

I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1312

(02/2016)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS

Telemeeting assessment

Method for the measurement of the communication effectiveness of multiparty telemeetings using task performance

Recommendation ITU-T P.1312

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Voice terminal characteristics	Series	P.30
		P.300
Reference systems	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of speech quality	Series	P.80
		P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than voice services	Series	P.1500

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T P.1312

Method for the measurement of the communication effectiveness of multiparty telemeetings using task performance

Summary

Recommendation ITU-T P.1312 describes a subjective test method for quantifying the effectiveness of telemeeting systems in conveying information in multiparty conversation scenarios. This method utilizes a predefined set of tasks designed to provoke rapid turn-taking and concurrent talking among participants. The goal of these tasks is to stress conferencing systems in order to clearly and measurably demonstrate their limits.

The method measures the rate at which multiple participants exchange information to assess the effectiveness of communication systems compared to face-to-face communication. This constitutes an objectively-measured performance metric providing stable and repeatable results across different laboratories and that is highly discriminative across various types of system properties.

This method may be used to evaluate individual subcomponents of a telecommunication system or the system as a whole. The proposed method is especially sensitive to the ability to transmit multiple concurrent voices and maintain their independence. The test has demonstrated sensitivity to system parameters including duplex capability, spatialization, bandwidth and delay. As such, it is suitable for the assessment of multiparty telemeeting systems (see Recommendation ITU-T P.1301 for general guidelines and additional methods).

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1312	2016-02-29	12	11.1002/1000/12751

Keywords

Double-talk, face-to-face reference, interactive test method, system effectiveness, task performance measurement, telemeeting.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2016

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope.....	1
2	References.....	1
3	Definitions	1
3.1	Terms defined elsewhere	1
3.2	Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	2
5	Conventions	2
6	Task effectiveness test method	2
6.1	Purpose	2
6.2	Test principle	3
6.3	Reference condition and test scenarios.....	4
6.4	Test facilities and equipment.....	6
6.5	Test administrators	6
6.6	Subjects.....	7
6.7	Experiment design	7
6.8	Task performance measurement.....	8
	Annex I – Construction of the pool of words for the word task	10
	Appendix I – Example of instructions	11
	Appendix II – Example approaches subjects may use to complete the tasks	13
	Appendix III – Example test results.....	14
	Bibliography.....	16

Recommendation ITU-T P.1312

Method for the measurement of the communication effectiveness of multiparty telemeetings using task performance

1 Scope

This Recommendation describes a method for quantifying the amount of information that multiple interlocutors correctly exchange over a system. The measure serves to assess the effectiveness of a voice communication system by comparing the amount of correct information exchanged over the system relative to the amount of correct information these interlocutors achieve in face-to-face communication. This metric qualitatively evaluates a multiparty conferencing system's ability to transmit multiple voices and maintain their independence. The method has demonstrated sensitivity to system parameters including duplex capability, spatialization, bandwidth and delay.

The method described herein applies only to voice communication; audiovisual communications is not covered by the scope of this Recommendation. The method does not include the measurement of technical properties of the systems.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [[ITU-T P.800](#)] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.
- [[ITU-T P.805](#)] Recommendation ITU-T P.805 (2007), *Subjective evaluation of conversational quality*.
- [ITU-R BS.1116] Recommendation ITU-T BS.1116-3 (2015), *Methods for the subjective assessment of small impairments in audio systems*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 concurrent talk [[b-ITU-T P.1311](#)]: When far-end speech of two or more talkers occurs simultaneously at a given point, typically the near end terminal.

3.1.2 multiparty [[b-ITU-T P.1301](#)]: More than two persons. Example: More than two persons are participating in a telemeeting, having a conversation, performing a test task together, etc. The term multiparty does not specify if the persons are distributed across two or more locations. If not explicitly stated differently, multiparty implicates that the persons are at two or more than two locations. When further specification is necessary, additional terms will be used or the number of locations will be explicitly stated.

3.1.3 one per site [[b-ITU-T P.1301](#)]: One person per connected location. Example: In a multiparty one-per-site telemeeting, more than two sites are connected with only one person present at each site.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 performance score: The number of correct sequences conveyed by an individual, team or group.

3.2.2 system effectiveness: The ratio, expressed as a percentage, between the performance score achieved on a given system and the performance score obtained in face-to-face communication.

3.2.3 reader: A participant for whom the role is to read out loud a sequence of letters or words to his or her partner(s).

3.2.4 responder: A participant for whom the role is to register the sequence of letters or words his or her partner communicates.

3.2.5 face-to-face meeting: A meeting for which all participants are collocated.

3.2.6 duplex: A communication channel for which signals can flow in both directions.

3.2.7 spatialization: The process of rendering audio sources in a three-dimensional environment.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

SuT System-under-Test

5 Conventions

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus this requirement need not be present to claim conformance.

The keywords "is prohibited from" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords "can optionally" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Recommendation.

6 Task effectiveness test method

6.1 Purpose

The purpose of the Recommendation is to enable multiparty conversational tests of voice-only communication systems that are sensitive to subtle changes of system parameters, to produce results that are repeatable in different laboratories with predictable test duration.

The test described in this Recommendation is based on the objectively-measured performance of participants when solving tasks over the communication system-under-test (SuT). This is unlike the conversational tests of [[ITU-T P.805](#)], which are based on subjective impressions reported by