
Assessment of outcomes of learning services — Guidance

Évaluation des acquis de la formation — Recommandations





COPYRIGHT PROTECTED DOCUMENT

© ISO 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Assessment planning	3
4.1 General	3
4.2 Needs analysis	3
4.3 Resource planning	3
4.4 Assessment framework	3
5 Assessment development	4
5.1 General	4
5.2 Timing of the assessment	4
5.3 Qualifications of assessment developers	5
5.4 Means of assessment	5
5.5 Assessment specifications	5
5.6 Objectivity, reliability and validity	6
5.6.1 General	6
5.6.2 Objectivity	6
5.6.3 Reliability	6
5.6.4 Validity	7
5.7 Item development	7
5.8 Scoring procedures	7
5.8.1 Hand scoring	7
5.8.2 Live scoring	7
5.8.3 Machine scoring	7
5.8.4 Scoring materials — procedures	7
5.9 Reporting of assessment results	8
5.9.1 Time frame	8
5.9.2 Information	8
5.9.3 Score expiration	8
5.10 Arbitration, grievances and appeal	8
5.11 Technical documentation	8
6 Administration of assessments	8
6.1 Guidelines	8
6.2 Assessment security plan	8
6.3 Proctors	9
6.3.1 Proctor qualifications	9
6.3.2 Proctor responsibilities	9
6.4 Qualifications of scorers/raters	9
7 Maintenance and revision	9
7.1 Assessment maintenance plan	9
7.2 Assessment revision plan	10
8 Fairness	10
8.1 Formal agreement	10
8.2 Non-discrimination	10
8.3 Rights of the examinees	10
8.4 Neutrality	11
8.5 Rules of conduct	11
8.6 Information provided to examinees	11
9 Ethics	11

9.1	Responsibilities	11
9.2	Assessment information	11
9.3	Information security	11
Annex A (informative) Technical documentation		13
Bibliography		15

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 232, *Learning services outside formal education*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Assessment is used for two fundamental purposes in learning services: (i) making a decision about the level of competence achieved by learners and (ii) identifying areas of improvement for learners. The intent of this document is to provide a framework for the development, implementation and use of results from assessments of learning outcomes in different contexts for different purposes and to address the needs of different stakeholders.

This document identifies the stages of assessment development, implementation and usage of results, and recommends criteria for each stage. It provides guidance on the selection, development, validation, planning, administration and use of assessments in a range of learning outcomes.

Assessment of outcomes of learning services — Guidance

1 Scope

This document provides guidance on the planning, development, implementation and review of assessments of the outcomes [knowledge, competence, performance] of learning services. It is intended for use by organizations providing learning services and organizations selecting, using or developing assessments.

This document is applicable to the development and use of assessments for the measurement of individual learners' outcomes and the use of assessments for determinations of learner progress.

The document does not apply to the direct evaluation of programs of instruction or the evaluation of learning service providers. It also excludes the technology requirements for the delivery of assessments.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1

assessment

test, examination, observation, or other process, designed to measure an examinee's knowledge, competence, or performance in a defined area against specific reference points or standards (educational or professional)

[SOURCE: ISO/IEC 23988:2007, 3.1, modified — “against specific reference points or standards (educational or professional)” added.]

3.2

assessor

person or organization responsible for interpreting an examinee's performance on the assessment tasks and providing appropriate reporting and feedback to examinees and the client

Note 1 to entry: Assessors are competent to make decisions about the use and interpretation of assessment procedures. In relation to psychological testing, in some countries they are referred to as “test users” to distinguish them from “test proctors”, “test administrators” or “monitors”. Assessors can be employees of the service provider or the client, or third parties contracted for the purposes of the assessment.

[SOURCE: ISO 10667-1:2011, 2.5, modified — aligned to assessment of learning services.]

3.3

competence

ability to apply knowledge and skills to achieve intended outcomes

Note 1 to entry: The ability to apply knowledge and skills means that the learner demonstrates appropriate attitudes and behaviour in different contexts or situations with responsibility and autonomy.

[SOURCE: ISO 29991:2014, 2.5, modified — Note 1 to entry added.]

3.4

diagnostic assessment

assessment which identifies what the learner already knows and/or the nature of difficulties that the learner might have

3.5

proctor

invigilator

authorised person who administers or supervises an assessment

3.6

scorer

rater

qualified and trained person who assigns a score to an examinee's performance based on defined criteria

3.7

hand scoring

human scoring of an examinee's response to an item using an answer key in which all acceptable responses are predetermined

3.8

live scoring

human scoring of an assessment during or immediately following the administration of an assessment

3.9

machine scoring

automatic scoring of an item using technology

Note 1 to entry: Machine scoring is an emerging technology where, for example, a computer program automatically scores a computer-based assessment.

3.10

assessment developer

organization or individual responsible for the planning, development, delivery and quality assurance of assessments

3.11

skill

ability to perform a task or activity with a specific intended outcome acquired through education, training, experience or other means

[SOURCE: ISO/IEC/TS 17027:2014, 2.74]

3.12

assessment sponsor

organization or individual accountable for the acquisition, development or selection of assessments

3.13

qualification

demonstrated education, training, and work experience, where applicable

[SOURCE: ISO/IEC 17024:2012, 3.7]

4 Assessment planning

4.1 General

Planning is a phase of the assessment life cycle that begins with needs analysis and resource planning and includes the assessment framework.

4.2 Needs analysis

The selection or development of an assessment should be based on the needs of the assessment sponsor, relative to the intended use of the assessment. To ensure that the assessment is appropriate for its intended use, the assessment sponsor should conduct a needs analysis before selecting or developing any assessments, and ensure that any selected assessments meet those needs.

The needs analysis should include consultation with stakeholders, to maximize opportunities for coordination and minimize duplication of effort. The input should be documented.

At a minimum, the needs analysis should cover the following:

- the goals and objectives of the organization(s) that will use the assessment and their clients;
- the field of knowledge, skills, capacities or competencies to be assessed;
- the timing of the assessment relative to the learning event;
- the type of decisions that will be made on the basis of assessment;
- the parties responsible for the assessment;
- the target population for the assessment and any accommodations required due to disability;
- the environment and location of the assessment;
- any constraints on the assessment;
- the duration of the validity of the scores of the assessment.

The needs analysis should be documented for future reference and revision.

4.3 Resource planning

Because there are so many components to assessment planning, assessment sponsors that wish to select or develop assessments should develop a plan for resource allocation in order to complete the steps recommended in this document, to include:

- selecting or designing and developing an assessment;
- administering the assessment;
- ensuring the security of the assessment;
- reviewing the assessment, as needed;
- ensuring the fair and ethical administration of the assessment.

4.4 Assessment framework

Assessment developers should develop a framework document in close coordination with the assessment sponsor and other relevant stakeholders, with input from outside assessment experts as

needed. The assessment framework is based on the needs analysis and provides the answers to the questions asked in [4.2](#). The assessment framework should contain the following:

- the scope of the assessment;
- the decisions to be made on the basis of assessment scores;
- the intended consequences of the assessment use;
- an explanation of the links between the goals of the assessment and the means of the assessment;
- an explanation of the links between assessment scores and their interpretations, uses and consequences;
- how the scores are reported to stakeholders.

The assessment framework should also contain:

- the conditions for developing and implementing assessments, including diagnostic assessments;
- timelines and costs;
- the environment of the assessment, whether a classroom, the working environment, an assessment centre, online or another environment;
- any prerequisites required for taking the assessment;
- standards to be applied, for example minimum or average scores to be achieved;
- tools and methods used for assessing learning outcomes;
- quality assurance mechanisms;
- qualifications parties involved in the development or administration of the assessment;
- the maintenance plan for the assessment;
- the rules of conduct of the assessment (see [8.5](#));
- the availability of existing assessments, and the degree to which they meet the requirements of the framework.

The assessment framework should be approved by the assessment sponsor.

5 Assessment development

5.1 General

The following subclauses provide a structure for reflection on the development of an assessment.

5.2 Timing of the assessment

The design phase should specify when the assessment should be administered, as shown in [Table 1](#).

Table 1 — Assessment timing

What is to be assessed?	Timing
Existing knowledge and abilities	Pre (before learning)
Growth in skills and knowledge of content areas	Ongoing during learning event at appropriate intervals
Attainment of (course) learning goals	End of learning event

Assessments may be independent of the learning event, in which case the recommended timing will not apply.

5.3 Qualifications of assessment developers

Assessment developers' main tasks are the planning, development and quality assurance of the assessment. They should be trained and qualified to undertake activities and tasks related to the development of assessments assigned to them.

Their roles, responsibilities and qualifications should be clearly defined and documented, whether they are internal or external resources for the organization developing the assessment.

They should guarantee impartiality and avoid conflicts of interest. They should:

- have appropriate knowledge and experience in the relevant field;
- be trained and qualified to analyse and interpret the data, give feedback and document the assessment process;
- be trained in validation of assessment processes;
- be familiar with different assessment methods and tools.

5.4 Means of assessment

The means of assessment should be specified and linked to the goals of the assessment, which have been determined through the needs analysis (see 4.2). The assessment may be performed in many ways, depending on what is to be assessed. Examples of means of assessments are given in [Figure 1](#); these are not exhaustive or mutually exclusive. Further examples may be found in the references in the Bibliography.

5.5 Assessment specifications

The assessment specifications document should provide detailed specifications regarding the design, content, scoring, reporting and intended use of the assessment.

For assessment administration and technological requirements, the descriptions should include required training and qualification information for any assessment administration personnel and any materials or technology needed to administer the assessment. Examples include, but are not limited to:

- physical assessment environment or setting;
- time allotted to assessment administration;
- assessment administration personnel (including training and qualification requirements);
- documents, materials and tools required by examinees or administrators;
- hardware and software (including version, bandwidth, security requirements);
- how long the assessment report will be retained.

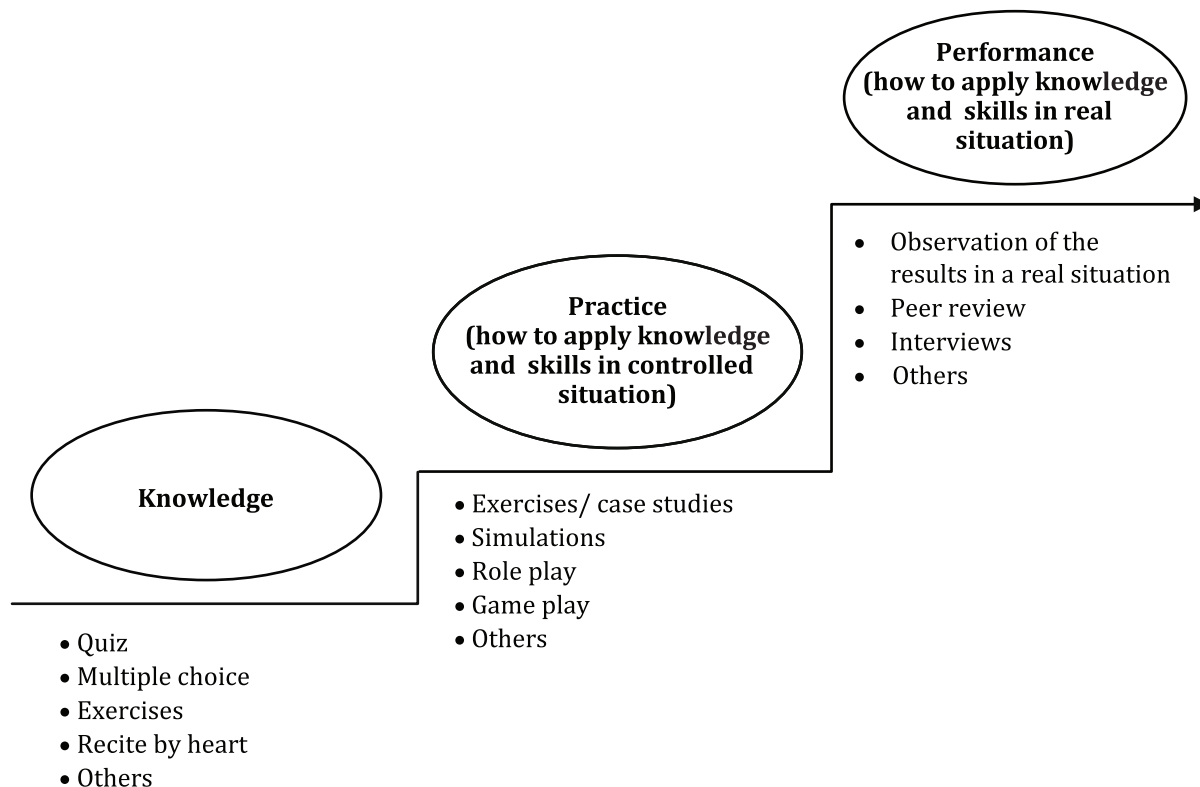


Figure 1 — Assessment process

5.6 Objectivity, reliability and validity

5.6.1 General

Any assessment needs to be objective, reliable and valid.

5.6.2 Objectivity

Assessments should be scored by means of agreed-upon criteria, so that different scorers will arrive at the same score for the same set of responses, to the degree possible.

5.6.3 Reliability

Assessments should be as reliable as possible. Reliability is an indicator of the extent to which examination scores are consistent across different examination times and locations, different examination forms and different examiners.

Reliability in reassessing: examinees retaking the same assessment after a reasonably short period should receive the same results.

Reliability in scoring: a completed assessment should receive the same score if scored again. For hand-scored tests (as opposed to machine-scored tests) the score should be the same if scored by more than one individual, or if scored more than once by the same individual.

Item reliability: the tasks, questions or prompts to which examinees respond should produce repeatable results across groups of examinees and, if hand scored, should produce reliable scores.

5.6.4 Validity

Assessments should be valid. Validity is a judgement about whether an assessment is appropriate for a specific group and purpose and includes considerations such as:

- whether the assessment really measures what it intends to measure;
- whether the results are similar to examinees' performance in other assessments, or in class or real-world activities;
- whether the use of assessment results has the intended effects.

Developers of an assessment should document evidence for the validity of the assessment. This evidence should be available to assist in the selection of assessments.

Care should be taken when comparing different assessments of the same subject. If comparisons are made, evidence for the basis of comparison, such as concurrent validity data, should be provided.

5.7 Item development

Items may include (but are not limited to) multiple-choice items; prompts, scoring rubrics and/or expected responses for constructed-response items; and tasks eliciting responses. Items should be coherent with the assessment specification (see [5.5](#)).

Assessment developers, whether writing items, prompts or tasks, should adhere to the assessment specifications and should submit all components of each item or prompt, including the expected response, to a rigorous review process.

5.8 Scoring procedures

5.8.1 Hand scoring

Hand scoring should use a clearly defined scoring protocol with predetermined, unambiguous criteria.

5.8.2 Live scoring

Live scoring should use a clearly defined scoring protocol.

5.8.3 Machine scoring

Machine scoring should meet the following criteria:

- The minimum technical requirements for the scoring technology should be specified.
- The accuracy of scoring machines should be periodically verified by hand scoring.

5.8.4 Scoring materials — procedures

Procedures for transporting assessment materials to the scoring locations should consider the security of the materials. The physical and electronic locations where materials will be stored before, during and after scoring should consider the security of the materials.

Procedures for how, where and when scorers record scores for individual items, sections and entire assessments should be established.

The assessment developers should provide an answer key, scoring guide and scale for the scorers, including lists of items required and lists of items prohibited.

5.9 Reporting of assessment results

5.9.1 Time frame

An appropriate time frame for reporting the results to the entity requesting the assessment should be established.

5.9.2 Information

The information in the report should be clear, concise, descriptive and usable by the requesting entity and the examinee.

5.9.3 Score expiration

Assessment developers should specify and justify score expiration.

5.10 Arbitration, grievances and appeal

The organization using the assessment should publish clearly defined policies describing the appeal process, the conditions under which an appeal can be requested, and the roles, responsibilities and timeline for the appeal process.

5.11 Technical documentation

The interested parties in the assessment should maintain technical documentation as described in [Annex A](#).

6 Administration of assessments

6.1 Guidelines

Guidelines for the administration of the assessment should be documented. These guidelines should address the delivery conditions, proctoring, scoring, procedures and materials, reporting of results, policies for grievances, arbitration and appeal, reassessment conditions, score expiration and records management.

Delivery conditions should be standardized to the greatest extent possible, with reasonable accommodations, to ensure an equitable experience for all examinees.

6.2 Assessment security plan

Assessment security encompasses all areas of development, production, administration, scoring and reporting. Every assessment should have a security plan, which should include, at a minimum:

- protection of personal data;
- a description of the roles and responsibilities of security personnel;
- a list of assessment security documents to be generated or appropriated to include instructions for development personnel, assessment security non-disclosure forms, instructions for proctors, examinees, scorers and policy statements;
- a description of training methods for personnel responsible for test security;
- a list of physical and electronic security requirements;
- a description of methods used to monitor for conformance with security policies;

- a security breach contingency plan;
- countermeasures to improper activities conducted by examinees, development of the manual on how to handle such activities, and measures to facilitate mutual understanding among staff and proctors.

6.3 Proctors

6.3.1 Proctor qualifications

If possible, proctors should not have conflicts of interest, such as being related to an examinee or the instructor of an examinee.

6.3.2 Proctor responsibilities

- Ensure integrity of assessment environment.
- Validate identification of examinees.
- Communicate with assessment administrators.
- Ensure the assessment process conforms with assessment procedures.

6.4 Qualifications of scorers/raters

Scorers/raters may be involved in the development of the assessment as well as the administration of the assessment. They should:

- be trained and qualified to determine the level of competence of the examinee against agreed criteria or standards;
- have appropriate knowledge and experience in the relevant field.

7 Maintenance and revision

7.1 Assessment maintenance plan

Maintenance means ensuring and documenting that the assessment remains valid and reliable. This plan should include the following elements:

- a list of documents comprising reliability and validity evidence;
- specifications for how assessment performance will be evaluated;
- a list of processes that will be used to review the items and assessment, conduct statistical analyses of operational items and assessments, and retrain and certify scorers;
- specifications of how often each of these processes will be performed over the life cycle of the assessment;
- metrics used to determine item or assessment life cycle or both;
- a recommendation on how to handle results of the maintenance review;
- an estimate of the resources needed to perform assessment maintenance (money, contracts, personnel).

7.2 Assessment revision plan

Assessments should be periodically reviewed based on the maintenance plan. In the event that revisions are recommended, the revision process should follow these steps:

- specification of the circumstances under which changes will be allowed and those under which changes will be mandatory;
- specification of the mechanisms for item refreshment (replacement of entire forms or certain percentages of items, how to generate new cut scores after item refreshment);
- specification of the circumstances under which cut scores may be changed without alteration of assessment composition;
- specification of how much change in the assessor pool is allowable;
- specification of the strategies and processes to be used to validate the new items;
- specification of resources needed to carry out revisions.

8 Fairness

8.1 Formal agreement

A formal written agreement prevents a misunderstanding among those concerned or allows them to easily settle any complaints. The formal agreement should include at least the following:

- examination fees;
- the duties and rights of the persons involved;
- any prerequisites required for taking the assessment;
- description of the assessment;
- personal data protection;
- time frame for reporting results to examinees.

8.2 Non-discrimination

All the examinees within the scope of the assessment should be provided with a fair opportunity to take the assessment regardless of their age, gender, nationality, religion or disabilities, including visual, auditory and physical impairment. It is also considered desirable to take into account access from the areas where they live.

Every effort should be made to provide reasonable accommodations for examinee disabilities in the development of assessments. For example, the following issues need to be considered:

- whether it is reasonable to give examinees with disabilities the same amount of time to complete an assessment as other examinees;
- whether the assessment result is fair if an alternative assessment is produced to accommodate disabilities.

8.3 Rights of the examinees

Interested parties should always consider the dignity and self-respect of examinees during an assessment. In particular, if the assessment result negatively impacts the examinee, his or her interests and rights should be protected as much as possible.

8.4 Neutrality

The assessment process should be neutral with respect to those involved in the assessment, whether examinees, assessors or others. Assessors should remain unbiased throughout the entire assessment process, including reporting of results.

Ideally, the assessor and proctors should not be the instructor of an examinee. If the personnel involved in the examinee's learning administers an assessment, measures should be taken to ensure the independence of the assessment.

8.5 Rules of conduct

Assessment developers should develop the rules of conduct that ensure fair conditions for examinees, to include identification requirements, materials allowed during the assessment and timing of departure on completion of the assessment. Assessment sponsors should ensure that the rules of conduct are observed.

8.6 Information provided to examinees

Assessment developers should provide the following information:

- the rules of conduct (see [8.5](#));
- the proctoring procedures;
- security requirements;
- the abilities that the assessment is intended to measure;
- the level of abilities the assessment is intended to measure;
- contact information in case examinees and other people involved need to make inquiries.

9 Ethics

9.1 Responsibilities

At all stages, all interested parties have ethical responsibilities. These organizations and individuals are responsible for determining, communicating and documenting any local responsibilities and obligations that may not be known to others involved in the development and administration of an assessment. In all phases of an assessment project, all interested parties are responsible for considering the ethical implications of their own and others' actions.

9.2 Assessment information

Assessment information provided to assessment sponsors, assessment developers, assessment users and examinees should be true and accurate. It is unethical to knowingly misrepresent information about the assessment.

9.3 Information security

The developers and users of assessment systems should guarantee:

- the confidentiality of the information collected;
- the protection and security of data that may be used in an assessment of knowledge or skills;
- the secure distribution, retrieval and transmission of information;

- the appropriate structure of the information management system, and development of a manual on the information management system;
- training for those in charge of information management, especially training to maintain up-to-date knowledge.

Annex A (informative)

Technical documentation

A.1 Documentation

A.1.1 Technical documentation

All assessments should include technical documentation that covers the assessment life cycle from initial planning and development through to ongoing assessment use.

The technical documentation should include sufficient information and evidence to evaluate the appropriateness and rigour of the approach, process, methodology, findings, decisions and deliverables as appropriate to each stage of the assessment life cycle.

Documentation of planning and development includes, but is not limited to:

- documents from the original needs analysis;
- the framework document;
- the assessment specification document;
- developer-provided materials: assessment administration manual, training process outline, rubrics and scorer forms.

Documentation of assessment use and performance includes, but is not limited to, the following, as applicable:

- documentation indicating the minimum qualifications and training criteria for all assessors, scorers, proctors and so forth; procedures for developing, reviewing and administering assessments; for scoring; signed agreements and non-disclosure agreements;
- documented test audit/reviews and validation since last evaluation of the same type;
- documentation indicating regular quality assurance records;
- documentation indicating that prepared test items for dynamic tests are being reviewed on a regular basis by the assessors administering them, individually for effectiveness, collectively for standardization;
- documentation indicating that static test items are analysed statistically on an ongoing basis to identify those inconsistent with applicable specifications;
- documentation indicating that constructed-response static items are reviewed by item writers and target language experts;
- documentation of periodic retraining of human scorers to ensure scorer reliability;
- the protocols used in hand scoring;
- the procedures for live scoring and documentation indicating the periodic reviews that show live scoring is following prescribed protocols;
- documentation verifying the accuracy of scoring machines;

- documentation of arbitration, grievance and appeal policies;
- documentation of reassessment conditions policy.

A.1.2 User support documentation

User support documentation should include providing support documents for various stakeholders of an assessment such as end users, assessment developers, scorers and proctors.

The documentation of assessment protocols and procedures, such as assessment samples, assessment manuals for various stakeholders or the assessment security instructions, should be provided and should include sufficient information for the intended audience to perform their roles and responsibilities.

Documentation should meet professional standards for presenting information and evidence as appropriate to the specific stage of the assessment life cycle.

The documentation can be provided as a series of individual reports for each stage or as a single report for the entire life cycle.

A.1.3 Retention of records/scores (records management)

All technical documents, including assessment review or audit results, should be retained, and should be accessible by respective stakeholders.

Documentation should be periodically updated and supplemented as the assessment is either modified or extended to additional uses, populations or contexts.

These updates can be provided as supplementary reports or updates to the original reports.

Bibliography

- [1] ISO 10667-1:2011, *Assessment service delivery — Procedures and methods to assess people in work and organizational settings — Part 1: Requirements for the client*
- [2] ISO/IEC 17024:2012, *Conformity assessment — General requirements for bodies operating certification of persons*
- [3] ISO/IEC/TS 17027:2014, *Conformity assessment — Vocabulary related to competence of persons used for certification of persons*
- [4] ISO 29990:2010, *Learning services for non-formal education and training — Basic requirements for service providers*
- [5] ISO/IEC 23988:2007, *Information technology — A code of practice for the use of information technology (IT) in the delivery of assessments*
- [6] ISO 29991:2014, *Language learning services outside formal education — Requirements*
- [7] American Psychological Association. *The Standards for Educational and Psychological Testing*. The American Psychological Association, Washington, DC, 2013
- [8] European Centre for the Development of Vocational Training. 2015. *European guidelines for validating non-formal and informal learning*

