الاتحـــــــــــاد الدولـــــــــــي للاتصـــــــــــالات

ITU-T

قطاع تقييس الاتصالات
في الاتحاد الدولي للاتصالات

P.800.2
(2016/07)

السلسلة P: المطاريف وطرائق التقييم الشخصي والموضوعي

أساليب للتقييم الموضوعي والشخصي لجودة الكلام والفيديو

تفسير وإبلاغ متوسط درجات الرأي

التوصية ITU-T P.800.2

الاتحاد الدولي للاتصالات

توصيات السلسلة P الصادرة عن قطاع تقييس الاتصالات

المطاريف وطرائق التقييم الشخصي والموضوعي

مفردات وتأثيرات معلمات الإرسال على رأي الزبائن في جودة الإرسال	السلسلة	P.10
خصائص مطراف الصوت	السلسلة	P.30
		P.300
الأنظمة المرجعية	السلسلة	P.40
أجهزة القياسات الموضوعية	السلسلة	P.50
		P.500
قياسات كهرسمعية موضوعية	السلسلة	P.60
قياسات متصلة بالجهارة الصوتية	السلسلة	P.70
أساليب للتقييم الموضوعي والشخصي لجودة الكلام	السلسلة	P.80
أساليب للتقييم الموضوعي والشخصي لجودة الكلام والفيديو	**السلسلة**	**P.800**
نوعية الأداء السمعي البصري في الخدمات متعددة الوسائط	السلسلة	P.900
أداء الإرسال وجوانب نوعية الخدمة في النقاط الانتهائية للشبكات القائمة على بروتوكول الإنترنت	السلسلة	P.1000
الاتصالات في المركبات	السلسلة	P.1100
نماذج وأدوات لتقييم جودة وسائط البث المتدفق	السلسلة	P.1200
الاجتماعات التي تعقد عن بُعد	السلسلة	P.1300
التحليل الإحصائي لقياسات الجودة والمبادئ التوجيهية لتقييمها وإعداد التقارير عنها	السلسلة	P.1400
أساليب للتقييم الموضوعي والشخصي لجودة خدمات غير خدمات الكلام والفيديو	السلسلة	P.1500

لمزيد من التفاصيل يرجى الرجوع إلى قائمة التوصيات الصادرة عن قطاع تقييس الاتصالات.

التوصية ITU-T P.800.2

تفسير وإبلاغ متوسط درجات الرأي

ملخص

تقدم التوصية ITU-T P.800.2 بعض الأنواع الأكثر شيوعاً من متوسط درجات الرأي (MOS)، وتصف الحد الأدنى من المعلومات التي يجب أن تصاحب قيم MOS كي يتسنى تفسيرها بشكلٍ صحيح.

التسلسل التاريخي

الطبعة	التوصية	تاريخ الموافقة	لجنة الدراسات	معرف الهوية الفريد*
1.0	ITU-T P.800.2	2013-05-14	12	11.1002/1000/11934
2.0	ITU-T P.800.2	2016-07-29	12	11.1002/1000/12973

مصطلحات أساسية

تصنيف الفئة المطلق، ACR، متوسط درجات الرأي، MOS، نموذج موضوعي، إبلاغ، تجربة شخصية

* للنفاذ إلى توصية، ترجى كتابة العنوان http://handle.itu.int/ في حقل العنوان في متصفح الويب لديكم، متبوعاً بمعرف التوصية الفريد. ومثال ذلك، http://handle.itu.int/11.1002/1000/11830-en.

تمهيد

الاتحاد الدولي للاتصالات وكالة متخصصة للأمم المتحدة في ميدان الاتصالات وتكنولوجيات المعلومات والاتصالات (ICT). وقطاع تقييس الاتصالات (ITU-T) هو هيئة دائمة في الاتحاد الدولي للاتصالات. وهو مسؤول عن دراسة المسائل التقنية والمسائل المتعلقة بالتشغيل والتعريفة، وإصدار التوصيات بشأنها بغرض تقييس الاتصالات على الصعيد العالمي.

وتحدد الجمعية العالمية لتقييس الاتصالات (WTSA) التي تجتمع مرة كل أربع سنوات المواضيع التي يجب أن تدرسها لجان الدراسات التابعة لقطاع تقييس الاتصالات وأن تُصدر توصيات بشأنها.

وتتم الموافقة على هذه التوصيات وفقاً للإجراء الموضح في القرار 1 الصادر عن الجمعية العالمية لتقييس الاتصالات.

وفي بعض مجالات تكنولوجيا المعلومات التي تقع ضمن اختصاص قطاع تقييس الاتصالات، تُعد المعايير اللازمة على أساس التعاون مع المنظمة الدولية للتوحيد القياسي (ISO) واللجنة الكهرتقنية الدولية (IEC).

ملاحظة

تستخدم كلمة "الإدارة" في هذه التوصية لتدل بصورة موجزة سواء على إدارة اتصالات أو على وكالة تشغيل معترف بها.

والتقيد بهذه التوصية اختياري. غير أنها قد تضم بعض الأحكام الإلزامية (بهدف تأمين قابلية التشغيل البيني والتطبيق مثلاً). ويعتبر التقيّد بهذه التوصية حاصلاً عندما يتم التقيّد بجميع هذه الأحكام الإلزامية. ويستخدم فعل "يجب" وصيغ ملزمة أخرى مثل فعل "ينبغي" وصيغها النافية للتعبير عن متطلبات معينة، ولا يعني استعمال هذه الصيغ أن التقيّد بهذه التوصية إلزامي.

حقوق الملكية الفكرية

يسترعي الاتحاد الانتباه إلى أن تطبيق هذه التوصية أو تنفيذها قد يستلزم استعمال حق من حقوق الملكية الفكرية. ولا يتخذ الاتحاد أي موقف من القرائن المتعلقة بحقوق الملكية الفكرية أو صلاحيتها أو نطاق تطبيقها سواء طالب بها عضو من أعضاء الاتحاد أو طرف آخر لا تشمله عملية إعداد التوصيات.

وعند الموافقة على هذه التوصية، لم يكن الاتحاد قد تلقى إخطاراً بملكية فكرية تحميها براءات الاختراع يمكن المطالبة بها لتنفيذ هذه التوصية. ومع ذلك، ونظراً إلى أن هذه المعلومات قد لا تكون هي الأحدث، يوصى المسؤولون عن تنفيذ هذه التوصية بالاطلاع على قاعدة البيانات الخاصة ببراءات الاختراع في مكتب تقييس الاتصالات (TSB) في الموقع http://www.itu.int/ITU-T/ipr/.

© ITU 2019

جميع الحقوق محفوظة. لا يجوز استنساخ أي جزء من هذه المنشورة بأي وسيلة كانت إلا بإذن خطي مسبق من الاتحاد الدولي للاتصالات.

جدول المحتويات

الصفحة

1 مجال التطبيق 1

2 المراجع 1

3 التعاريف 1

1.3 مصطلحات معرَّفة في مكان آخر 1

2.3 مصطلحات معرَّفة في هذه التوصية 1

4 المختصرات والمختزلات 2

5 الاصطلاحات 2

6 معلومات تمهيدية 2

7 قيم متوسط درجة الرأي الشخصي 3

8 تفسير قيم متوسط درجة الرأي 4

9 اعتبارات الفيديو 5

10 التحليل الإحصائي لمتوسط درجة الرأي 5

11 قيم متوسط درجة الرأي الموضوعي 6

12 الإبلاغ عن قيم متوسط درجة الرأي الشخصي 6

13 الإبلاغ عن قيم متوسط درجة الرأي الموضوعي 8

14 الترميز 8

بيبليوغرافيا 9

تفسير وإبلاغ متوسط درجات الرأي

1 مجال التطبيق

تقدم هذه التوصية بعض الأنواع الأكثر شيوعاً من متوسط درجات الرأي (MOS)، وتصف الحد الأدنى من المعلومات التي يجب أن تصاحب قيم MOS كي يتسنى تفسيرها بشكل صحيح.

ويجدر بالذكر أن هذا النص لا يهدف إلى تقديم دليل قطعي للاختبار الشخصي أو الموضوعي. وتقدم مراجع البيبليوغرافيا في ذيل هذه التوصية معلومات عن مواد ذات تفاصيل أوفى.

2 المراجع

تتضمن التوصيات التالية لقطاع تقييس الاتصالات وغيرها من المراجع أحكاماً تشكل من خلال الإشارة إليها في هذا النص جزءاً لا يتجزأ من هذه التوصية. وقد كانت جميع الطبعات المذكورة سارية الصلاحية في وقت النشر. ولما كانت جميع التوصيات والمراجع الأخرى تخضع إلى المراجعة، يُرجى من جميع المستعملين لهذه التوصية السعي إلى تطبيق أحدث طبعة للتوصيات والمراجع الأخرى الواردة أدناه. وتُنشر بانتظام قائمة توصيات قطاع تقييس الاتصالات السارية الصلاحية. والإشارة إلى وثيقة ما في هذه التوصية لا يضفي على الوثيقة في حد ذاتها صفة التوصية.

[ITU-T P.800.1] التوصية ITU-T P.800.1 (2006)، *مصطلح حساب متوسط درجات الرأي (MOS).*

3 التعاريف

1.3 مصطلحات معرَّفة في مكان آخر

لا توجد.

2.3 مصطلحات معرَّفة في هذه التوصية

تعرف هذه التوصية المصطلحات التالية:

1.2.3 الظرف (condition): هو واحدة من مجموعة حالات الاستخدام الجاري تقييمها في تجربة شخصية؛ وكثيراً ما يُشار إليه كدارة مرجعية افتراضية (HRC) في تجارب الفيديو.

2.2.3 الظرف الفرعي (sub-condition): هو مجموعة فرعية من ظرف معرف بخاصية محددة لحالة الاستخدام، من قبيل مادة كلام متحدث معين.

3.2.3 مساهم (subject): هو مشارك في تجربة معنية.

4.2.3 صوت (vote): هو رد المساهم على سؤال في مقياس تصنيف لعينة اختبار فردية أو تفاعل.

4 المختصرات والمختزلات

تَستعمل هذه التوصية المختصرات والمختزلات التالية:

ACR تصنيف الفئة المطلق *(Absolute Category Rating)*

DCR تصنيف فئة التردي *(Degradation Category Rating)*

DMOS متوسط درجة الرأي في التردي *(Degradation Mean Opinion Score)*

HRC دارة مرجعية افتراضية *(Hypothetical Reference Circuit)*

MOS متوسط درجة الرأي *(Mean Opinion Score)*

MUSHRA اختبار متعدد المنبهات بمرجع ومرتكز خافيين *(Multi-stimulus test with Hidden Reference and Anchor)*

QCIF ربع النسق المتوسط المشترك *(Quarter Common Intermediate Format)*

SSCQE تقييم الجودة المستمرة لمنبه واحد *(Single Stimulus Continuous Quality Evaluation)*

VGA صفيف رسومات فيديو *(Video Graphics Array)*

5 الاصطلاحات

لا توجد.

6 معلومات تمهيدية

تعبر الجودة السمعية والفيديوية عن كميات شخصية بطبيعتها. وهذا يعني أن أساس مقارنة الجودة السمعية والفيديوية هو رأي المستخدم. بيد أن رأي شخص ما فيما هو "جيد" قد يختلف تماماً عن رأي شخص آخر – وما من أحد منهما على صواب وما من أحد منهما على خطأ.

وقبل نشر تكنولوجيا إرسال سمعي أو فيديوي جديدة، من حسن الممارسات أن تقيَّم جودة الإرسال باستخدام تجربة شخصية واحدة أو أكثر. والغرض من التجربة الشخصية هو جمع آراء العديد من الأشخاص ("المساهمين") بشأن أداء النظام في عدد من حالات الاستخدام ("الظروف")[1] المحددة جيداً. أما متوسط درجة الرأي (MOS) في ظرف معين فهو ببساطة متوسط الآراء ("الأصوات") التي جُمعت بشأن حالة الاستخدام تلك.

وتهدف خوارزميات قياس الجودة الموضوعية إلى التنبؤ بقيمة متوسط درجة الرأي التي ستنتجها إشارة دخل معينة في تجربة شخصية. وبالتالي، عند تفسير قيمة متوسط درجة الرأي المشتقة بشكلٍ موضوعي، تقتضي الضرورة فهم التصميم الأساسي للتجربة التي يُنبأ بواسطتها.

وهناك عدة أنواع مختلفة من قيمة متوسط درجة الرأي والعديد من منهجيات الاختبار المختلفة لإنتاجها. والغرض من هذه التوصية هو إعطاء القارئ تقديراً للنقاط الرئيسية التي تتعين مراعاتها عند تفسير قيم متوسط درجة الرأي والحد الأدنى من المعلومات التي ينبغي أن تصاحب قيم متوسط درجة الرأي عند الإبلاغ عنها.

---

1 في تجارب الفيديو، كثيراً ما يُشار إلى الظروف باسم الدارات المرجعية الافتراضية (HRC).

7 قيم متوسط درجة الرأي الشخصي

أنواع متوسط درجة الرأي

يشيع اعتقاد خاطئ بأن قيم متوسط درجة الرأي تتعلق بالخدمات الصوتية حصراً، ولكن عملية سؤال الأشخاص عن تقييمهم للجودة يمكن تطبيقها على الخدمات الفيديوية والسمعية العامة بنفس سهولة تطبيقها على الخدمات الصوتية. ويمكن أيضاً الطلب من الأشخاص تقييم الجودة السمعية المرئية الإجمالية للخدمة. وقد أصدر الاتحاد معايير مختلفة تصف الجوانب المختلفة للاختبار الشخصي للتطبيقات الفيديوية والسمعية العامة بالإضافة إلى التطبيقات الصوتية، وهي مدرجة في البيبليوغرافيا.

ويمكن تقسيم التجارب الشخصية إلى نوعين: منفعلة وتفاعلية. ففي تجربة شخصية منفعلة، تُعرض على المساهمين عينات اختبار مسجلة مسبقاً تمثل ظروفاً تسترعي الاهتمام. ويُطلب من المساهمين الاستماع إلى و/أو مشاهدة مواد الاختبار بشكل منفعل وإبداء رأيهم باستخدام مقياس التصنيف المقدم. وفي تجربة تفاعلية، ينشط مساهمان أو أكثر في التحادث باستخدام معدات مصممة لمحاكاة حالات استخدام تسترعي الاهتمام. وكثيراً ما يكلَّف المساهمون بمهام لتحفيز المحادثة والتفاعل. ويغلب على التجارب أن تكون منفعلة في طبيعتها. ولكن هناك بعض الجوانب التي يعايشها المستخدم التي لا تظهر إلا في سيناريوهات المحادثة، كمؤثرات التأخير والصدى على سبيل المثال.

منهجية الاختبار ومقياس التصنيف

يُطلب من المساهمين في تجربة شخصية، إبداء آرائهم باستخدام "مقياس التصنيف". والغرض من المقياس هو أن يترجم المساهم تقييم الجودة إلى قيمة عددية يمكن حساب متوسطها بين المساهمين والعوامل التجريبية الأخرى.

ويشيع استخدام عدة مقاييس تصنيف، وتقع الفوائد النسبية لمختلف المقاييس خارج مجال تطبيق هذه التوصية. والمقياس الأكثر استخداماً هو مقياس تصنيف الفئة المطلق (ACR) المكون من 5 نقاط:

ممتاز	5
جيد	4
وسط	3
ضعيف	2
سيء	1

ومقياس تصنيف الفئة المطلق هو مقياس متمايز، مما يعني أن رد المساهم يقتصر على واحدة من القيم الخمس المذكورة أعلاه. غير أن عملية حساب المتوسط المستخدمة لدمج النتائج من مساهمين مختلفين تعني أن قيم متوسط درجة الرأي لا تقتصر على قيم عدد صحيح. وتحتوي بعض مقاييس التصنيف على أكثر من خمس تسميات منفصلة، بينما يسمح البعض الآخر للمساهم بتقديم ردود وسيطة عند نقاط بين تسميتين.

والجزء "المطلق" من تصنيف الفئة المطلق يتعلق بحقيقة أن المساهمين يطلب منهم تقييم كل عينة بشكل مستقل. وتسأل بعض مقاييس التصنيف، مثل مقياس تصنيف فئة التردي (DCR)، عن رأي المساهم في الفرق بين العينة المعالَجة عبر ظرف يسترعي الاهتمام وبين النسخة غير المعالجة من نفس العينة. وكثيراً ما تسمى قيمة متوسط درجة الرأي الناتجة في مثل هذه التجربة متوسط درجة الرأي في التردي أو DMOS.

وفي معظم التصاميم التجريبية، يُطلب من الأشخاص تقييم جودة عينات سمعية أو فيديوية قصيرة. وتتراوح مدة هذه العينات عادةً بين 6 و10 ثوانٍ، حيث يتيح ذلك وقتاً كافياً للمساهم كي يكون رأياً لا يشوبه أي تحيز قرب نهاية العينة. ويصعب على عينة واحدة بهذه المدة أن تمثل ظرفاً كاملاً، وبالتالي يُطلب من المساهمين عادة تقييم عينات اختبار متعددة مستمدة من حالة الاستخدام نفسها. فعلى سبيل المثال، يمكن، في تجربة صوتية، تمثيل كل ظرف شبكة قيد الاختبار بعينات كلامية من ثلاثة متحدثين وثلاث

متحدثات. وهذا يعني إمكانية إنتاج قيم متوسط درجة الرأي للظرف بأكمله، بحساب المتوسط عبر كل من المساهمين والمتحدثين، أو لظرف فرعي، مثل متحدث معين أو جنس المتكلم.

وتستخدم أساليب اختبار، مثل تقييم الجودة المستمرة لمنبه واحد (SSCQE)، عينات اختبار أطول من ذلك بكثير، وتتطلب من المساهم تحديث رأيه في الجودة باستمرار أثناء استعراض عينة الاختبار. وينتج عن ذلك تسلسل زمني لتصنيفات الجودة من كل مساهم، بدلاً من قيمة رأي واحدة.

وتتطلب بعض منهجيات الاختبار من المساهم الرد على أسئلة متعددة. ولا ينتج عن ذلك مزيد من المعلومات عن الظروف قيد الاختبار فحسب، بل يمكن أن يكون جزءاً ضرورياً من تصميم الاختبار. فعلى سبيل المثال، يتطلب أسلوب الاختبار في التوصية ITU-T P.835 أن يُبدي المساهم آراء منفصلة في جودة الكلام ونوعية الضوضاء بالعينة قبل تقديم درجة جودة إجمالية. وتبين أن هذه العملية تحقق نتائج أكثر استقراراً في أنظمة كبت الضوضاء من أسلوب اختبار تصنيف الفئة المطلق ذي السؤال الواحد.

ويجدر بالذكر أن بعض الأسئلة قد لا ترتبط بالجودة مباشرة، ولكنها قد تتناول جانباً مختلفاً من الاتصالات، فعلى سبيل المثال، تعرِّف التوصية [b-ITU-T P.800] مقياس جهد الاستماع للتجارب الصوتية. وبالمثل، تسأل بعض تجارب المحادثة المساهم عن تجربته عند التحدث، بدلاً منها عند الاستماع.

8 تفسير قيم متوسط درجة الرأي

تركز المناقشة التالية في البداية على قيم متوسط درجة الرأي الصوتية؛ إلا أن العديد من النقاط الواردة في الأقسام الفرعية تنطبق بالقدر نفسه على قيم متوسط درجة الرأي الفيديوية والسمعية والسمعية الفيديوية. ويرد توضيح الاختلافات الرئيسية بشأن الفيديو في الفقرة التالية.

إن الفكرة القائلة بأن لكودك صوتي معين متوسط درجة رأي معين هي فكرة خاطئة أخرى شائعة. ويعود أحد أسباب هذا الاعتقاد الخاطئ إلى الاستخدام الواسع لنماذج تقييم الجودة الموضوعي التي تنتج نتائج قابلة جداً للتكرار. فهذه النماذج صُممت للتنبؤ بمخرجات التجارب الشخصية أو لتقديرها؛ سوى أن قيمة متوسط درجة الرأي المحصَّلة في تجربة شخصية يمكن أن تختلف جوهرياً من تجربة إلى أخرى لأي كودك معين بمعدل بتات معين. وتتعدد أسباب ذلك.

أولاً، يمكن أن تتأثر قيم متوسط درجة الرأي الدقيقة المحصَّلة في ظرف معين في تجربة شخصية بعدد كبير من العوامل، بما فيها على سبيل المثال لا الحصر:

- التعليمات المقدَّمة للمساهمين وصياغة مقياس الرأي؛
- المعدات المستخدمة لعرض المواد (سماعة الهاتف، سماعات الرأس، مكبرات الصوت)؛
- عرض عبر أذن واحدة أو الأذنين معاً أو عرض استريو؛
- مستوى العرض؛
- البيئة السماعية؛
- إعداد المساهمين؛
- البيانات الوصفية للمساهم، مثل العمر والإلمام بالتكنولوجيا؛
- الاختلافات في تفسير واستخدام مقاييس التصنيف عبر الثقافات؛
- مواد الكلام (المحتوى النطقي وخصائص المتكلم)؛
- اللغة (وجود/غياب الأصوات وطبقاتها الخاصة، وانتشارها وأهميتها).

ثانياً، تعتمد قيمة متوسط درجة الرأي الدقيقة المحصَّلة لأي ظرف معين في تجربة شخصية على جودة الظروف الأخرى في التجربة. فعلى سبيل المثال، يمكن أن يسجل ظرف كودك الصوت ITU-T G.729 أكثر من 3,9 في تجربة تصنيف الفئة المطلق إذا كانت

معظم الظروف الأخرى بجودة أسوأ من تلك المحددة بالتوصية ITU-T G.729؛ وعلى العكس من ذلك، قد يسجل ظرف ITU-T G.729 درجة أقل بكثير من 3,9 إذا كانت معظم الظروف الأخرى تتمتع بجودة أفضل.

ثالثاً، في حال تشغيل تجربة باستخدام كودكات تعمل على عروض نطاق سمعية مختلفة، فإن وجود ظروف عرض نطاق أعلى ستقلل من متوسط درجة الرأي الناتج لظروف ذات عرض نطاق سمعي أقل. وكثيراً ما يطلق على أعلى عرض نطاق سمعي موجود في تجربة صوتية "سياق" التجربة. فعلى سبيل المثال، كثيراً ما يؤدي ظرف كودك الصوت ITU-T G.711 إلى الحصول على درجة أعلى من 4.0 في تجربة تصنيف الفئة المطلق ضيقة النطاق (300-3700 Hz)؛ في حين يرجح أن يحصل على درجة تتراوح بين 3,5 و3,7 في تجربة تصنيف الفئة المطلق واسعة النطاق (50-7000 Hz)، بسبب وجود عينات ذات جودة أعلى في النطاق الواسع.

وتبين هاتان النقطتان الأخيرتان حقيقة أن المساهمين في التجارب يميلون إلى تكييف استخدامهم لمقياس التصنيف مع محتوى التجربة. وفي الواقع، تتضمن التجارب المصممة جيداً فترة تدريب في بداية التجربة عندما يسمع المساهمون أمثلة لمجموعة من الظروف، بما فيها الأفضل والأسوأ.

ويتمثل أحد أهم مدلولات الاعتبارات الموضحة أعلاه في عدم جدوى مقارنة قيم متوسط درجة الرأي الناتجة مباشرة من تجارب منفصلة، ما لم تكن تلك التجارب مصممة بشكل صريح للمقارنة، وينبغي، حتى عندئذ، تحليل البيانات إحصائياً لضمان صلاحية هذه المقارنة.

9 اعتبارات الفيديو

إن العديد من الاعتبارات الموضحة أعلاه فيما يتعلق بالتجارب الشخصية الصوتية تنطبق أيضاً على تجارب الفيديو. والظروف التجريبية، التي كثيراً ما تسمى دارات مرجعية افتراضية (HRC)، تحدد نمطياً توليفات متنوعة من الكودك الفيديوي ومعدل البتات ومعدل الأطر وظروف الإرسال. وتشمل العوامل التي تؤثر على قيم متوسط درجة الرأي الدقيقة المحصَّلة لظرف معين، ما يلي، على سبيل المثال لا الحصر:

- المعدات المستخدمة لعرض المواد (تكنولوجيا العرض، معدل الإنعاش، التباين، وما إلى ذلك)؛
- بيئة المشاهدة (اللون ودرجة الحرارة ومستوى الإضاءة)؛
- مسافة المشاهدة (عادةً ما يعبَّر عنها كنسبة مسافة المشاهدة إلى ارتفاع الشاشة)؛
- محتوى الفيديو.

ولهذه النقطة الأخيرة أهمية خاصة في تجارب الفيديو. إذ يعد اختيار مادة الاختبار عاملاً أقوى بكثير في تجارب الفيديو مقارنة بالتجارب الصوتية. وذلك لأن محتوى تسلسل الفيديو يمكن أن يؤثر تأثيراً كبيراً على مدى الكفاءة التي يمكن بها تشفيره. فعلى سبيل المثال، يكون محتوى المعلومات في تسلسل رياضي سريع الحركة أعلى بكثير من تتابع مؤتمرات الفيديو الذي يُظهر الرأس والكتفين.

وفي تجارب الفيديو، يتحدد السياق الأساسي باستبانة صورة الفيديو. وبشكلٍ عام، لا تخلط التجارب الشخصية بين استبانات مختلفة، وبالتالي فإن قيم متوسط درجة الرأي للفيديو تتعلق باستبانة معينة، من قبيل 480p أو 1080p. وفي الحالات التي تُخلط فيها الاستبانات، يتحدد سياق التجربة بالاستبانة التي تحتوي على أكبر عدد من الخطوط. وفي هذه الحالة، من المهم ملاحظة ما إذا كانت الاستبانات الأصغر تُعرض كما هي أصلاً أم يغيَّر مقاسها إلى الاستبانة الأكبر في التجربة.

10 التحليل الإحصائي لمتوسط درجة الرأي

يقع التحليل الإحصائي لقيم متوسط درجة الرأي الشخصي خارج مجال تطبيق هذه التوصية. ولكن ينبغي أن تُرفَق قيم متوسط درجة الرأي بمعلومات كافية للسماح بإجراء تحليل إحصائي أساسي، من قبيل حساب فترة الثقة لكل ظرف. ولأي ظرف أو ظرف فرعي معين، تشتمل هذه المعلومات على عدد الأصوات ومتوسط الأصوات والانحراف المعياري للأصوات.

11 قيم متوسط درجة الرأي الموضوعي

يتمثل الغرض من نموذج الجودة الموضوعي في التنبؤ بقيمة متوسط درجة الرأي التي ستحصل عليها إشارة سمعية أو فيديوية في تجربة شخصية. وكما نوقش أعلاه، فإن قيمة متوسط درجة الرأي الدقيقة الناتجة في أي تجربة معينة لكودك معين أو سلسلة إرسال معينة تعتمد على العديد من الجوانب المختلفة لتصميم التجربة وتنفيذها. لذلك يتعين على مصممي النماذج الموضوعية التنبؤ بتجربة مثالية. تجرى هذه التجربة نمطياً وفق منهجية اختبار محددة، وعادةً ما تكون منهجية تصنيف الفئة المطلق، وتتضمن عينة متوازنة من التشوهات التي ستصادّف في مجال التطبيق الذي يسترعي الاهتمام.

فعلى سبيل المثال، يأخذ التقابل المحدد في التوصية [b-ITU-T P.862.1] المخرجات الخام من النموذج الموضوعي للتوصية ITU-T P.862 ويحدد ما يقابلها في مدى تحدد بحساب متوسط مخرجات عدد كبير من التجارب الشخصية التي أجريت وفقاً لأسلوب تصنيف الفئة المطلق على النحو الموضح في التوصية [b-ITU-T P.800]. ويضمَّن تقابل مماثل في مرحلة الخرج من نموذج التوصية ITU-T P.863.

وتتمثل إحدى مزايا النموذج الموضوعي في كون النتائج قابلة للتكرار، وبالتالي تمكن المقارنة المباشرة لقياسات أجريت في أوقات ومواقع مختلفة. ولكن ينبغي الاستمرار في توخي الحذر لأن عوامل، مثل اختيار مادة الاختبار وأي معالجة سابقة أو لاحقة، تظل قادرة على إدخال التحيز إلى النتائج.

ولأسباب ينبغي أن تكون واضحة الآن، يمكن أن تنتج نماذج موضوعية مختلفة قيماً متوقعة مختلفة لمتوسط درجة الرأي في الظروف نفسها. فعلى سبيل المثال، لا ينتج نموذجا ITU-T P.862.1 وITU-T P.863 نفس قيم متوسط درجة الرأي المتوقعة تماماً للكلام المشفر وفق التوصية ITU-T G.729، على الرغم من أن هذا الكودك يقع في مجال تطبيق كلا النموذجين. ويعود ذلك جزئياً إلى أن النموذجين قد دُربا وأعدا للأداء الأمثل باستخدام تجارب شخصية مختلفة. ولهذا السبب، عند مقارنة تنبؤات متوسط درجة الرأي الموضوعية ذات العتبات، مثلاً لمراقبة اتفاق مستوى الخدمة أو إطلاق إنذار، ينبغي اختيار هذه العتبات في سياق النموذج الذي ينتج التنبؤ.

12 الإبلاغ عن قيم متوسط درجة الرأي الشخصي

يصف الجدول 1 المعلومات التي يجب تقديمها عند الإبلاغ عن قيم متوسط درجة الرأي الشخصي، وماهية المعلومات الإضافية الموصى بتقديمها.

وإذا أجريت تجربة وفقاً لتوصية الاتحاد الدولي للاتصالات، يمكن عادةً تمثيل المعلومات المتعلقة بالمنهجية بإحالة بسيطة إلى المعيار ذي الصلة والأسلوب المعين المستخدم، علماً بأن الاختلافات عن الإجراءات المعيارية ينبغي ذكرها.

ومن المهم دائماً تقديم معلومات عن عينات الاختبار المستخدمة في التجارب المنفعلة. وفي حالة عينات الفيديو، لعل من المفيد تقديم معلومات أكثر تفصيلاً، من قبيل ما إذا كانت تتابعات معينة تحتوي على مسح بانورامي أو تغييرات في المشهد.

الجدول 1 – الحد الأدنى من المعلومات للإبلاغ عن قيم متوسط درجة الرأي الشخصي

المعلومات	نوع التجربة	الحكم
المنهجية منفعلة أو تفاعلية تقييم قائم على العينة أو مستمر تقييم مطلق أو نسبي للعينات التعليمات والسؤال (الأسئلة) المقدمة إلى المساهمين تسميات مقياس التصنيف هل مقياس التصنيف متمايز أم متواصل مدة العينة أو توصية الاتحاد الدولي للاتصالات والأسلوب المستخدم	كل الأنواع	إلزامي

الجدول 1 – الحد الأدنى من المعلومات للإبلاغ عن قيم متوسط درجة الرأي الشخصي

المعلومات	نوع التجربة	الحكم
خطة الاختبار الغرض من التجربة تاريخ ومكان إجراء الاختبار معالجة المعلومات التصميم التجريبي، من قبيل تصميم كتل التجربة عدد الجلسات ومدتها عدد المساهمين البيانات الوصفية للمساهمين وتوزع أعمارهم والتوزع بين الجنسين نوع المساهمين المستخدمين، من قبيل أغرار أم خبراء معلومات عن المعدات المستخدمة بيئة العرض، أي ضوضاء الخلفية، ومستوى الإضاءة، وما إلى ذلك	كل الأنواع	موصى به
معلومات عن الظرف/الدارة المرجعية الافتراضية (HRC) عدد الظروف قائمة الظروف متوسط الأصوات في كل ظرف (متوسط درجة الرأي) الانحراف المعياري للأصوات في كل ظرف عدد الأصوات في كل ظرف	كل الأنواع	إلزامي
معلومات عن الظرف الفرعي قائمة عوامل الظرف الفرعي قيم متوسط درجة الرأي في الظروف الفرعية عدد الأصوات والتغاير في كل ظرف فرعي	كل الأنواع	اختياري
عرض سمعي عرض (عروض) النطاق السمعي القنوات السمعية، من قبيل أحادية، أو ستيريو، وما إلى ذلك مستوى العرض السمعي أسلوب العرض السمعي من قبيل مكبرات الصوت، أو سماعات الرأس (عبر أذن واحدة أو الأذنين معاً)، وما إلى ذلك)	صوتية، سمعية، سمعية مرئية	إلزامي
عرض فيديوي استبانة (استبانات) الصورة الفيديوية (الملاحظتان 1 و2) مسافة المشاهدة حسب العلو، من قبيل 3 أمثال العلو نوع ومقاس الجهاز، مثل تلفزيون أو حاسوب لوحي أو هاتف، وما إلى ذلك. التطبيق، أي المهاتفة الفيديوية، أو الفيديو حسب الطلب، أو التلفزيون الخطي، وما إلى ذلك.	فيديوية، سمعية مرئية	إلزامي
اللغة	صوتية منفعلة، سمعية مرئية	إلزامي
عدد وجنس المتحدثين	صوتية منفعلة، سمعية مرئية	إلزامي
نوع المادة الفيديوية، من قبيل، رياضية أو رأس وجذع	فيديوية منفعلة، سمعية مرئية	إلزامي
نوع المادة السمعية، من قبيل، الموسيقى الكلاسيكية، أو الموسيقى الشعبية، أو الموسيقى التصويرية للأفلام	سمعية منفعلة، سمعية مرئية	إلزامي
الملاحظة 1 – يجب ذكر استخدام الصور المشذرة. **الملاحظة 2** – إذا كانت التجربة تحتوي على عدة استبانات للصور، فيجب تقديم معلومات بشأن ما إذا كانت استبانات الصور الأصغر قد عُرضت كما هي أصلاً أم رُفعت إلى الاستبانة الأكبر.		

13 الإبلاغ عن قيم متوسط درجة الرأي الموضوعي

عند الإبلاغ عن قيمة متوسط درجة الرأي التي ينتجها نموذج موضوعي لقطاع تقييس الاتصالات، يكتفى عموماً بالإبلاغ عن النموذج المستخدم وأي إعدادات غير مبدئية. وفي النماذج غير المقيَّسة، يجب تقديم المعلومات الموجودة في صف "المنهجية" في الجدول 1 لوصف التصميم التجريبي الجاري التنبؤ به. ويوصى أيضاً بتقديم معلومات عن نوع مادة الاختبار المستخدمة في التجارب المستخدمة لاختبار و/أو تدريب النموذج الموضوعي.

14 الترميز

تقدم التوصية [ITU-T P.800.1] ترميزاً عاماً وإجمالياً يمكن استخدامه للمساعدة في تحديد مصدر قيمة متوسط درجة الرأي. وعلى الرغم من أن الترميز الوارد في التوصية P.800.1 مفيد في تقديم نظرة عامة على السياق الذي تولدت بموجبه قيمة متوسط درجة الرأي، فإنه ليس بديلاً لوصف سياق مفصل مقدم وفقاً لهذه التوصية، والذي ينبغي تقديمه دائماً حيثما أمكن ذلك.

بيبليوغرافيا

قام الاتحاد بتقييس عدد من أساليب الاختبار الشخصية لتطبيقات مختلفة. ويرد أدناه بعض الأكثر استخداماً منها.

[b-ITU-T G.729] Recommendation ITU-T G.729 (2012), *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP).*

وتتضمن السلسلة ITU-T P.800 العديد من التوصيات المتعلقة بالتقييم الشخصي والموضوعي لجودة الصوت؛ ويُخَص بالذكر منها ما يلي:

[b-ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality.*

[b-ITU-T P.805] Recommendation ITU-T P.805 (2007), *Subjective evaluation of conversational quality*.

[b-ITU-T P.835] Recommendation ITU-T P.835 (2003), *Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm.*

[b-ITU-T P.862.1] Recommendation ITU-T P.862.1 (2001), *Mapping function for transforming P.862 raw result scores to MOS-LQO.*

[b-ITU-T P.863] Recommendation ITU-T P.863 (2014), *Perceptual objective listening quality assessment*.

وتتضمن السلسلة ITU-T P.900 توصيات بشأن تقييم الوسائط المتعددة:

[b-ITU-T P.910] Recommendation ITU-T P.910 (2008), *Subjective video quality assessment methods for multimedia applications.*

[b-ITU-T P.911] Recommendation ITU-T P.911 (1998), *Subjective audiovisual quality assessment methods for multimedia applications.*

[b-ITU-T P.912] Recommendation ITU-T P.912 (2016), *Subjective video quality assessment methods for recognition tasks.*

ونشر قطاع الاتصالات الراديوية أيضاً توصيات تتعلق بالتقييم الشخصي للجودة السمعية والفيديوية:

[b-ITU-R BS.1116-1] Recommendation ITU-R BS.1116-1 (1997), *Methods for the subjective assessment of small impairments in audio systems including multichannel sounds systems*.

[b-ITU-R BS.1534-1] Recommendation ITU-R BS.1534-1 (2003), *Method for the subjective assessment of intermediate quality levels of coding systems.*

[b-ITU-R BT.500-13] Recommendation ITU-R BT.500-13 (2012), *Methodology for the subjective assessment of the quality of television pictures*.

[b-ITU-R BT.710-4] Recommendation ITU-R BT.710-4 (1998), *Subjective assessment methods for image quality in high-definition television*.

ملاحظة – إن المنهجيات المقيَّسة التي وصفها قطاع الاتصالات الراديوية لا تقيس جميعها عادة متوسط درجات الرأي. وتقدم توصيات قطاع الاتصالات الراديوية المنشورة وثائق ومراجع كاملة لجميع المنهجيات ذات الصلة. وللاطلاع على وصف وتوضيح أفضل لمنهجيات الاختبار المحددة أعلاه، توجَّه عناية القارئ إلى فرادى التوصيات المنشورة المقدمة من قطاع الاتصالات الراديوية.

وترد في الكتيب الإرشادي التالي معالجة متعمقة لأساليب الاختبار والممارسات الفضلى الشخصية.

[b-ITU-T handbook] *Practical procedures for subjective testing* (2011).

سلاسل التوصيات الصادرة عن قطاع تقييس الاتصالات

السلسلة A	تنظيم العمل في قطاع تقييس الاتصالات
السلسلة D	المبادئ العامة للتعريفة
السلسلة E	التشغيل العام للشبكة والخدمة الهاتفية وتشغيل الخدمات والعوامل البشرية
السلسلة F	خدمات الاتصالات غير الهاتفية
السلسلة G	أنظمة الإرسال ووسائطه والأنظمة والشبكات الرقمية
السلسلة H	الأنظمة السمعية المرئية والأنظمة متعددة الوسائط
السلسلة I	الشبكة الرقمية متكاملة الخدمات
السلسلة J	الشبكات الكبلية وإرسال إشارات تلفزيونية وبرامج صوتية وإشارات أخرى متعددة الوسائط
السلسلة K	الحماية من التداخلات
السلسلة L	البيئة وتكنولوجيا المعلومات والاتصالات، وتغير المناخ، والمخلفات الإلكترونية، وكفاءة استخدام الطاقة، وإنشاء الكبلات وغيرها من عناصر المنشآت الخارجية وتركيبها وحمايتها
السلسلة M	إدارة الاتصالات بما في ذلك شبكة إدارة الاتصالات وصيانة الشبكات
السلسلة N	الصيانة: الدارات الدولية لإرسال البرامج الإذاعية الصوتية والتلفزيونية
السلسلة O	مواصفات تجهيزات القياس
السلسلة P	**المطاريف وطرائق التقييم الشخصية والموضوعية**
السلسلة Q	التبديل والتشوير
السلسلة R	الإرسال البرقي
السلسلة S	التجهيزات المطرافية للخدمات البرقية
السلسلة T	المطاريف الخاصة بالخدمات التليماتية
السلسلة U	التبديل البرقي
السلسلة V	اتصالات البيانات على الشبكة الهاتفية
السلسلة X	شبكات البيانات والاتصالات بين الأنظمة المفتوحة ومسائل الأمن
السلسلة Y	البنية التحتية العالمية للمعلومات، والجوانب الخاصة ببروتوكول الإنترنت وشبكات الجيل التالي وإنترنت الأشياء والمدن الذكية
السلسلة Z	اللغات والجوانب العامة للبرمجيات في أنظمة الاتصالات

طُبعت في سويسرا
جنيف، 2019