the basis for continuous in-service monitoring along the transmission chain both for digital and for mixed analogue-digital television transmission?

– What perceptual image/video quality assessment methods can be used to determine which tone-mapping operator maintains best the visual information of an HDR image or produces the highest-quality LDR image? What perceptual image/video quality assessment methods can be used to assess the quality of HDR content?

– What methods can be used to measure the visual fatigue in 3D video from the video capture, rendering and display?

### N.3 Tasks

Tasks include, but are not limited to:

– Quality assessment in multimedia services requires on the one hand the continuous updating of Recommendations under the responsibility of Study Group 12 and also the definition of new task oriented/application-dependent evaluation and subjective methods for the combined evaluation of audio and video signals.

– A new Recommendation utilizing expert viewers is expected. Three Recommendations defining objective methods for assessing audiovisual quality in multimedia services are expected to be approved.

– Initial work on quality assessment of interactive gaming applications will result in a new Recommendation.

– Maintenance and revision of Recommendations on 3D subjective methods.

– It is anticipated that new Recommendations will address: methods to characterize and select appropriately 3D displays for subjective evaluation of 3D picture quality; methods for HDR and UHDTV quality evaluation and methods to assess/characterize the impact of non-transparent displays on viewing experience.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### N.4 Relationships

Recommendations:

– P- and J-series

Questions:

– K/12

Study Groups:

– ITU‑T SG9, SG13, SG15, SG16

– ITU‑R SG6

Other bodies:

– ITU IRG-AVQA, VQEG, IETF and regional standardization bodies (e.g., ATIS)

WSIS Action Lines:

– C2

Sustainable Development Goals:

– SDG 9

## DRAFT QUESTION O/12 Perceptual and field assessment principles for quality of service (QoS) and quality of experience (QoE) of digital financial services (DFS)

(Continuation of Question 20/12)

### O.1 Motivation

QoE of Digital Financial Services (DFS)[[1]](#footnote-2) turns out to be one of the most critical in the developing digital society, there is an increasing importance to continue the support to the global community with the extension of appropriate methodologies for DFS quality assessment both in perceptual considerations and in the field assessment.

Work under this Question is carried out in response to:

– PP-22 Resolution 204 – Use of ICTs to bridge the financial inclusion gap

– WTSA-20 Resolution 89 – Promoting the use of information and communication technologies to bridge the financial inclusion gap

Three Recommendations and one Supplement related to DFS have been approved by SG12.

When several stakeholders, both of the financial and the telecom sectors work together to provide end-to-end DFS solutions or applications, each of the stakeholders needs to understand how to achieve the end-to-end performance objectives. Such objectives must be both adequate for the service being offered and feasible based on the available networking technologies.

A framework is needed to guide the development of Recommendations for performance aspects of Digital Financial Services including those supported by the emerging and heterogeneous infrastructure. Such a framework is also essential for relating performance.

Other Questions, ITU Study Groups, and some Standardization Bodies should expect that unique needs in the area of Digital Financial Services will be satisfied by this Question’s work, so that they can continue with their work plans without overlap.

The Question will provide the necessary support to produce field test and processing plans to execute appropriate tests of DFS.

The following major deliverables, in force at the time of approval of this Question, fall under its responsibility: G.1033, P.1502, P.1503, P-series Suppl. 30.

### O.2 Question

Study items to be considered include, but are not limited to:

– General and cross-technology performance studies

• How should the generic measurement points, reference events, communication functions, performance outcomes, and performance indicators be defined for different DFS scenarios and for different DFS implementations

• How can the measurement of DFS be coordinated, to address the issues and complexities associated with large network scale?

• Which layer(s) or other conventions have end-to-end significance in specifying performance of DFS?

• What reference events will be available to define performance indicators for these networks?

• Which scenarios, performance indicators and statistics should be standardized for such networks?

• How can complex topologies be assessed, e.g., topologies including multiple endpoints or solutions linking DFS with traditional banking scenarios such as checking accounts?

• Who are the different entities and interactions in use cases related to Digital Financial Services?

• What QoS levels will be needed by the services supported on these networks?

• How will the end-to-end QoS objectives for DFS be achieved when more than one network participates in the provision of communications?

– What new test plans are needed to evaluate (subjectively) end-to-end DFS over fixed and/or mobile networks?

### O.3 Tasks

Tasks include, but are not limited to:

– draft new Recommendation on new aspects of DFS QoE and QoS;

– new or revised Recommendation on DFS QoE and QoS;

– additions and updates to other existing Recommendations.

An up-to-date status of work under this Question is contained in the SG12 work programme at [https://itu.int/ITU-T/workprog/wp\_search.aspx?sg=12](https://www.itu.int/ITU-T/workprog/wp_search.aspx?sg=12).

### O.4 Relationships

Recommendations:

– P-series, G-series

Questions:

– J/12, K/12

Study Groups:

– ITU-T SG13

– ITU-D SG2

Other bodies:

– GSMA, ETSI TC STQ, 3GPP

WSIS Action Lines:

– C2, C7

Sustainable Development Goals:

– SDGs 5, 8, 9, 10

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

1. Methods to electronically store and transfer funds; to make and receive payments; to borrow, save, insure, and invest; and to manage a person's or enterprise's finances; and where these services are accessed via mobile communication devices. [↑](#footnote-ref-2)