www.pqsglobal.org
info@pqsglobal.org
+44 20 7126 8527

PQS FACT SHEET

Introduction

In compliance with ALTE's Code of Practice, PQS aims to meet the standards set by ALTE and holds itself accountable to the test takers in particular and the stakeholders in general. To this aim, as the manual (p.13) asserts, we are committed to:

• Defining what each examination assesses and what it should be used for.
• Describing the population(s) for which it is appropriate.
• Explaining relevant measurement concepts as necessary for clarity at the level of detail that is appropriate for the intended audience(s).
• Describing the process of examination development.
• Explaining how the content and skills to be tested are selected.
• Providing either representative samples or complete copies of examination tasks, instructions, answer sheets, manuals and reports of results to users.
• Describing the procedures used to ensure the appropriateness of each examination for the groups of different ethnic or linguistic backgrounds who are likely to be tested.
• Identifying and publish the conditions and skills needed to administer each examination.

Development of Exam Content

PQS covers a range of assessment tools ranging from A1 to C1 according to the Common European Framework of Reference (CEFR). The basic test cycle employed is the model proposed by ALTE (2011) as it is presented in Figure 1. The cycle starts with the justifications for providing each test. This is followed by provision of test specifications (see below) to help develop assessment tools. Upon the administration of the test, scoring the exams and grading the exams, the tests are validated.

The test developers have come up with the Assessment Use Argument (AUA) as proposed by Bachman and Palmer (2010). The arguments regarding the qualities of a test as well as the characteristics of the test takers are presented. PQS experts have considered warrants and rebuttals when it comes to presenting different data for the consistency, sufficiency, relevance, generalizability, impartiality, meaningfulness, values-sensitivity, equitability and beneficence of the exams. We further use the validation model proposed by Weir (2005).
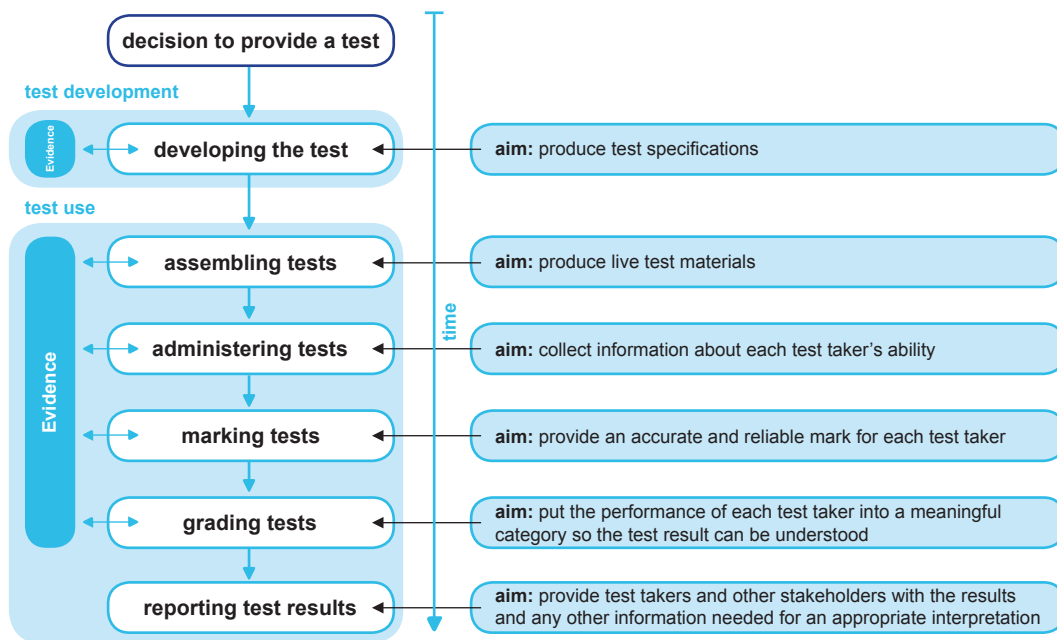
www.pqsglobal.org
info@pqsglobal.org
+44 20 7126 8527

Figure 1: The basic testing cycle proposed by ALTE

Background on Assessment Design

PQS exams go through some iterative steps. According to ALTE's Manual for *language Test Development and Examining*, after the decision to provide the test, the test development procedure starts.

Developing The Test

Developing a test starts with planning. When PQS prepares a test, we evaluate the resources available for test development. The chart below by ALTE condenses the activities, which are followed by PQS.
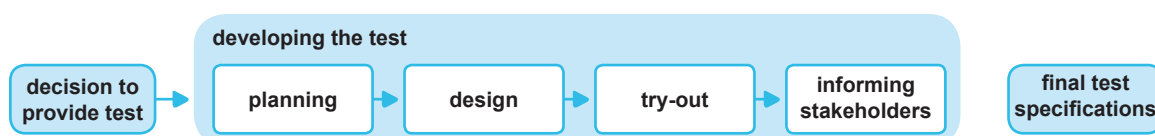

Figure 2: Developing the test

Following Bachman and Palmer's (1996) proposal, PQS team also focuses on the three phases below:

• Selecting elements from the areas of our knowledge (topical, language) for successfully completing the test task
• Formulating one or more plans for implementing these elements in a response to the test task
• Selecting one plan for initial implementation as a response to the test task

Also, when they talk about the **framework of language task characteristics,** they believe we should consider: setting, test rubric, input, expected response, and relationship between input and response.

Test Specifications

Test specification is a plan for what to assess and how to assess it. Although there are different models of test specification like those of Alderson, Clapham and Wall (1995), Davidson and Lynch (2002), Hughes (2002), and Popham (1978), we follow Bachman & Palmer's (2011) model to prepare our test specifications:
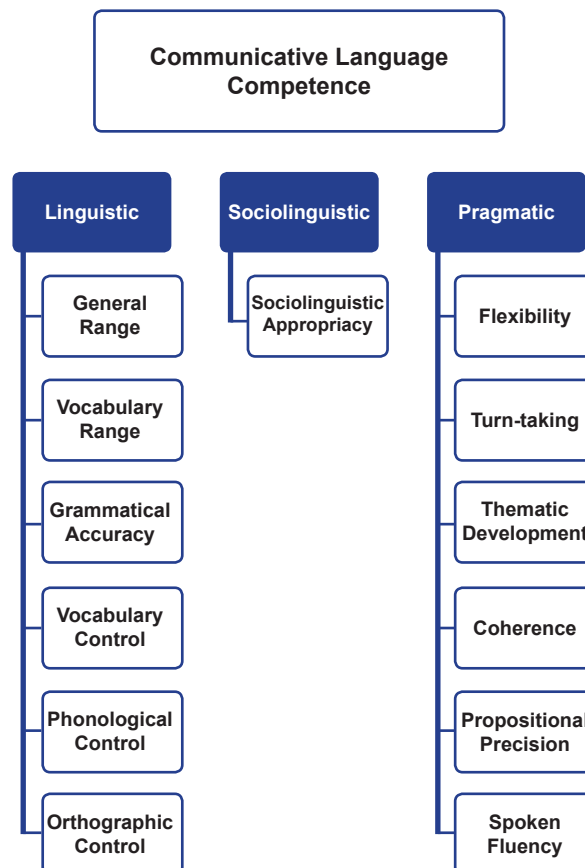
• Structure of the test
• How many parts or subtests; their ordering & relative importance; number of items/tasks per part
• Test task specifications
• Purpose & Definition of the construct
• Setting & Time allotment
• Instructions
• Characteristics of input & expected response
• Scoring method

When preparing tests, our item writers closely follow test specifications to improve the validity of our exams. As Bachman & Palmer's (2011) claim, the use of test specifications can further help our team by:

• Permitting the development of parallel tests;
• Evaluating the intentions of the test developers;
• Evaluating the correspondence between the test as developed and the blueprint from which it was developed;
• Evaluating authenticity of a test.

Communicative Language Competence

The can-do statements devised help practitioners by providing common grounds for teaching as well as assessment purposes. PQS also uses this framework for "transparency and clear reference points". The model adopted by PQs is that offered in the CEFR companion Model (2018) as it appears in Figure 3 below.

```
                    ┌──────────────────────────┐
                    │ Communicative Language   │
                    │       Competence         │
                    └──────────────────────────┘

      ┌───────────────┐   ┌─────────────────┐   ┌───────────────┐
      │  Linguistic   │   │ Sociolinguistic │   │   Pragmatic   │
      └───────────────┘   └─────────────────┘   └───────────────┘

      ┌───────────────┐   ┌─────────────────┐   ┌───────────────┐
      │    General    │   │  Sociolinguistic│   │  Flexibility  │
      │     Range     │   │    Appropriacy  │   └───────────────┘
      └───────────────┘   └─────────────────┘
                                              ┌───────────────┐
      ┌───────────────┐                       │  Turn-taking  │
      │   Vocabulary  │                       └───────────────┘
      │     Range     │
      └───────────────┘                       ┌───────────────┐
                                              │   Thematic    │
      ┌───────────────┐                       │  Development  │
      │   Grammatical │                       └───────────────┘
      │    Accuracy   │
      └───────────────┘                       ┌───────────────┐
                                              │   Coherence   │
      ┌───────────────┐                       └───────────────┘
      │   Vocabulary  │
      │    Control    │                       ┌───────────────┐
      └───────────────┘                       │  Propositional│
                                              │   Precision   │
      ┌───────────────┐                       └───────────────┘
      │  Phonological │
      │    Control    │                       ┌───────────────┐
      └───────────────┘                       │     Spoken    │
                                              │    Fluency    │
      ┌───────────────┐                       └───────────────┘
      │  Orthographic │
      │    Control    │
      └───────────────┘
```

Figure 3: Communicative Language Competence (adapted from CEFR Companion Volume with New Descriptors, p. 30)

As far as the communicative language activities and strategies are concerned, PQS includes mediation and interaction modes of communication into account while preparing its examinations by incorporating integrated task types especially at higher levels.

CEFR Alignment

Linking process serves the purpose of categorising the test takers as regards their level of proficiency such that regardless of the test they fall within the same level.

For instance, if one test taker is placed as A2 once, we have to make sure the test taker can perform the illustrative descriptor for the level and not those above or only below the level categorised. This can be reflected in the process of standard setting and CEFR alignment can guarantee this. To relate our examinations to the CEFR and build the arguments, we use four of the sets of procedures proposed by ALTE:

• Familiarisation
• Specification
• Standardisation training/benchmarking
• Standard setting

For the a posteriori validity argument, the model proposed by Weir (2005) will be used. Below the four interrelated procedures are briefly discussed:

Familiarisation

This stage is planned so that those involved in the process of linking our examinations to CEFR have an in-depth knowledge about the can-do statements including their descriptives and different levels. Familiarisation will take place at the level of standardization as well as well as specification. ALTE considers this stage as "a logical pre-requisite for effective linking". Before embarking on the specification and standardization processes, the familirisation stage can help those involved in the process of linking become more familiar with the descriptor of every level as well as different skills that are measured in our exams. The selected panel of experts will start the procedures as follows:

• Preparatory Activities before the Familiarisation Seminar
- Reading Section 3.6 in the CEFR
- Consideration of a selection of CEFR question boxes
- Doing exercises with the CEFR scales and tasks and performances
• Introductory Activities at the Seminar
- Sorting the text for the different levels
- Self-assessment using CEFR
- Qualitative Analysis of the CEFR Scales by sorting the individual descriptors from a CEFR scale
• Preparation for Rating the Productive Skills
- Reconstructing CEFR-based rating Grid to be used
- Illustrating the CEFR levels with student videoed performances

The findings will be documented and its degree of success will be reported.

Specification

This stage is "a self-audit of the coverage of the examination (content and tasks types) profiled in relation to the categories presented in CEFR. As well as serving a reporting function, these procedures also have a certain awareness-raising function that may assist in further improving the quality of the examination concerned. … Specification can be seen as a primarily qualitative method: providing evidence through "content-based arguments". (COE Learning, Teaching, Assessment Manual, p.10)

The specification procedures involve four steps:

• Assuring adequate familiarisation with the CEFR;
• Analysing the content of the examination or test in question in relation to the relevant categories of the CEFR; should an area tested not be covered by the CEFR, the user is asked to describe it;
• Profiling the examination or test in relation to the relevant descriptor scales of the CEFR on the basis of this content analysis;
• Making a first claim on the basis of this content analysis that an examination or test in question is related to a particular level of the CEFR.

The procedures involve three types of activity:

• Familiarisation activities as described in the familiarization section;
• Filling in a number of checklists with details about the content of the language examination;
• Using relevant CEFR descriptors to relate the language examination to the levels and categories of the CEFR.

This Specification process gives examination providers the opportunity to:

• Increase the awareness of the importance of a good content analysis of examinations;
• Become familiar with and use the CEFR in planning and describing language examinations;
• Describe and analyse in a detailed way the content of an examination or test;
• Provide evidence of the quality of the examination or test;
• Provide evidence of the relation between examinations/tests and the CEFR;
• Provide guidance for item writers;

• Increase the transparency for teachers, testers, examination users and test takers about the content and quality of the examination or test and its relationship to the CEFR. The forms to be filled in have an awareness-raising function (process) and are also sources of evidence to support the claim made (product).

The CEFR-based tools that are available in the COE Learning, Teaching, Assessment Manual, Appendix Section will be used for the specification. Content Analysis Grids that offer the possibility to work at the more detailed level of individual test tasks, classifying them by standard criteria.

Standardisation Training / Benchmarking

"The suggested procedures facilitate the implementation of a common understanding of the "Common Reference Levels", exploiting CEFR illustrative samples for spoken and written performance. These procedures deepen the familiarity with the CEFR levels obtained through the kinds of activities outlined in Familiarisation and assure that judgments taken in rating performances reflect the constructs described in the CEFR. It is logical to standardise – through sufficient training – the interpretation of the levels in this way before moving on to

(a) Benchmarking local performance samples and tasks/items, and
(b) Standard setting.

Successful benchmarking of local samples may be used to corroborate a claim based on Specification. If the result of the benchmarking process is that performance samples from the test are successfully benchmarked to the levels that were intended in designing the test, this corroborates the claim based on Specification." (COE Learning, Teaching, Assessment Manual, pp. 10-11)

Standard Setting

"The crucial point in the process of linking an examination to the CEFR is the establishment of a decision rule to allocate students to one of the CEFR levels on the basis of their performance in the examination. Usually this takes the form of deciding on cut-off scores, borderline performances. The preceding stages of Familiarisation, Specification and Standardisation can be seen as preparatory activities to lead to valid and rational decisions," (COE Learning, Teaching, Assessment Manual, p. 11). At PQS we use two different procedures to arrive at the final decision of setting cut scores depending on the suitability of context.

Assessment Validation Employed for PQS Exams

As Messick (1989) claims one validity aspect might not encompass different operations and conditions. Therefore, to ensure fairness, the validation method employed by PQS is that proposed by Weir (2005). PQS addresses:

• Context Validity
• Theory-based validity
• Scoring validity
• Consequential validity
• Criterion-related validity

As Weir (2005) claimed the first two types of validity will be carried out before a test is administered (a priori) which the other three will be addressed after the test administration. The validation is separately done for each skill; however, for the sake of brevity, a summary of the validation process is presented in Figure 2.
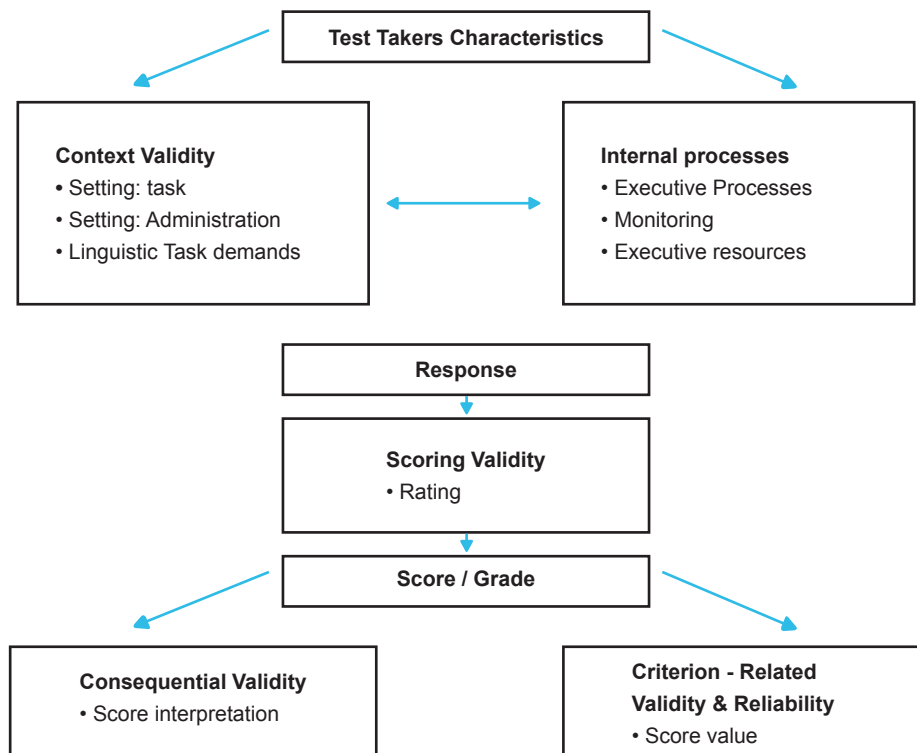

Test Takers Characteristics

Context Validity
• Setting: task
• Setting: Administration
• Linguistic Task demands

Internal processes
• Executive Processes
• Monitoring
• Executive resources

Response

Scoring Validity
• Rating

Score / Grade

Consequential Validity
• Score interpretation

Criterion - Related Validity & Reliability
• Score value

Figure 2: Validation Process for PQS Examinations, adapted from Weir (2005)

References

ALTE. (2011). Manual for language test development and examining. Retrieved from:
https://rm.coe.int/manual-for-language-test-development-and-examining-for-use-with-the-ce/1680667a2b

ALTE. (2020). Principles of good practice. Retrieved from:

https://alte.org/resources/Documents/ALTE%20Principles%20of%20Good%20Practice%20Online%20version%20Proof%204.pdf

Bachman, L. F., & Palmer, A. S. (2010). Language assessment in practice. Oxford University Press.

Common European framework of reference for languages: Learning, teaching, Assessment. Companion volume with descriptors. (2018). Retrieved from

https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Fulcher, G., & Davidson, F. (2007). Language testing and assessment. Routledge.
Martyniuk, W. (2011). Aligning tests with the CEFR. Cambridge University Press.
Messick, S. (1989). Validity. In R. Linn (Ed.), Educational Measurement (pp. 13–103). Macmillan.
O'Sullivan, B. (2011). Language testing: Theories and practices. Palgrave Macmillan.
Weir, C. J. (2005). Language testing and validation. Palgrave McMillan.