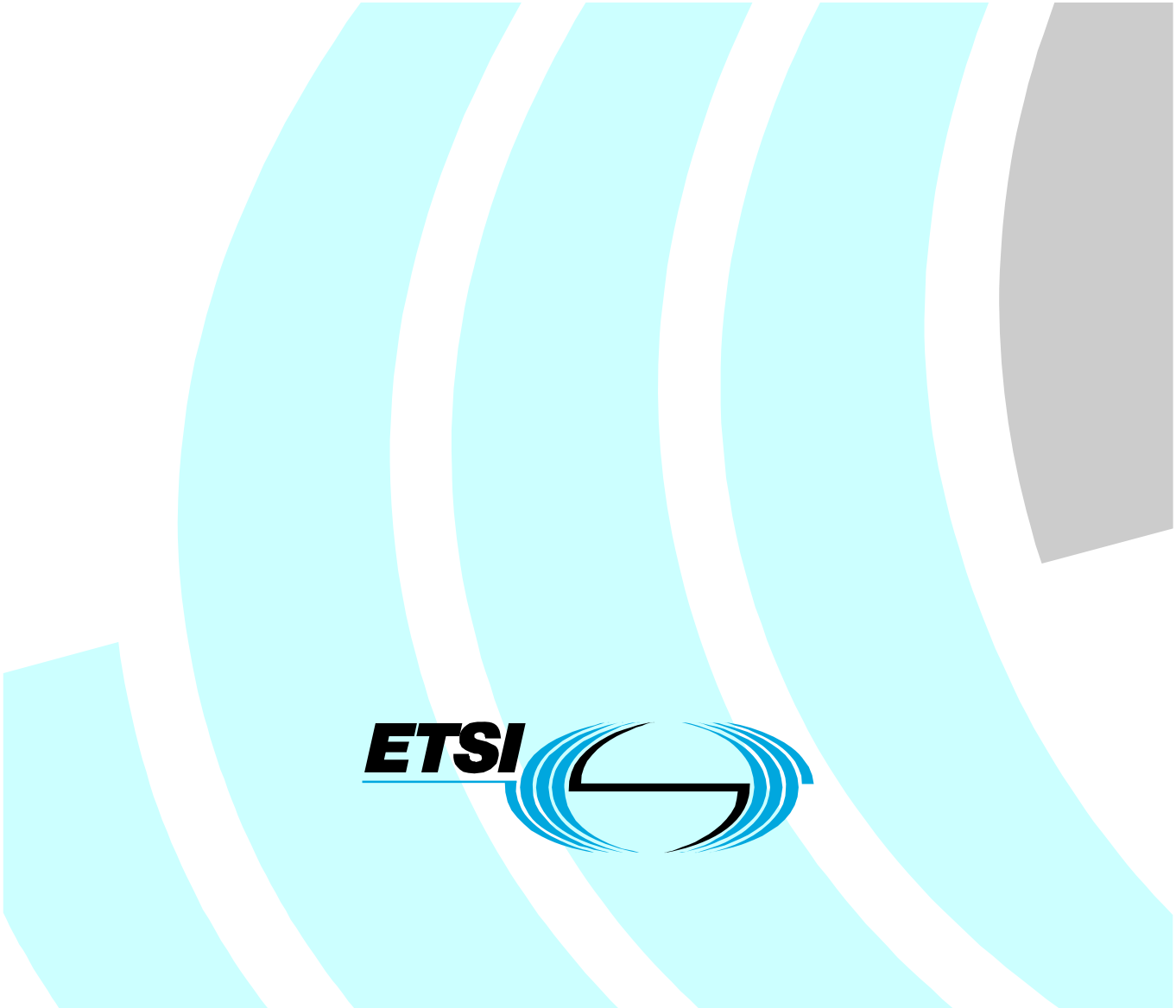
Final draft ETSI EG 202 670 V1.1.1 (2009-12)

ETSI Guide

Human Factors (HF); User Experience Guidelines for real-time communication services expressed in Quality of Service terms



Reference
DEG/HF-00112

Keywords
interaction, quality, service

ETSI

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

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

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

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

Foreword

This ETSI Guide (EG) has been produced by ETSI Technical Committee Human Factors (HF), and is now submitted for the ETSI standards Membership Approval Procedure.

The present document is primarily for the HF community.

The present document updates and replaces EG 202 534 [i.11]. New guidelines have been added and some guidelines from EG 202 534 [i.11] were not considered relevant for the present document (e.g. because of being out of date with technology developments). In addition, whereas EG 202 534 [i.11] focused on person-to-person (two-way) communication services, the present document extends to real-time person-to-machine (one-way) communication services.

Introduction

The real-time communication services that are currently available and in development place different demands on the communication channel and terminal equipment. Also they continue to evolve for both mobile and fixed usage. The services offer great potential but also complex choices regarding the most appropriate technologies and media that are suitable for different communication situations. Taking video communication as an example, there are advances at the two extremes of usage: for video calls with mobile telephones and conference-room 'Telepresence' systems and between these extremes video communication is increasingly available through personal computer-based 'Web-conferencing' services.

TR 102 353 [i.64] identified the need to develop guidelines on real-time communication services. The main aim is to provide guidelines for network operators, equipment manufacturers and service providers that address the user experience of different communication services in terms of either:

- the configuration and quality of service (QoS) of a particular communication service;
- the best choice between different communication services, where user experience data exists to enable recommendations on the basis of of set-up time, reliability, type of user tasks and user situations.

The general requirements for the guidelines contained in the present document are / main inclusion criteria for the guidelines are:

- **Empirical basis** - the guidelines are based on user test results where data has been collected on user performance and/or user opinion. The guidelines are therefore justified on the basis of either a laboratory experiment, field study, survey or an expert panel of people with empirical knowledge gained through experience with users or user tests.
- **Known QoS** - the test systems were of specified QoS parameter values, thereby enabling a statement of expected user experience for a given service with particular QoS.
- **Technology independent** - because the user experience results were related to test systems with specified QoS rather than a particular technology.

- **Vendor independent** - because the user experience results were related to test systems with specified QoS rather than a particular vendor.

- **User communication phase** - the guidelines concern user experience once communication is established, therefore the guidelines do not address interface design and call control;

- **Stakeholder interest** - priority has been given to developing guidelines in areas where stakeholder interest has been shown or is anticipated. Therefore Guidelines are provided for topics that have been identified as important for intended guideline users and for which user-based data existed or could be collected. For this reason, the development of guidelines from published literature is not exhaustive.

The guidelines are grouped into topics that include different user tasks (e.g. decision making, negotiation, persuasion), technical parameters (e.g. delay, packet loss, frame-rate) and special user groups (e.g. people with speech impairments, deaf and hearing impaired people, people with cognitive impairments). The user test results are derived primarily from industry and European Commission Framework Programme research reports, journal articles, conference proceedings and standards documents.

1 Scope

The present document provides guidelines for the user experience of real-time communication services. The services include person-to-person (two-way) communication and person-to-machine (one-way) communication.

The present document is revised from EG 202 534 [i.11] that was restricted to person-to-person communication services. The revision adds new guidelines that have been requested by stakeholders and omits some of the previous guidelines that are no longer considered relevant.

The guidelines are based on empirical data about user experience. Most of the data is obtained from scientific papers. A minority of guidelines are based on existing standardisation documents to ensure that relevant normative and informative standards material is clear amongst results available from journal articles, conference papers, research reports, etc.

The present document does not replace any existing standards.

Most of the guidelines derived from scientific papers are specific to a particular context, in that the original user tests were for specific tasks, users and technical parameters and therefore the results may not be generalisable. Although the guidelines provide information about the main user experience measure(s) and technical parameter(s) of a particular test result, it is beyond the scope of the present document to provide all of the variables concerned with each user test. However, the origin of each empirical source of a guideline is shown and listed in the References.

Those readers interested in the details of a particular guideline empirical source are also directed to a web-based system (http://portal.etsi.org/STFs/STF_HomePages/STF354/STF354.asp) that provides this information.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] Anderson, A., Smallwood, S., MacDonald, R., Mullin, J., Fleming, A. & O'Malley, C. (2000): "Video data and video links in mediated communication: What do users value? International Journal of Human-Computer Studies", 52(1), 165-187.

[i.2] ANSI T1.522-2000: "Quality of Service for Business Multimedia Conferencing".

[i.3] Bekkering, E. and Shim, J. (2006). "i2i Trust in Videoconferencing", July 2006/Vol. 49, No. 7 Communications of the ACM, pp. 103-107.

[i.4] Blythe, M.; Overbeeke, K., Monk, A., Wright, P. (Eds) (2004). "Funology: From Usability to Enjoyment" Kluwer Academic.

[i.5] Brennan, D., Georgeadis, A., Baron, C., Barker, L. (2004): "The Effect of Videoconference-Based Telerehabilitation on Story Retelling Performance by Brain-Injured Subjects and Its Implications for Remote Speech-Language Therapy", Telemedicine Journal and e-Health. June 1, 2004, 10(2): 147-154.

[i.6] Brooks, P., Schliemann, T., Hestnes, B., Frowein, H., Aaby, C., O'Malley, C. (2003): ""Final Report Project IST-1999-11577 Eye-2-Eye: Fitness-for-Purpose of Person-Person Communication Technologies". EC Deliverable IST11577/SEF/DIS/DS/Pub/008/b1, June 2003.

[i.7] Brooks, P. & Hestnes, B. (2003): "User-centred technical guidelines for real-time human communication services: Requirements and derivation". Proceedings of the 19th International Symposium on Human Factors in Telecommunication, Berlin, Germany, December 1-4 2003, pp. 11-18.

[i.8] Cullum, C., Weiner, M. Gehrmann, H. Hynan, L. (2006): "Feasibility of Telecognitive Assessment in Dementia. Assessment", Volume 13, No. 4, December 2006 385-390.

[i.9] Cuny, R., Li, M., Kristensson, M. (2006): "Mobile service applications and performance in UMTS", In Soldani, D., Li, M., Cuny, R. (Eds) "QoS and QoE Management in UMTS Cellular Systems", John Wiley & Sons, Ltd.

[i.10] Doherty-Sneddon, G., Anderson, A., O' Malley, C., Langton, S., Garrod, S. & Bruce, V. (1997): "Face-to-face and video mediated communication: A comparison of dialogue structure and task performance", Journal of Experimental Psychology: Applied , 3(2), 105-125.

[i.11] ETSI EG 202 534: "Human Factors (HF); Guidelines for real-time person-to-person communication services".

[i.12] ETSI ES 202 667 : "Speech and multimedia Transmission Quality (STQ); Audiovisual QoS for communication over IP networks".

[i.13] ETSI ES 202 737: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

[i.14] ETSI ES 202 738: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".

[i.15] ETSI ES 202 739: "Speech and multimediaTransmission Quality (STQ); Transmission requirements for wideband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

[i.16] ETSI ES 202 740: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".

[i.17] ETSI ETR 292: "Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D); Technical requirements specification; Network management".

[i.18] ETSI TR 102 535: "Human Factors (HF); Guidelines for real-time person-to-person communication services; Future requirements".

[i.19] ETSI ETR 334: "Human Factors (HF); The implications of human ageing for the design of telephone terminals".

[i.20] ETSI ETR 297: "Human Factors (HF); Human Factors in Videotelephony".

[i.21] ETSI TS 122 105: "Universal Mobile Telecommunications System (UMTS); Services and service capabilities (3GPP TS 22.105)".

[i.22] ETSI TS 123 107: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Quality of Service (QoS) concept and architecture (3GPP TS 23.107 Release 8)".

[i.23] Følstad, A., Brooks, P., Heim, J., Schliemann, T., Wiig., S., Hestnes, B., Heiestad, S., Ulseth, T., Frowein, H., Aaby, C., O'Malley, C., Brundell, P., Lonsdale, P. (2002): "Results of Field Experiments of Communication Media. IST Project 1999-11577. Eye-2-Eye: Fitness-for-purpose of Person-Person Communication Technologies", CEC Deliverable IST11577/SEF/DIS/DS/Pub/004/b1, October 2002.

[i.24] France, E. F., Anderson, A. H. & Gardner, M. (2001): "The impact of status and audio conferencing technology on business meetings", International Journal of Human Computer Studies, 54: 857-876.

[i.25] Frowein, H., Donker, W., Devoldere, P., Schliemann, T., Brooks, P., Følstad, A., Hestnes, B. et al. (2003): "Specification of Cost-Benefit Analysis Tool. Project IST-1999-11577 Eye-2-Eye: Fitness-for-Purpose of Person-Person Communication Technologies", EC Deliverable IST11577/IvD/RAD/DS/FP5/052/b1, May 2003.

[i.26] Ghinea, G., Thomas, J. (2005). "Quality of Perception: User Quality of Service in Multimedia Presentations", IEEE Transactions on Multimedia, Vol. 7, No. 4, August 2005, pp. 786-789.

[i.27] Grayson, D. & Coventry, L. (1998) "The Effects of Visual Proxemic Information in Video Mediated Communication", SIGCHI Bulletin Vol.30 No.3, July 1998.

[i.28] Hellström, G., Aström, L. (2008): "The value of total conversation call with emergency services, Proceedings of International Symposuim on Human Factors in Telecommunications", Kuala Lumpa, Malaysia, 17-20 March 2008.

[i.29] Hestnes, B., Brooks, P., Heiestad, S. (2004): "Mobile Eye-phone - a study of relevance, effectiveness and user-perceived suitability", Fornebu, Telenor R&D Report (Scientific Report R2/2004).

[i.30] Hestnes, B., Brooks, P., Heiestad, S., Ulseth, T., Aaby., C. (2003): "Quality of Experience in real-time person-person communication services - User based QoS expressed in technical network QoS terms", proceedings of the 19th International Symposium on Human Factors in Telecommunication, Berlin, Germany, December 1-4 2003, pp. 3-10.

[i.31] ISO 9241 (all parts): "Ergonomic requirements for office work with visual display terminals (VDTs)".

[i.32] ITU-R Recommendation BT.1359-1 (2003): "Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems".

[i.33] ITU-T Series H: "Audiovisual and multimedia systems, Supplement 1 (05/99) (1999) Application profile - Sign language and lip-reading real-time conversation using low bit-rate video communication".

[i.34] ITU-T Recommendation E.800 (2008): "Definitions of terms related to quality of service".

[i.35] ITU-T Recommendation F.700 (2000): Framework Recommendation for multimedia services.

[i.36] ITU-T Recommendation G.108 (1999): "Application of the E-model: A planning guide".

[i.37] ITU-T Recommendation G.114 (2003): "One-way transmission time", International.

[i.38] ITU-T Recommendation G.1010 (2001): "End-user multimedia QoS categories".

[i.39] ITU-T Recommendation G.1080: "Quality of experience requirements for IPTV services".

[i.40] ITU-T Recommendation P.10/G.100: "Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100", International Telecommunication Union, Geneva, Switzerland, 2008.

[i.41] ITU-T Recommendation Y.1541 (2006): "Network performance objectives for IP-based services", International Telecommunication Union, Geneva, Switzerland.

[i.42] Jumisko-Pyykkö, S., Vinod Kumar, M. V., Liinasuo, M., & Hannuksela, M. (2006): Acceptance of audiovisual quality in erroneous television sequences over a DVB-H channel. In Proceedings of the Second International Workshop in Video Processing and Quality Metrics for Consumer Electronics".

[i.43] Kilgore, R.M., and Chignell, M., (2006): "Listening to unfamiliar voices in spatial audio: Does visualization of spatial position enhance voice identification? Proceedings of the 20th International Symposium on Human Factors in Telecommunication", Sophia Antipolis, France, March 21-23 2006.

NOTE: Available from http://www.hft.org/HFT06/HFT_06_programme.htm).

[i.44] Knoche, H., McCarthy, J. D. and Sasse, M.A. (2005): "Can small be beautiful?: assessing image resolution requirements for mobile TV", Proc. ACM Multimedia 2005, ACM Press, 829-838.

[i.45] Knoche, H., McCarthy J., Sasse, M. A. (2006): "Reading the Fine Print: The Effect of Text Legibility on Perceived Video Quality in Mobile TV", ACM Multimedia, ACM, 727-730.

[i.46] Knoche, H.O., Sasse, M.A. (2008): "The sweet spot: how people trade off size and definition on mobile devices. In Proceeding of the 16th ACM international conference on Multimedia", Vancouver, British Columbia, Canada: ACM, pp. 21-30. .

[i.47] Mashima, P., Birkmire-Peters, D., Syms, M., Holtel, M., Burgess, L., Peters, L. (2003): "Telehealth: Voice Therapy Using Telecommunications Technology", American Journal of Speech-Language Pathology Vol.12 432-439.

[i.48] Mason, S. (2006): "Mobile TV - results from the DVB-H trial in Oxford", EBU TECHNICAL REVIEW - April 2006.

NOTE: http://www.ebu.ch/en/technical/trev/trev_306-mason.pdf

[i.49] McCarthy, J., Sasse, M. A. & Miras, D. (2004): "Sharp or Smooth? Comparing the effects of quantization vs. frame rate for streamed video. Proceedings of CHI 2004", Vienna, Austria, April 20-24, pp. 535-542.

[i.50] Mulbach, L., Bocker, M. & Prussog, A. (1995): "Telepresence in Videocommunications: A Study on Stereoscopy and Individual Eye Contact", Human Factors, 37(2), 290-305.

[i.51] Nokia (2004): "Quality of Experience (QoE) of mobile services: Can it be measured and improved?", White Paper, Nokia Corporation, Finland, 2004.

[i.52] O' Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G & Bruce, V. (1996): "Comparison of Face-to-Face and Video-Mediated Interaction. Interacting with Computers", 8(2), 177-192.

[i.53] O'Malley, C., Brundell, P., McFadzean, J., Lonsdale, P., Schliemann, T., Brooks, P. et al. (2002): "Results of Laboratory Experiments of Communication Media. IST Project 1999-11577. Eye-2-Eye: Fitness-for-purpose of Person-Person Communication Technologies", CEC Deliverable IST11577/UON/SOP/DS/Pub/003/b1, December 2002.

[i.54] Post & Telestyrelsen (2005): "Mobile video communications for people who are deaf: Report on trial operations with broadband for people with disability", Swedish National Post and Telecom Agency Report No. PTS-ER-2005:14.

[i.55] Raake, A. (2006). Short- and Long-Term Packet Loss Behavior: Towards Speech Quality Prediction for Arbitrary Loss Distributions, IEEE Transactions on Audio, Speech and Language Processing, Vol. 14, No. 6.

[i.56] Schliemann, T., Asting, T., Brooks, P., Følstad, A., Heim, J., Skjetne, J.H., Hestnes, B., Heiestad, S., Ulseth, T., Frowein, H., Devoldere, P., Aaby, C., O'Malley, C., Brundell, P., Lonsdale, P. (2001): "Results of Baseline Communication Experiments. Project IST-1999-11577 Eye-2-Eye: Fitness-for-Purpose of Person-Person Communication Technologies".

[i.57] Takahashi, A. (2009): "Framework and Standardization of Quality of Experience (QoE) Design and Management for Audiovisual Communication Services, NTT Technical Review", April 2009 Vol. 7 No. 4, pp. 1-5.

[i.58] Vienott, E., Olson, J., Olson, G. & Fu, X. (1999): "Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other", Proc. CHI 1999, ACM.

[i.59] Watson & M. A. Sasse (2000): "The Good, the Bad, and the Muffled: The Impact of Different Degradations on Internet Speech", proceedings of the 8th ACM International Conference on Multimedia, Oct. 30- Nov. 3, Marina Del Rey, CA; pp. 269-302.

[i.60] Watts, L., Monk, A. & Daly-Jones, O. (1996): "Inter-personal awareness and synchronization: Assessing the value of communication technologies", International Journal of Human-Computer Studies, 44, 849-873.

[i.61] Williams; D., Caplan, S., & Xiong, L. (2007): "Can You Hear Me Now? The Impact of Voice in an Online Gaming Community", Human Communication Research 33 (2007) 427-449.

[i.62] Yankelovich, N., Kaplan, J., Provino, J., Wessler, M., DiMicco, J. (2006): "Improving Audio Conferencing: Are Two Ears Better than One? Proceedings of CSCW '06", November 4-8, 2006, Banff, Alberta, Canada, pp. 333-342.

[i.63] ITU-T Recommendation E.860: "Framework of a service level agreement".

[i.64] ETSI TR 102 353: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM); Guidelines for the Satellite Independent Service Access Point (SI-SAP)".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

audio: all signals that are audible to human beings, including speech, music and background noise

audio/video asynchrony: when audio and video information that leaves one communicating party is received by the other communicating party at different times (e.g. typically the audio information arrives before the video information in an asynchronous situation)

NOTE: It is calculated as audio delay subtracted from video delay (e.g. if audio delay is 50 ms and video delay is 200 ms, then asynchrony is 150 ms; if audio delay is 100 ms and video delay is 50 ms, then asynchrony is -50 ms).

audio delay: mean time required for an audio signal to reach the listener's ear

audio protocol: set of rules defining the way audio information is represented in a network

avatar communication: use of a service that transmits voice or text in real-time over a telecommunication network in combination with a graphical (human) representation of the speaker

CD quality: audio quality with a 44,1 kHz sampling rate without compression

communication service: user-initiated service that is provided via a telecommunication network for people to share information

NOTE: Examples are speech telephony, email, videoconferencing, avatar-telephony, audio conferencing.

communication situation: combination of task, motive, content and user (group) characteristics

communication task: what the end-users (want to) do with a communication service (e.g. social chatting, buying or selling shares, conducting a job interview, etc.)

communicative behaviour: end-user behaviour while using a communication service, including turn taking, interruptions, verbal and non-verbal back-channels and gaze

conversational text: See real-time text.

data communication: use of a service that transmits personal computer-based information (e.g. presentation slides)

data conferencing: See data communication.

duration: length of time of the communication task

dyadic communication: (distance) communication between two people

effectiveness: accuracy and completeness with which specified users can achieve specified goals in particular environments

NOTE: See ISO 9241 [i.31] definition.

efficiency: resources expended in relation to the accuracy and completeness of goals achieved

NOTE: See ISO 9241 [i.31] definition.

end-users: people who use a communication service

face-to-face (videoconferencing): the use of video communication to see the person who is talking

frame-rate: frequency by which a full frame is updated, in the case of video frame-rate sometimes called video temporal resolution or image frequency

group: (distance) communication between three or more people

NOTE: Either in a point-to-point or a multi-point configuration.

high quality videoconferencing: video communication using an analogue simulation of PAL quality, with technical parameter values: delay < 40 ms; frame rate 25 fps; resolution 4CIF (PAL); no packet loss

NOTE: A laboratory and field study set-up used for the user tests described in annex B

instruction task: communication task to between two or more people collaborate in order to transfer information. The communication may be more one-way and unequal with respect to expertise

media effects: effect a particular communication medium has on an end-users task outcome, communicative behaviour, attitudes and beliefs

monitor size: number in inches of the diagonal of the image screen on a screen

multimedia communication: use of a service that transmits voice, video and data signals in real-time over a telecommunication network

multimedia conferencing: service for transmitting voice, video and data signals over a telecommunication network

multi-point: distance communication between three or more locations

negotiation task: communication task between people in order to achieve an agreement

network performance: the ability of a network or network portion to provide the functions related to communications between users

NOTE 1: Network performance applies to the network provider's planning, development, operations and maintenance and is the detailed technical part of QoSO.

NOTE 2: Network performance parameters are meaningful to network providers and are quantifiable at the part of the network which they apply.

NOTE 3: From ITU-T Recommendation E.800 [i.34].

network quality of service: degree of conformance of the service delivered to a user by a provider with an agreement between them

NOTE: From ITU-T Recommendation E.860 [i.63].

packet loss: loss of one or more packet that can be described using a certain statistical model

packet size: magnitude of a unit of data transmitted over a packet switching network as part of a message transferred in number of Bytes

personal involvement: extent to which the communication parties are committed to the outcome of the task or perform the task more on behalf of another party than themselves

person perception: extent to which the perception of the other person's attributes (how likeable, intelligent, friendly, etc.) is positive or negative

persuasion task: communication task in which one person attempts to make one or more other do or believe something that previously they would probably not do or believe.

NOTE: The communication involves giving another person a good reason to do something or making someone believe something.

point-to-point: communication between two locations

problem solving task: communication task where the primary goal is for two or more people to collaborate and share relatively equal but different expertise to find a solution to a problem

quality of experience (QoE): overall acceptability of an application or service, as perceived subjectively by the end-user

NOTE 1: Quality of experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.).

NOTE 2: Overall acceptability may be influenced by user expectations and context.

NOTE 3: ITU-T Recommendation P.10 [i.40]/G.100 Ammendment 2 definition.

quality of service: totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service

NOTE: ITU-T Recommendation E.800 [i.34] definition.

quality of service experienced/perceived by customer/user (QoSE): statement expressing the level of quality that customers/users believe they have experienced

NOTE 1: The level of QoS experienced and/or perceived by the customer/user may be expressed by an opinion rating.

NOTE 2: QoSE has two main man components: quantitative and qualitative. The quantitative component can be influenced by the complete end-to-end system effects (network infrastructure).

NOTE 3: The qualitative component can be influenced by user expectations, ambient conditions, psychological factors, application context, etc.

NOTE 4: QoSE may also be considered as QoSD (QoS delivered/achieved by service provider) received and interpreted by a user with the pertinent qualitative factors influencing his/her perception of the service.

NOTE 5: ITU-T Recommendation E.800 [i.34] definition.

real-time: describes information and communication technologies that are able to generate and deliver information in a time-frame similar to the real-life process that it is assisting

EXAMPLE 1: Real-time charging and billing information is to be generated, processed, and transported to a desired conclusion in less than 1 second [i.30].

EXAMPLE 2: Refers to the generation of network management information in a time frame comparative to the real life process that it is controlling or monitoring [i.17].

real-time communication service: service with which users expect to share information instantly and continuously with one or more other user

NOTE 1: A real-time communication service generates and delivers either text, audio (such as speech), graphics, video and data or some combination of these communication media.

NOTE 2: The information sharing process occurs either by: (1) a person interacting via technology directly to another person (person-to-person) or; (2) a person interacting with a machine (person-to-machine).

EXAMPLE: An example real-time person-to-person communication service is videoconferencing and an example real-time person-to-machine communication service is Live TV.

real-time text: service for transmitting alpha-numeric characters in real-time over a telecommunication network

real-time transport protocol: standardized packet format for delivering audio and video over the Internet

NOTE 1: It was developed by the Audio-Video Transport Working Group of the IETF and first published in 1996 as RFC 1889, and superseded by RFC 3550 in 2003.

NOTE 2: RTP is frequently used in streaming media systems (together with the RTSP) as well as in videoconferencing and push to talk systems. For these it carries media streams controlled by H.323 or Session Initiation Protocol (SIP) signaling protocols, making it the technical foundation of the Voice over IP industry.

remote inspection videoconferencing: use of video communication to observe an object or environment while also engaging in person-to-person communication using speech communication

NOTE 1: It enables a person to see what is talked about rather than who is talking. The minimal configuration for remote inspection is one-way video with two-way speech.

NOTE 2: The service is sometimes also called Tele-inspection, Tele-data, See what I See (SWIS).

resolution: term denoting the degree of detail which can be created by a particular visual display system expressed in pixels in x- and y-directions

satisfaction: comfort and acceptability of the work system to its users and other people affected by its use

NOTE: ISO 9241 [i.31] definition.

situation formality: relative amount of ceremonious or conventional communication versus casual or unconstrained communication

task outcome: extent to which task performance dependent on the medium

task: what users of *communicative technology* actually do in order to accomplish some *task goal*

NOTE: In experiments tasks may be described to the participants or they are embedded in scenarios as a part of a *situation*.

telephony: service for transmitting voice signals in real-time over a telecommunication network

text communication: use of a service that transmits alpha-numeric characters in real-time over a telecommunication network

NOTE: Also known as real-time text and conversational text.

up-time: the time that a communication service provides a connection to a user to enable the intended use of the service (rather than having an unplanned break in connection)

urgency: extent to which a task is particularly urgent or under particular time pressure

usability: *effectiveness*, *efficiency*, and *satisfaction* with which specified users achieve specified goals in particular environments

NOTE: See ISO 9241 [i.31] definition.

video communication: use of a service that transmits voice and video signals in real-time over a telecommunication network, i.e. use of videotelephony or videoconferencing

NOTE 1: Sometimes called audiovisual communication when including transfer of information between user and machine/host in addition to from user to user

NOTE 2: For the present document the communication involves a loud-speaking audio system and not a handset.

videoconferencing: service for transmitting voice and video signals in real-time over a telecommunication network for group communication

NOTE: In the present document the audio system is considered loud-speaking and not with a handset or headset.

video delay: time between the input of the first pixel of a particular picture at the sending end encoder and the output of the pixel from the decoder at the receiving end

video protocol: set of rules defining the way video information is represented in a network

videotelephony: service for transmitting voice and video signals in real-time over a telecommunication network for dyadic communication

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

3G	3rd Generation
AAC	Advanced Audio Coding
ADPCM	Adaptive Delta Pulse Code Modulation
BC	Book Chapter
CIF	Common Intermediate Format

NOTE: A video format defined by ITU-T.

CP	Conference proceedings
DVI	Digital Video Interface
EFR	Enhanced Full Rate
EP	Expert panel
ETSI ES	ETSI Standard
ETSI TR	ETSI Technical Report
GSM	Global System for Mobile (telephony)
H	picture Height
HD	High Definition
ICT	Information and Communmnication Technology
IP	Internet Protocol
IPR	Intellectual Property Rights
IPTD	Internet Protocol Packet Transfer Delay
IST	Information Society Technologies

ISDN	Integrated Services Digital Network
ITU-R	International Telecommunication Union Radiocommunication Sector
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
ITU-T (info)	ITU-T Recommendation where the source of a guideline is contained in an appendix or a supplement rather than the body or the annexe(s) of the Recommendation
JA	Journal Article
MAUT	Multi-Attribute Utility Technique
PAL	Phase Alternating Line - a TV standard
PLC	Packet Loss Concealment
PPD	Pixels per Degree
Q	Quantization
QCIF	Quarter CIF
QoE	Quality of Experience
QoS	Quality of Service
QoSD	QoS delivered/achieved by service provider
QoSO	QoS offered/planned by service provider
QVGA	Quarter Video Graphics Array
RA	Research Article or technical review
RR	Research Report
RTP	Real-time Transport Protocol
SQCIF	Sub Quarter CIF
STF	Specialist Task Force
SVGA	Super Video Graphics Adapter/Array
SWIS	See What I See
UDP	User Datagram Protocol
VoIP	Voice over Internet Protocol
WP	Workshop proceedings
WVGA	Wide Video Graphics Array
XVGA	eXtended Video Graphics Array
VTC	Video TeleConferencing
SLA	Service Level Agreement

4 Overview of the real-time communication services addressed in the present document

A real-time telecommunication system generates and delivers information in a time-frame similar to the real-life process that it is assisting. Depending on the real-life process being assisted, the time-frame may include seconds or be perceived by end-users as immediate. These systems therefore have real-time technical requirements and constraints for their implementation associated with real-time computing and the real-time transport protocol (or RTP).

One group of real-time telecommunication systems concerns real-time communication services that enable communication between people. In the absence of a previous accepted definition, the present document defines a real-time communication service as:

- A service with which users expect to share information instantly and continuously with one or more other user. These real-time communication services generate and deliver either:
  - text;
  - audio, such as speech;
  - graphics, such as a computer animation to create a moving image (e.g. an avatar);
  - video;
  - data, mainly concerning the transmission of shared visual information such as a 'White board' or presentation document (e.g. Microsoft PowerPoint); or
  - some combination of these communication media.

Some real-time communication services enable primarily two-way communication, whilst others enable primarily one-way communication (Figure 1). Some services support both one-way and two-way communication. Two-way communication services enable people in different places to share information by speaking, writing, moving their bodies or using other signals in a way intended to be similar to communication when in the same place at the same time. An example of a two-way communication service is videoconferencing. With one-way communication a person interacts with a machine as an end-point rather than directly to another person. Therefore, a person generating the information does not necessarily interact directly with the end-user of the service and the end-user of the service may not necessarily interact with the information being delivered in a way that influences or directs the actions of the person generating the information. An example of a one-way communication service is Live TV. A real-time communication service that promotes both one-way and two-way communication is Real-time Gaming, as end-users may interact with both machine-generated content and other people through text, voice or person-generated avatars.

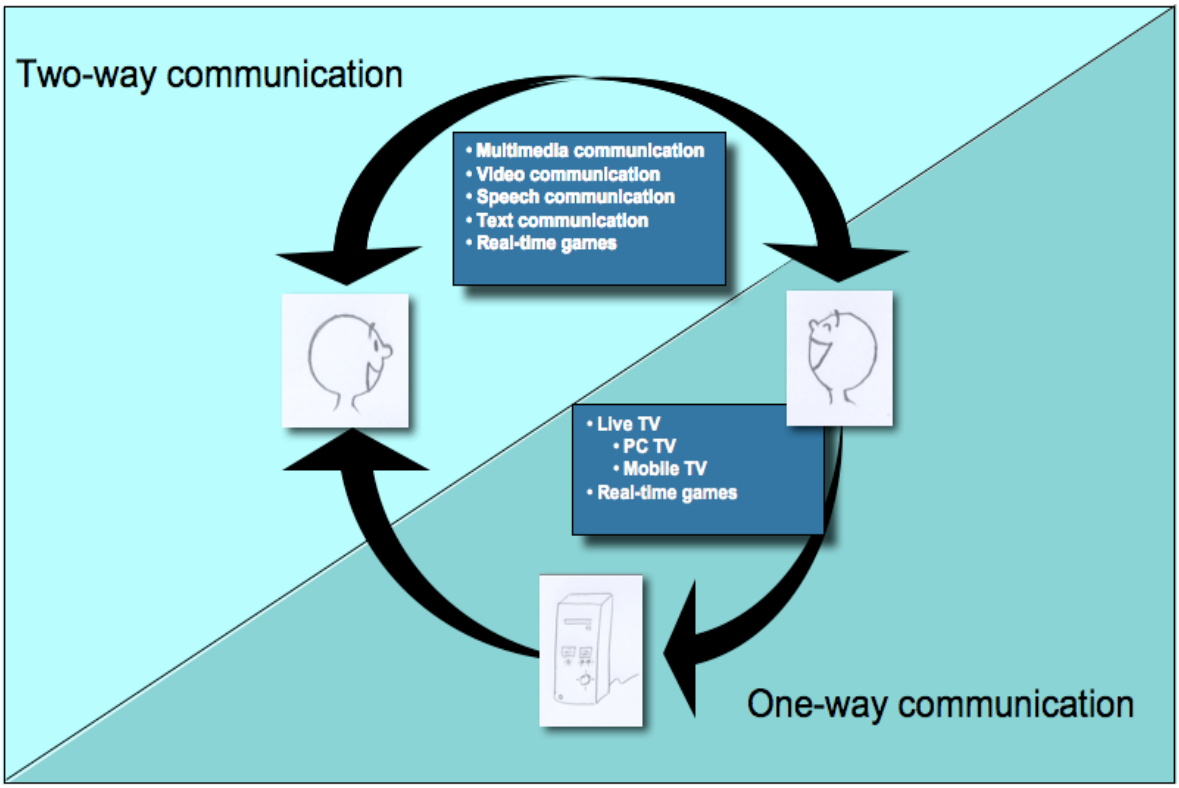


Figure 1: Real-time communication services enable two-way and one-way communication

Whether the real-time communication service is videoconferencing, Live TV, Real-time Gaming or some other service, the generation and delivery of the information is in a time-frame of 'immediate'. People expect to talk during a videoconferencing as if communicating face-to-face and people view a live TV broadcast as if they are watching 'as it happens'. Non-real-time communication services (such as e-mail, voice-mail, video-mail, web content and media on demand) usually do not place the same requirements on the users or technology for the generation and delivery of the communication media. For example, the content of a web site or an e-mail may be created significantly in advance of the end-users' perception and interaction with the material.

User expectations for real-time communication services can be illustrated with examples from related work:

- TS 123 107 [i.22] defines four QoS classes for 3G ('conversational', 'streaming', 'interactive' and 'background') for which the main distinguishing factor concerns the delay sensitivity of the traffic. Conversational and Streaming classes are mainly intended to be used to carry real-time traffic flows.

- ITU-T Recommendation G.1010 [i.38] defines four user-centric QoS categories on the basis of tolerance to packet loss and one-way delay: 'interactive' (delay <<1 s), 'responsive' (delay ~2 s), 'timely' (delay ~10 s) and 'non-critical' (delay >>10 s).

- ITU-T Recommendation Y.1541 [i.41] defines two classes of network Quality of Service (QoS) for real-time, jitter sensitive, high interaction applications (VoIP and 'video teleconferencing'): 'Class 0' applies for constrained routing and distance while 'class 1' applies for less constrained routings and distances.

- ES 202 667 [i.12] classifies audiovisual applications for IP networks into the two groups of 'delay sensitive' and 'delay insensitive' applications.

In addition to one-way delay of a medium, the delay of arrival between media, such as audio and video, is another parameter related to user expectations for instant and continuous information. For example, the potential different arrival time of audio and video in videoconferencing and live TV concerns the lip-asynchrony of the speaker.

Some example end-to-end delay and audio-video synchrony recommendations for real-time communication services are provided in table 1. Table 1 illustrates current performance targets for services that aim to meet user expectations for instant and continuous information. As table 1 draws on a range of sources employing different test methodologies (such as acceptability threshold, perception threshold and task performance) different targets may be proposed in some cases.

Table 1: Examples of end-to-end delay and audio-video synchrony recommendations for real-time services

Medium Service or application		End-to-End one-way delay (seconds) Note. For some services it is more appropriate to consider these values as start-up delay		Audio-Video asynchrony or Lip-asynchrony (seconds)	Audio or Video arriving first	Source
		Preferred	Limit			
Text	Real-time text	< 1 s	2 s	Not applicable	Not applicable	ITU-T Rec. F.700 [i.35] (see note)
Audio	Speech communication	< 0,15 s	0,4 s	Not applicable	Not applicable	TS 122 105 [i.21] ITU-T Rec. G.114 [i.37]
	High Quality streaming	< 10 s		Not applicable	Not applicable	ITU-T Rec. G.1010 [i.38] (see note)
	Speech, mixed speech and music, medium and high quality music	< 10 s (start-up delay)		Not applicable	Not applicable	TS 122 105 [i.21]
Video	Video communication	Not specified	Not specified	0,04 s	Not specified	ETR 297 [i.20]
		Not specified	Not specified	< 0,08 s	Video or audio	ANSI TI.552 [i.65]
		0,1 s	0,4 s	0,1 s	Not specified	ITU-T Series H, Supplement 1 [i.33] (see note)
		< 0,15 s	0,4 s	< 0,1 s	Not specified	TS 122 105 [i.21]
		< 0,15 s	0,4 s	< 0,08 s	Video or audio	ITU-T Rec. G.1010(see note) [i.38]
		Not specified	Not specified	<0,2 s	Audio	EG 202 534 [i.11]
	Television	Not specified	Not specified	<0,185 s	Video	ITU-R. Rec. BT.1359-1 [i.32]
		Not specified	Not specified	<0,09 s	Audio	ITU-R. Rec. BT.1359-1 [i.32]
	« One-way »	<10 s	Not specified	Not specified	Not specified	ITU-T. Rec. G.1010 [i.38] (see note)
	Movie clips, surveillance, realtime video	< 10 s (start-up delay)		Not specified	Not specified	TS 122 105 [i.21]
Data	Real-time games	< 75 ms	Not specified	Not applicable	Not applicable	TS 122 105 [i.21]
	Interactive games	< 200 ms		Not applicable	Not applicable	ITU-T Rec. G.1010(see note) [i.32](see note)
	Telnet	< 200 ms	Not specified	Not applicable	Not applicable	ITU-T Rec. G.1010 [i.38] (see note)
		< 250 ms	Not specified	Not applicable	Not applicable	TS 122 105 [i.21]
	Telemetry - two-way control	< 250 ms	Not specified	Not applicable	Not applicable	TS 122 105 [i.21]
	Bulk data transfer/retrieval, still image	< 15 s	< 60 s	Not applicable	Not applicable	ITU-T Rec. G.1010 [i.38] (see note)
	Bulk data transfer/retrieval, playout and synchronization information, still image	< 10 s (start-up delay)		Not applicable	Not applicable	TS 122 105 [i.21]
NOTE:	The source is contained in an appendix or a supplement (informative) rather than the body or annex(es) of the Recommendation document.					

The real-time communication services addressed in the present document concern services that are initiated by users in order to deliberately share information. The present document excludes systems that act in automatic or artificially intelligent ways on behalf of people but without deliberate initiation by people. For example, it does not include agent, sensor and ubiquitous systems that undertake and execute actions without reference to people in either the choice of decision or the course of action.

Therefore the key elements of real-time communication services are considered to be:

- user expectations to share information instantly and continuously with one or more other user;
- real-time technical requirements and constraints;
- users deliberately initiates their use.

The real-time communication services addressed in the present document are described briefly below and illustrated in Figure 1 on the basis of whether they enable two-way or one-way communication.

4.1 Text communication

A text communication service enables people to use alpha-numeric characters for writing a message. To be considered a real-time service the alpha-numeric characters should be delivered to the receiver at one or more remote site as they are created by the user (e.g. within 1 second of being typed according to ITU-T Recommendation F.700 [i.35]). Services such as 'text chat' are not considered real-time services even if they may deliver text almost immediately. With 'text chat' the text information may not be transmitted instantly and continuously, because it is the sender who decides when to transmit the text (e.g. by pressing the 'enter' key or clicking on a 'send' button to send a sentence or paragraph).

4.2 Speech communication

A speech communication service enables people to use their voice to communicate through a handset or headset (e.g. telephony between two people) or a loud-speaking function (e.g. audio conferencing between groups of people at two or more locations). This service may also be known as conversational voice, voice telephony or communication by speech.

4.3 Video communication (including speech)

A video communication service combines voice and video information. It is a person-to-person (two-way) communication service which can be considered within more general audiovisual communication including TV (see clause 4.6). There are two main categories of video communication that are distinguished primarily on the basis of the video information that is transmitted. These two types of video communication are 'face-to-face video communication' and 'remote inspection / observation / viewing video communication'.

4.3.1 Face-to-face video communication

A face-to-face video communication service is typically referred to as videotelephony when the communication is between two locations or videoconferencing when the communication is between groups of people at three or more locations. This service is also sometimes known as video teleconferencing (VTC).

4.3.2 Remote inspection or 'see what I see' (SWIS) video communication

A remote inspection video communication service enables observation of an object or environment while also engaging in person-to-person communication using speech communication. It enables a person to see what is talked about rather than who is talking. The minimal configuration for remote inspection is one-way video with two-way audio. The service is sometimes also called Tele-inspection and Tele-data.

4.4 Multimedia communication

A multimedia communication service combines two or more media. A typical situation is the use of voice, video and computer presentation graphics (e.g. Microsoft PowerPoint). Another example is the combination of text and video without audio as in a 'total conversation call' for hearing-impaired people.

A video communication service is actually a special case of multimedia communication because it combines video and audio. However, it is treated as a separate, fundamental category.

4.5 Real-time games

Real-time games use multiple media for interactive, imaginative experience. The media are predominantly graphics, text and audio. Real-time games are a special category of multimedia communication.

4.6 TV including PC TV and Mobile TV

A Live TV service televises an event at the same time that it happens. It is a special category of multimedia communication using predominantly video and audio and that is primarily one-way communication, whereas multimedia communication (clause 4.4) is a person-to-person (two-way) communication service.

5 Overview of user experience guidelines expressed in Quality of Service terms

When implementing real-time communication services there are many network, codec and environment characteristics that may interfere with human communication and affect the user experience [i.30]. For example, in a packet switched network the main interfering characteristics are bandwidth, packet size, delay, delay jitter, packet loss, burst packet loss and sequencing. The main codec characteristics are the media protocols (e.g. G.7xx, AAC, H.26x, MPEGx), video space resolution, video time resolution, delay, distortion and monitor size. The environmental characteristics may include lighting conditions, background patterns, colour and reflex, acoustics, audio echo degradation, viewing distance, camera position and camera parameters.

These technical characteristics of a communication service that bear on its ability to satisfy stated and implied needs of the user of the service are a topic of Quality of Service (QoS) [i.34].

During the 1980s and 1990s most user-centred work focused on the usability of information and communication technology (ICT). Since 1998 usability has been defined in ISO 9241 [i.31], p. 2 as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

In recent years the use of ICT has extended from the workplace to the home and for applications that support leisure and social activities in addition to work. Consequently, the concerns of human-computer interaction have evolved from a focus on effectiveness and efficiency to user experience factors such as enjoyment, engagement and the appeal of using and owning ICT e.g [i.4].

The concept of user experience is related to established concepts such as QoS perceived by the user e.g [i.13], [i.14], [i.15] and [i.16] and the relatively new concept of Quality of Experience (QoE) e.g. [i.30]; [i.51]; [i.54]; [i.11] and [i.39]. The concept of QoE is relatively new and evolving and there are several published definitions; in particular:

Within ETSI TC HF deliverables QoE has been defined (since 2006) as [i.11] and [i.18]:

The performance of users when using what is presented by a communication service or application user interface.

NOTE: It takes into account the individual Quality of Services and measures the acceptability of a service or application by including factors such as usability, utility, fidelity and level of support from the application or service provider (e.g. sales, delivery, error corrections).

EXAMPLE: A service provider may conclude that a particular communication service with a certain level of Quality of Service used for a particular communication situation offers users good or very good Quality of Experience as measured by user satisfaction, task efficiency and task effectiveness.

Within ITU-T QoE has been defined (since 2008) as [i.40]:

The overall acceptability of an application or service, as perceived subjectively by the end-user.

NOTE 1: Quality of experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.).

NOTE 2: Overall acceptability may be influenced by user expectations and context.

An overview of the QoE assessment of audiovisual communication services within ITU-T and ITU-R is provided by [i.57] with a focus on quality models that estimate the user-perceived acceptability of a system from the physical characteristics of that system.

The ITU-T also offers a related definition of QoS experienced/perceived by customer/user (QoSE) [i.34]:

A statement expressing the level of quality that customers/users believe they have experienced.

NOTE 1: The level of QoS experienced and/or perceived by the customer/user may be expressed by an opinion rating.

NOTE 2: QoSE has two main man components: quantitative and qualitative. The quantitative component can be influenced by the complete end-to-end system effects (network infrastructure).

NOTE 3: The qualitative component can be influenced by user expectations, ambient conditions, psychological factors, application context, etc.

NOTE 4: QoSE may also be considered as QoSD (QoS delivered/achieved by service provider) received and interpreted by a user with the pertinent qualitative factors influencing his/her perception of the service.

Deriving a harmonised definition of QoE requires further work and is beyond the scope of the present document. For current purposes the key aspects are that QoE should be:

- Primarily based on data from user tests of actual usage, which imply:
  - Not just results from quality models but user-based measures such as communication task effectiveness, task efficiency, user satisfaction and user enjoyment;
  - Not just results concerning perception of media quality, but also data that is often linked to a particular user task or communication purpose;
- Expressed in relation to technical parameters such as QoS.

Therefore, user-centred guidelines for real-time communication services should succeed where possible to combine knowledge of both user experience and technical parameter values; for example, to provide a statement about user experience with a particular communication service with known levels of QoS. In recognition of this, the present document has been developed by using a systematic approach for extracting and combining user experience and technical parameters from user test results. A database of structured user test results was derived from which the guidelines were abstracted. The user test results are structured according to the clause:

IF <communication situation>;

USING <service prescription>;

WITH <technical parameters>;

THEN <user experience>.

The attributes of this format can be further broken down as follows.

<Communication situation>

The <communication situation> has three main parameters:

<user task>

e.g. give instruction, negotiate with a sales person;

<user group>

e.g. business people, people with a speech impairment;

<user environment>.

e.g. e.g. conference room, in parked car.

Example communication situations are:

- A salesperson reading email in a car.
- A commuter watching TV while waiting at a bus stop.
- A personnel manager interviewing an overseas job candidate from a videoconferencing room.

<Service prescription>

The <Service prescription> is the service or services used in the <communication situation>.

<Service prescription> has two main parameters:

<service type>

e.g. video call, audio call, Live TV;

<terminal type>

e.g. lap-top computer, mobile handset.

<Technical parameters>

The <Technical parameters> attribute includes the main QoS parameters of the communication service concerned. There can be many parameters and some examples are:

<bit-rate>

e.g. 1 Mbps, 64 kbps;

<media protocol>

e.g. H.264, MPEG2, AAC;

<network protocol>

e.g. TCP-IP, UDP-IP, RTSP;

<delay>

e.g. 50 ms, 500 ms, 1 s;

<audio-video asynchrony>

e.g. 0, 50 ms, -100 ms;

<jitter>

e.g. 50 ms, 100 ms, 1 s;

<packet loss>

e.g. 0,5 %, 1 %, 5 %;

<video frame rate>

e.g. 7 fps, 25 fps, 30 fps;

<video resolution>

e.g. CIF, 1920x1080, XGA.

<User experience>

The <user experience> is the measure of end-user behaviour with the communication service. The user experience may be linked with the communication situation, such as the user task, the type of service and the technical parameters. There are many potential measures of user experience that will depend on the particular user test, but four main parameters may be considered, into which most measures can be grouped:

<task effectiveness>

e.g. task accuracy, value of negotiated agreement;

<task efficiency>

e.g. task time, number of speech interruptions;

<user satisfaction>

e.g. acceptability of the service, level of social presence;

<user enjoyment>

e.g. level of engagement in use, level of fun.

With this essential information collected and structured, guidelines can be abstracted that aim to state the principal messages of relevance to the intended guideline users [i.7]. It is these abstracted guidelines (Figure 2) that form the main content of the present document.

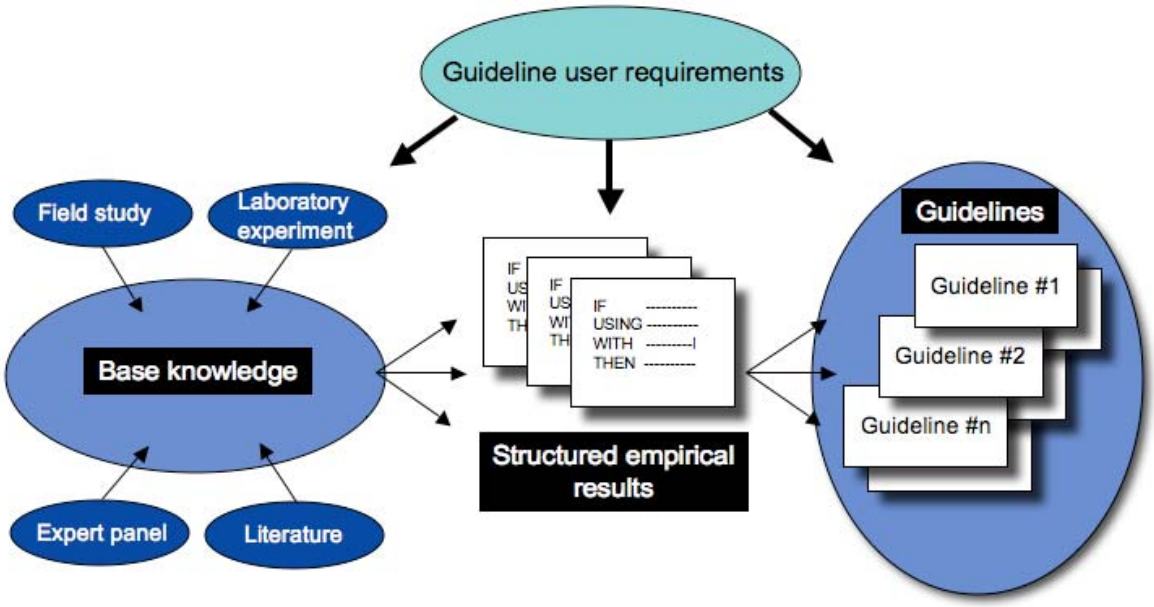


Figure 2: Guidelines are derived from base empirical knowledge through a process of structuring user test results and mapping user and technical variables

6 Advice on using the Guide

The present document provides informative elements for the HF community with the aim to bring together user test data as guidelines about user experience for known levels of QoS.

The present document does not replace any existing standards.

Those readers interested in the details of a particular guideline empirical source are also directed to a web-based system (http://portal.etsi.org/STFs/STF_HomePages/STF354/STF354.asp) that provides this information or to the original literature from which the guidelines were abstracted.

This remainder of this clause provides an overview of requirements for the guidelines, a description of the structure of the guidelines and consideration of the current level of maturity of the guidelines given the scope and restrictions of empirical work to date.

6.1 Requirements for guidelines

The guidelines concern particular usage situations for real-time communication services and provide advice for:

- optimising the QoS of a particular service;
- selecting between services.

The guidelines are intended to assist by bringing together user experience data, expressing these data in relation to key QoS parameters and enabling a guideline user to make a judgement about the potential value and application of the guidelines based on their own particular expertise and experience.

Clause A.2 summarises the requirements of intended guideline users identified during interviews and workshops with stakeholders performed by STF 284 and STF 354.

There are many people in network operator, service provider and equipment manufacturer organisations who have a professional interest to understand and apply knowledge about user experience. These people have diverse roles and include persons from more technical to more financial work areas, and from more design-oriented to management functions. More technically oriented roles include product strategists, strategic service developers, system integrators, Service Level Agreement (SLA) negotiators, technical researchers, interaction designers, development engineers, audio/video codec engineer. More customer oriented roles include marketing, service portfolio specialists, sales persons, sales support, customer support, human factors researchers and service hosts. Also, many senior management roles in marketing, research, operations, development, etc., could expect to benefit from knowledge of user experience for strategic decision making. Where possible all of these types of people have been included as stakeholders for user experience guidelines.

6.2 The structure of the guidelines

Each guideline is:

- stated in a single sentence intended to be concise;
- accompanied by:
  - one or more summary of the empirical source on which that guideline is based;
  - a reference to each empirical source.

Each summary of an empirical source is stated as a principle and where possible the following consistent construction is employed:

> <Communication service> *with* <QoS> *for* <Communication situation> *gives* <Usage outcome> *(compared with* <communication service or QoS>*).*

Therefore a summary of an empirical source provides an overview of the:

1) Communication service
2) Main QoS or other technical parameters
3) Communication situation
4) Usage outcome
5) Communication service or QoS with which a comparison is made (if applicable)

6.3 Maturity and restrictions of the guidelines

The guideline set is considered initial because there is a very large range of potential communication situations, service prescriptions and QoS parameters for which user experience data could be collected. To date a relatively small number of empirical studies has been performed and new usage situations and (potential) services for testing continually emerge. Further study is required to increase the range of the guidelines.

Whereas the guidelines drawn from related standardisation work may be mature and based on many user tests, others may be immature if they are drawn from an isolated small-scale study. As stated in clause 6.1, a guideline user may make a judgement about the potential value and application of a guidelines based on their own particular expertise and experience. This may include, for example, weighing the result achieved through a standardisation process with a result reported in a scientific publication.

The validity of some guidelines remains open for further study, because:

- Some of the user tests on which the guidelines are based should be replicated and extended to different user groups and task types. There may be cross-cultural issues concerned with real-time communication services, and the available user test data is mainly restricted to samples within particular countries (e.g. the UK and Norway).
- Some of the user tests should be repeated with different technical parameters to be more relevant to current service developments. For example, whereas there is data that with CD stereo audio quality users may rate social presence higher than 3,1 kHz mono or stereo, there is no known similar experiment done with wideband speech/audio.
- Most of the empirical tests involve condition comparisons that reveal where significant differences between conditions or groups exist but they were not designed to identify precise thresholds for user experience.
- Most of the laboratory results for two-way communication concern dyadic communication (i.e. between two people) that is point-to-point (i.e. between two locations). Some field data exists for group communication (i.e. between three or more people) that is point-to-point. There is currently little data available for multi-point communication (i.e. between three or more locations).
- Different user tests may have there different strengths and weaknesses, due to particular research designs and choice of more objective or more subjective user measures.
- Some guidelines are based on conclusions by expert panels rather than user tests, although this is relatively few (less than 10 %).

7 The guidelines

The guidelines are grouped according to whether they concern:

- A specific service
- Selection between services

Within each group guidelines are clustered under topics, consisting of:

- A guideline number
- The guideline statement
- One or more empirical source summary (with its reference)

The empirical source summary references correspond to the reference list in clause 2.2 and the type of reference is indicated using the following classification and abbreviations:

- BC Book chapter
- CP Conference proceedings

- EP Expert panel
  Further information about the expert panels is provided in annex B
- ETSI ETSI document
  ES ETSI Standard
  TR Technical Report
- ITU-T ITU-T Recommendation
  In addition, the text 'info' is added if the source is contained in an appendix or a supplement rather than the body or the annexe(s) of the Recommendation
- JA Journal article
- RA Research article or Technical Review
- RR Research report
  Further information about research conducted deliberately to input to the guidelines is provided in annex B
- WP Workshop proceedings

7.1 Text communication

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Delay			
	7.1.1	***1-2 second delay when used for text communication is usable but not good***	
		• Text communication conversation with 1-2 second delay when used for conversation by general users is usable but not good	[i.35] (ITU-T info)
	7.1.2	***Less than 1 second delay when used for text communication offers good quality***	
		• Text communication conversation with less than 1 second delay when used for conversation by general users offers good quality	[i.35] (ITU info)
Duration			
	7.1.3	***Text communication may take three-times longer than talking***	
		• Text communication with character-by-character transmission for negotiation can take three-times longer to reach consensus for no extra gain (compared with speech communication and when actually face-to-face)	[i.56] (RR)
		• Text communication with character-by-character transmission for negotiation can take three-times longer to reach consensus for no extra gain (compared with Video communication and when actually face-to-face)	[i.56] (RR)

Topic	Guideline number	***Guideline*** • Empirical source	Ref.
Person perception			
	7.1.4	***Text communication can lead to users being perceived as less formal, compared with speech communication***	
		• Text communication with character-by-character transmission when used for persuasion when persons are arguing when used for a third party can lead to users being perceived as less formal (compared with speech communication)	[i.56]] (RR)

Negotiation task		
	7.1.5 ***Sellers may obtain a less-favourable outcome with text communication, compared with video communication***	
	• Text communication with character-by-character transmission when used for sales negotiation can lead to sellers obtaining a less-favourable outcome (than with Video communication)	[i.56] (RR)
	7.1.6 ***Sellers may offer the truth less to buyers with text communication, compared with when face-to-face***	
	• Text communication with character-by-character transmission when used for sales negotiation can lead to sellers offering the truth less to buyers (than when actually face-to-face)	[i.56] (RR)
	• Video communication with high quality when used for sales negotiation can lead to sellers offering the truth more to buyers (than when using Text communication)	[i.56] (RR)
	7.1.7 ***Sellers may tell the truth less with text communication, compared with video communication***	
	• Text communication with character-by-character transmission when used for sales negotiation can lead to sellers telling the truth less (compared with Video communication)	[i.56] (RR)

7.2 Speech communication

Topic	Guideline number	***Guideline*** • Empirical source	Ref.
Delay			
	7.2.1	***Speech delay should be less than 150 ms for good interactivity***	
		• Highly interactive tasks (e.g. some speech, video conferencing and interactive data applications) may be affected by delays below 100 ms. ITU-T Recommendation G.114 [i.37] states that although a few applications may be slightly affected by end-to-end (i.e. "mouth-to-ear" in the case of speech) delays of less than 150 ms, if delays can be kept below this figure, most applications, both speech and non-speech, will experience essentially transparent interactivity.	[i.37] (ITU-T)
		• Regardless of service or application, a one-way delay of < 150 ms is recommended by ITU-T to enable users to experience essentially transparent interactivity	[i.37] (ITU-T)
		• With speech communication with full-duplex VoIP and with delay < 200 ms (and no special task dependency on high interactivity) the vast majority of users should be satisfied	[i.9] (BC)

…topic continued on next page

Delay (continued)	7.2.2	***Audio delay should not be more than 400 ms for any speech communication***	
		• While it is recommended that a one-way delay of 400 ms should not be exceeded for general network planning, it is important to appreciate that highly interactive tasks (e.g. many voice calls, interactive data applications, video conferencing) can be affected by much lower delays.	[i.37] (ITU-T)
		• In speech communication one-way delay of 400 ms is recommended as the limit by ITU-T	50 (ITU-T info)
		• With speech communication with full-duplex VoIP with delay up to 300 ms (and no special task dependency on high interactivity) many users should be satisfied	[i.9] (BC)
		• With speech communication and VoIP with delay up to 400 ms communication is possible, but not interactive	[i.9] (BC)
		• Regardless of service or application, a one-way delay of < 400 ms is recommended by ITU-T to prevent an excessive number of unacceptable user experiences	[i.37] (ITU-T)
Stereo			
	7.2.3	***With CD stereo audio quality users may rate social presence higher than 3,1 kHz mono or stereo***	
		• For speech communication with CD stereo audio quality, users may rate social presence higher (compared with 8k mono audio or 8k stereo audio (3,1 kHz))	[i.62] (CP)
	7.2.4	***3,1 kHz stereo audio may not provide higher social presence than 3,1 kHz mono***	
		• For speech communication with 8k (sampling) stereo audio (3,1 kHz), users may not rate social presence higher (compared with 8k mono audio quality (3,1 kHz))	[i.62] (CP)

…topic continued on next page

Stereo (continued)	7.2.5	***With 8k mono or stereo audio (3,1 kHz) users may not make more listening errors compared with CD stereo audio***	
		• For speech communication with CD stereo audio, users may rate listening quality higher (compared with 8k mono audio or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k mono audio or 8k stereo audio quality (3,1 kHz), users may require more effort for listening (compared with CD quality stereo audio)	[i.62] (CP)
	7.2.6	***With 8k mono or stereo audio (3,1 kHz) users may not make more listening errors compared with CD stereo audio***	
		• For speech communication with 8k mono audio or 8k stereo audio (3,1 kHz), users may not make more listening errors (compared with CD stereo audio)	[i.62] (CP)
	7.2.7	***With 3,1 kHz mono audio, users may not make more listening errors compared with 3,1 kHz stereo***	
		• For speech communication with 8k mono audio (3,1 kHz), users may not make more listening errors (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
	7.2.8	***3,1 kHz mono audio may be perceived as good as 3,1 kHz stereo***	
		• For speech communication with 8k mono audio (3,1 kHz), users may not rate listening effort higher (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k mono audio (3,1 kHz), users may not rate listening quality lower (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k (sampling) stereo audio (3,1 kHz), users may not rate social presence higher (compared with or 8k mono audio quality (3,1 kHz))	[i.62] (CP)
	7.2.9	***CD stereo audio quality may provide higher social presence than 3,1 kHz mono or stereo***	
		• For speech communication with CD stereo audio quality, users may rate social presence higher (compared with 8k mono audio or 8k stereo audio (3,1 kHz))	[i.62] (CP)

Spatial speaker recognition			
	7.2.10	***Speaker identification between five or more people can be enhanced by a visual representation of speaker location that supplements spatial audio***	
		• Speaker identification during speech communication between five or more people is enhanced by a visual representation of speaker location in addition to spatial audio when some of the voices are unfamiliar (compared with mono audio and no spatial-visual representation)	[i.43] (CP)
		• Speaker identification during speech communication between five or more unfamiliar people is enhanced by a visual representation of speaker location in addition to spatial audio (compared with mono audio and no spatial-visual representation)	[i.43] (CP)
		• Speaker identification during speech communication between five or more slightly familiar people is enhanced by a visual representation of speaker location in addition to spatial audio (compared with mono audio and no spatial-visual representation)	[i.43] (CP)
Packet loss			
	7.2.11	***Audio packet loss should be less than 3 % for good audio quality***	
		• In speech communication a packet loss ratio of <3 % is a provisional figure from ITU-T	[i.43] (ITU-T info)
	7.2.12	***Audio with stable packet loss is preferred compared with high packet loss in short periods***	
		• For speech communication with VoIP with short-term periods of higher packet loss, users seem to weigh the bad periods much more critically (compared with moderate loss variations over the longer-term)	[i.55] (JA)
	7.2.13	***Speech communication with less than 5 % packet loss may have the same perceived quality as no packet loss for listening to conversation***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss for listening to conversation may not improve the perceived quality (compared with 5 % packet loss)	[i.59] (CP)
		NOTE: Any contradiction with other results may be due to differences in user test methodology	

Media Quality			
	7.2.14	***Speech and audio terminals can strongly influence the speech communication quality***	
		• A set of standards provide terminal equipment requirements which enable manufacturers and service providers to enable good quality end-to-end speech performance as perceived by the user in narrowband (3,4 kHz) and wideband (7 kHz)	[i.13], [i.14], [i.15], [i.16] (ETSI ES)
	7.2.15	***Speech communication with each word understandable offers very good quality***	
		• Speech communication with every word spoken being understandable offers very good quality (compared with speech communication with quality that not every word is understandable)	EP
	7.2.16	***Speech communication with CD quality is required for good listening quality, rather than 3,1 kHz (mono or stereo)***	
		• For speech communication with CD stereo audio (44,1 kHz sampling without compression), users may rate listening quality higher (compared with 8k mono audio or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k mono audio or 8k stereo audio quality (3,1 kHz), users may require more effort for listening (compared with CD stereo audio (44,1 kHz sampling without compression))	[i.62] (CP)
		• For speech communication with 8k mono audio or 8k stereo audio (3,1 kHz), users may not make more listening errors (compared with CD stereo audio (44,1 kHz sampling without compression))	[i.62] (CP)
	7.2.17	***Speech communication with 3,1 kHz mono may be rated similar to stereo for listening***	
		• For speech communication with 8k mono audio (3,1 kHz), users may not rate listening effort higher (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k mono audio (3,1 kHz), users may not make more listening errors (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
		• For speech communication with 8k mono audio (3,1 kHz), users may not rate listening quality lower (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)

…topic continued on next page

Media Quality (continued)	7.2.18	***Speech communication with CD quality may be rated higher for social presence than 3,1 kHz mono or stereo***	
		• For speech communication with CD stereo audio (44,1 kHz sampling without compression) quality), users may rate social presence higher (compared with 8k mono audio or 8k stereo audio (3,1 kHz))	[i.62] (CP)
	7.2.19	***Speech communication with 3,1 kHz stereo for social presence may not be rated higher than 3,1 kHz mono***	
		• For speech communication with 8k (sampling) stereo audio (3,1 kHz), users may not rate social presence higher (compared with or 8k mono audio quality (3,1 kHz))	[i.62] (CP)
	7.2.20	***Speech communication with no packet loss with one person at a too high volume may have a higher perceived quality than 5 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss and normal volume may have a higher perceived quality (compared with one voice at a high volume and no packet loss)	[i.59] (CP)
	7.2.21	***Speech communication with no packet loss and one person at too high volume may have a similar perceived quality as 20 % packet loss with normal volume***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may have a similar perceived quality (compared with normal volume and 20 % packet loss)	[i.59] (CP)
	7.2.22	***Speech communication with no packet loss with a poor quality microphone may have a similar perceived quality as 5 % packet loss and a good quality microphone***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss and a good quality microphone may have a higher perceived quality (compared with one person using a poor quality microphone and no packet loss)	[i.59] (CP)

…topic continued on next page

Media Quality (continued)	7.2.23	***Speech communication with no packet loss and a bad quality microphone may be more stressful for listeners than 5 % packet loss if using a good quality microphone***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and a bad quality microphone may be more stressful for listeners (compared with a good quality microphone and 5 % packet loss)	[i.59] (CP)
	7.2.24	***Speech communication with no packet loss and one voice at a low volume may be more stressful for listeners than 5 % packet loss with a normal volume***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a low volume may be more stressful for listeners (compared with normal volume and 5 % packet loss)	[i.59] (CP)
	7.2.25	***Speech communication with no packet loss and one voice at a high volume may be more stressful for listeners than 5 % packet loss with a normal volume***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may be more stressful for listeners (compared with normal volume and 5 % packet loss)	[i.59] (CP)
	7.2.26	***Speech communication with no packet loss and one voice at a low volume may be less stressful for listeners than normal volume and 20 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a low volume may be less stressful for listeners (compared with normal volume and 20 % packet loss)	[i.59] (CP)
	7.2.27	***Speech communication with no packet loss and one voice at a high volume may be more stressful for listeners than low volume and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may be more stressful for listeners (compared with low volume and no packet loss)	[i.59] (CP)

Business communication			
	7.2.28	***Speech communication may lead to higher-status members dominating over low-status members, compared with face-to-face meetings***	
		• Loud speaking speech communication when used for discussions can exaggerate the tendency for high-status members of a business organisation to dominate in meetings (compared to actual face to face communication)	[i.24] (JA)
Listening task			
	7.2.29	***With 3,1 kHz mono audio, users may not make more listening errors than with 3,1 kHz stereo***	
		• For speech communication with 8k mono audio (3,1 kHz), users may not make more listening errors (compared with or 8k stereo audio (3,1 kHz))	[i.62] (CP)
	7.2.30	***With 8k mono or stereo audio (3,1 kHz) users may not make more listening errors than with CD stereo audio***	
		• For speech communication with 8k mono audio or 8k stereo audio (3,1 kHz), users may not make more listening errors (compared with CD stereo audio (44,1 kHz sampling without compression))	[i.62] (CP)
	7.2.31	***Speech communication with 5 % packet loss may have a similar perceived quality as no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss for listening to conversation may not improve the perceived quality (compared with 5 % packet loss)	[i.59] (CP)
	7.2.32	***Speech communication with 5 % packet loss may not reduce the perceived quality compared to when there is no packet loss and one person speaks at a low volume***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss for listening to conversation may not reduce the perceived quality (compared with one person speaking at a low volume and no packet loss)	[i.59] (CP)
	7.2.33	***Speech communication with 5 % packet loss may have a higher perceived quality than when there is echo and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss may have a higher perceived quality (compared with echo and no packet loss)	[i.59] (CP)

…topic continued on next page

Listening task (continued)	7.2.34	***Speech communication with 5 % packet loss may have a higher perceived quality than when there is one voice at a high volume and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss and normal volume may have a higher perceived quality (compared with one voice at a high volume and no packet loss)	[i.59] (CP)
	7.2.35	***Speech communication with 5 % packet loss and a good quality microphone may have a higher perceived quality than when there is one person using a poor quality microphone and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 5 % packet loss and a good quality microphone may have a higher perceived quality (compared with one person using a poor quality microphone and no packet loss)	[i.59] (CP)
	7.2.36	***Speech communication with 20 % packet loss may have a similar perceived quality than when there is echo and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 20 % packet loss and no echo may have a similar perceived quality (compared with echo and no packet loss)	[i.59] (CP)
	7.2.37	***Speech communication with no packet loss and one voice at a high volume may have a similar perceived quality to when there is 20 % packet loss and a normal volume***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may have a similar perceived quality (compared with normal volume and 20 % packet loss)	[i.59] (CP)
	7.2.38	***Speech communication with no packet loss and a bad quality microphone may be more stressful for listeners than when there is a good quality microphone and 5 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and a bad quality microphone may be more stressful for listeners (compared with a good quality microphone and 5 % packet loss)	[i.59] (CP)
	7.2.39	***Speech communication with no packet loss and one voice at a low volume may be more stressful for listeners than when there is normal volume and 5 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a low volume may be more stressful for listeners (compared with normal volume and 5 % packet loss)	[i.59] (CP)

…topic continued on next page

Listening task (continued)	7.2.40	***Speech communication with no packet loss and one voice at a high volume may be more stressful for listeners than when there is normal volume and 5 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may be more stressful for listeners (compared with normal volume and 5 % packet loss)	[i.59] (CP)
	7.2.41	***Speech communication with no packet loss and one voice at a high volume may be more stressful for listeners than when there is low volume and no packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a high volume may be more stressful for listeners (compared with low volume and no packet loss)	[i.59] (CP)
	7.2.42	***Speech communication with 20 % packet loss may be more stressful for listeners than 5 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with 20 % packet loss may be more stressful for listeners (compared with 5 % packet loss)	[i.59] (CP)
	7.2.43	***Speech communication with no packet loss and one voice at a low volume may be less stressful for listeners than when there is normal volume and 20 % packet loss***	
		• Speech communication with DVI (ADPCM) at 8 kHz sampling rate with no packet loss and one voice at a low volume may be less stressful for listeners (compared with normal volume and 20 % packet loss)	[i.59] (CP)
Negotiation task			
	7.2.44	***Speech communication may allow less monitoring of the other person's concentration, compared with video communication***	
		• Speech communication when used for negotiation allows less monitoring of the other person's level of concentration (compared with video communication with sufficient quality)	[i.60] (JA)
		• Speech communication when used for information transfer allows less monitoring of the other person's level of concentration (compared with video communication with sufficient quality)	[i.60] (JA)

Problem solving task			
	7.2.45	***Speech communication may reduce outcomes when speaking a foreign language, compared with video communication***	
		• Speech communication when used for joint problem solving may reduce task outcomes when the users' communication abilities are low (compared with video communication)	[i.58] (CP)
		• Speech communication when used for joint problem solving may not reduce task outcomes when the users' communication abilities are high (compared with video communication)	[i.58] (CP)
	7.2.46	***Dialogues may be shorter with speech communication than with video communication***	
		• Speech communication when used for joint problem solving produces shorter dialogues (compared with video communication)	[i.10] (JA)
	7.2.47	***Users may interrupt less with speech communication than with video communication***	
		• Speech communication when used for joint problem solving produces less interrupted dialogues (compared with video communication)	[i.10] (JA)
Instruction task			
	7.2.48	***Speech communication may allow less monitoring of the other person's concentration, compared with "face-to-face" video communication***	
		• Speech communication when used for information transfer allows less monitoring of the other person's level of concentration (compared with video communication with sufficient video quality)	[i.60] (JA)
Person perception			
	7.2.49	***Speech communication can lead to users being perceived as more formal***	
		• Speech communication with 7 kHz bandwidth when used for persuasion when arguing for a third party can lead to users being perceived as more formal (compared with Video communication)	[i.56] (RR)
		• Speech communication with 7 kHz bandwidth when used for persuasion when arguing for a third party can lead to users being perceived as more formal (compared with Text communication)	[i.56] (RR)
		• Speech communication with 7 kHz bandwidth when used for persuasion when arguing for a third party can lead to users being perceived as more formal (compared with actual face-to-face communication)	[i.56] (RR)

Elderly people	7.2.50	***Background noise should be minimal***	
		• Speech communication when used for general communication by elderly people should have background noise kept to a minimum	[i.19] (ETSI ETR)
	7.2.51	***A receiving amplifier may improve communication***	
		• Speech communication with receiving amplifiers when used for general communication by elderly hearing impaired persons can improve communication (compared with standard speech communication)	[i.19] (ETSI ETR)
Deaf or hearing impairment	7.2.52	***An induction loop may improve communication***	
		• Speech communication with an induction loop when used for general communication by persons with hearing aids can improve communication (compared with standard speech communication)	[i.19] (ETSI ETR)

7.3 Video communication (including speech)

Guidelines on videoconferencing are clustered within three different service areas:

- face-to-face communication;
- remote inspection of an object or environment with real-time audio (sometimes also called "see what I see" (SWIS) communication);
- multi-point and heterogeneous networks.

7.3.1 Video communication: Face-to-face

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Packet loss	7.3.1.1	***Video packet loss should be less than 1 % for good video quality***	
		• With video communication a packet loss ratio of <1 % is a preliminary value from ITU-T	[i.38] (ITU-T info)

Audio-video asynchrony			
	7.3.1.2	***For joint problem solving 200 ms asynchrony offers good quality***	
		• Video communication with 200 ms asynchrony when used for joint problem solving may not significantly affect task outcome compared with video communication with synchrony)	[i.53] (RR)
	7.3.1.3	***For joint problem solving 200 ms asynchrony offers quality similar to speech communication***	
		• Video communication with 200 ms asynchrony when used for joint problem solving can lead to communications similar to speech communication	[i.53] (RR)
	7.3.1.4	***Video communication with lip asynchrony of 200 ms has a higher perceived utility than speech communication***	
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with loud speaking speech communication)	[i.25] (RR)
	7.3.1.5	***Less than 80 ms is recommended for audio-video lip asynchrony***	
		• With video communication audio-video (lip) asynchrony of < 80 ms is a preliminary figure from ITU-T	[i.38] (ITU-T info)
Delay			
	7.3.1.6	***No delay may improve user performance for problem solving tasks***	
		• Video communication with no delay significantly improves joint problem solving performance (compared with video communication with 500 ms delay)	[i.52] (JA)
	7.3.1.7	***No delay may lead to less interruptions between users during problem solving tasks***	
		• Video communication with no delay when used for joint problem solving can lead to less interruptions (compared with video communication with 500 ms delay)	[i.52] (JA)
	7.3.1.8	***500 ms may reduce user performance when used for problem solving***	
		• Video communication with 500 ms delay when used for joint problem solving significantly reduces performance (compared with video communication with no delay)	[i.52] (JA)

…topic continued on next page

Delay (continued)	7.3.1.9	***500 ms may give more interruptions between users during problem solving tasks***	
		• Video communication with 500 ms delay when used for joint problem solving can lead to more interruptions (compared with video communication with no delay)	[i.52] (JA)
	7.3.1.10	***650 ms may give shorter communication for negotiation tasks than when there is no delay***	
		• Video communication with 650 ms delay when used for negotiation can lead to shorter communication (compared with video communication with no delay)	[i.53] (RR)
	7.3.1.11	***650 ms may not reduce user performance when used for negotiation tasks, compared with a shorter delay***	
		• Video communication with 650 ms delay when used for negotiation can have no significant difference in negotiation outcomes (compared to video communication with a delay of 200 ms)	[i.53] (RR)
	7.3.1.12	***650 ms may not give more interruptions between users during negotiation tasks, compared with a shorter delay***	
		• Video communication with 650 ms delay when used for negotiation can have no significant difference in interruptions (compared to video communication with a delay of 200 ms)	[i.53] (RR)
	7.3.1.13	***650 ms may not effect turn taking during negotiation tasks, compared with shorter delay***	
		• Video communication with 650 ms delay when used for negotiation can have no significant difference in turn taking (compared to video communication with a delay of 200 ms)	[i.53] (RR)
	7.3.1.14	***650 ms may improve outcomes when used for a seller, compared to no delay***	
		• Video communication with 650 ms delay when used in a sales negotiation game can improve results for the seller (compared with video communication with no delay and those people acting as buyers)	[i.53] (RR)
	7.3.1.15	***650 ms may reduce outcomes when used for a buyer, compared with no delay***	
		• Video communication with 650 ms delay when used in a sales negotiation game can reduce results for the buyer (compared with video communication with no delay and those people acting as sellers)	[i.53] (RR)

Packet loss			
	7.3.1.16	***Video communication with audio protocol G.722 and video protocol H.263 should offer better than 1 % packet loss on a fixed line to be perceived as good quality***	
		• Video communication with audio protocol G.722 and video protocol H.263 and a random (not burst) packet loss of 1,5 % is usable	EP
		• Video communication with audio protocol G.722 and video protocol H.263 and a random (not burst) packet loss of 1 % offers good quality	EP
		• Video communication with audio protocol G.722 and video protocol H.263 and a random (not burst) packet loss of 0,5 % offers very good quality	EP
	7.3.1.17	***Video communication with 1 % packet loss on a fixed line offers good quality for the video channel***	
		• Video communication with 1 % random packet loss (not burst) offers good quality for the video channel	EP
	7.3.1.18	***Video communication with 3 % packet loss on a fixed line offers good quality for the audio channel***	
		• Video communication with 3 % random (not burst) packet loss offers good quality for the audio channel	EP
	7.3.1.19	***Video communication with 5 % packet loss on a fixed line with the PLC (Packet Loss Concealment) audio protocol offers good quality for the audio channel***	
		• Video communication with 5 % random (not burst) packet loss with G.729 with PLC (Packet Loss Concealment) offers good quality for the audio channel	EP
		• Video communication with 5 % random (not burst) packet loss with G.723.1 with PLC (Packet Loss Concealment) offers good quality for the audio channel	EP
Resolution			
	7.3.1.20	***Video communication with a 29" screen with CIF when used for general communication offers good quality***	
		• Video communication when used between videoconferencing rooms with 29" screens and CIF for mainly business communication offers good quality	EP
	7.3.1.21	***Video communication with a 29" screen with QCIF when used for general communication is usable***	
		• Video communication when used between video conferencing rooms with a 29" screen and QCIF for mainly business communication is usable in some situations	EP

…topic continued on next page

Resolution (continued)	7.3.1.22	***Mobile video communication with SQCIF when used for general communication offers good quality***	
		• Video communication with a 2,5" screen with SQCIF when used for general communication offers good quality	EP
Media Quality			
	7.3.1.23	***If people have high quality video communication on their desktop, they are likely to use it***	
		• Video communication with high quality from a personal office can become integrated with daily work communications (compared with videoconference rooms)	[i.23] (RR)
	7.3.1.24	***High quality desktop video communication may be preferred to a videoconference room***	
		• Video communication with high quality from one's own office when used for managerial work is judged to be an improvement relative to travelling to a dedicated videoconference room	[i.23] (RR)
	7.3.1.25	***High quality video communication is considered suitable if actual face-to-face communication is not possible***	
		• Video communication with high quality is perceived to be a good way to conduct a broad range of communication tasks when actual face-to-face communication is not an option	[i.23] (RR)
	7.3.1.26	***High quality video communication may be suitable for the majority of managerial communication***	
		• Video communication with high quality when used for managerial work tasks between colleagues is perceived to be suitable when used for all communication tasks	[i.23] (RR)
	7.3.1.27	***High quality video communication is preferred when used for process-oriented communications***	
		• Video communication with high quality is preferred for managerial communications that involve a process (compared with loud speaking speech communication, speech communication with a handset and avatar communication)	[i.23] (RR)
	7.3.1.28	***High quality video communication is preferred for non-urgent communication***	
		• Video communication with high quality from a personal office for non-urgent communication is judged to be an improvement relative to using speech communication with a handset	[i.23] (RR)
	7.3.1.29	***High quality video communication is preferred for communication of long duration***	
		• Video communication with high quality when used for managerial work tasks is selected more often if the duration of the call is long (compared with speech communication)	[i.23] (RR)

…topic continued on next page

Media Quality (continued)	7.3.1.30	***More is said with high quality video communication for the same outcome, compared with speech communication***	
		• Video communication with high quality when used for negotiation can increase the amount users say in order to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
	7.3.1.31	***High quality video communication can lead to more interruptions than with speech communication***	
		• Video communication with high quality when used for negotiation can increase users' interruptions to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
	7.3.1.32	***An easy to set-up high quality conference does not stop actual face-to-face meetings***	
		• Video communication with high-quality and easy and low-cost access when used for managerial work tasks does not change the pattern of actual face-to-face meetings	[i.23] (RR)
Screen size			
	7.3.1.33	***Video communication with a 29" screen with CIF when used for general communication offers good quality***	
		• Video communication when used between videoconferencing rooms with 29" screens and CIF for mainly business communication offers good quality	EP
	7.3.1.34	***Video communication with a 29" screen with QCIF when used for general communication is usable***	
		• Video communication when used between video conferencing rooms with a 29" screen and QCIF for mainly business communication is usable in some situations	EP
	7.3.1.35	***Mobile video communication with SQCIF when used for general communication offers good quality***	
		• Video communication with a 2,5" screen with SQCIF when used for general communication offers good quality	EP
	7.3.1.36	***Using a mobile screen for negotiating can work (in a similar way as with a large screen)***	
		• Video communication with 3,5" screen when used for negotiation may not have a significantly negative effect on task outcome (compared with a 29" screen)	[i.53] (RR)
		• Video communication with 3,5" screen when used for negotiation may not have a significantly negative effect on communicative process (compared with a 29" screen)	[i.53] (RR)
		• Video communication with 3,5" screen when used for negotiation may not lead to significantly different dialogue content (compared with a 29" screen)	[i.53] (RR)

Screen size (continued)	7.3.1.37	***QCIF works as well as CIF on a mobile screen***	
		• Video communication with 3,5" screen and QCIF resolution when used for negotiation can result in no significant difference in task outcome (compared with video communication with 3,5" screen and CIF resolution)	[i.53] (RR)
		• Video communication with 3,5" screen and QCIF resolution when used for negotiation can result in no significant difference in communicative process (compared with video communication with 3,5" screen and CIF resolution)	[i.53] (RR)
		• Videoconferencing with a mobile screen and QCIF resolution can result in no significant difference in usage outcome (compared with videoconferencing with a mobile screen and CIF resolution)	EP
Reliability			
	7.3.1.38	***Video communication set-up is likely to be abandoned by people after 3-10 failed attempts***	
		• An intended video communication session may be abandoned by people after 3 failed set-up attempts	EP
		• An intended video communication session can be expected to be abandoned by people during up to 10 failed set-up attempts	EP
	7.3.1.39	***Video communication with an up-time less than 90 % is perceived as unacceptable by users***	
		• Video communication between videoconferencing rooms with an up-time of less than 90 % is perceived by mainly business users as unacceptable	EP
	7.3.1.40	***Video communication with more than 5 connection terminations per hour is perceived by users as unacceptable***	
		• Video communication between videoconferencing rooms with more then 5 connection terminations per hour is perceived by mainly business users as unacceptable	EP
	7.3.1.41	***Video communication set-up attempts that fail for 5-10 minutes usually result in users switching to speech communication***	
		• Video communication has a set-up time-slot of 5-10 minutes before people switch to speech communication, based mainly on service providers' experiences with booked meetings in videoconference rooms	EP

Cost-benefit			
	7.3.1.42	***If cost matters, high quality video communication may be chosen over face-to-face when used for relatively simple communication between managers***	
		• Video communication with high quality is chosen more for group managerial work tasks that involve planning and task distribution when cost of meeting matters (compared with actual face-to-face communication)	[i.23] (RR)
	7.3.1.43	***Video communication can be expected to be used in business situations where the benefit is obviously high***	
		• Video communication when travel is not an option can be expected to be used in business situations when the benefit is obviously high	EP
Urgency			
	7.3.1.44	***High quality video communication may be preferred for non-urgent communication***	
		• Video communication with high quality from a personal office for non-urgent communication is judged to be an improvement over using speech communication with a handset	[i.23] (RR)
		• Video communication with high quality when used for managerial work tasks is selected more often if the duration of the call is long (compared with speech communication)	[i.23] (RR)
	7.3.1.45	***With high quality video communication more may be said for the same outcome, compared with speech communication***	
		• Video communication with high quality when used for negotiation can increase the amount users say in order to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.23] (RR)
	7.3.1.46	***High quality video communication may lead to more interruptions during negotiation than with speech communication***	
		• Video communication with high quality when used for negotiation can increase users' interruptions to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)

Negotiation task			
	7.3.1.47	***With high quality video communication more is said for the same outcome during a negotiation task, compared with speech communication***	
		• Video communication with high quality when used for negotiation can increase the amount users say in order to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
	7.3.1.48	***With high quality video communication users may interrupt more frequently during negotiation tasks***	
		• Video communication with high quality when used for negotiation can increase users' interruptions to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
	7.3.1.49	***Video communication may allow greater monitoring of the other person's concentration***	
		• Video communication with sufficient video quality when used for negotiation may allow greater monitoring of the other person's level of concentration (compared with speech communication)	[i.60] (JA)
		• Video communication with sufficient video quality when used for information transfer can allow greater monitoring of the other person's level of concentration (compared with speech communication)	[i.60] (JA)
	7.3.1.50	***Eye-contact may not lead to higher satisfaction***	
		• Video communication with direct eye-contact when used for negotiation does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)
		• Video communication without direct eye-contact when used for negotiation does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication with direct eye contact)	[i.50] (JA)
Problem solving task			
	7.3.1.51	***200 ms asynchrony offers good quality***	
		• Video communication with 200 ms asynchrony when used for joint problem solving may not significantly affect task outcome compared with video communication with synchrony)	[i.53] (RR)
	7.3.1.52	***200 ms asynchrony offers quality similar to speech communication***	
		• Video communication with 200 ms asynchrony when used for joint problem solving can lead to communications similar to speech communication	[i.53] (RR)

…topic continued on next page

Problem solving task (continued)	7.3.1.53	***Video communication can lead to looking more at the other person's face, compared with face-to-face communication***	
		• Video communication when used for joint problem solving can lead to looking more at the other person's face (compared with actual face-to-face communication)	[i.10] (JA)
	7.3.1.54	***Video communication may improve outcomes when used for people using a foreign language, compared with speech communication***	
		• Video communication when used for joint problem solving may improve task outcomes when the users' communication abilities are low (compared with speech only)	[i.58] (CP)
		• Video communication when used for joint problem solving may not improve task outcomes when the users' communication abilities are high (compared with speech only)	[i.58] (CP)
	7.3.1.55	***Dialogues may be longer with video communication than with speech communication***	
		• Video- communication when used for joint problem solving produces longer dialogues (compared with speech communication)	[i.10] (JA)
	7.3.1.56	***Users may interrupt more with video communication than with speech communication***	
		• Video- communication when used for joint problem solving produces more interrupted dialogues (compared with speech communication)	[i.10] (JA)
	7.3.1.57	***Users may interrupt less with speech communication than with video communication***	
		• Speech communication when used for joint problem solving produces less interrupted dialogues (compared with video communication)	[i.10] (JA)
Instruction task			
	7.3.1.58	***Video communication can allow greater monitoring of the other person's concentration when giving information***	
		• Video communication with sufficient video quality when used for information transfer can allow greater monitoring of the other person's level of concentration (compared with speech communication)	[i.60] (JA)

Decision making task			
	7.3.1.59	***Eye-contact may not create higher satisfaction***	
		• Video communication with direct eye-contact when used for decision making does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)
		• Video communication without direct eye-contact when used for decision making does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication with direct eye contact)	[i.50] (JA)
Group video communi-cation			
	7.3.1.60	***If cost matters, high quality video communication may be chosen over actual face-to-face when used for straight-forward managerial communication***	
		• Video communication with high quality is chosen more for group managerial work tasks that involve planning and task distribution when cost of meeting matters (compared with actual face-to-face communication)	[i.23] (RR)
	7.3.1.61	***High quality video communication may be preferred for enabling group processes and active involvement from the participants***	
		• Video communication with high quality is chosen more for group managerial work tasks that involve group processes and active involvement from the participants (compared with speech communication)	[i.23] (RR)
		• Video communication with high quality is chosen more for group managerial work tasks that involve planning and task distribution (compared with speech communication)	[i.23] (RR)
	7.3.1.62	***Video communication rooms or equipment shared by occasional users is perceived as usable if human assistance arrives within 5 minutes when requested***	
		• Video communication for planned meetings using rooms or equipment shared by occasional users is perceived as usable if human assistance arrives within 5 minutes after being requested	EP
Human support			
	7.3.1.63	***Video communication rooms or equipment shared by occasional users is perceived as usable if human assistance arrives within 5 minutes when requested***	
		• Video communication for planned meetings using rooms or equipment shared by occasional users is perceived as usable if human assistance arrives within 5 minutes after being requested	EP

View of person			
	7.3.1.64	***With a head-only view between strangers interaction may become less formal***	
		• Video communication with a head-only view between strangers can result in less formal interaction (compared to video communication with a head-and-torso view)	[i.27] (JA)
	7.3.1.65	***With a head-and-torso view between strangers interactions may be more formal***	
		• Video communication with a head-and-torso view between strangers can result in more formal interaction (compared to video communication with a head-only view)	[i.27] (JA)
	7.3.1.66	***With a head-only view between strangers, communication can become more like that between familiar people***	
		• Video communication with a head-only view when used for problem solving between strangers can result in behaviour that is more like interaction between familiar people (compared to video communication with a head-and-torso view)	[i.27] (JA)
Eye contact			
	7.3.1.67	***Parallax differences of less than 8° are acceptable***	
		• Video communication with non-eye contact systems should provide a parallax differences of less than 8° to be within the acceptable range according to the European Standard ETR 297 [i.20]	[i.20] ETSI ETR
		• Video communication in general conditions does not require the optical axes of camera and monitor coinciding	EP
		• The major video communication equipment manufacturers and service providers do not provide eye-contact systems on the market as users are not asking for them	EP
		• Video communication in a conference room with non-exact eye contact is reported as satisfactory for general use by users and service providers	EP
	7.3.1.68	***Eye contact during negotiation may not lead to higher satisfaction***	
		• Video communication with direct eye-contact when used for negotiation does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)
	7.3.1.69	***Eye contact during decision making may not lead to higher satisfaction***	
		• Video communication with direct eye-contact when used for decision making does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)

Person perception			
	7.3.1.70	***High quality video communication can lead to users being perceived as less formal, compared with speech communication***	
		• Video communication with high quality when used for persuasion when persons are arguing when used for a third party can lead to users being perceived as less formal (compared with speech communication)	[i.53] (RR)
	7.3.1.71	***Eye contact may improve trust (compared with a 20° view angle with no eye contact)***	
		• For video communication a centre viewing angle may increase trust in the communication partner (compared with a viewing angle either slightly at the side or from the top at approximately 20°)	[i.3] (JA)
		• For video communication a viewing angle slightly at the side by approximately 20° may decrease trust in the communication partner (compared with a centre viewing angle)	[i.3] (JA)
		• For video communication a viewing angle slightly at the top by approximately 20° may decrease trust in the communication partner (compared with a centre viewing angle)	[i.3] (JA)
Deaf or hearing impairment			
	7.3.1.72	***A frame-rate of 20 fps offers good quality***	
		• Video communication with 20 fps when used for sign language and lip-reading by deaf persons offers good quality	[i.33] (ITU-T info)
	7.3.1.73	***A frame-rate of 12 fps to 15 fps is possible to use***	
		• Video communication with 12 fps to 15 fps when used for sign language and lip-reading by deaf persons is possible to use by experienced persons when used for short conversations	[i.33] (ITU-T info)
	7.3.1.74	***100 ms delay offers preferred quality***	
		• Video communication with maximum 0,1 second delay when used for sign language and lip-reading by deaf persons offers preferred quality	[i.33] (ITU-T info)
	7.3.1.75	***400 ms delay offers acceptable quality***	
		• Video communication with 0,4 second delay when used for sign language and lip-reading by deaf persons offers acceptable quality	[i.33] (ITU-T info)
	7.3.1.76	***800 ms delay makes good sign conversation difficult***	
		• Video communication with over 0,8 second delay when used for sign language and lip-reading by deaf persons hinders good conversation	[i.33] (ITU-T info)

…topic continued on next page

Deaf or hearing impairment (continued)	7.3.1.77	***The hands and face should have the highest resolution***	
		• Video communication with different resolutions for different parts of the picture when used for sign-language by deaf persons should give the hands and face the highest resolution	[i.33] (ITU-T info)
	7.3.1.78	***QCIF resolution is usable***	
		• Video communication with QCIF resolution when used for sign language by deaf persons is usable	[i.33] (ITU-T info)
	7.3.1.79	***CIF resolution is good***	
		• Video communication with at least CIF resolution when used for sign-language by deaf persons is good	[i.33] (ITU-T info)
	7.3.1.80	***QCIF and 10 fps to14 fps video communication can aid everyday life, compared with text communication***	
		• Mobile video communication with QCIF and 10 fps to14 fps when used for sign-language provides people who are deaf with more opportunities when used for increased spontaneity and flexibility in everyday life (compared with text communication)	[i.54] (RR)
	7.3.1.81	***QCIF resolution and a frame-rate of 10 fps to 14 fps enables users to benefit greatly from increased communication***	
		• Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language enables people who are deaf to benefit greatly from increased communication	[i.54] (RR)
	7.3.1.82	***3G mobile video communication has reduced utility because calls may fail***	
		• Mobile video communication on a 3G network when used for sign-language by people who are deaf has reduced utility because 13 % to 17 % of calls either fail to connect or become disconnected	[i.54] (RR)
	7.3.1.83	***100 ms asynchrony is acceptable for voice-supported lip-reading***	
		• Video communication with up to 100 ms asynchronisation between audio and video when used for voice-supported lip-reading by deaf persons is acceptable	[i.33] (ITU-T info)
	7.3.1.84	***QCIF resolution is adequate***	
		• Video communication with QCIF resolution when used for lip-reading by deaf persons is adequate	[i.33] (ITU-T info)

…topic continued on next page

Deaf or hearing impairment (continued)	7.3.1.85	***QCIF and 10 fps to14 fps offers acceptable quality for sign-language interpretation and communication relay***	
		• Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language interpretation and communication relay by deaf persons and interpreters can improve quality of life	[i.54] (RR)
		• Mobile video communication with QCIF when used for sign-language interpretation and communication relay by deaf persons and interpreters functions beyond expectations	[i.54] (RR)
		• Mobile video communication with 10 fps to 14 fps when used for sign-language interpretation and communication relay by deaf persons and interpreters functions beyond expectations	[i.54] (RR)
Speech impairment			
	7.3.1.86	***Video communication for voice therapy may reduce the number of patients that stop therapy prematurely, compared with face-to-face therapy***	
		• Video communication (hard-wired camera and monitor) for voice therapy may provide the same positive outcomes as conventional therapy in the same room	[i.47] (JA)
	7.3.1.87	***Video communication for speech-language therapy for brain injured people may provide the same positive outcomes as face-to-face therapy***	
		• Video communication (hard-wired camera and monitor) for voice therapy may reduce the number of patients that discontinuing therapy prematurely (compared with conventional therapy in the same room)	[i.47] (JA)
	7.3.1.88	***Video communication for speech-language therapy may provide the same positive outcomes as face-to-face therapy***	
		• Video communication for speech-language therapy for brain injured people may provide the same positive outcomes as face-to-face therapy	[i.5] (JA)
	7.3.1.89	***Video communication for speech-language therapy may have a high level of acceptance by brain injured people***	
		• Video communication for speech-language therapy may have a high level of acceptance by brain injured people	[i.5] (JA)

Medical interview task			
	7.3.1.90	***Video communication for voice therapy may provide the same positive outcomes as face-to-face therapy***	
		• Video communication (hard-wired camera and monitor) for voice therapy may provide the same positive outcomes as conventional therapy in the same room	[i.47] (JA)
	7.3.1.91	***Video communication for voice therapy may reduce the number of patients that stop therapy prematurely, compared with face-to-face therapy***	
		• Video communication (hard-wired camera and monitor) for voice therapy may reduce the number of patients that discontinuing therapy prematurely (compared with conventional therapy in the same room)	[i.47] (JA)
	7.3.1.92	***Video communication for speech-language therapy may provide the same positive outcomes as face-to-face therapy***	
		• Video communication for speech-language therapy for brain injured people may provide the same positive outcomes as face-to-face therapy	[i.5] (JA)
	7.3.1.93	***Video communication for speech-language therapy may have a high level of acceptance by brain injured people***	
		• Video communication for speech-language therapy may have a high level of acceptance by brain injured people	[i.5] (JA)
	7.3.1.94	***Video communication for diagnosis of dementia may be as accurate as face-to-face diagnosis***	
		• Video communication for diagnostic interviewing has shown good agreement with conventional face-to-face diagnosis of dementia	[i.8] (JA)

7.3.2 Video communication: Remote inspection

Topic	Guideline number	***Guideline*** • Empirical source	Ref.
Audio-video asynchrony			
	7.3.2.1	***Remote inspection with 500 ms asynchrony offers good quality*** • Remote inspection with 500 ms asynchrony when used for giving advice on a procedure does not affect task performance (compared with remote inspection with no delay)	[i.53] (RR)
Resolution			
	7.3.2.2	***Remote inspection with CIF and 15 fps offers good quality for moving the camera to show the environment*** • Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for moving the camera to show the environment offers good quality	EP
	7.3.2.3	***Remote inspection with CIF and 10 fps offers acceptable quality for moving the camera to show the environment*** • Remote inspection with asymmetric video (from a mobile device to personal computer) with CIF and 10 fps when used for moving the camera to show the environment offers acceptable quality	EP
	7.3.2.4	***Remote inspection with CIF and 5 fps is usable for moving the camera to show the environment*** • Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for moving the camera to show the environment is usable	EP

…topic continued on next page

Resolution (continued)	7.3.2.5	***Remote inspection with CIF and 15 fps offers good quality for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for recognising objects offers good quality	EP
	7.3.2.6	***Remote inspection with CIF and 5 fps is unusable for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for recognising objects offers acceptable quality	EP
	7.3.2.7	***Remote inspection with QCIF and 25 fps is usable for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 25 fps when used for recognising objects is usable	EP
	7.3.2.8	***Remote inspection with QCIF and 10 fps offers acceptable quality for reading 10-point text***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 10 fps when used for reading 10-point text offers usable quality	EP
	7.3.2.9	***Remote inspection with QCIF and 5 fps is unusable for reading 10-point text***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 5 fps when used for reading 10-point text hinders good communication	EP
	7.3.2.10	***Remote inspection with SQCIF and 10 fps offers good quality for recognising an environment***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with SQCIF and 10 fps when used for showing an environment offers good quality	EP
Frame-rate			
	7.3.2.11	***Remote inspection with CIF and 15 fps offers good quality for moving the camera to show the environment***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for moving the camera to show the environment offers good quality	EP

…topic continued on next page

Frame-rate (continued)	7.3.2.12	***Remote inspection with CIF and 10 fps offers acceptable quality for moving the camera to show the environment***	
		• Remote inspection with asymmetric video (from a mobile device to personal computer) with CIF and 10 fps when used for moving the camera to show the environment offers acceptable quality	EP
	7.3.2.13	***Remote inspection with CIF and 5 fps is usable for moving the camera to show the environment***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for moving the camera to show the environment is usable	EP
	7.3.2.14	***Remote inspection with CIF and 15 fps offers good quality for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for recognising objects offers good quality	EP
	7.3.2.15	***Remote inspection with CIF and 5 fps is unusable for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for recognising objects offers acceptable quality	EP
	7.3.2.16	***Remote inspection with QCIF and 25 fps is usable for recognising objects***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 25 fps when used for recognising objects is usable	EP
	7.3.2.17	***Remote inspection with QCIF and 10 fps offers usable quality for reading 10-point text***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 10 fps when used for reading 10-point text offers usable quality	EP
	7.3.2.18	***Remote inspection with QCIF and 5 fps is unusable for reading 10-point text***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 5 fps when used for reading 10-point text hinders good communication	EP
	7.3.2.19	***Remote inspection with SQCIF and 10 fps offers good quality for recognising an environment***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with SQCIF and 10 fps when used for showing an environment offers good quality	EP

Packet loss	7.3.2.20	***Remote inspection video communication with 1 % packet loss on a fixed line offers good quality for the video channel***	
		• Video communication with 1 % random packet loss (not burst) offers good quality for the video channel	EP
Cost-benefit	7.3.2.21	***Remote inspection video communication can be expected to be used for business applications in situations where the benefit is obviously high***	
		• Remote inspection video communication can be expected to be used for business applications in situations where the benefit is obviously high (such as in road construction, oil extraction or ship operations)	EP
Self view	7.3.2.22	***Remote inspection with self-view improves communication efficiency (compared with remote inspection without self-view)***	
		• Remote inspection with both sites seeing the same video images improves communication efficiency (compared with remote inspection without self-view)	EP
		• Remote inspection when the users at the site sending the video images do not see the images that are transmitted offers acceptable quality	EP
Instruction task	7.3.2.23	***Remote inspection may allow less monitoring of the other person's concentration, compared with "face-to-face" video communication***	
		• Speech communication when used for information transfer allows less monitoring of the other person's level of concentration (compared with video communication with sufficient video quality)	[i.60] (JA)
Problem solving task	7.3.2.24	***Video communication used for remote inspection rather than "face-to-face" video communication may be preferred for problem solving***	
		• Remote inspection when used for problem solving may be preferred (compared to face-to-face video)	[i.1] (JA)

Showing surroundings task	7.3.2.25	***Remote inspection with CIF and 15 fps when used for moving the camera to show the environment offers good quality***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for moving the camera to show the environment offers good quality	EP
	7.3.2.26	***Remote inspection with CIF and 10 fps offers acceptable quality for moving the camera to show the environment***	
		• Remote inspection with asymmetric video (from a mobile device to personal computer) with CIF and 10 fps when used for moving the camera to show the environment offers acceptable quality	EP
	7.3.2.27	***Remote inspection with CIF and 5 fps is usable for moving the camera to show the environment***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for moving the camera to show the environment is usable	EP
Object recognition task	7.3.2.28	***Video communication used for remote inspection rather than "face-to-face" video communication may be preferred for object selection***	
		• Remote inspection when used for object selection tasks may be valued more than "face-to-face" video communication	[i.1] (JA)
	7.3.2.29	***Remote inspection with CIF and 15 fps when used for recognising objects offers good quality***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 15 fps when used for recognising objects offers good quality	EP
	7.3.2.30	***Remote inspection with CIF and 5 fps when used for recognising objects is unusable***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with CIF and 5 fps when used for recognising objects offers acceptable quality	EP
	7.3.2.31	***Remote inspection with QCIF and 25 fps when used for recognising objects is usable***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 25 fps when used for recognising objects is usable	EP

…topic continued on next page

Object recognition task (continued)	7.3.2.32	***Remote inspection with QCIF and 10 fps when used for reading 10-point text offers acceptable quality***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 10 fps when used for reading 10-point text offers usable quality	EP
	7.3.2.33	***Remote inspection with QCIF and 5 fps when used for reading 10-point text is unusable***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with QCIF and 5 fps when used for reading 10-point text hinders good communication	EP
	7.3.2.34	***Remote inspection with SQCIF and 10 fps when used for recognising an environment offers good quality***	
		• Remote inspection with asymmetric video (from a mobile device to a personal computer) with SQCIF and 10 fps when used for showing an environment offers good quality	EP
Blind or visual impairment			
	7.3.2.35	***Some blind persons may prefer to use mobile video communication rather than a guide dog or accompanying person***	
		• Mobile video communication with high quality on a 6" screen when used for a service provider to assist a blind user is preferable to a guide dog or an accompanying person when used for some blind people	[i.29] (RR)
	7.3.2.36	***Remote assistance is a service that blind users are likely to use***	
		• Mobile video communication with high quality on a 6" screen when used for a service provider to assist a blind user is likely to be used if offered as a real service	[i.29] (RR)
		• Mobile video communication with high quality on a 6" screen when used for a service provider to assist a blind user to observe someone or something can be very important when used for the blind user	[i.29] (RR)
	7.3.2.37	***At least CIF and 5 fps may be required when used for reading text***	
		• Mobile video communication with at least CIF resolution and 5 fps may be required for reading text with adequate effectiveness and efficiency	[i.29] (RR)
		• Mobile video communication with QCIF or SQCIF and 2 fps to 3 fps may not be adequate for reading text with effectiveness and efficiency (compared with mobile video communication with CIF resolution and 5 fps)	[i.29] (RR)

…topic continued on next page

Blind or visual impairment (continued)	7.3.2.38	***At least CIF and 10 fps may be required for identifying an object while moving***	
		• Mobile video communication with at least CIF resolution and 10 fps may be required for identifying an object while moving with adequate effectiveness and efficiency	[i.29] (RR)
		• Mobile video communication with QCIF or SQCIF and 5 fps to 6 fps may be inadequate for identifying an object while moving with effectiveness and efficiency (compared with mobile video communication with CIF resolution and 10 fps)	[i.29] (RR)
	7.3.2.39	***GSM audio offers good enough quality for assisting blind users***	
		• Mobile video communication with GSM audio quality when used for assisting blind users is good enough quality	[i.29] (RR)
	7.3.2.40	***Assisting a blind user to verify an object or information can have high effectiveness***	
		• Mobile video communication with high quality on a 6" screen when used for a service provider to assist a blind user to verify an object or information can be 100 % effective	[i.29] (RR)
	7.3.2.41	***Assisting a blind user to search for information or an object can have high effectiveness***	
		• Mobile video communication with high quality on a 6" screen when used for a service provider to assist a blind user to search when used for an object or information can be 98 % to 100 % effective	[i.29] (RR)

7.3.3 Video communication: Multi-point and heterogeneous networks

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Window configuration			
	7.3.3.1	***Multipoint mobile video communication with continuous presence via a 4-spilt windows matrix offers good quality for general meetings***	
		• Multipoint mobile video communication with continuous prescience with window(s) covering 1/4 of the screen when used for general meetings offers good quality	EP
	7.3.3.2	***Multipoint mobile video communication with continuous presence via a 9-split windows matrix is usable for general meetings***	
		• Multipoint mobile video communication with continuous prescience with window(s) covering 1/9 of the screen for general meetings is usable	EP

…topic continued on next page

Window configuration (continued)	7.3.3.3	***If users of multi-point videoconferencing can chose between audio-switching and continuous presence, they generally prefer continuous presence display of the other sites***	
		• If the users of multi-point videoconferencing can chose between audio-switching and continuous presence, they generally prefer continuous presence display of the other sites, even though continuous presence costs more	EP

7.4 Multimedia communication

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Audio-video asynchrony			
	7.4.1	***200 ms asynchrony offers good quality***	
		• Video communication with 200 ms asynchrony when used for joint problem solving may not significantly affect task outcome compared with video communication with synchrony)	[i.53] (RR)
	7.4.2	***200 ms asynchrony offers quality similar to speech communication***	
		• Video communication with 200 ms asynchrony when used for joint problem solving can lead to communications similar to speech communication	[i.53] [i.25] (RR)
	7.4.3	***Video communication with lip asynchrony of 200 ms may have a higher perceived utility than speech communication***	
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with loud speaking speech communication)	[i.25] (RR)
View of person			
	7.4.4	***With a head-only view between strangers interaction may become less formal***	
		• Video communication with a head-only view between strangers can result in less formal interaction (compared to video communication with a head-and-torso view)	[i.27] (JA)

…topic continued on next page

View of person (continued)	7.4.5	***With a head-and-torso view between strangers interactions may be more formal***	
		• Video communication with a head-and-torso view between strangers can result in more formal interaction (compared to video communication with a head-only view)	[i.27] (JA)
	7.4.6	***With a head-only view between strangers, communication can become more like that between familiar people***	
		• Video communication with a head-only view when used for problem solving between strangers can result in behaviour that is more like interaction between familiar people (compared to video communication with a head-and-torso view)	[i.27] (JA)
Eye contact			
	7.4.7	***Parallax differences of less than 8° are acceptable***	
		• Video communication with non-eye contact systems should provide a parallax differences of less than 8° to be within the acceptable range according to the European Standard ETR 297 [i.20]	[i.20] ETSI ETR
		• Video communication in general conditions does not require the optical axes of camera and monitor coinciding	EP
		• The major video communication equipment manufacturers and service providers do not provide eye-contact systems on the market as users are not asking for them	EP
		• Video communication in a conference room with non-exact eye contact is reported as satisfactory for general use by users and service providers	EP
	7.4.8	***Eye contact during negotiation may not lead to higher satisfaction***	
		• Video communication with direct eye-contact when used for negotiation does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)
	7.4.9	***Eye contact during decision making may not lead to higher satisfaction***	
		• Video communication with direct eye-contact when used for decision making does not lead to higher kevels of satisfaction or acceptance of the technology (compared to video communication without direct eye contact)	[i.50] (JA)

Media Quality			
	7.4.10	***Multimedia communication with high quality may be considered the most useful service for managerial communication***	
		• Multimedia communication with high quality when used for managerial work communication is considered the most useful new service (compared with video communication, speech communication and avatar communication)	[i.23] (RR)
		• Multimedia communication with high quality when used for managerial work tasks supports a real need when used for this type of communication	[i.23] (RR)
	7.4.11	***Multimedia communication with high quality can become integrated into daily work communications if on the users' desktop***	
		• Multimedia communication with high quality from a personal office can become integrated into daily work communications (compared with a video communication room)	[i.23] (RR)
	7.4.12	***Multimedia communication with high quality is often preferred because of the ability to present information on the screen***	
		• Multimedia communication with high quality when used for managerial work is preferred because of the ability to present information on the screen (compared with video communication, loud speaking speech communication, speech communication with a handset & avatar communication)	[i.23] (RR)
	7.4.13	***An easy to set-up and high quality conference does not stop actual face-to-face meetings***	
		• Multimedia communication with high-quality and easy and low-cost access when used for managerial work tasks does not change the pattern of actual face-to-face meetings	[i.23] (RR)
Urgency			
	7.4.14	***High quality video communication may be chosen for non-urgent communication***	
		• Video communication with high quality from a personal office for non-urgent communication is judged to be an improvement relative to using speech communication with a handset	[i.23] (RR)
		• Speech communication with a handset is regarded the best way to conduct urgent communication for managerial work tasks (compared with office-based loud speaking speech communication, video communication, multimedia communication and avatar communication)	[i.23] (RR)

…topic continued on next page

Urgency (continued)	7.4.15	***With high quality video communication more may be said for the same outcome, compared with speech communication***	
		• Video communication with high quality when used for negotiation can increase the amount users say in order to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
	7.4.16	***High quality video communication can lead to more interruptions than with speech communication***	
		• Video communication with high quality when used for negotiation can increase users' interruptions to agree a negotiated outcome (compared with speech communication and actual face-to-face communication)	[i.53] (RR)
Deaf or hearing impairment			
	7.4.17	***Multimedia communication for people with communication-related disabilities may improve the usability of emergency calls***	
		• Multimedia communication combining video communication with text communication for people with communication-related disabilities may improve the usability of emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.18	***Multimedia communication with a relay service may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls***	
		• Multimedia communication with video and text and a relay service (3-party multimedia) may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.19	***Multimedia fixed-line communication for sign language and remote inspection for emergency calls from people with communication-related disabilities gives excellent media quality***	
		• Multimedia communication with fixed broadband video and text for sign language and remote inspection for making emergency calls by people with communication-related disabilities gives excellent quality media (compared to emergency calls with a 3G handset)	[i.28] (CP)

…topic continued on next page

Deaf or hearing impairment (continued)	7.4.20	***Multimedia 3G communication for sign language and remote inspection for emergency calls from people with communication-related disabilities may be successful but less appreciated, compared to a fixed broadband terminal***	
		• Multimedia communication with mobile 3G video communication for sign language and remote inspection for making emergency calls by people with communication-related disabilities is successful but less appreciated (compared to emergency calls with a fixed broadband terminal)	[i.28] (CP)
	7.4.21	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may be very efficient, compared to traditional voice-only relay***	
		• Multimedia communication with a fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may be very efficient (compared to emergency calls with a traditional voice-only relay service)	[i.28] (CP)
	7.4.22	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may give less stress and a better feeling of contact than with instant messaging***	
		• Multimedia communication with fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may give less stress in moments of urgency and a better feeling of contact (than a traditional instant messaging method of text transmission in completed sentences)	[i.28] (CP)
Cognitive impairment	7.4.23	***Video communication for diagnosis of dementia may be as accurate as face-to-face diagnosis***	
		• Video communication for diagnostic interviewing has shown good agreement with conventional face-to-face diagnosis of dementia	[i.8] (JA)
	7.4.24	***Multimedia communication appears to be a valid way to conduct neuropsychological evaluation of cognitive impairment***	
		• Multimedia communication appears to be a valid way to conduct neuropsychological evaluation of older adults with cognitive impairment (compared with face-to-face examination)	[i.8] (JA)

…topic continued on next page

Cognitive impairment (continued)	7.4.25	***Using multimedia communication for neuropsychological evaluation of cognitive impairment may not present difficulties or concerns for patients***	
		• Multimedia communication for neuropsychological evaluation of older adults with cognitive impairment may not present difficulties or concerns for patients	[i.8] (JA)
	7.4.26	***If using multimedia communication for neuropsychological evaluation of cognitive impairment, a minority of patients may express a preference for personal contact***	
		• With multimedia communication for neuropsychological evaluation of older adults with cognitive impairment, a minority of patients may express a preference for personal contact	[i.8] (JA)
	7.4.27	***Multimedia communication for people with communication-related disabilities may improve the usability of emergency calls***	
		• Multimedia communication combining video communication with text communication for people with communication-related disabilities may improve the usability of emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.28	***Multimedia communication with a relay service may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls***	
		• Multimedia communication with video and text and a relay service (3-party multimedia) may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.29	***Multimedia fixed-line communication for sign language and remote inspection for emergency calls from people with communication-related disabilities gives excellent media quality***	
		• Multimedia communication with fixed broadband video and text for sign language and remote inspection for making emergency calls by people with communication-related disabilities gives excellent quality media (compared to emergency calls with a 3G handset)	[i.28] (CP)

…topic continued on next page

Cognitive impairment (continued)	7.4.30	***Multimedia 3G communication for sign language and remote inspection for emergency calls from people with communication-related disabilities may be successful but less appreciated, compared to a fixed broadband terminal***	
		• Multimedia communication with mobile 3G video communication for sign language and remote inspection for making emergency calls by people with communication-related disabilities is successful but less appreciated (compared to emergency calls with a fixed broadband terminal)	[i.28] (CP)
	7.4.31	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may be very efficient, compared to traditional voice-only relay***	
		• Multimedia communication with a fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may be very efficient (compared to emergency calls with a traditional voice-only relay service)	[i.28] (CP)
	7.4.32	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may give less stress and a better feeling of contact than with instant messaging***	
		• Multimedia communication with fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may give less stress in moments of urgency and a better feeling of contact (than a traditional instant messaging method of text transmission in completed sentences)	[i.28] (CP)
Medical interview task			
	7.4.33	***Multimedia communication appears to be a valid way to conduct neuropsychological evaluation of cognitive impairment***	
		• Multimedia communication appears to be a valid way to conduct neuropsychological evaluation of older adults with cognitive impairment (compared with face-to-face examination)	[i.8] (JA)
	7.4.34	***Using multimedia communication for neuropsychological evaluation of cognitive impairment may not present difficulties or concerns for patients***	
		• Multimedia communication for neuropsychological evaluation of older adults with cognitive impairment may not present difficulties or concerns for patients	[i.8] (JA)

…topic continued on next page

Medical interview task (continued)	7.4.35	***If using multimedia communication for neuropsychological evaluation of cognitive impairment, a minority of patients may express a preference for personal contact***	
		• With multimedia communication for neuropsychological evaluation of older adults with cognitive impairment, a minority of patients may express a preference for personal contact	[i.8] (JA)
Emergency calls	7.4.36	***Multimedia communication for people with communication-related disabilities may improve the usability of emergency calls***	
		• Multimedia communication combining video communication with text communication for people with communication-related disabilities may improve the usability of emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.37	***Multimedia communication with a relay service may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls***	
		• Multimedia communication with video and text and a relay service (3-party multimedia) may improve the confidence and efficiency of people with communication-related disabilities for making emergency calls (compared to emergency calls without video and real-time text)	[i.28] (CP)
	7.4.38	***Multimedia fixed-line communication for sign language and remote inspection for emergency calls from people with communication-related disabilities gives excellent media quality***	
		• Multimedia communication with fixed broadband video and text for sign language and remote inspection for making emergency calls by people with communication-related disabilities gives excellent quality media (compared to emergency calls with a 3G handset)	[i.28] (CP)
	7.4.39	***Multimedia 3G communication for sign language and remote inspection for emergency calls from people with communication-related disabilities may be successful but less appreciated, compared to a fixed broadband terminal***	
		• Multimedia communication with mobile 3G video communication for sign language and remote inspection for making emergency calls by people with communication-related disabilities is successful but less appreciated (compared to emergency calls with a fixed broadband terminal)	[i.28] (CP)

…topic continued on next page

Emergency calls (continued)	7.4.40	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may be very efficient, compared to traditional voice-only relay***	
		• Multimedia communication with a fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may be very efficient (compared to emergency calls with a traditional voice-only relay service)	[i.28] (CP)
	7.4.41	***Multimedia fixed-line communication with a text relay service for emergency calls by people with communication-related disabilities may give less stress and a better feeling of contact than with instant messaging***	
		• Multimedia communication with fixed broadband video and text communication and a text relay service for making emergency calls by people with communication-related disabilities may give less stress in moments of urgency and a better feeling of contact (than a traditional instant messaging method of text transmission in completed sentences)	[i.28] (CP)

7.5 Real-time games

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Delay			
	7.5.1	***One-way delay should be less than 200 ms for interactive games***	
		• With interactive games a one-way delay of <200 ms is a preliminary figure from ITU-T	[i.38] (ITU-T info)
Medium selection			
	7.5.2	***For game-play including social conversations requiring fewer cues and less speed, text may be preferred rather than speech***	
		• For real-time games private and group conversation requiring fewer cues and less speed may take place via text (compared with more complex and time-sensitive tasks performed by speech)	[i.61] (JA)
	7.5.3	***For game-play including complex and time-sensitive tasks, speech may be preferred rather than text***	
		• For real-time games more complex and time-sensitive tasks may be performed by speech (compared with private and group conversation requiring fewer cues and less speed taking place via text)	[i.61] (JA)

Background noise	7.5.4	***Voice communication during team-play may aid sociability if there is low background noise and the players know who is talking, compared with high background noise and when players do not know who is talking***	
		• For real-time games when background noise is low, voice may aid sociability if the players know who is talking in their team (compared with when background noise is high and players do not know who is talking on their team)	[i.61] (JA)
	7.5.5	***Voice communication during team-play may not aid sociability if there is high background noise, compared with low background noise and when players know who is talking***	
		• For real-time games when background noise is high, voice may not aid sociability among players in the same team (compared with when background noise is low and players know who is talking on their team)	[i.61] (JA)
Person perception	7.5.6	***Voice communication may increase liking between team members, compared with text communication***	
		• For real-time games use of voice communication may increase liking between team members (compared with text-only play)	[i.61] (JA)
	7.5.7	***Voice communication may increase trust between team members, compared with text communication***	
		• For real-time games the addition of VoIP may increase trust between team members (compared with text-only play)	[i.61] (JA)
	7.5.8	***With text communication, liking between team members may not increase, compared with text and voice communication***	
		• For real-time games with text-only, liking between team members may not increase (compared with text and VoIP play)	[i.61] (JA)
Social wellbeing	7.5.9	***With games offering text-only team-play, users may feel generally less happy, compared with users having text and voice communication***	
		• For real-time games with text-only play in teams, players may feel less happy (compared with text plus VoIP)	[i.61] (JA)

…topic continued on next page

Social wellbeing (continued)	7.5.10	***With games offering both text and voice communication, players may feel generally happier, compared with users having text-only communication***	
		• For real-time games with text and VoIP play in teams, players may feel happier (compared with text-only play)	[i.61] (JA)
	7.5.11	***With games offering text-only team-play, players may feel more lonely, compared with users having both text and voice communication***	
		• For real-time games with text-only play in teams, players may feel more lonely (compared with text plus VoIP play)	[i.61] (JA)
	7.5.12	***With games offering both text and voice communication, users may feel less lonely, compared with users having text-only communication***	
		• For real-time games with text and VoIP play in teams, players may feel less lonely (compared with text-only play)	[i.61] (JA)

7.6 TV

Guidelines on television are clustered within two different areas that primarily distinguish the size of the device for watching:

- PC TV (e.g. watching on a lap-top or desk-top computer);
- Mobile TV.

7.6.1 PC TV

Topic	Guideline number	***Guideline*** • Empirical source	**Ref.**
Frame-rate			
	7.6.1.1	***Reducing both colour depth (from 24-bit to 8-bit) and frame-rate (from 25 fps to 5 fps) at the same time may reduce user satisfaction, compared with when just one parameter is changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if both colour depth and framerate are simultaneously changed to 8-bit and 5 fps within a single presentation, user satisfaction is very likely to be affected (compared with when just one parameter is changed)	[i.26] (JA)

…topic continued on next page

Frame-rate (continued)	7.6.1.2	***If colour depth is reduced from 24-bit to 8-bit, user satisfaction may not be affected, compared with when both colour depth and frame rate are changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if colour depth is changed to 8-bit, user satisfaction is unlikely to be affected (compared with when both colour depth and frame rate are changed)	[i.26] (JA)
	7.6.1.3	***If frame rate is reduced from 25 fps to 5 fps, user satisfaction may not be affected, compared with when both frame rate and colour depth are changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if frame rate is changed to 5 fps, user satisfaction is unlikely to be affected (compared with when both colour depth and frame rate are changed)	[i.26] (JA)
Colour depth			
	7.6.1.4	***Reducing both colour depth (from 24-bit to 8-bit) and frame-rate (from 25 fps to 5 fps) at the same time may reduce user satisfaction, compared with when just one parameter is changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if both colour depth and framerate are simultaneously changed to 8-bit and 5 fps within a single presentation, user satisfaction is very likely to be affected (compared with when just one parameter is changed)	[i.26] (JA)
	7.6.1.5	***If colour depth is reduced from 24-bit to 8-bit, user satisfaction may not be affected, compared with when both colour depth and frame rate are changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if colour depth is changed to 8-bit, user satisfaction is unlikely to be affected (compared with when both colour depth and frame rate are changed)	[i.26] (JA)
	7.6.1.6	***If frame rate is reduced from 25 fps to 5 fps, user satisfaction may not be affected, compared with when both frame rate and colour depth are changed***	
		• For one-way video with CIF and 24-bit colour depth and 25 fps, if frame rate is changed to 5 fps, user satisfaction is unlikely to be affected (compared with when both colour depth and frame rate are changed)	[i.26] (JA)

7.6.2 Mobile TV

Topic	Guideline number	***Guideline*** • Empirical source	Ref.
Frame-rate			
	7.6.2.1	***With 12,5 fps for mixed-content viewing, 80 % of users may be satisfied***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, a large majority of users (about 80 %) may be satisfied	[i.48] (RA)
	7.6.2.2	***For watching football, higher resolution may be preferred over higher frame rate***	
		• For Mobile TV for watching football, perceived quality may be relatively insensitive to changes in frame rate and users prefer high-resolution images to high frame rate	[i.61] (CP)
Resolution			
	7.6.2.3	***An image size of 168x126 pixels, regardless of content type, may be less acceptable than 208x156***	
		• Mobile TV with an image size of 168x126, regardless of content type, may be considered less acceptable (compared with the larger images of 208x156 and 240x180)	[i.56] (CP)
	7.6.2.4	***An image size of 120x90 pixels, regardless of content type, may be less acceptable than 208x156***	
		• Mobile TV with an image size of 120x90, regardless of content type, may be considered less acceptable (compared with the larger images of 208x156 and 240x180)	[i.44] (CP)
	7.6.2.5	***An image size of 208x156 pixels, regardless of content type, may be as acceptable as 240x180***	
		• Mobile TV with an image size of 208x156, regardless of content type, may be considered acceptable (no difference in acceptability when compared with the larger image of 240x180)	[i.44] (CP)
	7.6.2.6	***An image size of 208x156 pixels, regardless of content type, may be more acceptable than 168x126***	
		• Mobile TV with an image size of 208x156, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)

…topic continued on next page

Resolution (continued)	7.6.2.7	***An image size of 240x180 pixels, regardless of content type, may be more acceptable than 168x126***	
		• Mobile TV with an image size of 240x180, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)
	7.6.2.8	***An image size of 240x180 pixels, regardless of content type, may be more acceptable than 168x126***	
		• Mobile TV with an image size of 240x180, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)
	7.6.2.9	***An image size of 240x180 pixels for viewing cartoon animation, may be more acceptable than 120x90***	
		• For Mobile TV with an image resolution of 240x180 for viewing cartoon animation, acceptability may increase (compared with an image resolution of 120x90)	[i.44] (CP)
		• For Mobile TV with an image resolution of 168x126 for viewing cartoon animation acceptability may not decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	7.6.2.10	***An image resolution of 208x156 pixels for viewing News may be less acceptable than 240x180***	
		• For Mobile TV with an image resolution of 208x156 for viewing News, acceptability may decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	7.6.2.11	***An image resolution of 240x180 pixels for viewing Music videos may not increase acceptability, compared with 208x156***	
		• For Mobile TV with an image resolution of 240x180 for viewing Music videos, acceptability may not increase (compared with an image resolution of 208x156)	[i.44] (CP)
	7.6.2.12	***An image resolution of 168x126 for viewing Music videos may not increase acceptability, compared with 120x90***	
		• For Mobile TV with an image resolution of 168x126 for viewing Music videos, acceptability may not increase (compared with an image resolution of 120x90)	[i.44] (CP)
	7.6.2.13	***An image resolution of 168x126 for viewing Music videos may not decrease acceptability, compared with 208x156***	
		• For Mobile TV with an image resolution of 168x126 for viewing Music videos, acceptability may not decrease (compared with an image resolution of 208x156)	[i.44] (CP)

…topic continued on next page

Resolution (continued)	7.6.2.14	***An image resolution of 208x156 for viewing Music videos may not decrease acceptability, compared with 240x180***	
		• For Mobile TV with an image resolution of 208x156 for viewing Music videos, acceptability may not decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	7.6.2.15	***An image resolution of 240x180 for viewing Sport may not increase acceptability, compared with 208x156***	
		• For Mobile TV with an image resolution of 240x180 for viewing Sport, acceptability may not increase (compared with an image resolution of 208x156)	[i.44] (CP)
	7.6.2.16	***An image resolution of 208x156 for viewing Sport may increase acceptability, compared with 168x126***	
		• For Mobile TV with an image resolution of 208x156 for viewing Sport, acceptability may increase (compared with an image resolution of 168x126)	[i.44] (CP)
	7.6.2.17	***An image resolution of 168x126 for viewing Sport may increase acceptability, compared with 120x90***	
		• For Mobile TV with an image resolution of 168x126 for viewing Sport, acceptability may increase (compared with an image resolution of 120x90)	[i.44] (CP)
	7.6.2.18	***With 120x90 resolution users may require angular sizes of at least 4°***	
		• For Mobile TV with 120x90 resolution users may require angular sizes of at least 4° (or 14H)	[i.46] (CP)
	7.6.2.19	***With 120x90 resolution users may prefer an angular size of 5,4°, compared with an angular size of 6,5° for 168x126***	
		• For Mobile TV with 120x90 resolution users may prefer an angular size of 5,4° (10,6H) (compared with an angular size of 6,5° (8,7H) for 168x126 resolution)	[i.46] (CP)
	7.6.2.20	***With 168x126 resolution users may prefer an angular size of 6,5°, compared with an angular size of 5,4° for 120x90***	
		• For Mobile TV with 168x126 resolution users may prefer an angular size of 6,5° (8,7H) (compared with an angular size of 5,4° (10,6H) for 120x90 resolution)	[i.46] (CP)
	7.6.2.21	***A screen size of 4 cm height and QCIF resolution should result in very high user acceptability***	
		• Mobile TV devices with a screen size of 4 cm height and QCIF resolution should result in very high user acceptability	[i.46] (CP)

…topic continued on next page

Resolution (continued)	7.6.2.22	***Watching News with 168x126 or QCIF may be acceptable, compared with 120x90***	
		• Mobile TV for watching News with video resolution of at least 168x126 or QCIF format (176x144) may be acceptable (compared with 120x90 and regardless of whether text is big enough to read or is delivered separately from the video)	[i.45] (CP)
	7.6.2.23	***For watching football, perceived quality may be relatively insensitive to changes in frame-rate and users may prefer high-resolution images to high frame rate***	
		• For Mobile TV for watching football, perceived quality may be relatively insensitive to changes in frame rate and users prefer high-resolution images to high frame rate	[i.48] (CP)
	7.6.2.24	***Watching sport with quantization (Q) levels ranging from 12 to 24 may result in a sharp drop in acceptability, compared with Q ranging from 2 to 8***	
		• With IPTV with quantization (Q) levels ranging from 12 to 24 for watching Sport may result in a sharp drop in acceptability (compared with Q ranging from 2 to 8)	[i.48] (CP)
Packet loss			
	7.6.2.25	***For time-slice error rates the quality acceptance threshold may be between of 6,9 % and 13,8 %***	
		• For Mobile TV the quality acceptance threshold lies somewhere between time-slice error rates of 6,9 % and 13,8 %	[i.42] (WP)

Bit-rate		
7.6.2.26	***With 16 kbps for audio, users may rate the image quality as more acceptable compared with 32 kbps***	
	• For Mobile TV with a video bit rate of 224 kbps and audio bitrate of 16 kbps users may be more likely to rate image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 192 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 160 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 128 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 96 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 64 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
	• For Mobile TV with a video bit rate of 32 kbps and audio bitrate of 16 kbps users may be more likely to rate the image quality as acceptable (compared with an audio bit rate of 32 kbps and regardless of different image resolutions and content types)	[i.44] (CP)
Screen size		
7.6.2.27	***A screen height of 4 cm and QCIF resolution should result in very high user acceptability***	
	• Mobile TV devices with a screen size of 4 cm height and QCIF resolution should result in very high user acceptability	[i.46] (CP)

Content type		
7.6.2.28	***Regardless of content type, 168x126 pixels may be less acceptable than 208x156***	
	• Mobile TV with an image size of 168x126, regardless of content type, may be considered less acceptable (compared with the larger images of 208x156 and 240x180)	[i.44] (CP)
7.6.2.29	***Regardless of content type, 120x90 pixels may be less acceptable compared with of 208x156***	
	• Mobile TV with an image size of 120x90, regardless of content type, may be considered less acceptable (compared with the larger images of 208x156 and 240x180)	[i.44] (CP)
7.6.2.30	***Regardless of content type, 208x156 pixels may be as acceptable as 240x180***	
	• Mobile TV with an image size of 208x156, regardless of content type, may be considered acceptable (no difference in acceptability when compared with the larger image of 240x180)	[i.44] (CP)
7.6.2.31	***Regardless of content type, 208x156 may be more acceptable compared with 168x126***	
	• Mobile TV with an image size of 208x156, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)
7.6.2.32	***Regardless of content type, an image size of 240x180 may be more acceptable than 168x126***	
	• Mobile TV with an image size of 240x180, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)
7.6.2.33	***Regardless of content type, 240x180 may be more acceptable than 168x126***	
	• Mobile TV with an image size of 240x180, regardless of content type, may be considered more acceptable (compared with the smaller image of 168x126)	[i.44] (CP)
7.6.2.34	***For viewing cartoon animation, with 240x180 pixels acceptability may increase compared with 120x90***	
	• For Mobile TV with an image resolution of 168x126 for viewing cartoon animation acceptability may not decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	• For Mobile TV with an image resolution of 240x180 for viewing cartoon animation, acceptability may increase (compared with an image resolution of 120x90)	[i.44] (CP)

…topic continued on next page

Content type (continued)	7.6.2.35	***For viewing News, with 208x156 pixels acceptability may decrease compared with 240x180***	
		• For Mobile TV with an image resolution of 208x156 for viewing News, acceptability may decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	7.6.2.36	***For viewing Music videos, with 240x180 pixels acceptability may not increase compared with 208x156***	
		• For Mobile TV with an image resolution of 240x180 for viewing Music videos, acceptability may not increase (compared with an image resolution of 208x156)	[i.44] (CP)
	7.6.2.37	***For viewing Music videos, with an image resolution of 168x126 acceptability may not increase compared with 120x90***	
		• For Mobile TV with an image resolution of 168x126 for viewing Music videos, acceptability may not increase (compared with an image resolution of 120x90)	[i.44] (CP)
	7.6.2.38	***For viewing Music videos, with 168x126 acceptability may not decrease compared with 208x156***	
		• For Mobile TV with an image resolution of 168x126 for viewing Music videos, acceptability may not decrease (compared with an image resolution of 208x156)	[i.44] (CP)
	7.6.2.39	***For viewing Music videos, with 208x156 acceptability may not decrease compared with 240x180***	
		• For Mobile TV with an image resolution of 208x156 for viewing Music videos, acceptability may not decrease (compared with an image resolution of 240x180)	[i.44] (CP)
	7.6.2.40	***For viewing Sport, with 240x180 acceptability may not increase compared with 208x156***	
		• For Mobile TV with an image resolution of 240x180 for viewing Sport, acceptability may not increase (compared with an image resolution of 208x156)	[i.44] (CP)
	7.6.2.41	***For viewing Sport, with 208x156 acceptability may increase compared with 168x126***	
		• For Mobile TV with an image resolution of 208x156 for viewing Sport, acceptability may increase (compared with an image resolution of 168x126)	[i.44] (CP)

…topic continued on next page

Content type (continued)	7.6.2.42	***For viewing Sport, with an image resolution of 168x126 acceptability may increase compared with 120x90***	
		• For Mobile TV with an image resolution of 168x126 for viewing Sport, acceptability may increase (compared with an image resolution of 120x90)	[i.44] (CP)
	7.6.2.43	***For mixed content, with 12,5 fps a large majority of users should be satisfied***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, a large majority of users (about 80 %) may be satisfied	[i.48] (RA)
	7.6.2.44	***For watching News, having text presented in a separate 'ticker window' may increase the perceived quality of the video, compared to having text in the main window***	
		• Mobile TV for watching News that has text presented in a separate 'ticker window' may lead to users rating the video quality as higher (compared to presentation of text in the main window)	[i.45] (CP)
	7.6.2.45	***For watching News, text presented in a separate 'ticker window' with a text height of 8 pixels may increase the perceived quality of the video, compared with text height of 6 pixels***	
		• Mobile TV for watching News with text height of 8 pixels presented in a separate 'ticker window' may lead to higher acceptability of video quality (compared to text that has a height of 6 pixels)	[i.45] (CP)
	7.6.2.46	***For watching News, text presented in a separate 'ticker window' with a text height of 6 pixels may decrease the perceived quality of the video, compared with text height of 8 pixels***	
		• Mobile TV for watching News with text height of 6 pixels presented in a separate 'ticker window' may lead to lower acceptability of video quality (compared to text that has a height of 8 pixels)	[i.45] (CP)
	7.6.2.47	***For watching News, 168x126 or QCIF may be acceptable, whereas 120x90 may be unacceptable***	
		• Mobile TV for watching News with video resolution of at least 168x126 or QCIF format (176x144) may be acceptable (compared with 120x90 and regardless of whether text is big enough to read or is delivered separately from the video)	[i.45] (CP)
	7.6.2.48	***For watching football, perceived quality may be relatively insensitive to changes in frame-rate and users may prefer high-resolution to high frame rate***	
		• For Mobile TV for watching football, perceived quality may be relatively insensitive to changes in frame rate and users prefer high-resolution images to high frame rate	[i.48] (CP)

…topic continued on next page

Content type (continued)	7.6.2.49	***Watching Sport with quantization (Q) levels ranging from 12 to 24 may result in a sharp drop in acceptability, compared with Q ranging from 2 to 8***	
		• With IPTV with quantization (Q) levels ranging from 2 to 8 for watching Sport may result in a large increase in acceptability (compared with Q ranging from 12 to 24)	[i.48] (CP)
Pattern of use			
	7.6.2.50	***For mixed-content viewing, the average view time per session is about 25 minutes***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, the average view time per session may be about 25 minutes	[i.48] (RA)
	7.6.2.51	***For mixed-content viewing, the typical number of session may be 1-2 per day***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, the typical number of session may be 1-2 per day	[i.48] (RA)
	7.6.2.52	***For mixed-content viewing, the average viewing duration may be 3 hours per week***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, the average viewing duration may be 3 hours per week	[i.48] (RA)
	7.6.2.53	***For mixed-content viewing, the most frequent viewing locations are at Home (most frequent), at Work/University and 'on a bus'***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, the main viewing locations may be the Home (most frequent), at Work/University (second most frequent) and the Bus (third most frequent)	[i.48] (RA)
	7.6.2.54	***For mixed-content viewing, the most suitable type of programme to watch may be News (most suitable), Soaps and Music***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, the three content types considered most suitable to watch may be News (most suitable), Soaps and Music	[i.48] (RA)
	7.6.2.55	***For mixed-content viewing, a channel-change time of 5-6 seconds may be acceptable***	
		• For Mobile TV with 12,5 fps for mixed-content viewing, a channel-change time of 5-6 seconds may be acceptable	[i.48] (RA)

Topic	Guideline number	*Guideline* • Empirical source	Ref.
Viewing distance			
	7.6.2.56	***Services should be designed for viewing distances between 25 cm to 50 cm***	
		• Mobile TV services should be designed for close viewing distances between 25 cm to 50 cm	[i.46] (CP)
	7.6.2.57	***With 120x90 resolution users may require angular sizes of at least 4°***	
		• For Mobile TV with 120x90 resolution users may require angular sizes of at least 4° (or 14H)	[i.46] (CP)
	7.6.2.58	***With 120x90 resolution users may prefer an angular size of 5,4°, compared with 6,5° for 168x126***	
		• For Mobile TV with 120x90 resolution users may prefer an angular size of 5,4° (10,6H) (compared with an angular size of 6,5° (8,7H) for 168x126 resolution)	[i.46] (CP)
	7.6.2.59	***With 168x126 resolution users may prefer an angular size of 6,5°, compared with 5,4° for 120x90 resolution***	
		• For Mobile TV with 168x126 resolution users may prefer an angular size of 6,5° (8,7H) (compared with an angular size of 5,4° (10,6H) for 120x90 resolution)	[i.46] (CP)

7.7 Service selection

Topic	Guideline number	*Guideline* • Empirical source	Ref.
Set-up time			
	7.7.1	***Speech communication with fast call set-up may be preferred to video communication with a long set-up time***	
		• Speech communication with fast call set-up is preferred when used for managerial work tasks (compared with speech communication with high-quality and 7 second call set-up)	[i.23] (RR)
		• Speech communication with 3,1 kHz bandwidth and fast call set-up and call forwarding is chosen more when used for managerial communication (compared with speech communication with 7 kHz bandwidth and 7 second call set-up and no call forwarding)	[i.23] (RR)

Urgency			
	7.7.2	***With video communication more is said for the same outcome, compared with speech communication***	
		• Speech communication with a handset is regarded the best way to conduct urgent communication for managerial work tasks (compared with office-based loud speaking speech communication, video communication, multimedia communication and avatar communication)	[i.23] (RR)
Negotiation task			
	7.7.3	***Loud-speaking speech communication may have a higher perceived utility for negotiation tasks than speech with a handset***	
		• Loud speaking speech communication when used for negotiation has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)
	7.7.4	***Speech communication may have a higher perceived utility for negotiation tasks than video communication with 650 ms delay***	
		• Speech communication when used for negotiation has a lower perceived utility (compared with video communication with 650 ms delay)	[i.25] (RR)
	7.7.5	***Video communication may have a higher perceived utility for negotiation tasks than speech communication with a handset***	
		• Video communication with no delay when used for negotiation has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)
	7.7.6	***Video communication may have a higher perceived utility for negotiation tasks than loud-speaking speech communication***	
		• Video communication when used for negotiation has a higher perceived utility (compared with loud speaking speech communication)	[i.25] (RR)
	7.7.7	***Video communication with 650 ms delay may have a higher perceived utility for negotiation tasks than speech communication***	
		• Video communication with 650 ms delay when used for negotiation has a higher perceived utility (compared with speech communication)	[i.25] (RR)
	7.7.8	***Mobile video communication may have a higher perceived utility for negotiation tasks than speech communication***	
		• Video communication with a small (3,5") screen when used for negotiation has a high perceived utility (compared with speech communication)	[i.25] (RR)

…topic continued on next page

Negotiation (continued)	7.7.9	***Sellers may obtain a more-favourable outcome with video communication than with speech communication***	
		• Video communication with high quality when used by people acting as sellers in a sales negotiation game can lead to obtaining a more-favourable outcome (compared with speech communication and people acting as buyers)	[i.56] (RR)
	7.7.10	***Sellers may obtain a more-favourable outcome with video communication than with text communication***	
		• Video communication with high quality when used by people acting as sellers in a sales negotiation game can lead to obtaining a more-favourable outcome (than with Text communication and people acting as buyers)	[i.56] (RR)
	7.7.11	***Negotiating an agreement can be reached in similar times with speech communication, video communication and when face-to-face***	
		• Speech communication with 7 kHz bandwidth when used for negotiation may not lead to a significant difference in time to reach consensus (with efficiency comparable with video communication and when actually face-to-face)	[i.56] (RR)
		• Video communication with high quality when used for negotiation may not lead to a significant difference in time to reach consensus (with efficiency comparable with speech communication and when actually face-to-face)	[i.56] (RR)
Problem solving task			
	7.7.12	***Video communication with lip asynchrony of 200 ms may have a higher perceived utility for problem solving tasks than speech communication***	
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)
		• Video communication with lip asynchrony of 200 ms when used for joint problem solving has a higher perceived utility (compared with loud speaking speech communication)	[i.25] (RR)
	7.7.13	***Audio conferencing may have a higher perceived utility for problem solving tasks than speech telephony***	
		• Loud speaking speech communication when used for joint problem solving has a higher perceived utility (compared with speech communication with a handset)	[i.25] (RR)

Persuasion task			
	7.7.14	***High quality video communication when used for persuasion may be considered suitable by potential users without direct experience***	
		• Video communication with high quality when used for persuasion is considered suitable by potential users without direct experience (unlike Speech communication)	[i.56] (RR)
	7.7.15	***Speech communication for persuasion may not be considered suitable by potential users without direct experience***	
		• Loud speaking speech communication with 7 kHz bandwidth is considered by people without direct experience as unsuitable for persuasion	[i.56] (RR)
	7.7.16	***High quality video communication may offer an advantage over speech communication when arguing a case***	
		• Video communication with high quality when used for persuasion can have an advantage if arguing a case with personal involvement (compared with Speech communication)	[i.56] (RR)
		• Speech communication with 7 kHz bandwidth when used for persuasion can have a disadvantage if arguing a case with personal involvement (compared with Video communication)	[i.56] (RR)
		• Speech communication with 7 kHz bandwidth when used for persuasion when arguing when used for a third party can lead to users being perceived as more formal (compared with Video communication)	[i.56] (RR)
	7.7.17	***Text communication may offer an advantage over speech communication when arguing a case***	
		• Speech communication with 7 kHz bandwidth when used for persuasion can have a disadvantage if arguing a case with personal involvement (compared with Text communication)	[i.56] (RR)

Managerial work			
	7.7.18	***If cost matters, high quality video communication may be chosen over actual face-to-face for straight-forward communication***	
		• Video communication with high quality is chosen more for group managerial work tasks that involve planning and task distribution when cost of meeting matters (compared with actual face-to-face communication)	[i.23] (RR)
	7.7.19	***High quality video communication may be good for enabling group processes and active involvement from the participants***	
		• Video communication with high quality is chosen more for group managerial work tasks that involve group processes and active involvement from the participants (compared with speech communication)	[i.23] (RR)
		• Video communication with high quality is chosen more for group managerial work tasks that involve planning and task distribution (compared with speech communication)	[i.23] (RR)
Using a foreign language			
	7.7.20	***Video communication may improve outcomes when speaking a foreign language, compared with speech communication***	
		• Video communication when used for joint problem solving may improve task outcomes when the users' communication abilities are low (compared with speech only)	[i.58] (CP)
	7.7.21	***Speech communication may reduce outcomes when speaking a foreign language, compared with video communication***	
		• Speech communication when used for joint problem solving may reduce task outcomes when the users' communication abilities are low (compared with video communication)	[i.58] (CP)

Deaf or hearing impairment			
	7.7.22	***Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language may enable people with low written language competence to communicate more than with text communication***	
		• Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language provides children who are deaf with more opportunities to communicate (compared with text communication and SMS)	[i.54] (RR)
		• Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language provides migrant people who are deaf with more opportunities to communicate (compared with text communication and SMS)	[i.54] (RR)
	7.7.23	***Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language can lead to a reduction in the use of text communication***	
		• Mobile video communication with QCIF and 10 fps to 14 fps when used for sign-language by people who are deaf can lead to a reduction in the use of text communication	[i.54] (RR)
Elderly people			
	7.7.24	***Video communication for providing psycho-social and physical health care can be perceived to have higher utility than speech communication***	
		• Video communication when used for providing consultancy care services to elderly persons at home is perceived to have high utility (compared with speech communication)	[i.25] (RR)
		• Speech communication when used for providing consultancy care services to elderly persons at home is perceived to have lower utility (compared with home visits)	[i.25] (RR)
	7.7.25	***Video communication for providing psycho-social consultancy care can be perceived to have higher utility than home visits***	
		• Video communication when used for providing psychological and social consultancy care services to elderly persons at home is perceived to have higher utility (compared with home visits)	[i.25] (RR)

Deaf or hearing impairment		
	7.7.26 ***Text communication may not meet the need for spontaneous and flexible communication***	
	• Text communication does meet the need for day-to-day spontaneous and flexible communication by people who are deaf (compared with mobile video communication for sign-language)	[i.54] (RR)
	7.7.27 ***Text communication may not provide sufficient opportunities for people with low written language competence to communicate, compared with mobile video communication***	
	• Text communication may not provide sufficient opportunities for children who are deaf to communicate (compared with mobile video communication with QCIF and 10 fps to 14 fps for sign-language)	[i.54] (RR)
	• Text communication may not provide migrant people who are deaf with sufficient opportunities to communicate (compared with mobile video communication for sign-language)	[i.54] (RR)

Annex A: Requirements for guidelines and a web-based guideline and tutorial system

This annex provides a summary of the requirements identified for guidelines and a web-based guideline and tutorial system. More detailed information is provided in TR 102 535 [i.18]. The web-based system is developed by ETSI STF 354.

A.1 Requirement derivation process

Identifying the requirements for guidelines in this area involved an iterative procedure. The guidelines of EG 202 534 [i.11] have subsequently been revised and extended based on further study with a larger sample of potential guideline users. Interviews and workshops have been performed with over 300 persons from network operator, equipment manufacturer and service provider organizations.

The precise procedure performed within a particular organization varied primarily due to the time available to the participants. However, the general approach was to:

- Present an overview of the:
  - issues concerning The user experience of real-time communication services;
  - approach being used for guideline development and example guidelines;
  - expected ways guidelines could be used;
  - web-based system.
- Obtain feedback from the participants concerning the:
  - draft guidelines;
  - web-based system;
  - topics for which guidelines should exist.

Results from this activity are summarized below.

A.2 Requirements for guidelines

TR 102 535 [i.18] identified initial requirements for content and format of guidelines. Together with further interviews, workshops and questionnaire surveys with over 150 persons working in network operator, equipment manufacturer and service provider organizations the main requirements for information that could enhance their work. Are described below, along with the main implications of each requirement.

A.2.1 Provide information on key topics of concern that will aid development choices

The issues that concern developers are primarily based on their knowledge of current and foreseeable future technology and market areas. These issues can therefore be considered as topics for which guidelines should be developed. An issue is usually related to a particular communication service with which an individual developer is most concerned.

In addition to requiring user experience data for the optimum design of a particular communication service, data is also required to aid the selection between candidate communication services (e.g. reasons for choosing between audio conferencing and video conferencing). Guidelines that address some of these topics could be derived from existing empirical results available in the literature, or from dedicated user tests and expert panels (described further in clause B.1.4).

Some requested topics could not be addressed by the present document because user-based test data or expert opinion is currently too provisional or lacking.

A.2.2 Provide information on related concepts

All system and service developers deal with continually evolving technology and applications. This requires all persons to apply a certain amount of multidisciplinary knowledge, incorporating for example knowledge of technologies and knowledge of users. Therefore, information on related concepts should be available in order to help guideline users understand and apply the information on key topics.

This implies the development of tutorial information that explains the key concepts to which the guidelines refer.

A.2.3 Provide user experience data that can be used from different perspectives

The application of user experience data will be different between different guideline users. For example, it will depend on their particular role in the development of a system or service. It is possible that particular user experience data can be useful for different developers dealing with apparently different, though related, issues. It is also possible that particular user experience data can be used by the same person differently at different times, depending on a particular project at hand.

The implication is that the development of guidelines from base knowledge of user behaviour should be topic related and allow for different abstractions to different guidelines. Base knowledge should be made accessible in a format that promotes abstraction to concise and applicable conclusions.

A.2.4 Link QoS and user experience variables

Although developers typically appreciate the need for user-based knowledge, their main reference points and decisions usually concern the technical QoS characteristics of a service. Therefore, whilst user experience embodies psychological measures of user behaviour it is should also be expressed in relation to technical QoS. Any guidelines should succeed where possible to combine both user experience and QoS measures to provide an expression of the usage outcome when performing a particular communication task with a particular communication service with known levels of QoS.

The approach used to develop the present document was to derive a database of structured user test results from which more concise guidelines could be abstracted. As summarized in clause 5 of the present document, the structured user test results are constructed according to the clause:

- IF <communication situation>;
- USING <service prescription>;
- WITH <technical parameters>;
- THEN <user experience>.

The technical parameters concern QoS and network performance measures such as network delay and packet loss.

A.2.5 Provide information about user behaviour that is feasible to apply

The <user experience> attribute of an intermediate guideline could include many measures of user behaviour. For example, in the area of real-time person-to-person communication and leisure-oriented services such as Mobile TV and real-time games the traditional usability variables of (communication) effectiveness, efficiency and satisfaction can be supplemented with measures of enjoyment, person perception and social presence.

Depending on the original user tests, all of these measures have the potential to consist of multiple variables. However, as most users of user experience data come from technical and business backgrounds and require concise, understandable advice it may be necessary to produce summary statements of user behaviour that may differ from the original used in a user test. An example would be the derivation of global user experience ratings for a particular service, task type, user groups, etc.

A.3 Requirements for a web-based system

In addition to providing guidelines in the present document, it has been found to be important to develop a web-based system in order to enhance three fundamental aspects:

- **Navigation:** To enhance traversing the information that is in the present document (e.g. with hyperlink properties).
- **Education:** To tutor important aspects of the present document; this is necessary because the Guidelines cover a complex and continually evolving area (e.g. guideline users deal with an increasingly wide technological area) and because the guidelines deliberately combine multidisciplinary knowledge (e.g. from more technical QoS to more psychological user experience perspectives).
- **Dissemination:** To make the content available in an alternative way for a wider potential audience, thereby maximizing the spread of knowledge and good practice.

The main requirements for implementation have been identified as:

- Provide two main facilities:
  - "Find a guideline";
  - "Take a tutorial".
- Enable search for guidelines within:
  - Service areas;
  - Guideline topics.
- With each guideline, offer users the option to also receive:
  - Justification of the guideline;
  - Empirical source reference(s) for the guideline;
  - More detail.
- Provide multi-media tutorials with default and user selection of media combinations (text, audio, graphics, video);
- Provide option to take a tutorial as:
  - Entire lesson;
  - Selected parts (e.g. overview, definition, frequently asked questions).

A web-based system is developed by ETSI STF 354 to meet these requirements.

Annex B:
Sources for the guidelines

A number of methodologies were applied for specifically deriving user-based data:

- laboratory experiments;
- field studies;
- surveys;
- expert review and panels.

The expert reviews and expert panels were set-up in cases where there was previously no known empirical data existing for a topic area and when it was not feasible to perform a user test.

Some base data were derived from existing literature. In order to be considered valid for the current guidelines it was necessary that published results were derived from user tests for which the report provided sufficient information on the technical parameters of the equipment or service used.

B.1 Bespoke user studies

The following user studies were specifically conducted to provide data to input to the guidelines.

B.1.1 Laboratory experiments

Comparison tests between text, speech, avatar and video communication have used a persuasion task (N = 100), a task involving negotiated outcome based on trust (N = 142) and a joint problem-solving task (N = 82) [i.56].

Further laboratory experiments have tested speech, avatar and video communication [i.53]. Seven main experiments were conducted to investigate effects of:

- changes in task type and its impact on performance and attitudes when communication services were varied between speech-telephony and video-telephony (N = 66);
- different tasks on opinion measures such as social presence and person perception (N = 44);
- manipulating discrepancies between audio and video delays (asynchrony) in a problem solving task (N = 48);
- different image sizes in a task involving negotiation to address questions about the use of small screens for mobile videotelephony (N = 48);
- variations in resolution for small screens and their impact on performance, communication and attitudes with a negotiation task (N = 86);
- delay and their impact on performance, communication and attitudes with a negotiation task (N = 42);
- variations in packet loss and delay in a remote inspection task aimed at simulating mobile communication (N = 48).

Theses laboratory experiments have been subject to a Multi-Attribute Utility Technique (MAUT) analysis to derive conclusions on the user-perceived utility of the communication services examined [i.25].

B.1.2 Field studies

Several field studies provided test results for the guidelines:

- Two longitudinal studies of mobile videotelephony applied to the support of blind users from a service centre. The two blind participants undertook a trip for either leisure or business purposes specially developed test equipment with high quality video [i.29].
- A study of five key usage situations for the application of mobile videotelephony to the support of blind users from a service centre. The study involved 10 blind or severely visually impaired persons in addition to the service operator [i.29].
- An investigation of communication service choice in an administrative work setting for five persons in a distributed organization [i.23]. The participants were provided audio conferencing, avatar telephony, videoconferencing and multimedia conferencing for point-to-point communication. All communication services were equally accessible on the participants' desktop and the participants had an established pattern of communication before the field study. In addition, The field study [i.23] has been subject to a Multi-attribute Utility Technique (MAUT) analysis to derive conclusions on the user-perceived utility of the communication services examined [i.25].
- A study by the EC IST Eye-2-Eye project in collaboration with the EC IST project IST@Home that piloted how a service content provider could assess the utility of a service for elderly persons provided with videoconferencing in addition to speech telephony at home [i.25].

B.1.3 Survey

Participants (N = 53) from different demographic groups in Norway were asked to rate their preferences for different communication services and face-to-face communication for each of a set of communicative scenarios, after having acquired hands-on experience with the services [i.48]. The communication services were implemented as optimal quality and involved text communication, speech communication with a handset, avatar communication and video communication. The demographic groups were:

- Youths: 6 males and 6 females, 14 years to 15 years of age.
- Young adults: 8 males and 9 females, aged 19 to 31 (mean 23).
- Senior citizens: 9 males and 3 females, aged 68 to 85 (mean 77).
- Business professionals: 9 males and 3 females, aged 27 to 55 (mean 39).

The different services were demonstrated by having the participants solve simple exercises through using video telephone, hands-free telephone, and text chat. The participants did not use the avatar telephone themselves. Instead they watched a demonstrational video of two persons interacting by the use of this technology.

B.1.4 Expert review and expert panels

Expert reviews involve a single expert making a judgement that could be used for a guideline. Expert panels involve two or more experts working interactively to derive a judgement.

Guidelines were derived from the following use of experts:

- Expert panels of two or more members were specially convened by ETSI STF 284 and drew on the following specialities:
  - real-time services on current and future telecommunication networks;
  - telecommunication standardization and QoS;
  - development of fixed, mobile and multiparty conferencing;
  - hosting (service point-of-contact) point-to-point videoconferencing services;

- principle operator of a videoconferencing service multi-party bridge;
- psychology of real-time person-person communication.

- An experienced service centre professional for the support of blind persons at distance by videotelephony was employed to assess the quality of mobile videotelephony for reading text and finding an object when moving. For these two situations video quality was reduced in terms of screen resolution (CIF, QCIF, SQCIF) and frame rate (2 fps to 3 fps, 5 fps to 6 fps, 10 fps to 15 fps, 20 fps to 25 fps) until the tasks could not be performed with the same speed and accuracy [i.29].
- Members of the EC project IST-999-11577 Eye-2-Eye [i.6] formed a multidisciplinary expert panel to derive guidelines on the issue of eye contact between videoconferencing participants. The panel applied knowledge from human factors research literature in addition to experience from equipment manufacturer and videoconferencing service provider perspectives.

B.2 Existing literature

Known published results are included only if they address a topic that was identified as important for the intended guideline users. Therefore, the development of guidelines from literature is not exhaustive.

Each guideline contains a link to its source reference(s) listed in clause 2.2.

History

Document history		
V1.1.1	December 2009	Membership Approval Procedure MV 20100219: 2009-12-21 to 2010-02-19