

QMS as a Continuous Process Of Self-Evaluation and Quality Improvement for Testing Bodies

Henk Kuiper, CITO

Recent research into language courses for immigrants in Holland has pointed to the poor results achieved on these courses. The construction of tests of Dutch as a second language for these courses has raised a number of questions regarding the quality, use and political consequences of these tests:

- Are the tests really as good as we thought they were?
- Are they suitable for all types of immigrants, ranging from well-educated people to people with hardly any education?
- Are the interpretations of the test results in the research mentioned correct?
- Is it ethical to administer tests after courses that we know are far too short for the majority of immigrants?

In short, questions are being raised as to our social and political responsibility. In the near future the Dutch Government will request the design of a test for naturalisation. Then the questions we are confronted with will become much more urgent.

This case is a good example of the issues that have arisen in ALTE now that we are setting up a Quality Management System. I am very happy to be a member of ALTE, because it is an excellent forum for discussing these urgent issues and reflecting on them with colleagues from Europe who are dealing with similar difficulties and finding different solutions to their problems. It provides evidence that the activities in the context of the Quality Management Systems are far from theoretical. In fact, the work we have done from 2000 onwards has turned out to be very inspiring and fruitful - a view shared by many ALTE colleagues.

ALTE is an association with a wide variety of cultures, although maybe the diversity is limited in the sense that all members do share a common European tradition. Kunnan (May,2002) has outlined an interesting difference between government-ruled organisations and private "not-for loss" companies. Both of these types are represented in ALTE - as well as many variations in between.

Having said this, it may not be surprising to note that differences between the ALTE members are great with respect to the organizational, linguistic, educational and cultural contexts within which the examinations are being developed and used. Furthermore, there are huge differences in knowledge and traditions with respect to statistical and empirical issues like data gathering, data analysis, equating different examinations, etc. Yet the members of ALTE share a commonly felt need for the fairness of their examinations.

Introducing a system of quality assurance could be very threatening for some members who know for themselves that they do not meet standards at the moment - particularly when compared with other members of the group. However, the QMS was designed to lower anxiety and is meant to be a supportive tool; it should allow members:

- to enhance the quality of their examinations in the perspective of fairness for the candidates;
- to engage in negotiations with their senior management and sponsors in a process of organisational change (e.g. to ensure that resources are made available to support on-going improvements).

In 2000, after several years focusing on other projects, we began discussions within ALTE about how to implement the Code of Practice in our own examination processes. After sometimes long but very open discussions, a working group was established which is now playing a leading role in the making the Code of Practice a central issue within ALTE.

Time schedule: from self-monitoring to the possibility of external verification

ALTE members have accepted a time schedule to steer the members through the diverse stages of quality improvement. A very important aspect of this schedule is that stress is placed on the processes of quality enhancement and awareness raising and that the ultimate goal of a quality mark - or Q mark - is deferred and is a longer term objective. Involvement and awareness of the importance of quality is actually a necessary precondition to making QMS really workable. After all, *including* and not excluding members that have not yet achieved acceptable standards is fundamental at this moment in our development. It is of great value for test takers that the tests ALTE members develop be of good quality. It is equally important that the use of tests is ethically acceptable. These are the aims of QMS. The result of this approach is that the active involvement of all ALTE members is of great importance. The activities carried out so far have led to a growing awareness by ALTE members of the strong and weak points in their own examinations. This awareness is far from being imposed from outside but is engaged with by the members themselves.

We are now at the point where most members have used checklists as part of a self-evaluation exercise. Members have held intensive discussions on the outcomes of this exercise and the results of the discussions can be summarised as follows:

- Trend analysis
- The Code of Practice as the leading document in determining future ALTE activities
- Data gathering and data analysis for validation and reporting

Trend analysis

The revised checklists now exist in 4 main parts or modules:

- Test design and Construction
- Administration - including conduct of the exams
- Processing - including marking, grading and issuing of results
- Analysis and review of the examinations

As far as test design and construction is concerned, ALTE members in general follow standardised procedures in developing tests, based largely on communicative language constructs and careful content descriptions of test items, tasks and components. Despite differences, test construction generally meets minimum standards - which of course does not mean that improvement is not needed in some areas.

Administration reveals pronounced organisational differences between the ALTE members, e.g. depending on factors like the number of test takers, the question of whether the examinations are administered in-country or abroad, and so on. Many members, though, feel that their procedures have to be reconsidered and improved in the near future and they have set themselves goals to accomplish this.

Processing and Analysis - are the areas which many members feel are in greatest need of improvement. In this respect ALTE can function as a forum to provide opportunities for these areas to be discussed and improved through, for example, workshops, exchange of methodologies, peer consultation and monitoring.

The Code of Practice as a leading document for workshops and presentations within ALTE meetings

An important result of QMS activities to date is the agreement of all ALTE members that the future activities of ALTE will be centred on the issues that seem to be most in need of improvement based on the completed checklists. Furthermore, engendered by the Code of Practice, work will be addressed at forthcoming ALTE meetings. The following areas can be distinguished:

1. Developing routine procedures for data gathering, data entry and data analysis:

- ALTE members will have to gather data and analyse their examinations by means of pre-testing or post-examination analysis in order to be able to demonstrate the quality of their examinations and the fairness of decisions made based on examination results.
- These procedures will play a role in providing information on validity and reliability in relation to the use of tests.

2. Responsibilities of test developers for the social outcomes of their tests: the discussions within ALTE as a consequence of introducing the Code of Practice have already led to discussions about the responsibilities of test developers in the following areas:

- Tests in the context of immigration
- Tests for citizenship
- Dealing with candidates with special circumstances and needs.

Action for the short term

At the end of 2002 the first steps for setting up routine procedures in data analysis were taken by those ALTE members who were in the first phase of this area of language testing. Members executed post-examination analysis of at least one of their item based examinations and in the near future will be able to present figures on difficulty and reliability. To help them achieve this goal, a workshop on data analysis using the program "TIA plus" developed by Cito was held in April 2002. ALTE members were able to use this program free of charge. This is just one of the many examples of ALTE as a forum for exchange of know-how and assistance and for peer evaluation.

ALTE members are in the process of up-dating their responses to the checklists; reviewing and updating their answers in the four modules describing their examination development, administration, grading and analysis.

The fact that many members are involved in designing tests in the immigration process and tests to be used in naturalisation and acquiring citizenship has led to an initiative to set up a working group to discuss how the different ALTE members deal with this issue.