Previous training in different language testing and assessment activities:  
- Were there differences between regions of Europe?

Detailed findings regarding the respondents’ previous training in different assessment activities are reported below. Included are only those respondents who worked in Europe and replied to the questions in Part 1 of the questionnaire (Teachers’ Questionnaire) are included here (n = 491 – 566, depending on the question / activity).

Classroom-focused testing / assessment

Preparing your own classroom tests

Most of the respondents (80-90%) had received some training in this activity, and quite often the training was more extensive than just a couple of days in length (over 40% on average).

Significant differences were found between the regions (chi square = 25.4, df = 12, p = .013). In particular, the Northern European and Baltic respondents differed from the other respondents.

Using ready-made tests from textbook packages or from other sources
The respondents were quite split as far as the amount of training in this activity was concerned: typically, they had either received no training at all or they had received more extensive training. Significant differences were found between the regions (chi square = 24.7, df = 12, p = .016). In particular, the Eastern and Northern European respondents different from the others: the Eastern European respondents were better trained than the others whereas the Northern European respondents less well trained.

Interpreting test results (classroom or external tests)

This was one of the activities where most respondents had received at least some training; however, over 20% had had no training. About half of those who had some training had only a little bit of it, and another half had more extensive education in this activity. No statistically significant differences were found between the regions.
Giving feedback to students based on information from tests / assessment

The respondents were typically split into three equally big groups: about a third had received no training, a third had a little bit of training and a third had more extensive training.
No statistically significant differences were found between the regions.

Using self / peer assessment

About 40% of the respondents reported having received no training in this activity. Those who had received some training were often split rather equally between those who had a little training and those who had more extensive training.
Significant differences were found between the regions (chi square = 22.1, df = 12, p = .037). In particular, the Central European respondents were better trained than the others. More Northern European respondents had received ‘a little training’ than could be expected and fewer Baltic respondents had received ‘more extensive training’ than could be expected.

Using informal, continuous, non-test type of assessment

![Bar chart showing training distribution across regions of Europe.]

A significant portion, 30-40%, of the respondents had received no training in this activity. Those who had were often split into two rather equal groups in terms of the length of their training. Significant differences were found between the regions (chi square = 27.1, df = 12, p = .008). The Central and Western European respondents were better trained than the others. More Northern European and Baltic respondents than could be expected had received ‘a little training’ and fewer ‘more extensive training.’

Using the European Language Portfolio, an adaptation of it or some other portfolio
Most respondents (about 60%) had received no training in using portfolio type of assessment, although this varied somewhat from region to region. A minority only had received any extensive training. Significant differences were found between the regions (chi square = 24.1, df = 12, p = .020). The Southern and Western European respondents were better trained than the others. More Northern European respondents than could be expected had received ‘a little training’.

Purposes of testing

To give grades
Perhaps not surprisingly, most respondents (80%) had received at least some training in this very common assessment activity; usually, the training had been more extensive. However, 20% had had no training. No statistically significant differences were found between the regions.

To find out what needs to be taught / learned next

The training background for this activity is very similar to the previous one (grade giving). However, a little bit more respondents reported having had no training. No statistically significant differences were found between the regions.

To place students onto courses, programmes, etc

The amount training varied considerably from region to region, but typically up to half of the respondents had had no training in this activity.
Very significant differences were found between the regions (chi square = 35.4, df = 12, p = .000). The Western, Central and Eastern European respondents were better trained in this than the others. The Northern European and Baltic respondents were less well trained in this activity than the others.

To award final certificates (from school / programme; local, regional or national level)

![Graph showing training levels by region](image)

The amount of previous training varied considerably for this activity, too, depending on the region. Typically, 40-50% of the respondents had had no training. When they had, it tended to be more extensive than just a couple of days.

Very significant differences were found between the regions (chi square = 36.2, df = 12, p = .000). The Northern and South-Eastern European respondents were less trained in this activity than the others, whereas the Western, Central and Southern European were better trained than the others.

**Content and concepts**

Receptive skills (reading/ listening)
This was one of the assessment activities where the great majority of the respondents had received some training, often more than just a couple of days. No statistically significant differences were found between the regions.

Productive skills (speaking/writing)

This was one of the assessment activities where the great majority of the respondents had received some training, often more than just a couple of days. No statistically significant differences were found between the regions.

Microlinguistic aspects (eg grammar/vocabulary)
This was one of the assessment activities where the great majority of the respondents had received some training, often more than just a couple of days. No statistically significant differences were found between the regions.

Integrated language skills

Compared with training in the specific skills above, the training of assessing language skills in an integrated fashion was slightly less common / extensive: some 20% had received no training in it. On the whole, however, the majority had been trained, often extensively. No statistically significant differences were found between the regions.

Aspects of culture
Compared with the other content areas above, the assessment of different aspects of culture was something that the respondents had not very often been trained on: 40-50% had had no training at all.
Significant differences were found between the regions (chi square = 32.1, df = 12, p = .001). The Western European respondents were better trained and the South-Eastern and Southern European respondents less well trained than the others.

Establishing reliability of tests / assessment

The respondents were split into three groups of roughly equal size for this activity, with 30-40% of them having had no training.
No statistically significant differences were found between the regions.

Establishing validity of tests / assessment
As with the reliability above, the respondents were split into three groups of roughly equal size for this activity, with 30-40% of them having had no training. No statistically significant differences were found between the regions.

Using statistics to study the quality of tests / assessment

About half of the respondents had received no training in the use of statistics, and for only 20% of them the training had been more extensive. No statistically significant differences were found between the regions.

**External tests and examinations (regional or national)**

It should be noted that the training on activities related to external tests and examinations is often the responsibility of the examination board or organization in question rather than e.g. the teacher-training programmes which are typically responsible for the general
assessment education reflected in the activities listed above. Looking at the amount of training reported for activities related to external examinations is more informative when compared with the actual need to carry out these tasks (i.e., a respondent may only function in one role in the external examination, e.g. that of a rater of performances, or an item writer and therefore may not have received training focusing on tasks that he/she does not need in his/her capacity).

Taking part in rating oral or written performances

![Graph showing distribution of training among regions of Europe for rating oral or written performances.]

Most respondents (70-80%) had received training in this activity. Significant differences were found between the regions (chi square = 24.5, df = 12, p = .018). The Central, Southern and Western European respondents were somewhat better trained and the South-Eastern European respondents somewhat less well trained than the others.

Using statistics to study the quality of tests / assessment

![Graph showing distribution of training among regions of Europe for using statistics to study the quality of tests.]

Using statistics to study the quality of tests / assessment
Most of the respondents were not trained in this (60%) which is probably not surprising as it can be assumed that only some of those who are involved in school-external examinations and tests actually carry out statistical analyses on them. No statistically significant differences were found between the regions.

Writing items/test tasks for an examination body

About half of the respondents had not received training in this activity. No statistically significant differences were found between the regions.

Reviewing items/test tasks for an external examination

About 50-60% of the respondents had not received training in this activity.
Significant differences were found between the regions (chi square = 22.0, df = 12, p = .038). The Western European respondents were better trained and the Northern European respondents less well trained than the others.

Acting as an interviewer / interlocutor in an oral test or examination

Together with ‘rating oral and written performances’ (and ‘defining assessment criteria’ below) this activity was one of the activities related to external examinations where most of the respondents had received some training – presumably because they need it in their role.

No statistically significant differences were found between the regions.

Defining assessment criteria
As with the activity above, about a half or more of the respondents had received some training in this activity. No statistically significant differences were found between the regions.