

Foreign Language Assessment Cultures

Policies and Practices in European Union Countries

Final Report

Groningen, January 2005

Peter Edelenbos

Foreign Language Assessment Cultures Policies and Practices in European Union Countries

The Dutch Ministry of Education, Culture and Science has initiated this study

ISBN 90 329 2217 3

© National Bureau for Modern Foreign Languages (NaB-MVT), Enschede, the Netherlands

Reference to free use of this document. No part of this publication may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission of the publisher.

© Nationaal Bureau Moderne Vreemde Talen (NaB-MVT), Enschede

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enige andere manier zonder voorafgaande toestemming van de uitgever.

Contents

	Page
Executive Summary	iii
1. Introduction	1
1.1 Rationale for the Study	1
1.2 Objectives	2
2. Research questions	3
3. Backgrounds to conceptualisations of Foreign Language Assessment Cultures	5
3.1 Introduction	5
3.2 Types of assessments	5
3.3 Dimensions of a Foreign Language Assessment Culture	7
4. Method	11
4.1 Process	11
4.2 Research Groups	13
4.3 Research Instruments	14
5. Results	19
5.1 Introduction	19
5.2 Context: Languages and Educational Systems	19
5.3 Objectives and Types of Assessments	23
5.4 Reasoning behind Current Assessment Cultures	27
5.5 Intensity of Assessments	35
5.6 Structure of Tests	40
5.7 Organisation and Implementation	49
5.8 Use of Results	54
5.9 A brief Exploration of the Concept of Foreign Language Assessment Culture	58

6.	Conclusions	63
6.1	Introduction	63
6.2	Objectives and Assessments	63
6.3	Reasoning behind Current Assessment Cultures	65
6.4	Structure of Tests	66
6.5	Organisation and Implementation	67
6.6	Use of Results	68
6.7	Areas for future study	69

References	71
-------------------	----

Appendices (available in a separate document)

Appendix 1:	Questionnaire for Eurydice Units	73
Appendix 2:	Questionnaire for Policy Makers	81
Appendix 3:	Questionnaire for Test Experts	93
Appendix 4:	Official and co-official languages, minority and regional languages	109
Appendix 5:	Descriptions of Foreign Language Assessment Cultures in 25 EU-Countries	131

Executive Summary

In June 2004, the Netherlands began its six-month European Union presidency. During this period and the years to come, the Barcelona strategy must be further developed. This requires quality indicators for foreign language education (strategic aim 3.3) that reflect the various European nations' test and assessment cultures, highlighting their diversity. The Dutch Ministry of Education, Culture and Science has therefore initiated a study into foreign language assessment cultures, policies and practices in European Union countries.

The research has two focuses. Firstly, it provides a description of the broad outlines of the test cultures in 25 different countries. This includes determining the educational objectives within each culture and the types of tests used for assessing these objectives, describing the structure of tests, the organisation and implementation of the tests, as well as the way in which the results are used. Secondly, the study examines the reasoning behind current test cultures and the expectations attached to them by policy makers, key influencers and test developers.

A study into the similarities/differences between policies and practices of foreign language assessment at the conclusion of compulsory education among countries in the European Union is, by nature, primarily descriptive. The methodology is therefore a mixture of qualitative and quantitative techniques. These techniques have been applied in a predefined sequence so as to ensure clarity.

It is virtually impossible to establish a direct relationship between countries' stated objectives regarding foreign language learning and the types of assessments used. In some countries, several types of assessment are integrated, while in others the same types are used independently of each other. The diversity in types of assessments of foreign languages must be due to varying perceptions of and rationales for assessments, as well as contrasting views on the functions of assessments. Beneath these differences are each country's historical backgrounds, including specific economical, political and socio-cultural needs that have been incorporated into an operational educational policy regarding foreign language assessments. Therefore, utmost attention should be given to adequate quality indicators for foreign language education reflecting European nations' assessment cultures and their diversity. Also, a well defined path should be provided to inform the audience how quality indicators are chosen and to which extent these are important for every country.

Each assessment implemented has a specific function within its own educational context. Examination and certification are the most important functions of nation-wide assessments. System monitoring is the prevailing function of standardised assessments,

while other types of assessments are more focused on examination and certification, as well as providing guidance to pupils and information to parents.

Proficiency in the four skills is an important aim for most countries, with some rating oral skills above the others. Many respondents also cited communicative competencies as an important aim. Across the EU, there is a shared practical and conceptual approach to the skills to be tested and a common focus on communicative competencies. This should benefit future development of tests for measuring students' foreign language skills at the end of secondary education. Each European country has formulated a unique set of aims and goals for the attainment level of foreign languages in secondary education. However, an in-depth analysis of the complete set of countries' stated objectives, similar to comparative studies as for example TIMSS still has to be conducted. The shared skills that can be tested and the common focus on communicative competencies can only be seen as a challenge. It is a fertile starting point for increasing cooperation between organisations and countries in Europe. Multi-national and bi-national cooperation will support mutual understanding of the aims and goals that are stated and the levels of competence that are strived for in each country. Within a multitude of SOCRATES-programmes and cross border programmes, possibilities for funding are available.

English and French are the two most widely assessed languages at the end of the second cycle of secondary education throughout all EU-countries. However, German, Spanish and Italian are prominent in assessments in a wide range of EU-countries. There is common ground regarding the communicative competencies that are assessed most often: reading comprehension and writing. This does not mean that other competencies are less highly valued. The assessment tasks applied for the different skills do not seem to differ that much between countries. A wider set of tentative try outs of bi-national or multi-national assessments of listening and speaking and languages could provide tentative developmental knowledge to support the more scientific approach in the context of the Barcelona strategy.

Each EU-country has implemented an organisational model to facilitate the development and administration of tests at the end of the first or second cycle of secondary education. A number of countries employ a top-down model. Despite the similar approach, though, there are key variations with respect to the organisations involved and the kinds of tasks they are assigned. Even among countries where schools are empowered with the administration of assessments, one finds stark differences in the design and implementation of organisational models for realising the main objectives: examination and certification.

1. Introduction

1.1 Rationale for the Study

In June 2004, the Netherlands began its six-month European Union presidency. During this period and the years to come, the Barcelona strategy must be further developed. This requires quality indicators for foreign language education (strategic aim 3.3) that reflect the various European nations' test and assessment cultures, highlighting their diversity. The Dutch Ministry of Education, Culture and Science has therefore initiated a study into foreign language assessment cultures, policies and practices in European Union countries.

The Barcelona strategy calls for a test development phase during the period 2005-2007, which will concern tests for measuring students' foreign language skills at the end of secondary education (16-18 yrs.). The tests will reveal each student's performance, but must also be used to give each country an indication of the quality of foreign language education given at its schools. Before any such test development, however, a number of choices must be made regarding the tests' content and form. This study aims to gather the information necessary for forming the tests and for determining the context in which the tests should take place.

The research has two focuses. Firstly, it provides a description of the broad outlines of the test cultures in 25 different countries, determining the educational objectives within each culture and the types of tests used for assessing these objectives. This also includes describing the structure, organisation and implementation of the tests, as well as the way in which the results are used. Secondly, the study examines the reasoning behind current test cultures, as well as the expectations attached to them by policy makers, key influencers and test developers.

The research is part of a broader initiative of related projects aimed at developing a theoretical framework and an item bank (Alderson et al, 2004); Preparatory Documents to the Scheveningen Conference, October 2004). A group of high-profile, language test experts have been given a Socrates grant and are laying down the foundation for an item bank. This will comprise sets of highly valued items found using a theoretical framework, and it is likely that a number of these non-biased items will act as key items in future test development. The item bank should be based on the Common European Framework for Languages and should contain national- and culture-based items in addition to its key items. This will make it possible to determine the quality of foreign language education in each individual country and allow for cross-national comparisons.

1.2 Objectives

Studies into the relationship between assessment and instruction in foreign language teaching have indicated important differences in various countries' test cultures (see Lynch 2003, McNamara 2001, Rea-Dickins 2001, 2003). Also, studies have already revealed the differences between backgrounds in foreign language teaching (i.e. policy, system and trends) for a wide range of countries (see for example Dickson & Cumming, 1996). However, given that the European Union has expanded to include 10 new countries, it is necessary to expand on this research and determine the exact differences between test cultures for foreign languages. The focus of this study is to address the question:

What are the similarities/differences between the policies and practices of foreign language assessment at the conclusion of compulsory education among countries in the European Union?

Studies carried out by Dickson & Cumming (1996), West, Edge and Stokes (2000) and by Euridyce (2001), have shown that policies regarding the conclusion of compulsory education differ among the E.U. countries. Some countries conclude compulsory education with national exams that include foreign language assessment. Other countries, such as Sweden and Germany, do not administer national exams. Also, the situation in the Netherlands has proven that national exams can exist without laws that make education compulsory. It is important to note the differences between types of schools and the ages at which students finish their secondary education. The organisation of secondary education and the policies regarding compulsory schooling can have a direct relationship to a country's test culture. However, a country's test culture may have characteristics that are inconsistent with the exams given at the end of compulsory education. This study is concerned with the way in which students conclude their secondary education. Knowledge of these procedures is important for an adequate understanding of each country's assessment culture, even for countries that do not require students to finish secondary school.

2. Research Questions

The study's focus "What are the similarities/differences between the policies and practices of foreign language assessment at the conclusion of compulsory education among countries in the European Union?" has been subdivided into four research questions.

Objectives and Types of assessments

Research question 1 Which objectives for foreign language education are tested and which types of foreign language assessments are used?

This study incorporates nationwide tests administered during or at the end of secondary education. Examinations such as the European or International Baccalaureate and the GCSE, however, are not included. Furthermore, this study aims to gather information about the types of assessment implemented, the specified objectives and, if applicable, the specified contents.

Structure of Tests

Research question 2 How are foreign language tests structured- i.e., which languages and skills are tested, what types of questions are asked and what is the mode of delivery?

Answering this question does not require an extensive inventory. However, the distinction between languages and skills tested is important for providing data on the types of testing used for evaluation purposes and in day-to-day practice. Also, information about the structure of a test is necessary for determining the test's quality and the extent to which it complies with (inter)national standards such as the CEF. Another important aspect is the relationship between a test's content and the education offered. Finally, it is necessary to estimate the influence variables such as the school, teacher, class and student have on the results of a country's foreign language education.

Organisation and implementation

Research question 3 Who decides to organise foreign language tests and who is in charge of the implementation?

This study aims to gather information on the organisations involved in foreign language testing, and those that have specific tasks within the process of testing.

Use of Results

Research question 4 How and to what extent are the results of foreign language tests used by different groups?

The use of foreign language test results varies per user group. Here, the distinction between results at class/school level and those at a national level is relevant: In what ways do students, teachers and schools use test results, and how are the results used for the evaluation and development of local and national educational policies? The study looks at the media in which the results are presented, the interpretations of aggregated results and the reactions from the public, professionals and important decision makers. The study also takes into account the organisations that report test results and the levels at which data are interpreted (e.g. school-level, groups of schools, nationally, etc.).

3. Background to Conceptualisations of Foreign Language Assessment Cultures

3.1. Introduction

This study aims to describe and analyse the various European nations' foreign language test and assessment cultures, highlighting their similarities and differences. A foreign language assessment culture, however, is not a concept rooted in a linguistic or social-political theory, a nomological framework or a long-standing tradition of research (Byram, 2000).

A growing number of publications have already dealt with the differences in foreign language teaching in primary and secondary education in the European Union (Eurydice, 1992, 2001; West, Edge and Stokes, 1999; Blondin, et al. 1998). These included information on examinations and assessments of foreign languages. Eurydice has developed *Eurybase* [http://www.eurydice.org/Eurybase/frameset_eurybase.html], a comprehensive database of factual knowledge on educational systems in all EU-countries. A substantial amount of information from *Eurybase* was used for the production of the national reports (see appendix 4 and 5). Descriptions of general backgrounds, legislation, structures, grading and certification can be found in these reports. However, *Eurybase* does not include details on objectives for foreign language learning, types of assessments, the structure of tests, the organisation and implementation of assessments and the use of results.

3.2 Types of Assessments

Tests, assessments and examinations are concepts that have distinct meanings and connotations in different contexts, countries and cultures. The use of these terms, then, can create confusion when dealing with 25 EU-countries. Currently, no comprehensive typology of tests, assessments and examinations can overcome these differences, making a practical solution necessary. For the purposes of this study, then, a very general distinction has been made between nation-wide assessments, standardised assessments and other assessments. Of course, the latter type has also been subdivided.

In a nation-wide assessment, instruments are used that have been developed from shared aims. Students meet the same conditions for the assessment and must fulfil the same requirements. Still, the curricula students have followed can differ, and students

may have followed different syllabuses or courses. The significance of the outcomes (scores), though, is the same for all students. In this sense, the term nation-wide assessment covers examinations and tests of skills at a national level without a pass/fail decision attached to them. Examples of nation-wide assessments are: examinations at the end of the first or second cycle of secondary education, tests offered to all schools in a country to monitor the progress of students at fixed intervals, as well as the DESI-project in Germany.

A nation-wide assessment can function as an examination. In several countries, an examination is divided into an external part and a school-based part. In an external assessment, instruments are used that have been developed from shared aims. Students meet the same conditions for the assessment and need to fulfil the same requirements. Still, the curricula students have followed can differ. The outcomes (scores) can have the same significance for all students. That will be the case when the external assessment, for example, is a reading test that counts for half of the overall grade a candidate receives for his or her exams. However, the outcomes of an external assessment can also have a different use or significance for each student. Schools can call upon an external institute that offers well-established tests for listening or speaking. In the resulting account of the examination, schools can have the freedom to value the outcomes as they see fit. A nation-wide assessment is always external, but an external assessment is not necessarily a nation-wide assessment.

In a standardised assessment, instruments are used that are developed from shared aims. Students meet the same conditions for the assessment and need to fulfil the same requirements. The curricula students have followed can differ, but this is not always the case. The outcomes (scores) are based upon pre-set standards and can have a different use or significance for each student. For governments, the aggregated outcomes of these standardised assessments can be used to keep track of changes in pupils' performances, over longer periods of time. Data from standardised assessments can enable system control.

An example of a standardised assessment is a listening test that provides information, comparable over the years, in a certain educational sector for a specified student age or level. Another example of a standardised assessment is sample based testing, every two or three years, to monitor the outcomes of a newly introduced foreign language.

The category 'other assessments' is subdivided into school-based assessments, general teacher assessments, portfolio assessments and continuous assessments. School-based assessments, in general, can materialize in two ways. The practice of education is based on schools trying to achieve a set of objectives, and testing is generally used to generate feedback on the skills and attitudes of students during this process. The

emphasis is then on administering grade-based assessments of student performance. This type of school-based assessment is used in all European countries. School-based assessments can make up the foreign language part of a nation-wide assessment or exam that concludes the first or second cycle of secondary education.

General teacher assessment has been added to complete the range of possible other types of assessments. However, general teacher assessment is a trivial concept since in all secondary schools teachers will frequently assess progress of pupils through observations, tests, questions and several other kinds of informal, diagnostic techniques (Edelenbos & Kubanek-German, 2004; Rea-Dickins 2001, 2004).

Continuous assessment is an ambiguous concept. On the one hand it can mean keeping a record of accomplishment of what pupils do and how they perform in a foreign language, on the other a means keeping an in-depth record of performances in several foreign languages, continually assessing pupil performance with different types of tests.

3.3 Dimensions of a Foreign Language Assessment Culture

A foreign language assessment culture comprises several complementary mechanisms, five of which are reviewed here.

Bottom-up versus Top-down

In everyday educational practice, there is usually a bottom-up approach to testing. The practice of education is based on schools trying to achieve a set of objectives, and testing is generally used to generate feedback on the skills and attitudes of students during this process. The emphasis can be on administering grade-based assessments of student performance, and grades for individual students can be aggregated to reflect performance at higher levels (class, school, region, and country). In contrast to the bottom-up approach, a nation or a state can favour a top-down approach to testing. In a top-down approach, a policy originates from a set of objectives that is formulated at an overarching level (regional, national, and supra-national). Testing is mandatory at various points within the national school system (e.g. at the end of primary or secondary schooling) to measure the extent to which these objectives have been achieved. These two mechanisms are relatively independent of each other; it is possible for the two to clash but also for them to operate in harmony (see Dickson & Cumming, 1996 and Euridyce, 2001).

Formative versus Summative

National test cultures also can reflect a choice between two basic functions: formative and summative assessment. The practice of education is based on schools trying to achieve a set of objectives and schools use tests as a formative procedure to generate feedback on the skills and attitudes of students during the learning process (Cronbach, 1987). These grades for individual students may also be aggregated to provide information on achievement at higher levels (class, school, region, country). This is part of formative assessment, which is probably mainly operational in countries where teachers are solely responsible for the assessment in school-based exams (Dickson & Cumming, 1996).

By contrast, summative assessment consists of testing a set of national objectives at various points within the national school system (Philips, 1999, 2000). However, the countries that employ this approach may differ in the way they provide centralised exams. A mixture of both approaches is taken when school-leaving exams consist of both a school-based and a central part. In this case, the results from both parts are combined.

Centralisation versus Decentralisation

The organisation and development of assessments can take on many forms. A country can centralise authority for organising assessments by placing the Ministry of Education in charge. In a much centralised state, the Ministry of Education can develop assessments and take responsibility for the whole process of collecting, grading, marking and reporting. The Ministry of Education can operate with the assistance of a closely affiliated institute for test development, experts, teachers unions and universities. In contrast, this responsibility can also be passed on to a group of specialised organisations. When these organisations collaborate with several interest groups in the field, this type of organisation can be typified as decentralised.

User-oriented versus System-oriented

User-oriented assessments are highly focused on the users of the system, i.e. pupils and parents. This places emphasis on providing guidance for students' educational careers and providing information to parents. When the focus is on the system, the functions of the assessment will be perceived in terms of system monitoring, leverage for educational change and determining school quality.

Low versus High Intensity of Assessments

From a pupil's perspective, this is an essential characteristic of a foreign language assessment culture. An extreme, low intensity approach would be where a student takes a test just once, at the end of secondary education. This would be a once in a lifetime,

'high stakes' experience, but then again the pupil would have had a very low test taking profile during the rest of his or her educational career. At the other end of the scale would be cultures that assess pupils during and at the end of lower secondary education and again during and at the end of upper secondary education. This high intensity approach to assessment involves an entirely different organisation and development of tests, as well as distinctive provisions at the school level and mechanisms for feedback.

4. Method

4.1 Process

A study into the similarities/differences between policies and practices of foreign language assessment at the conclusion of compulsory education among countries in the European Union is, by nature, primarily descriptive. The methodology is therefore a mixture of qualitative and quantitative techniques for obtaining the necessary data. These techniques have been applied in a predefined sequence so as to ensure clarity, which is especially beneficial in an international study. The individual steps in the process are described below.

1. Approaching Members of Networks

Established networks were asked to participate in the study and small, informal networks were expanded for the same purpose. In both cases, the aim was to reach suitable respondents and to fill the gaps created by the expansion of the European Union.

2. Research

This study involves researching the foreign language test cultures in the 25 European Union countries. In 2001, the European Commission's Eurydice network published the extensive report, "Foreign language teaching in schools in Europe", which covered the 15 member countries of the European Union, the three EFTA/EEA countries and the 11 pre-accession countries. This publication included cross-national studies as well as reports on individual countries; here it forms the basis for a short background description of each country's foreign language assessment culture. Each description was subject to approval by a representative from the network of test developers.

3. Survey

Three groups of respondents were asked to participate in this study: policy makers, test experts and employees from national Eurydice agencies. Per country, questionnaires were sent to one representative from each group. Exceptions were Belgium and the U.K. (three representatives from each group), and Germany, where four states were asked to participate (a total of four representatives per group). While each group received questions specific to its field, a large number of questions were identical. Further details on the study's research groups, research instruments and procedures are provided below.

4. Data Analyses

Two types of analysis were used: Qualitative analysis for providing an extensive case-study for each country (Madaus & Kellaghan 1996; Yin, 1997) and quantitative analysis for producing tables and descriptions.

5. Consultation

The results from both types of data-analysis were presented to six experts in language testing. Six topics were discussed:

Method

1. How has the study been carried out?

Description of Assessment Cultures

2. To what extent is the description of assessment cultures of foreign languages in Europe complete?

Developing perspectives

3. Which parts/aspects of the study (per research question) are valuable?

Preparing decisions

4. Which data offer adequate possibilities to work with in order to formulate conclusions; which data are less suitable?

Developing insights

5. What are the most important trends that emerge from the data? Which data still pose problems?

Suggestions

6. What should be included when drafting the conclusions? Which conclusions and recommendations can be clearly made based on the data?

6. National Reports

The comments and suggestions from the consultation group were used in creating a report for each of the 25 countries. Each report is limited to 5 pages, focusing on background information and the most important findings for each of the four research questions. Also, each report has been offered for validation to representatives from the government officials and / or test experts.

7. Description of Trends in Test Culture

The final report on major trends in European assessment cultures is based on the 25 national reports as well as the data obtained from test experts, policy makers and experts from Eurydice.

4.2 Research Groups

25 EU Countries

The EU is composed of 25 countries, but within some countries there are historical, legal or linguistic differences that have to be accounted for. In the case of Belgium there is a division between the Dutch, French and German speaking communities. From the 15 German states and independent cities, four states were selected based upon their location: Baden Wurttemberg, Saxony, Lower Saxony and North Rhine- Westphalia. For the United Kingdom, the study differentiates between England, Scotland and Wales.

Composition of Three Research Groups

Initially there were four research groups, but some alterations were made during the course of the study. One originally conceived research group, school headmasters, was omitted after the first experiments (see section 4.3) showed they had insufficient knowledge of foreign language assessments. Also, it soon became clear that the test developers and test experts could be combined into one group. Another research group, the national representatives from Eurydice units, was added after approaching the Dutch national unit. The Eurydice units formed the third research group, and proved to be very useful in providing general data on foreign language learning in their respective countries.

The first research group consisted of policy makers from the national governments who have been working together since the Lisbon-summit and perform functions within the framework of the Barcelona-strategy. However, a full list of key policy makers in all EU-countries was far from complete. In some countries, policy makers specialising in foreign language education were on leave, while other countries were not yet represented by such policy makers in an EU context. Through informal networks, advice from experts, a number of phone calls to Ministries of Education and the Internet it became possible to contact policy makers in all EU countries, with the exception of Wales.

The second research group is made up of test developers and of experts and researchers in the field of modern foreign languages. There are a number of related European networks, the most important of which are actively involved in the OECD – PISA study, ALTE and EALTA. Using informal networks and the advice of the consultation group, a total of 32 test experts were included. The third research group consisted of representatives from the national Eurydice units in the European Union. This is an established network and all 29 representatives received a questionnaire.

Response rates

Questionnaires were sent by mail and e-mail to 31 policy makers, 22 of which responded. The raw response rate is 71.0%, and the information provided covers 19 countries (76.0%). A total of 32 test experts were invited to cooperate in this study. 21 test experts responded, resulting in a raw response rate of 65.6% and an EU-country response of 80.0%. A response was received from 18 of the 28 Eurydice agencies, giving a raw response of 67.9% and an EU-country response of 72.0%. In nearly all instances the questionnaires were fully completed. Data for the research on the basic variables were available for 24 countries (96.0% of the 25 EU-countries).

4.3 Research Instruments

Developmental Process

A draft version of the questionnaire was written after the initial desk analyses were carried out. The desk analyses were absolutely necessary, as they made the notable differences in terminology and systems between the 25 EU-countries clear (see also appendix 5). The draft version was then discussed with a small group of professionals, and this led to a complete revision of the questionnaire. The second version was sent to two Dutch test experts and one former policy maker. Their responses were recorded and used for a revision of a large section of the questions. The questionnaire was then filled in by Danish and German test experts and a German policy maker. In the interviews afterwards, they commented on the complexity of the questionnaire as a whole, the difficulty of the questions and their problems providing context-specific information to rather general questions. The latter was important, because the questions had to cover the diverse conceptualisations and systems of the 25 EU-countries.

The questionnaires were discussed with the consultation group during a one-day meeting, in which the general set up of each of the three questionnaires was reviewed and all of the questions were meticulously revised. The session resulted in the final versions of the questionnaires for Eurydice representatives, test experts and policy makers; these can be found in appendices 1, 2 and 3.

Research Questions and Questionnaires

Three questionnaires were used for obtaining the basic information necessary for answering the four research questions. The questionnaires included both closed and open-ended questions. Table 4.1 presents an overview of each questionnaire's content in relation to the research questions.

Table 4.1 Questionnaire Content in Relation to Research Questions and Additional Topics

		Research Questions				Additional topics	
		1	2	3	4		
Research groups	Back-grounds	Assess. types	Structure of tests	Organi-sation	Use of results	Use of CEF	Changes in policy
Eurydice	x	x	x		x		
Test experts		x	x	x	x	x	
Policy Makers	x	x		x		x	x

Below is an overview of the subjects that were addressed for each of the four research questions. The groups that answered questions related to the individual topics are cited in brackets behind each topic. Each group received some questions specific to its field, increasing the study's range and efficiency. However, in several instances identical questions were posed to more than one group, so as to ensure the data's reliability.

Research question 1 Objectives and types of assessments

- Overall aims of foreign language learning in secondary education [policy makers]
- Nation-wide assessments of foreign languages [policy makers, test experts and Eurydice]
- Standardised assessments incorporating samples, certain age-groups or grades in a country/region/school [policy makers, test experts and Eurydice]
- Other types of assessment, such as teacher assessments, continuous assessment, school-based assessments, portfolio assessments [policy makers, test experts and Eurydice]
- Rationale for nation-wide assessments [test experts, policy makers]
- Functions of assessments [policy makers, test experts and Eurydice]

Research question 2 Structure of foreign language tests

- Languages tested: how many; which ones and why [test experts, policy makers, Eurydice]
- Scheduling assessments: primary, lower secondary, upper secondary [test experts, policy makers]
- Communicative competencies tested and rationale [test experts, policy makers, Eurydice]
- Types of questions and modes of delivery [test experts, policy makers]
- Choice of content [test experts]
- Assessments related to the CEF, empirical verification and impact [test experts, policy makers]

Research question 3. Organisation and implementation

- Institutions actively involved in the development and implementation of foreign language tests [test experts, policy makers]
- Institutions actively responsible for the development and implementation of foreign language tests [test experts, policy makers]
- Factors influencing the design of assessments [test experts, policy makers]
- Organisations in charge of collecting, marking, grading and reporting tests and their results [test experts, policy makers]

Research question 4 Use of foreign language test results

- Types of feedback received by participants [test experts, policy makers]
- Types of results published [test experts, policy makers]
- Form in which results are published [test experts]
- Audiences who receive results [test experts]
- Examples of concrete changes in policy or practice resulting from language testing [policy makers]

Validity

The instruments were developed in two stages and in close cooperation with the consultation group mentioned above. The validity of the initial stage was enhanced by studying and analysing the available documents on the specific traits and characteristics of foreign language assessment cultures in the 25 European Union member states. Publications of the European Commission's Eurydice network, cross-national studies as well as reports on individual countries were used to design questions and formats. These

publications were also used as a validated basis for short background descriptions of each country's assessment culture within foreign language education.

Similar structures were implemented in all three questionnaires. Each respondent answered questions on nation-wide assessments, standardised assessments and other types of assessments. Furthermore, in a later stage of the study respondents were asked to validate the summaries included in this final report.

The test experts and policy makers who participated in this study are both capable and willing to provide criticism when necessary, and five of the respondents did so. The respondent from France made justifiable objections to the limited scope of the questionnaire, pointing out that it did not cover the complexity of the assessments in France. The policy maker from the Netherlands stated that he was not equipped with enough factual knowledge to give an adequate response. However, he did answer the questions and provided the necessary information. The test-expert from Wales defined the questionnaire as professional but said that it did not meet 'my own framework'. The test-expert from Finland found some questions difficult to answer because of their phrasing and because of the complex situation in his country. Finally, the test-expert from Estonia mentioned that the last part of the questionnaire was difficult to answer, because of the questions' general phrasing. All in all the criticism was mild, and a number of respondents were enthusiastic about the study's approach to the problem of foreign language assessment cultures.

Though one must always be critical, it is safe to say that the validity of the questionnaires' content was reasonably high.

Reliability

Different groups of professionals answered questions related to the same topic, and nearly always these questions were identical. On the one hand this satisfies the need for different perspectives from different professional points of view, but on the other hand it raises the question of reliability (i.e. does a test-expert's answer to a particular question conforms to that of his country's policy maker?). The test experts and policy makers from Sweden, Hungary, Malta, the Netherlands, Slovakia, Portugal, Finland and Latvia all answered the questionnaires independently. Cohen's Kappa was calculated using data obtained from identical questions posed to test-experts and policy makers, producing the satisfactory inter-rater reliability of .72.

5. Results

5.1 Introduction

This study addresses the question: What are the similarities/differences between the policies and practices of foreign language assessment at the conclusion of compulsory education among countries in the European Union? Three focuses can be distinguished. Firstly, the study identifies the contexts of assessments that are given during and at the end of secondary education (5.2), as well as the educational objectives within each culture and the types of assessments used (5.3). These descriptions form a starting point for the study's second focus, which is to determine the reasoning behind current assessment cultures, as well as the expectations attached to them by policy makers and test developers (5.4). This entails looking at backgrounds to the current situation in EU-countries: functions of assessments, rationales for assessments and relevant changes in government policy that have occurred during the last five years or are expected. Thirdly, a description of current E.U. test cultures is provided, including information on the intensity of testing (5.5), the test structure (5.6), the organisation and implementation of tests (5.7) and the manner in which results are used (5.8).

5.2 Context: Languages and Educational Systems

A 5 to 7 page report is provided for each of the 25 EU-countries in appendix 5, outlining some of the essential features of each country's foreign language assessment culture. These include descriptions of:

1. The educational system (languages, structure, legislation, recent developments in educational policy);
2. Foreign language education (languages taught, goals, teaching time);
3. Grading and evaluation (lower and upper secondary education, national exams);
4. Certification (for lower and upper secondary education);
5. Assessments of foreign languages (types, structure, organisation and implementation, use of results and relationships with the CEF).

The national reports provide a wealth of detailed information. This section summarises some aspects of the reports and it aims to provide a rich context for the next paragraphs.

Table 5.1 Official Languages and Number of Co-Official, Minority and Regional Languages per Country

Country	Abbr.	First official language	Co-Official Languages	Minority and Regional Languages
Austria	AT	German	3	
Belgium	BE	Dutch, French and German		
Czech Republic	CZ	Czech		3
Denmark	DK	Danish		6
Estonia	EE	Estonian		4
Finland	FI	Finnish and Swedish	1	
France	FR	French		7
Germany	DE	German		5
GB England	GB (E)	English		
GB Scotland	GB (S)	Scottish-English		1
GB Wales	GB (W)	English and Welsh		
Greece	GR	Greek		1
Hungary	HU	Hungarian	6	6
Ireland	IE	Irish language English		
Italy	IT	Italy	2	9
Latvia	LV	Latvian	1	1
Lithuania	LT	Lithuanian	4	3
Luxembourg	LU	Letzeburgesch, French and German	2	
Malta	MT	Maltese	1	
Netherlands	NL	Dutch	1	
Poland	PL	Polish	1	8
Portugal	PT	Portuguese	1	
Republic of Cyprus	CY	Greek and Turkish		5
Slovakia	SK	Slovak	3	2
Slovenia	SI	Slovene	2	
Spain	ES	Spanish	5	
Sweden	SE	Swedish	2	5

Quite a few countries have co-official languages. Denmark and Estonia have none, while the Netherlands (Frisian) and Malta (English) have one. Hungary has six co-official languages: Roma, German Slovak, Croatian, Serbian and Slovenian. These minorities have a fairly developed education system of their own with instruction in their mother tongue.

Countries with the most minority or regional languages are France (Breton, Catalan, Corsican, Occitan, Basque, Alsatian, and Flemish), Italy (Albanian, Catalan, German, Greek, Croat but also languages as Provençal-French, Friulan, Languedoc and Sardinian) and Poland

(Kashubian, Byelorussian and Ukrainian, but also Lithuanian, Slovak, Lemkov, Greek and Hebrew).

An elaborate description of co-official languages, minority and regional languages can be found in appendix 4.

Duration of Compulsory Education and Secondary Education

Policies regarding the conclusion of compulsory education differ among the E.U. countries. Some countries conclude compulsory education with national exams, which include foreign language assessment. Other countries do not administer national exams. Also, the situation in the Netherlands has proven that national exams can exist without laws that make education compulsory. It is important to note the differences between types of schools and the ages at which students finish (that part of) compulsory education and secondary education. The duration of compulsory education can differ from the duration of secondary education.

Table 5.2 Duration and Ages for the Completion of Compulsory Education and Secondary Education

Country	Compulsory education		Secondary education	
	Duration	Until the age of	Duration ¹	Until the age of
Austria	9	14	9	19
Belgium	9	18	6	18
Cyprus	9	14	6	18
Czech	9	15	9	19
Denmark	9	17	6-7	17
Estonia	9	16	6	18
Finland	9	15	6-7	18/19
France	10	15	7	18/19
Germany	9	14	9	19
GB England	11	16	8	19
GB Scotland	11	16	8	19
GB Wales	11	16	8	19
Greece	9	14	6-7	17/18
Hungary	13	18	7-8	17/18
Ireland	9	16	6-7	17/18
Italy	9	14	6-8	17/19
Latvia	9	15	8	18/19
Lithuania	10	15	7-8	18/19
Luxembourg	9	14	7	19
Malta	11	16	6-8	16/18
The Netherlands	12	16	5-6	17/18
Poland	9	16	6-7	18/19
Portugal	9	14	6	18
Slovakia	10	15	7	18/19
Slovenia	9	14	7-8	18/19
Spain	10	15	6	18
Sweden	9	16	7	19

The range of the number of years of compulsory education is from 9 – 13. In most countries it is nine years, while Hungary (13 years) and The Netherlands (12) have the longest span of compulsory education. In most countries, compulsory education lasts until the age of 14, 15 or 16, with the notable exceptions of Belgium, Denmark and Hungary.

¹ The number of years refers to the total time span in lower and upper secondary education or first and second cycle of secondary education. In some countries the first cycle of what can be seen as secondary education is legally part of primary education or basic education. Details of each country can be found in the reports in appendix 5.

The duration of secondary varies considerably between the EU-countries. Some countries clearly define the number of years allocated to secondary education, for example in Luxembourg (7 years), Portugal (6 years) and the Czech Republic (9 years). Other countries allocate three or four years to lower secondary education, while upper secondary education varies in length. Besides length in several EU-countries the types of upper secondary education differ considerably. In Hungary pupils begin upper secondary education in the 9th grade, after the age of 14, and continue up to the upper secondary school leaving examination or until leaving the secondary school system. Upper secondary education is organised in separate school types, varying per educational programme. The types of institutions providing education at upper secondary level in Hungary are: the general secondary school, the vocational secondary school, the short-term vocational secondary school, the trade school and the special trade school. In Swedish upper secondary education is from the age of 16 to 19 and is not compulsory, although 97% (or so) of the pupils chooses to do the not compulsory part as well. In the upper secondary school all education is organised in 17 national programmes. All the upper secondary school programmes contain the same eight core subjects: Swedish or /Swedish as a second language, English, civics, religion, mathematics, science studies, physical education and health as well as artistic activities. In addition to these core subjects, pupils study subjects specific to a programme. There is both vocational and general upper secondary education.

The end of the first cycle of secondary education is not the equivalent of the end of compulsory education. In most EU-countries, the first cycle of secondary education or lower secondary education is part of compulsory education. On the other hand, the duration of the second cycle or 'upper' secondary education is very diverse and, depending on the country, upper secondary education can be compulsory, partly compulsory or not compulsory at all.

5.3 Objectives and Types of Assessments

Objectives

Policy makers and test experts from 16 countries generously provided the overall aims for foreign language learning in secondary education.

Eight possible types of aims for foreign language learning in secondary education were distinguished: socio-cultural knowledge, knowledge of literature, knowledge of a foreign language, proficiency (in one of the four skills), communicative skills, transfer of learning, learning strategies and attitudes.

Proficiency in the four skills was mentioned in each of the responses. In some countries, these skills are mentioned for both types of secondary education. An example from **Denmark**:

“Lower Secondary: The aim is that pupils acquire knowledge and proficiency so that they can understand spoken and written English/ French /German and can express themselves orally and in writing. The teaching shall also develop the pupils’ awareness of the English/French /German language and usage as well as of language learning.

Upper Secondary: Here there are two levels: A, advanced and B, obligatory. Listening, speaking, reading and writing are common to both levels. In the advanced level there is a special focus on language and culture (for English, French or German)”

In other countries the four skills are integrated into a wider set of skills/objectives.

Sweden: “the overall aim is to develop an all-round communicative ability and language skills necessary for international contacts and should provide pupils with opportunities to use the English language and develop their ability. In the syllabus five main areas are in focus: a) receptive skills, b) productive skills, c) interactive skills, d) cultural awareness, e) reflection and responsibility for one’s own learning.

Sometimes oral skills were rated higher than the other skills. The overall aims provided by the Czech Republic, Denmark, Finland and France explicitly refer to the levels specified in the CEF. **Czech Republic:** “the ‘new curriculum’ for both gymnasiums and vocational and technical secondary schools defines language learning goals as goals to achieve a communicative competence of students at either B1 or B2 CEF-level”. The overall aim of foreign language learning in Finnish secondary schools is heading in the same direction. **Finland:** “According to the new national framework curriculum for general secondary education the level of competence of an average student should, at the end of secondary education, be B1-B2 (long syllabus) and A2-B1 (short syllabus) on the CEF scale. Within vocational education, foreign language studies are less comprehensive and the attainment level has not been defined on the CEF scale”.

Achieving communicative competence is another aim included in all sets of aims. **The Netherlands:** “To use language to communicate, receptively and productively, verbally and in writing. To acquire insight in societies and cultures”. The (meta) knowledge of a foreign language and socio-cultural knowledge are mentioned in some of the aims presented. In **Hungary** aims for the foreign languages are described in the core curriculum: “The fundamental aim of teaching modern foreign languages is to develop communicative competences in the given language. This refers to the ability to use a language in the manner demanded by the situation at hand. Language proficiency can be measured and assessed through the four basic language skills (listening, speaking,

reading and writing). The development of communicative competence in a language involves the following:

- (1) By the end of compulsory education, pupils must be able to use one or two foreign languages appropriately in personal, educational, public and professional contexts;
- (2) While learning foreign languages, pupils should develop and maintain a positive, motivated attitude to language learning, the learnt language, the people speaking that language, their culture and learning about other languages and cultures in general;
- (3) Pupils must learn to be able to maintain the acquired level of language proficiency themselves throughout their lives and to learn other foreign languages efficiently and effectively."

Sometimes the overall aim of foreign language learning in secondary education is split up into different types of goals. In the case of **Cyprus** the transfer of knowledge and learning to learn is among them:

- "Proficiency goals: general competency, mastery of the four skills;
- Cognitive goals: mastery of linguistic and cultural knowledge;
- Affective goals: achieving positive attitudes about the foreign language;
- Transfer goals: learning how to learn and life long learning;
- Communicative goals: for everyday situations in both the oral and the written mode."

The transfer of learning and attitudes towards foreign language learning were mentioned in just three sets of aims. Knowledge of the literature was not mentioned once, in the objectives that were offered to this study.

Types of Assessments

As mentioned in chapter 3, this study distinguishes between three major types of assessments, i.e. standardised assessments, nation-wide assessments and other assessments. The category 'other assessments' is subdivided into school-based assessments, portfolio assessments and continuous assessments. The category 'general teacher assessments' has been omitted, as it can be safely assumed that teachers assess students generally and on a very regular basis in each of the 25 EU countries.

Below is an overview of the different types of assessments in use in secondary education in EU-countries.

Table 5.3 Type of Assessments Used in Secondary Education in Europe

Types of Assessment	N	Percentage the total of responses
Nation-Wide assessment	22	78.6%
Standardised assessment	4	14.3%
Portfolio assessment	12	42.9%
School-Based assessment	24	85.7%
Continuous assessment	21	75.0%

Clearly, three types of assessment stand out as the most frequently used: school-based, nation-wide and continuous assessment.

School-based assessments, in general, can materialize in two ways. The practice of education is based on schools trying to achieve a set of objectives, and testing is generally used to generate feedback on the skills and attitudes of students during this process. The emphasis, then, is on administering grade-based assessments of student performance. This type of school-based assessment is applied in all European countries. School-based assessments can also be part of a nation-wide assessment or exam that concludes the first or second cycle of secondary education. In this case, schools can be made individually responsible for assessing students' foreign language skills as part of the nationally administered exam. Also, throughout Europe, there are several educational systems in which only school-based assessments are administered, for example in Belgium (Flemish, French and German speaking communities), the Czech Republic and Germany (all Bundesländer and independent cities). There are ten EU-countries in which school-based assessments are conducted during or at the end of secondary education, in combination with external assessments.

Nation-wide assessments are carried out in twenty-two EU countries. Nations that do not use these assessments, as mentioned above, include Germany, the Czech Republic and Belgium. Nation-wide assessments can take different shapes and forms. Firstly, they can take place at the end of secondary education or at another given moment. Secondly, they may consist of series of tests issued to schools, of which participation may be voluntary. Thirdly, nation-wide assessments can be a combination of tests for certain skills at the school level and external assessments provided at the national level.

No more than four countries regularly employ standardised assessments. Standardised assessments use instruments developed from shared aims. While students may have

followed different curricula, they participate in a standardised assessment under the same conditions and must fulfil the same requirements. The scores are based on predetermined standards, but the use and significance of the results can vary for each student. An example of a standardised assessment is that carried out in Finland. Though it only incorporates a sample of pupils in grade nine, the assessment covers a number of languages: English, French, Finnish, Swedish, German and Russian. In addition, the full range of communicative competencies is assessed, from listening comprehension to writing and integrated skills. Finland holds this assessment every year. France also uses standardised assessments. These are sample-based, standardised tests used for system monitoring every four or five years. English, German and Spanish are assessed using a sample of pupils from grade nine, and the assessment focuses on listening, reading, writing and integrated skills. In Scotland the national Assessment of Achievement Programme (AAP) is in operation, in which a large representative sample of pupils in Primary 4, Primary 7 and Secondary 2. In the last AAP for English there was a Modern Languages study at Primary 7 and Secondary 2.

The concept of continuous assessment seems to be open to interpretation. For some test experts that were consulted, continuous assessment means keeping a track record of what pupils do and how they perform in a foreign language. For others, it explicitly involves keeping an in-depth record of performances in several foreign languages, continually assessing pupil performance with different types of tests. While it sometimes constitutes broad sets of informal techniques, continuous assessment is one of the most applied types of assessment in Europe.

Portfolio assessments are carried out in twelve countries. In most cases, however, respondents stated that portfolio assessment was still at a very preliminary stage. The use of this type of assessment is not wide spread, and one can safely conclude that it is more likely to be found in innovation driven schools.

5.4 Reasoning behind Current Assessment Cultures

Functions of Different Types of Assessments

Each educational system, including its approach to assessing foreign languages, is the result of long historical and political processes. The roots of modern foreign language teaching, however, are not the subject of this study. Each country's legislation accurately describes how its educational system should operate, but the reasoning behind its foreign language assessment culture is harder to ascertain. This study looked to national test experts and policy-makers for this information. Using open questions would have resulted

in a number of methodological problems, such as scope, time, interpretation and comparability; therefore experts and policy-makers were asked to give the main and subsidiary functions of assessments and the rationales behind the assessments.

Table 5.4 Main Functions of Assessments in Secondary Education

Main Functions	Types of Assessments		
	Nation-Wide N = 17	Standardised N = 4	Other types N = 24
Examination	16 (94.1%)	1 (25.0%)	15 (62.5%)
Certification	15 (88.2%)	1 (25.0%)	13 (54.2%)
Transparency	7 (41.2%)	1 (25.0%)	2 (8.3%)
System monitoring	5 (29.4%)	2 (50.0%)	3 (12.5%)
Guidance for students' career	6 (35.3%)	1 (25.0%)	14 (58.3%)
Determining school quality	2 (11.8%)	0 (0.0%)	1 (4.2%)
Leverage for educational change	0 (0.0%)	1 (25.0%)	1 (4.2%)
Information for parents	4 (23.5%)	0 (0.0%)	16 (66.7%)

The two most important functions of nation-wide assessments of foreign languages are examination and certification. Transparency, system monitoring, guidance for students' educational career and information for parents are important functions for a little less than half of all countries. In six countries transparency is the main function of a nation-wide assessment of foreign languages and in five countries it is a subsidiary function. Austria is the only country in which the main function of nation-wide assessments is to determine school quality.

By far, if this can be said considering the small amount of cases, the most important function of standardised assessments of foreign languages is system monitoring. In some countries standardised assessment also have the function of providing guidance for the students' career and leverage for educational change.

The main function of the other types of assessments is in over half of the countries examination and certification. As can be seen from the nation reports, in these countries school-based assessment are mostly part of national examinations. In addition to that it is rather obvious that in nearly two-third of all countries the guidance for students educational career is the main function of these other type of assessments. And providing

information to the parents is in two-third of all countries, using other types of assessments a main function.

Table 5.5 Subsidiary Functions of Assessments in Secondary Education

Subsidiary Functions	Types of Assessments		
	Nation-Wide N = 17	Standardised N = 4	Other types N = 24
Examination	1 (5.9%)	0 (0.0%)	0 (0.0%)
Certification	1 (5.9%)	0 (0.0%)	1 (4.2%)
Transparency	5 (29.4%)	1 (25.0%)	6 (25.0%)
System monitoring	10 (58.9%)	2 (50.0%)	8 (33.3%)
Guidance for students' career	7 (41.2%)	1 (25.0%)	2 (8.3%)
Determining school quality	4 (23.5%)	1 (25.0%)	5 (20.8%)
Leverage for educational change	7 (41.2%)	0 (0.0%)	3 (12.5%)
Information for parents	7 (41.2%)	0 (0.0%)	5 (20.8%)

System monitoring is the most frequently mentioned subsidiary function in nation-wide assessments. But guidance for students' career, leverage for educational change and providing information for parents are also mentioned by more than 40% of the respondents as subsidiary functions. System monitoring is also the most frequently mentioned subsidiary function for the other types of assessments, while transparency, determining the school quality and providing information for parents are also mentioned by (a relatively few) respondents.

Rationales

What became clear from the above is that nation-wide assessment is in many perceptions the equivalent of a national examination at the end of lower or upper secondary education. However, it has been shown that this is not always the case. Tests are used that have been developed from shared aims. Students meet the same conditions and must fulfil the same requirements. Still, the curricula students have followed can differ. The significance of the scores is the same for all students. Since nation-wide assessments fulfil several functions and have an impact on a broad audience it is opportune to look into the rationales a bit deeper. What are the most important rationales for nation-wide assessments?

Table 5.6 Rationales for Nation-Wide Assessments in Secondary Education from the Perceptions of Test Experts and Policy Makers

Rationales	Test Experts N = 17	Policy Makers N = 14
Utilitarian		
Importance of the skill	13 (76.4%)	11 (78.6%)
Political		
Comparability of outcomes	8 (47.1%)	9 (64.3%)
System control	5 (29.1%)	5 (35.7%)
Sociological		
Fairness	3 (17.6%)	6 (42.9%)
Equality of opportunities	7 (41.1%)	4 (28.6%)
Technocratic		
Consistency over the years	8 (47.1%)	4 (28.6%)
Quality of the assessment	7 (41.1%)	8 (57.1%)
Practical		
Necessity for further exams	2 (11.8%)	3 (21.4%)
Comply with national curriculum	1 (5.9%)	3 (21.4%)

There is a wide set of rationales within each country to hold nation-wide assessments. There does not seem to be a single 'leit motif' to operate a nation-wide assessment. Just below eighty percent of the respondents placed "the importance of the skill" among the most important rationales for assessing the communicative competencies in nation-wide assessments. A utilitarian motive seems to prevail the most. For the test experts a political motive, comparability of outcomes, and a technocratic motive, consistency over the years, are of importance as well. Policy makers find that political motive important as well, but also the technocratic motive of quality of the assessment. Sociological motives, such as fairness and equality of opportunities are less highly valued. Especially a few policy makers mention very practical motives, such as that the assessment needs to comply with the goals and content specified in the national curriculum.

Side Effects

So far the results have focused on the main and subsidiary functions of assessments and the rationales behind the nation-wide assessments. Test experts and policy makers had the option of stating what kinds of side effects they perceived.

Table 5.7 Side Effects of Assessments in Secondary Education

Side Effects	Types of Assessments		
	Nation-Wide N = 18	Standardised N = 4	Other types N = 24
Examination	0 (0.0%)	0 (0.0%)	1 (4.2%)
Certification	0 (0.0%)	0 (0.0%)	1 (4.2%)
Transparency	3 (20.0%)	0 (0.0%)	7 (29.2%)
System monitoring	2 (11.8%)	0 (0.0%)	3 (12.5%)
Guidance for students' career	5 (27.8%)	0 (0.0%)	3 (12.5%)
Determining school quality	11 (64.7%)	0 (0.0%)	3 (12.5%)
Leverage for educational change	7 (38.9%)	2 (50.0%)	6 (25.0%)
Information for parents	6 (33.3%)	0 (0.0%)	7 (29.2%)

The most important side effect of a nation-wide assessment is determining school quality. Guidance for students' career, leverage for educational change and providing information for parents are also mentioned frequently. In the context of standardised assessments, the only side effect that is reported is that the information can be used as leverage for educational change. The three most frequently mentioned side effects from other types of assessments are transparency, leverage for educational change and providing information for parents.

Changes in Government Policy on the Assessment of Foreign Languages

Of course, changes in policy towards the assessment of foreign languages are always part of a broader nation-wide change in policy towards (secondary) education. Still, policy makers were asked if and what kind of changes in government policy on the assessment of foreign languages had occurred over the past five years and if there were changes to be expected.

Table 5.8 Changes in Government Policy on the Assessment of Foreign Languages

Changes	Over the past 5 years	In the next 5 years
No changes	6	7
Minor changes	1	0
Major changes	12	11

No changes in government policy on the assessment of foreign languages have occurred in Finland, France, Latvia, Portugal, The Netherlands or Sweden. Minor changes were reported by the Flemish speaking community of Belgium. Major changes over the past five years have taken place in Belgium (French speaking part), Cyprus, the Czech Republic, Denmark, France, in the German Bundeslander Baden-Wurtemberg and Saxony, Hungary, Lithuania, Malta, Slovenia and Scotland. In Scotland and Hungary no major changes are expected in the next five years. The other countries that have undergone major changes expect more change to come, or are in the midst of the process of change. Similarly, countries that reported no changes in the last five years expect no changes in the foreseeable future, with the exception of Sweden. It is still very uncertain, but in Sweden, changes in the function of assessments and the system of quality control are probably being discussed.

What types of changes have occurred over the past five years? The introductions of new goals and of other types of assessments have been at the centre of the changes and innovations.

Table 5.9 Changes in Government Policy on the Assessment of Foreign Languages in the Past Five Years (N = 16)

	Number of countries	Countries
No changes	6	FI, FR, LV, NL, PT, SE
Function of assessments has changed	6	CY, CZ, DK, LT, MT, SK
Another type of assessment has been proposed and / or implemented	8	CY, CZ, DE (S), DE (Bw), DK, HU, LT, M
New goals for foreign language education have been implemented	9	BE (Ds), BE (Fs), CY, CZ, DE (S), DK, HU, MT, SK
Intensity of assessments has increased	3	CZ, MT, SK
Intensity of assessments has decreased	0	
A new system of quality control through assessments has been implemented	5	CZ, DE (S), DE (Bw), MT, SK
Other	5	CZ, DE (Bw), DK, FR, GB (SCO)

The function of assessments has changed in Cyprus, the Czech Republic, Denmark, Lithuania, Slovakia and Malta. This is connected to the fact that in these countries, with the exception of Slovakia, another type of assessment has been proposed and / or implemented. This is also the case in Germany. New goals for foreign language education were drafted and implemented in the Dutch and French speaking communities of Belgium, as well as in Cyprus, the Czech Republic, Denmark, Hungary, Lithuania, Malta and Slovakia. Only three countries (the Czech Republic, Malta and Slovakia) reported an increase in assessments and no country indicated a decrease. The Czech Republic, two German Bundesländer, Malta and Slovakia reported the implementation of a new system of quality control through assessments.

Other types of changes differ very much from country to country. In France a standardised assessment at the primary level was prepared and will be implemented. Scotland reports that a new qualifications system takes a modular based approach in courses. This will bring an assessment through “unit assessments undertaken within schools and national exams”. In Denmark transparency is a key factor. “The current government in Denmark wants to establish coherence in the assessment – systems at all levels. This in order to clarify expectations and thus enhance quality in education in general.” In Germany, work has been done towards introducing State Exams. All changes in the Czech Republic with regard to the assessment of foreign languages bear a

relationship to the implementation of the new school act that will come into force on the 1st of January 2005. However, what about the future? What kinds of changes in government policy on the assessment of foreign languages are expected?

Table 5.10 Expected Changes in Government Policy on the Assessment of Foreign Languages

	Number of countries	Countries
No changes perceived	7	FI, FR, GB (SCO), HU, LV, PT, NL
Function of assessments will change	3	CY, CZ, DK
Another type of assessment will be proposed and / or implemented	5	BE (Ds), CY, CZ, DK, M
New goals for foreign language education will be implemented	4	CY, CZ, DK, SK
The intensity of assessments will increase	4	CZ, DK, MT, SK
The intensity of assessments will decrease	0	
A new system of quality control through assessments will be implemented	6	BE (Ds), BE (Fs), CZ, DE (Bw) MT, SK
Other	3	CY, LT, SE

Belgium (the Dutch speaking community), Cyprus, the Czech Republic, Denmark, and Malta will implement a new type of assessment. The Czech Republic and Denmark will do that with new goals and an increase in the intensity of assessments. Six European countries will implement a new system of quality control through assessments. Cyprus and Lithuania are preparing to incorporate the European Language Portfolio, while in Lithuania initiatives are also being taken to establish relationships between exams and the CEF.

Resumé

The two most important functions of nation-wide assessments and standardised assessments of foreign languages are examination and certification. Transparency, system monitoring, guidance for students' educational career and information for parents are important functions for a little less than half of all countries. Transparency is the main function of a nation-wide assessment of foreign languages in six countries and in five countries it is a subsidiary function. In countries where school-based assessments are part of national

examinations, examination and certification are the main functions for other types of assessments. In addition, predictably, guidance for students' educational careers is a main function of other types of assessments in nearly two-thirds of all countries. Another main function of other assessments is to provide parents with information.

Responses about the rationales behind assessments produced no clear-cut picture. "The importance of the skill" is among the most important rationales for assessing the communicative competencies in nation-wide assessments. It seems that a utilitarian motive prevails. Test experts also cited a political motive- comparability of outcomes -and a technocratic motive- consistency over the years. Policy makers find comparability of outcomes important as well, but they also mention the technocratic motive of quality of the assessment. Sociological motives, such as fairness and equality of opportunities are less highly valued.

In interpreting these results, one must keep in mind that there is a group of about eight EU-countries in which there have been no recent changes in educational policy towards the assessment of foreign languages. A larger group of 11 EU-countries reported definite changes over the past few years, many of which are ongoing, and this can result in changes in functions of assessments, systems of quality control, and other aspects.

A direct relation between the objectives stated and the types of assessment used can not be established. On the one hand there are to little similarities between objectives stated in different countries. On the other hand it has been shown that within countries there are several types of assessment in operation that are in some instances complementary and sometimes independent. This diversity in objectives and configuration of types of assessments within countries makes every kind of analysis superficial.

5.5 Intensity

In an ideal world, available data would show the total number of foreign language assessments per school, as well as the amount of time put into their development and the skills they assess. This is not the case. Due to financial and organisational restraints, no data on the total number of assessments of foreign languages in secondary education could be gathered. Nor was it possible to consult databases containing this information. In the absence of such data, this study uses four indicators to determine the intensity of foreign language assessment in secondary education:

the number of types of assessments used;

the number of nation-wide assessments during secondary education;

the number of foreign languages assessed in secondary education, and;

the number of assessments of different communicative competencies.

This section presents results using these indicators. In particular, the number of assessments of different communicative competencies is analysed in detail. This indicator covers several aspects of indicators 1, 2 and 3, and supplies enough information for determining each country's position on the scale from low to high intensity of assessment.

Number of Assessments Used

Test experts and policy makers were asked about the different types of assessments used. The minimum score is theoretically 1, when only one type of assessment (for example just school-based assessment or just continuous assessment) is used. The maximum score is 5, when nation-wide assessments, standardised assessments, school-based assessments, portfolio assessments and continuous assessment are carried out. These different types of assessments are not easily compared. Their purpose, influence and scope are rather different. The purpose of the analysis is exploration.

Table 5.11 Total Number of Types of Assessments Used (N = 26)

Number of assessments	Number of Countries	Countries
1. Minor	0 (0.0%)	
2. Some	9 (34.6%)	AT, BE (Ds), BE (Fs), DE (Bw), GB (E), LU, MT, PT, SK
3. Several	7 (26.9%)	CZ, EE, FR, HU, IE, NL, SI
4. A Wide range	10 (38.5%)	DE (S), DE (Ns), DK, FI GB (SCO), IT, LV, LT, PL, SE
5. All types	0 (0.0%)	

On average, three different types of assessments are used. In countries that only employ 'some' types (two), the emphasis is generally on school-based assessment in combination with continuous assessment. However, in England, school-based assessments are conducted as well as nation-wide assessments at different stages (key stage 3, GCSE, optional foreign languages and A level exams). A wide range of types of assessments is used in ten countries. However, these countries do not necessarily use the same types of assessments. France combines continuous, nation-wide and standardised assessments, but does not use school-based or portfolio assessments. In contrast, Denmark carries out

nation-wide assessments in combination with continuous, portfolio and school-based assessments.

Number of Nation-Wide Assessments during Secondary Education

Section 5.3 showed that 22 EU-countries hold nation-wide assessments. The scheduling of these assessments varies per country (see paragraph 5.5). Three quarters of the countries that hold nation-wide assessments administer them at the end of the second cycle of secondary education. In most cases this is nearly equal to the end of compulsory education. Ten countries administer nation-wide assessments of foreign languages at the end of the first cycle of secondary education. The total number of nation-wide assessments per country can be seen below.

Table 5.12 Number of Nation-Wide Assessments during Secondary Education (N = 21)

Number of Nation-Wide Assessments	Number of Countries	Countries
One Assessment	12 (57.1%)	AT, CY, DK, FI, GR, HU, LT, MT NL, PL, PT, SK
Two Assessments	3 (14.3%)	FR, IE, SI
Three or more Assments	6 (28.6%)	EE, GB (E), GB (SCO), LU, LV, SE

The majority of EU-countries that conduct nation-wide assessments in secondary education hold just one such assessment. Three or more nation-wide assessments are held during secondary education in six countries, though the natures of these differ per country. Sweden's nation-wide assessments serve the sole function of guiding pupils' educational careers, and teachers and schools may choose not to use these tests. Nation-wide assessments in Luxembourg are mandatory. In England, assessments are offered at several stages and schools can choose the organisation that provides the assessment.

Number of Foreign Languages Assessed in Secondary Education

Some countries offer just one foreign language in the first cycle of secondary education, but add several more to the curriculum in the second cycle. The number of foreign languages offered varies per country, as outlined in section 5.2. Below is an overview of

the number of foreign languages assessed in lower (1st cycle) and upper (2nd cycle) secondary education.

Table 5.13 Number of Foreign Languages Assessed in the First and Second Cycles of Secondary Education (N = 28)

Number of foreign languages assessed	Secondary Education			
	First cycle (lower)	Countries	Second cycle (upper)	Countries
no foreign language	7 (36.8%)	CY, FI, HU, LT PL, PT, SI	1 (5.6%)	AT
1 foreign language	2 (10.5%)	AT, SE	0 (0.0%)	
2 foreign languages	2 (10.5%)	LU, SI	0 (0.0%)	
3 foreign languages	2 (10.5%)	DK, GB (E)	1 (5.6%)	LU
4 foreign languages	2 (10.5%)	EE, LV	8 (44.4%)	CY, EE, FI, GB (E) HU, LT, LV, PT
5 foreign languages or more	4 (21.1%)	FR, IE, NL, MT	8 (44.4%)	DK, FR, IE, MT NL, PL, SI, SK

Seven EU-countries do not hold nation-wide assessments that include foreign languages. Five countries offer assessments for more than 5 foreign languages: France, The Netherlands, Ireland and Malta. There is more uniformity in the second cycle of secondary education, in which most countries assess 4 or more foreign languages. Notable exceptions are Austria and Luxembourg.

Assessments of Different Communicative Competencies

The latter section showed that there is little difference between the EU-countries as far as the number of foreign languages assessed is concerned. However, this is complicated when taking into account the assessments of communicative competencies. One point is assigned for each communicative competency incorporated into the one of the various types of assessment. The different types of assessment are (when relevant in the first and second cycle of secondary education):

1. Nation-wide assessment;
2. Standardised assessment;
3. Continuous assessment;

4. Portfolio assessment;
5. School-based assessment;

A sum score is given for each country, representing the number of assessments of each of the communicative competencies.

Table 5.14 Number of Assessments of Communicative Competencies per EU-Country

Countries	Listening	Speaking	Spoken interaction	Reading	Writing	Integrated skills	Sum score
Austria	7	7	7	7	7	6	41
Belgium (Ds)	6	6	6	6	6	6	36
Belgium (Fs)	6	6	6	6	6	0	30
Cyprus	8	6	6	7	7	7	41
Czech Republic	3	7	4	6	6	3	29
Denmark	3	6	6	6	6	6	33
Estonia	8	6	8	8	8	6	44
Finland	7	6	6	7	7	6	39
France	8	7	7	9	9	7	47
Germany (Bw)	3	3	3	3	3	0	15
Germany (Ns)	6	6	6	6	6	6	36
GB (England)	2	2	0	2	2	0	8
GB (Scotland)	4	4	4	4	4	4	24
Hungary	0	6	6	6	6	6	30
Ireland	5	1	1	8	8	1	24
Italy	0	6	0	6	6	0	18
Latvia	8	8	7	8	8	3	42
Lithuania	1	1	0	1	1	6	10
Luxemburg	7	7	7	8	8	7	44
Malta	3	3	0	7	7	7	27
Netherlands	8	6	8	8	8	0	38
Poland	7	7	7	7	7	7	42
Portugal	0	3	3	4	4	0	14
Slovakia	7	7	7	7	7	7	42
Slovenia	1	4	1	4	4	0	14
Sweden	8	8	8	8	8	8	48

The average score of the total number of communicative competencies that are assessed is 31.4 (s.d. 12.2). Based upon this average score and the standard deviation, the countries can be divided into four groups:

Score 1 – 18; a low intensity of assessment of foreign language skills (England, Lithuania, Portugal, Slovenia, Germany (Baden Wurtemberg) and Italy);

Score 19 – 30; a moderate intensity of assessment of foreign language skills (Ireland, Scotland, Malta, Czech Republic, Belgium (French speaking community), Hungary);

Score 31 – 40; an above average intensity of assessment of foreign languages skills (Denmark, Belgium (Dutch speaking community), Germany (Niedersachsen), The Netherlands, Finland);

Score 41 – 60; a high intensity of assessment of foreign language skills (Austria, Cyprus, Latvia, Poland, Slovakia, Estonia, Luxemburg, France and Sweden).

5.6 Structure of Tests

Foreign Languages Offered

The most recent comparative study on foreign languages in primary and secondary education in the EU is that published by Eurydice in 2001. This publication, however, uses data and descriptions stemming from 1998/1999. In the report, various official documents (usually curricula) indicate a limited range of options for students choosing a compulsory first foreign language, with most EU countries offering no more than six languages. Exceptions include Austria (up to eight languages), France (up to 12 languages) and the United Kingdom (England and Wales – up to 19 languages). None of the pre-accession countries, however, are said to offer more than six languages. Precise lists of foreign languages are not given for Spain, Finland and Hungary. The study finds that most EU countries do not offer a broader range of languages to students choosing a second compulsory foreign language, with the exception of the Netherlands (where pupils may, in theory, choose between seven languages) and France (15 languages). Again, this number does not exceed six in the pre-accession countries. In most cases, the choices for first and second foreign languages are the same.

As the report mentions, first foreign languages are often set by the national curriculum, with English being the required first foreign language in Denmark, Greece, the Netherlands, Sweden, Liechtenstein, Norway, Cyprus and Latvia. Similarly, French is the first foreign language in Belgium (German-speaking Community), German in Luxembourg, and Danish in Iceland. The Eurydice study shows this can also be the case with second foreign languages: French is the required second foreign language in Luxembourg, Liechtenstein, and Cyprus (in the *Gymnasio*), whereas English is required in Iceland. Also,

English is found to be the only official EU language offered as a first and/or second compulsory foreign language in all of the countries that provided lists of languages – with the exceptions of Belgium’s German-speaking Community and Luxembourg. French and German are also among the languages offered in a large number of countries. Spanish and/or Italian are offered in a number of EU countries (including France, Italy, the Netherlands, Austria, Portugal, Sweden, and Great Britain) as well as in some the pre-accession countries, i.e. Bulgaria, the Czech Republic, Romania, and Slovakia. Russian is also found in all pre-accession countries (with the exception of Cyprus and Slovenia), as it is in France, Italy, the Netherlands, Austria, and Great Britain (E/W, SCO).

It is important to note that since the Eurydice report, the European Union has expanded to include 10 new countries. Also, six years have passed, and some policies and practices for foreign language teaching have changed. It is not within the scope of this study to simulate the immense research carried out by Eurydice. However, the following accounts of the foreign languages offered in secondary education are provided as an example of the differences between now and then.

In addition to English, Swedish secondary schools offer French, German and Spanish as a foreign language. Schools are required by law to make at least two of the languages available to pupils. Danish secondary schools offer English, French, Spanish and Russian.

Hungarian schools offer English, French, Spanish and several languages from neighbouring countries, such as Croatian, Slovak, Slovene and German. Slovakian schools for secondary education must offer at least one foreign language, which one depends on the circumstances at each school. Recommended languages are English, French, German, Russian and Spanish.

Latvian schools also offer English as the first foreign language, and pupils choose a second starting in the 6th form. The most popular foreign languages are German, Russian and French.

German secondary schools may offer English, French, Spanish and Italian, and have the option of adding languages from one of the neighbouring countries: Danish, Dutch or Polish. In the Netherlands pupils may, in theory, choose between seven languages. In practice English is predominant, while French and German are also widely offered. Foreign languages offered in the first cycle of French secondary education include English and languages from neighbouring countries, namely Spanish, German, Italian and Dutch. Other options are Chinese, Danish, Japanese, Modern Greek, modern Hebrew, Polish, Portuguese, Russian, standard Arabic and Turkish. This adds up to a total of 15 options, and this number even increases to 44 in the second cycle of French secondary education.

In Spain, schools for secondary education are required by law to offer English, Italian, German or a language from a neighbouring country such as French or Portuguese. Until

now, English was the only compulsory foreign language in Italian secondary schools. Starting with the school year 2004/2005, a second compulsory foreign language from the European Community (generally French, German or Spanish) will be introduced in the first cycle of secondary education.

In the United Kingdom, the framework of the statutory curriculum does not require any specific foreign language to be taught. However, schools are required to offer at least one of the official working languages of the European Union. Starting in September 2004, modern foreign languages will no longer be a required subject for pupils aged 14-16, though it will be a statutory 'entitlement area' i.e. schools must provide the opportunity for students to study a modern foreign language.

Foreign Languages Assessed

Below are two overviews showing the foreign languages assessed in or at the end of the first cycle of secondary education (table 5.15) and at the end of the second cycle of secondary education (table 5.16).

Table 5.15 Number of Foreign Languages in Nation-Wide Assessments in the First Cycle of Secondary Education per EU-country

EU-country	Foreign Languages						Total		
	English	French	German	Spanish	Italian	Russian		Other	
Austria		x					0	1	
Belgium	Ds	No Nation-wide assessments							
	Fs	No Nation-wide assessments							
Cyprus		Not in the first cycle of secondary education							
Czech Republic		No Nation-wide assessments							
Denmark		x	x	x			0	3	
Estonia		x	x	x			x	0	4
Finland		x	x	x			x	0	4
France		x		x	x		x	10	15
Germany		No Nation-wide assessments							
Great Britain	S	No Nation-wide assessments							
	E	Not in the first cycle of secondary education							
Greece		Not in the first cycle of secondary education							
Hungary		Not in the first cycle of secondary education							
Ireland		x	x	x	x	x		1	6
Italy		x						0	1
Latvia		x	x	x			x	0	4
Lithuania		Not in the first cycle of secondary education							
Luxembourg			x	x				0	2
Malta		x	x	x	x			0	5
Netherlands		x	x	x				0	3
Poland		Not in the first cycle of secondary education							
Portugal		Not in the first cycle of secondary education							
Slovakia		Not in the first cycle of secondary education							
Slovenia		Not in the first cycle of secondary education							
Spain		x	x	x		x		0	4
Sweden		x						0	1

Table 5.16 Number of Foreign Languages in Nation-Wide Assessments in the Second Cycle of Secondary Education per EU-country

EU-country	Foreign Languages							Total	
	English	French	German	Spanish	Italian	Russian	Other		
Austria	Not in the second cycle of secondary education								
Belgium	Ds	No Nation-wide assessments							
	Fs	No Nation-wide assessments							
Cyprus		x	x	x		x		4	
Czech Republic		No Nation-wide assessments							
Denmark		x	x	x	x	x	1	7	
Estonia		x	x	x			x	0	4
Finland		x	x	x	x		x	4	9
France		x		x	x	x	x	39	44
Germany		No Nation-wide assessments							
Great Britain	S	No Nation-wide assessments							
	E	Any ²							
Greece		x	x					0	2
Hungary		x	x	x	x	x	x	1	7
Ireland		x	x	x	x	x		1	6
Italy		x	x					0	2
Latvia		x	x	x			x	0	4
Lithuania		x	x	x			x	1	5
Luxembourg		x	x	x				0	3
Malta		x	x	x	x	x		0	5
Netherlands		x	x	x	x		x	3	8
Poland		x	x	x	x	x	x	3	8
Portugal		x	x	x	x			0	4
Slovakia		x		x				0	2
Slovenia		x	x	x	x	x	x	0	6
Spain		x	x	x	x	x		0	5
Sweden		x	x	x	x			0	4

Countries that carry out nation-wide assessments in or at the end of the first cycle of secondary education generally incorporate English into that assessment, with the only exception being Luxembourg. French and German are often included in these

² At age 17 and 18, pupils take GCE, AS and A2 qualifications. Like GCSEs, these are single-subject, externally set and marked qualifications. These qualifications are available in a wide range of modern foreign languages, but only a small minority of pupils include one or more modern language in the three, four or five subjects they study, which are freely chosen from those offered by their school or college.

assessments as well. In most cases the number of languages offered in the first cycle of secondary education is limited to about three to five. France, however, offers assessments for 15 foreign languages.

Nation-wide assessments at the end of secondary education all offer English. French is offered in the nation-wide assessments in all but one or two countries. German, and to a lesser extent Spanish and Italian, are often included in these assessments. Russian is assessed in quite a few countries (10). Other than English, French, German, Italian, Spanish and Russian, few foreign languages are assessed, and when another language is assessed it is usually one from a neighbouring country.

Scheduling of Assessments

EU countries hold nation-wide assessments at various stages of secondary education, as indicated already in tables 5.15 and 5.16.

Table 5.17 Scheduling of Nation-Wide (N = 21)

When	Nation-Wide Assessments	
During the first cycle	2	(9.5%)
At the end of the first cycle	10	(47.6%)
During the second cycle	2	(9.5%)
At the end of the second cycle	16	(76.2%)
At the end of compulsory education	7	(33.3%)

Three quarters of the countries that hold nation-wide assessments administer them at the end of the second cycle of secondary education. There are ten countries that administer nation-wide assessments of foreign languages at the end of the first cycle of secondary education. Portugal, Slovakia and Lithuania, are examples of countries that only hold nation-wide assessments at the end of secondary education. France, Luxembourg and Malta carry out nation-wide assessments at the end of both the first and second cycle of secondary education. In Sweden, nation-wide assessments are held during secondary education, though these tests are not obligatory and teachers may choose not to use them. Also, upper secondary education in Sweden is course-oriented, which means that the assessment at the end of a course is dissimilar to, for example, a nation-wide assessment held in Luxembourg. In France the nation-wide assessments are national

examinations held at the end of grade nine (brevet des colleges) and at the end of grade twelve (baccalaureate).

Communicative Competencies

The Common European Framework of Reference for Languages offers a widely accepted set of communicative competencies. A summary of the number of countries that assess each of these communicative competencies is presented in table 5.18.

Table 5.18 Communicative Competencies Assessed

Communicative Competencies	Secondary Education			
	First Cycle N=13		Second Cycle N=21	
Listening	10	(76.9%)	16	(76.2%)
Speaking	8	(61.5%)	14	(66.7%)
Spoken interaction	8	(61.5%)	13	(61.9%)
Reading	12	(92.3%)	19	(90.5%)
Writing	13	(100%)	19	(90.5%)
Integrated Skills	3	(23.1%)	8	(38.1%)

Reading and writing are nearly always integral components of nation-wide assessments at the end of the first cycle of secondary education. A total of 21 European countries organise nation-wide assessments at the end of the second cycle of secondary education. Nearly all incorporate the assessments of reading and writing, while 16 countries include listening and 14 countries assess speaking. In standardised assessments, reading and writing is incorporated. In two out of four standardised assessments, listening comprehension is part of the communicative competencies assessed. In other types of assessments, the full range of communicative competencies is taken into account. While in nation-wide assessments and standardised assessments the integrated skills are covered in just a few countries, in continuous assessment and portfolio assessment this is done in the majority of the countries.

Types of Assessment Tasks and Modes of Delivery

There is a wide range of assessment tasks available. When test experts were asked which types of assessment tasks were employed in a nation-wide assessment a restriction was made to seven types. This classification was taken from the manual relating language examinations to the CEF (DGIV/EDU/LANG,2003).

Table 5.19 Assessment Tasks Employed in the First and Second Cycle of Secondary Education

Types of assessment tasks	Secondary Education	
	First Cycle N=13	Second Cycle N = 16
Closed questions / multiple choice	10 (76.9%)	12 (75.0%)
Fill in the gap	9 (69.2%)	13 (81.3%)
Short answer (short sentence / completion)	10 (76.9%)	12 (75.0%)
More than one short sentence	6 (46.2%)	9 (56.3%)
Interaction with examiner or peers	7 (53.8%)	12 (75.0%)
Portfolio	1 (7.7%)	0 (0.0%)
Essays	5 (38.5%)	13 (81.3%)

Three types of assessment tasks stand out as the most frequently used in nation-wide assessments in the first cycle of secondary education: closed questions/multiple choice, fill in the gap and short answer (short sentence/completion). Other types frequently used are more than one short sentence, essays and interaction with examiners. In the second cycle of secondary education four assessment tasks stand out: closed questions/multiple choice, fill in the gap, essays and short answer (short sentence/completion).

Obviously the use of assessment tasks is, to a large extent, related to the type of communicative competency that is being assessed. Reading and writing are the most frequently assessed competencies in nation-wide assessments, and this is reflected in the types of assessment tasks chosen. Interaction with examiners is more likely to be used when speaking or spoken interaction is assessed.

Table 5.20 Modes of Delivery Used in Nation-Wide Assessments for each of the Communicative Competencies (N = 15)

Communicative competencies	Audio aid	Visual aid	Computer delivery	Written instructions
Listening comprehension	10	3	0	5
Speaking	0	4	0	5
Spoken interaction	1	4	0	6
Reading comprehension	0	4	0	12
Writing	1	2	0	14
Integrated skills	1	1	0	4

Listening comprehension in nation-wide assessments is most often (in two-third of all countries) carried out using audio aids. A visual aid is used in another three instances. Visual aids are sometimes used for assessing speaking or spoken interaction, though not as often as written instructions. Written instructions are nearly always used for testing reading and writing. When assessing integrated skills mostly written instructions are used.

Choice of Content

Nation-wide assessments are mostly repetitive, as this serves their main function of examination and certification. Aims and objectives are very important for the types of assessment tasks and the chosen modes of delivery. The choice of content remains vitally important for regulating the difficulty of assessment tasks and for ensuring the assessment's continuity over the years. Students and teachers rely on this continuity when preparing for the examination.

Table 5.21 Procedures for Choosing Content for the Assessment of Communicative Competencies (N=18)

Procedures	Number of Countries
Using previous experience	15 (83.3%)
Using checklists	2 (11.1%)
Using similar examples of reliable and valid tasks	6 (33.3%)
Using guidelines	12 (66.7%)
Descriptions of curricula	15 (83.3%)
Calibration to level descriptions	4 (22.2%)

Three procedures for choosing content for the assessment of communicative competencies clearly stand out. Previous experience and descriptions of curricula are used most often, while guidelines are employed in two-thirds of the countries. The use of previous experience is not surprising given the fact that several countries have been holding nation-wide assessments for decades. The curricula describe the essential skills that must be taught, and are a very important source for developers choosing assessment content. The use of guidelines is important in two-thirds of all countries. One-third of all countries use similar examples of valid and reliable tasks, while in four countries content chosen for assessments is calibrated to level descriptions. Each country's choice of content for nation-wide assessments is, of course, different and subject to national preferences.

5.7 Organisation and Implementation

Involvement and Responsibility

The table below indicates the number of countries that involve various organisations in the development of assessments. It was not within the scope of this study to describe the underlying relationships, the precise extent of involvement or the size of each organisation's investment in terms of people, resources or expertise.

Table 5.22 Organisations Actively Involved in the Development of Assessments (N = 27)

Organisations	Types of Assessments		
	Nation-wide	School-based	Other
Ministry of Education	20 (74.1%)	5 (18.5%)	0 (0.0%)
Other Ministries	2 (7.4%)	0 (0.0%)	0 (0.0%)
Universities	9 (33.3%)	3 (11.5%)	4 (14.8%)
Inspectorate of Education	7 (25.9%)	7 (25.9%)	3 (11.1%)
Teachers' unions or associations	3 (11.1%)	5 (19.2%)	0 (0.0%)
Schools or groups of schools	5 (18.5%)	14 (51.9%)	3 (11.1%)
Institute for test development	10 (37.7%)	4 (14.8%)	1 (3.7%)
Other (experts, teachers etc)	7 (25.9%)	5 (18.5%)	5 (18.5%)

The Ministry of Education is, in most cases (74.1%), actively involved in the development of nation-wide assessments. Universities are involved in nine countries and institutes for test-development play an active role in the development of nation-wide assessment in ten countries. Other groups involved include experts, teachers and the Inspectorate of Education.

The most important organisations for school-based assessments are the schools or groups of schools themselves, while the inspectorate of education and teachers are also involved, though not nearly as often. Other types of assessments are developed primarily by schools and teachers. Two or three organisations are involved in the development of nation-wide assessments, while fewer organisations are involved in the development of school-based assessments and other types of assessments.

Institutions Responsible for Development and Implementation

Involvement in the development of assessments can vary in degree and intensity. Bearing responsibility for them, however, could be something quite different. Of course this all depends on the organisational setting that is chosen.

Table 5.23 Organisations Responsible for the Production of Assessments (N = 28)

Organisations	Types of Assessments		
	Nation-wide	School-based	Other
Ministry of Education	18 (64.3%)	4 (14.8%)	0 (0.0%)
Research centres	3 (10.7%)	0 (0.0%)	1 (3.6%)
Schools	0 (0.0%)	23 (85.2%)	2 (7.4%)
Test Development Centres	9 (32.1%)	2 (7.4%)	1 (3.6%)
Universities	6 (21.4%)	2 (7.4%)	2 (7.4%)
Teacher association	1 (3.6%)	2 (7.4%)	1 (3.6%)
Other (experts, teachers etc.)	8 (28.6%)	4 (14.8%)	2 (7.4%)

Again, the Ministries of Education were mentioned in two-thirds of all countries. Other groups and organisations responsible for the production of nation-wide assessments include test-development centres and experts or teachers.

As could be expected, schools are generally responsible for school-based assessments (23 out of 28 countries), though in a few countries experts and teachers play a vital role.

On average are there two organisations responsible for the production of nation-wide assessments (N = 22, s.d. 0.8). For school-based assessments, an average of 1.2 organisations are responsible (N=21, s.d. 0.9) and generally either schools or teachers are responsible for other types of assessments.

Factors Influencing the Design of Assessments

Assessments are not produced within a vacuum. They have specific functions within the educational system of a country and, accordingly, are influenced by several factors. For this research, however, a sweeping account was avoided: the emphasis here is on educational factors that have a direct influence on the design and development of assessments.

Table 5.24 Influence on the Design and Development of Assessments (N = 28)

Influenced by	Types of Assessments		
	Nation-wide	School-based	Other
A needs analysis	7 (25.0%)	7 (25.0%)	2 (7.1%)
Description of aims	14 (50.0%)	10 (35.7%)	5 (17.9%)
A syllabus or curriculum	20 (71.4%)	17 (60.7%)	5 (17.9%)
Candidate profiles	1 (3.6%)	8 (28.6%)	2 (7.1%)
Other,	0 (0.0%)	0 (0.0%)	0 (0.0%)

In most European countries, a combination of the curriculum and either a needs analysis or a description of aims influences the design and development of nation-wide assessments. When a needs analysis is involved in designing a nation-wide assessment, this must go hand in hand with a set of aims, usually laid out the curriculum. The profile of candidates is not considered influential, but it should be taken in consideration that candidates' characteristics of are accounted for in curricula.

School-based assessments are mostly influenced by the curriculum, frequently in combination with either a needs-analysis, candidate profiles or a description of aims. The design and development of other types of assessments are usually influenced by the curriculum and a description of aims.

An important step in the development of assessments is, among others, setting up standards. This is especially true for nation-wide assessments, as they often function as final examinations.

Table 5.25 Procedures for Setting up Standards for Assessments (N=27)

Procedures for setting up standards	Assessments		
	Nation-wide	School-based	Other
Elaborate prescriptions	11 (40.7%)	2 (7.7%)	2 (7.7%)
Pre-testing	8 (29.6%)	4 (15.4%)	0 (0.0%)
Consensus in assessment committees	19 (70.4%)	6 (23.1%)	1 (3.7%)
Procedures for consistency	14 (51.9%)	3 (11.5%)	0 (0.0%)
Other	0 (0.0%)	2 (7.4%)	2 (7.4%)

There is no discernible tendency for setting up standards for the "other types of assessments" (portfolio, general teacher and continuous assessment). For nation-wide assessments, however, a consensus in assessment committee is the most important procedure. Standards for these assessments are also set using explicit procedures for consistency and elaborate prescriptions. These prescriptions are usually combined with pre-testing, while consensus in assessment committees is frequently accompanied by procedures for consistency.

Procedures for setting up standards are unusual for school-based assessments. Consensus in assessment committees, used in little more than one-fifth of all cases, is the most applied procedure for school-based assessments.

Collecting, Marking, Grading and Reporting

Schools hold assessments of foreign languages, and are nearly always actively involved. This includes collecting, marking, grading and reporting results. These are important tasks to be taken seriously. The organisations in charge of collecting, marking, grading and reporting results in the EU-countries are listed below.

Table 5.26 Organisations in Charge of Collecting, Marking, Grading, and Reporting Results (N=28)

Organisations	Number of Countries
Ministry of Education	18 (64.3%)
Research centres	4 (14.3%)
Universities	5 (17.9%)
Schools	16 (57.1%)
Test Development Centres	4 (14.3%)
Others	6 (21.4%)

Ministries of Education are responsible for collecting, marking, grading and reporting results in two-thirds of all countries. Schools, as stated, also play an important role in the grading, marking and reporting of results. Here we can see the difference made when countries use nation-wide assessments as national exams. When school-based assessments are part of national-exams, the schools play an active role in grading, marking and reporting. Unsurprisingly, schools play an important role in instances where only school-based exams are held. Development centres in Estonia, Latvia and the Netherlands are also put in charge of collecting, marking, grading and reporting.

While possibly trivial, it is still notable that countries put either the Ministry of Education or a test development centre in charge, indicating well-defined parameters for each organisation's responsibilities.

5.8 Use of Results

Types of Feedback Received by Participants

Pupils generally receive some information on their achievements and results. This can take on many forms: from entirely undefined, general impressions to profiles of performances per subtest and relative positions in the population.

Table 5.27 Types of Results from Nation-Wide, School-Based and Other Assessments (N = 26)

Type of result	Types of Assessments		
	Nation-wide	School-based	Other
No grades, general impression	0 (0.0%)	2 (7.7%)	0 (0.0%)
Basic grade or pass / fail	13 (50.0%)	17 (65.4%)	2 (7.7%)
Basic grade or pass / fail per subtest	3 (11.5%)	6 (23.1%)	1 (3.8%)
Basic grade plus profile across subtests	5 (19.2%)	4 (15.4%)	1 (3.8%)
Profile of performance per subtest	1 (3.8%)	5 (19.2%)	1 (3.8%)
Relative position in the population	1 (3.8%)	4 (15.4%)	0 (0.0%)

Results from nation-wide assessments are generally given as basic grades or a pass / fail decision. Added feedback can include basic grades and a pass / fail per subtest or a basic grade plus a profile across subtests. In nearly two-third of the European countries, school-based assessments also provide a basic grade or a pass / fail decision. Various other forms of feedback given to students include a basic grade or pass / fail per subtest (six countries), a profile of performance (five countries), a basic grade plus a profile across subtests and information about their relative position in the population (both in four countries). Nearly always one single type of feedback is provided. In just a few countries two types of feedback were indicated.

Table 5.28 Form of Reporting Results for Nation-Wide, School-Based and Other Assessments (N = 26)

Form of report	Types of Assessments		
	Nation-wide	School-based	Other
Raw scores	13 (50.0%)	10 (38.5%)	0 (0.0%)
Undefined grades (for example 'D')	7 (26.9%)	9 (34.6%)	1 (3.8%)
Level on a defined scale	5 (19.2%)	6 (23.6%)	3 (11.5%)
Diagnostic profile	1 (3.8%)	2 (7.7%)	0 (0.0%)
Relative position in the population	2 (7.7%)	3 (11.5%)	0 (0.0%)

Results from nation-wide assessments are mainly reported as raw scores (50% of all countries), as undefined grades or as levels on a defined scale. The same three types are used for school-based assessments, though with less frequent use of raw scores.

Documents Used for Publishing Results

The documents students receive are the most tangible materialisation of their efforts. The most classic documents are the report card and the certificate or diploma.

Table 5.29 Documents for Publishing Results from Nation-Wide, School-Based and Other Assessments (N = 25)

Documents	Nation-wide	Types of Assessments	
		School-based	Other
Letter or email	6 (24.0%)	0 (0.0%)	1 (4.0%)
Report card	2 (8.0%)	17 (68.0%)	1 (4.0%)
Certificate / Diploma	19 (76.0%)	6 (24.0%)	2 (8.0%)
Diagnostic profile in school database	2 (8.0%)	5 (20.0%)	0 (0.0%)

Since the most important functions of the nation-wide assessments are examination and certification, it is unsurprising that in three-quarters of the countries the results are reported on a diploma or certificate. Candidates in six countries receive a letter or e-mail reporting the results.

When a country uses school-based assessments as part of the examination, the results are given on a report card. Results are also given on a certificate/diploma in countries where the school-based assessment makes up the entire examination. In five countries, the results of the school-based assessment form diagnostic profiles to be entered into a school database.

Audiences

Here, two aspects of the release of information are addressed. Table 5.30 indicates the types of information released, while table 5.31 is concerned with the groups who receive that information.

Table 5.30 Information Released to the Public Concerning the Assessment of Communicative Competencies in Nation-Wide Assessments (N = 15)

Information	Released to the public	
	Beforehand	Afterwards
No information	3 (20.0%)	2 (13.3%)
The difficulty of items	0 (0.0%)	3 (20.0%)
The difficulty of sub scales	0 (0.0%)	2 (13.3%)
The range of scores	1 (6.7%)	7 (46.7%)
Grades	0 (0.0%)	8 (53.3%)
Pass marks	4 (26.7%)	5 (33.3%)
The validity of the test	0 (0.0%)	2 (13.3%)
The reliability of the test	0 (0.0%)	2 (13.3%)
The presentation	4 (26.7%)	2 (13.3%)
The content	2 (13.2%)	4 (26.7%)
Scoring	2 (13.2%)	5 (33.3%)

Three countries do not release information beforehand and two countries release none afterwards. Two topics are addressed in information to the public beforehand: the content and the way the test is presented to the candidates. However, one should keep in mind that this is done only in four countries. In Ireland, Poland, Portugal and Slovakia the public receives information on the pass marks, while information on the presentation is released in Estonia, Poland, Portugal and Sweden.

After nation-wide assessments take place, the public generally receives information about grades, the range of scores, pass marks and scoring. Taken at face value, this shows that all information pertaining to the final scores is released.

In Ireland, Poland, Portugal, Slovakia and Estonia two or more types of information are released before the nation-wide assessment is held. In other countries this is either one type or no information at all. A great amount of information is released in some countries, especially in Sweden where eight of the types mentioned above are released. Five to six of these types of information are released in Slovenia, Holland, Hungary and Latvia.

Table 5.31 Recipients of Results from Nation-Wide and Other assessments (N = 28)

Recipients of Results	Types of Assessments	
	Nation-wide	Other
Ministry of Education	19 (67.9%)	8 (28.6%)
Newspapers	11 (39.3%)	3 (10.7%)
Schools	21 (75.0%)	14 (50.0%)
Parents	9 (32.1%)	10 (35.7%)
Parent associations	2 (7.1%)	1 (3.6%)
Pupils	17 (60.7%)	14 (50.0%)
Teacher associations	2 (7.1%)	0 (0.0%)
Universities	4 (14.3%)	2 (7.1%)
The public, on the internet	7 (25.0%)	1 (3.6%)
Supra national institutions	1 (3.6%)	0 (0.0%)

The three most important audiences for the results from nation-wide assessments are schools, Ministries of Education, and pupils. Parents and the wider public (via the Internet) are also mentioned frequently.

Results from other types of assessments are, as could be expected, directed toward schools, parents and pupils. Those involved directly are thus the first to know. Ministries of Education are informed of these results in eight countries.

5.9 A brief Exploration of the Concept of Foreign Language Assessment Culture

In section 3.3 a brief description was given of several complementary mechanisms that are possibly involved in determining a foreign language assessment culture. Without the presumption to be complete, five mechanisms were discussed:

1. Bottom-up versus Top-down approach
2. Centralisation versus Decentralisation in organisation
3. Formative versus summative types of assessment
4. User-oriented versus System-oriented
5. Low versus High Intensity of Assessments

In this study for each of these mechanisms so-called indicators are chosen. One of the hallmarks of a bottom-up approach (1) is the shared responsibility for the process of assessing foreign languages. In the context of this study, this would be a large group of organisations that is involved in the development of the foreign language assessment. A centralised approach (2) becomes visible in a restricted number of organisations that are responsible for the development and administration of the assessment. The basic data on the number of organisations that are involved (bottom up) and the number of organisations that are responsible (centralised) are shown in table 5.32. The data in table 5.32 are just provided to illustrate these two indicators.

Table 5.32 Number of Organisations Involved and Responsible for the Development of Nation-Wide Assessments

Number of Organisations	Involved	Responsible
1	4 (19.0%)	8 (34.8%)
2	6 (28.6%)	11 (47.8%)
3	7 (33.3%)	3 (13.0%)
4	1 (4.8%)	1 (4.3%)
5	1 (4.8%)	0 (0.0%)
6	2 (9.5%)	0 (0.0%)

On average 2.8 organisations are involved (standard deviation 1.5) and 1.9 organisations bear responsibilities (standard deviation 0.8). In most countries, it is the Ministry of Education together with universities, teacher unions, schools, or an institute for test development that is involved. However, when it comes to responsibilities, the government (Ministry of Education) and a maximum of one or two delegated institutes bear responsibility.

An examination at the end of the first or second cycle of secondary education is essentially summative (3) by nature. In 5.3 it was already shown that Nation-Wide assessments were nearly always held to function as examinations; there is hardly any differentiation across Europe. A nation-wide assessment can be summatively used when system monitoring is intended. A user-orientation (4) regarding assessment becomes apparent when its prime function is to provide information to parents, while a system-orientation becomes clear when, amongst others, determining the school quality is of importance. The basic data on these two dimensions are shown below in table 5.33.

Table 5.33 Summative versus formative functions and system orientation versus user orientation of of Nation-Wide Assessments

	Main Function	Subsidiary function	Side effect
Summative			
System monitoring	5 (29.4%)	10 (58.8%)	2 (11.8%)
Formative			
Guidance for students' career	6 (33.3%)	7 (38.9%)	5 (27.9%)
System oriented			
Determining school quality	2 (11.8%)	4 (23.5%)	11 (64.7%)
User oriented			
Information for parents	4 (23.5%)	7 (41.2%)	6 (35.3%)

The intensity of assessments (5) was already discussed in section 5.4, where four groups of countries were discerned with a low intensity of testing, a moderate intensity of testing, an above average intensity and a high intensity of testing.

In order to explore the relationships between the five possible mechanisms that can constitute a foreign language assessment culture, correlations were calculated first of all. Between none of the five mechanisms, there is a positive or negative correlation that is significant at the level of $p < .10$. To push the exploration a bit further, it was calculated if there were any differences in the number of organisations that are involved and the number of organisations that bear responsibilities in countries with a summative orientation versus a formative orientation. The same analysis was done for countries with a bottom-up approach versus a top-down approach. In countries with a strong summative orientation, the number of organisations involved and the number of organisations bearing responsibilities don't differ from the number in the countries with a formative approach. Nor were there differences between countries with a user orientation versus countries with primarily a system orientation in number of organisations involved and the number of organisations bearing responsibilities.

A conclusion that can be drawn, and this is done with the utmost care, is that the mechanisms and indicators chosen are not coherent in constituting a foreign language assessment culture. The mechanisms and indicators seem to operate independently of each other and bear no inner relationships. The data from this study have revealed that there is a great diversity in the types of assessments of foreign languages used. This

must be due to varying perceptions of and rationales for assessments, as well as contrasting views on the functions of assessments. The study has not yet uncovered the prevailing mechanisms that constitute empirically a foreign language assessment culture.

6. Conclusions

6.1 Introduction

The essential feature of policies and practices of foreign language assessment in the EU is diversity. This study's focus has been to identify the wide range of possibilities countries have for organising and implementing foreign language assessments. It has led to seven principal conclusions, and these are outlined in the sections below. What is clear is that each EU country has designed an educational infrastructure that corresponds with its views on the functions of assessments, and perceptions of its economical, political and socio-cultural needs. This is to say that objectives, test-structures, the organisation and implementation of tests, as well as the way in which results are used, have all been fine-tuned by the expectations politicians, policy makers and key influencers have.

6.2 Objectives and Assessments

Eight types of aims for foreign language learning in secondary education were distinguished: socio-cultural knowledge, knowledge of literature, knowledge of a foreign language, proficiency (in one of the four skills), communicative skills, transfer of learning, learning strategies and attitudes.

Proficiency in the four skills was an important aim for most countries, with some rating oral skills above the others. The Czech Republic, Denmark, Finland and France explicitly refer to the levels of proficiency specified in the CEF. Many respondents also cited communicative competencies as an important aim. The transfer of learning and attitudes towards foreign language learning, however, were mentioned in just three sets of aims, and knowledge of literature was not mentioned once. These results lead to this study's first principal conclusion: Across the EU, there is a shared practical and conceptual approach to the skills to be tested and a common focus on communicative competencies. This should benefit future development of tests for measuring students' foreign language skills at the end of secondary education. Each European country has formulated a unique set of aims and goals for the attainment level of foreign languages in secondary education. The shared skills that can be tested and the common focus on communicative competencies can only be seen as a challenge. It is a fertile starting point for increasing cooperation between organisations and countries in Europe. Those projects will support

mutual understanding of the aims and goals that are stated and the levels of competence that are strived for in each country.

School-based, nation-wide and continuous assessments stand out as the most frequently used types of assessment. School-based assessments, in general, can materialize in two ways. The practice of education is based on schools trying to achieve a set of objectives, and testing is used to generate feedback on the skills and attitudes of students during this process. The emphasis, then, is on administering grade-based assessments of student performance. This type of school-based assessment is applied in all European countries. On the other hand, school-based assessments can be used as part of a nation-wide assessment or exam that concludes the first or second cycle of secondary education. In this case, schools can be made individually responsible for assessing students' foreign language skills as part of the nationally administered exam. Also, throughout Europe, there are several educational systems in which only school-based assessments are administered. This is the case in Belgium (Flemish, French and German speaking communities), the Czech Republic and Germany (all Bundesländer and independent cities). There are ten EU-countries in which school-based assessments are conducted in combination with external assessments during or at the end of secondary education.

Nation-wide assessments are carried out in twenty-two EU countries. Nations that do not use these assessments, as mentioned above, include Germany, the Czech Republic and Belgium. Nation-wide assessments can take different shapes and forms: Firstly, they can take place at the end of secondary education or at another given moment. Secondly, they may consist of series of tests issued to schools, for which participation may be voluntary.

There are just four countries that regularly employ standardised assessments. These assessments use instruments developed from shared aims, and while students may have followed different curricula, they participate in a standardised assessment under the same conditions and must fulfil the same requirements. The scores are based on predetermined standards, but the use and significance of the results can vary for each student. Continuous assessment, while sometimes only a broad set of informal techniques, is one of the most widely used types of assessment in Europe. Portfolio assessments are carried out in twelve countries, though its use within these countries is not wide spread, and one can safely conclude that it is more likely to be found in innovation driven schools.

A second principal conclusion is that it is virtually impossible to establish a direct relationship between countries' stated objectives regarding foreign language learning and the types of assessments used. Though there was some coherence in the stated objectives (especially on skills and communicative competencies), the very general

descriptions prohibit direct comparisons and analysis. This is also hindered by the fact that in some countries, several types of assessment are integrated, while in others the same types are used independently of each other. This diversity in objectives and the configuration of types of assessments shows, in fact, that a direct relationship does not exist. The sheer size of the EU, with 25 countries, makes it impossible to arrive at a descriptive, statistical relationship, between goals and the types of assessment used, let alone a causal relationship. This leads to a third conclusion: The diversity in types of assessments of foreign languages must be due to varying perceptions of and rationales for assessments, as well as contrasting views on the functions of assessments. Beneath these differences are each country's historical backgrounds, including specific economical, political and socio-cultural needs that have been incorporated into an operational educational policy regarding foreign language assessments.

6.3 Reasoning behind Foreign Language Assessment Cultures

The two most important functions of nation-wide assessments and standardised assessments of foreign languages are examination and certification. Transparency, system monitoring, guidance for students' educational career and information for parents are important functions for a little less than half of all countries. Transparency is the main function of a nation-wide assessment of foreign languages in six countries and a subsidiary function in five countries. Examination and certification are the main functions for other types of assessments in countries where school-based assessments are part of national examinations. Predictably, guidance for students' educational careers is a main function of other types of assessments in nearly two-thirds of all countries. Another main function of other assessments is to provide parents with information. The most important function of standardised assessments of foreign languages is system monitoring.

Responses concerning the rationales behind assessments produced an interesting but disparate impression. "The importance of the skill" is among the most important rationales for assessing communicative competencies in nation-wide assessments. It seems, then, that a utilitarian motive prevails. Test experts also cited a political motive-comparability of outcomes -and a technocratic motive- consistency over the years. Policy makers find comparability of outcomes important as well, but they also mention the technocratic motive of quality of the assessment. Sociological motives, such as fairness and equality of opportunities, are less highly valued.

The fourth conclusion is that each assessment implemented has a specific function within its own educational context. Examination and certification are the most important

functions of nation-wide assessments. System monitoring is the prevailing function of standardised assessments, while other types of assessments are more focused on examination and certification, as well as providing guidance to pupils and information to parents.

6.4 Structure of tests

Nation-wide assessments at the end of secondary education all offer English. French is offered in the nation-wide assessments of all but one or two countries, but also German and Spanish and Italian, are often included in these assessments. Russian is still assessed in ten countries. Few other foreign languages are assessed, and when one is it is usually a language from a neighbouring country. The fifth conclusion is that English and French are the two most widely assessed languages at the end of the second cycle of secondary education throughout all EU-countries. However, German, Spanish and Italian have gained prominence in assessments at the end of secondary education in a wide range of EU-countries.

Three quarters of the countries that hold nation-wide assessments administer them at the end of the second cycle of secondary education. In several countries, though, this is not the equivalent of the end of compulsory education. In most EU-countries, the first cycle of secondary education or lower secondary education is part of compulsory education. On the other hand, the duration of the second cycle or 'upper' secondary education is very diverse and, depending on the country, upper secondary education can be compulsory, partly compulsory or not compulsory at all.

Half of the EU-countries administer nation-wide assessments of foreign languages at the end of the first cycle of secondary education. In most cases, reading and writing are integral components of the assessment. 21 European countries organise nation-wide assessments at the end of the second cycle of secondary education. All incorporate the assessments of reading and writing, while 16 countries include listening and 14 assess speaking. Reading and writing are also incorporated in all standardised assessments. Listening comprehension is assessed in two of the four standardised assessments cited in this report. Other types of assessments cover the full range of communicative competencies. Integrated skills are incorporated into nation-wide and standardised assessments in just a few countries, but are included in the majority of continuous and portfolio assessments.

The three types of assessment tasks most often used in nation-wide assessments in the first cycle of secondary education are closed questions/ multiple choice, fill in the gap

and short answer (short sentence/completion). Other types frequently used are more than one short sentence, essays and interaction with examiners. In the second cycle of secondary education four assessment tasks stand out: closed questions/ multiple choice, fill in the gap, essays and short answer (short sentence/completion).

Listening comprehension in nation-wide assessments is most often (in two-third of all countries) carried out using audio aids. A visual aid is used in another three instances. Visual aids are sometimes used for assessing speaking or spoken interaction, though not as often as written instructions. Unsurprisingly, written instructions are nearly always used for testing reading and writing.

There are various procedures for choosing content for the assessment of communicative competencies, and three of these clearly stand out. Previous experience and descriptions of curricula are used most often, while guidelines are employed in two-thirds of the countries. Curricula describe the essential skills that must be taught, and are a very important source for developers choosing assessment content.

The sixth conclusion is that there is common ground regarding the communicative competencies that are assessed most often: reading comprehension and writing. This does not mean that other competencies are less highly valued. Practical reasons (feasibility) and restraints (time consumption) can prohibit national assessments of speaking and spoken interaction. Goals stated in the curriculum are the most operational and recognisable starting point for choosing content to be used in an assessment.

6.5 Organisation and Implementation

Ministries of Education are often actively involved in the development of nation-wide assessments, as are Universities. Institutes for test-development play an active role in the development of nation-wide assessment in ten countries. Responsibility for the production of nation-wide assessments is given to Ministries of Education in two-thirds of all countries. Other groups and organisations responsible for the production of nation-wide assessments include test-development centres and experts or teachers. As could be expected, schools are generally responsible for school-based assessments, though in a few countries experts and teachers play a vital role as well.

The seventh conclusion is that each country has implemented an organisational model to facilitate the development and administration of tests at the end of the first or second cycle of secondary education. A number of countries employ a top-down model. Despite the similar approach, though, there are key variations with respect to the organisations involved and the kinds of tasks they are assigned. Even among countries where schools

are empowered with the administration of assessments, one finds stark differences in the design and implementation of organisational models for realising the main objectives: examination and certification.

In most European countries, a combination of the curriculum and either a needs analysis or a description of aims influences the design and development of nation-wide assessments. When a needs analysis is involved in designing a nation-wide assessment, this must go hand in hand with a set of aims, which is usually laid out the curriculum. The profile of candidates is not considered influential, but it should be taken in consideration that candidates' characteristics are accounted for in curricula.

There is no discernible tendency for setting up standards for portfolio, general teacher and continuous assessment. Consensus in an assessment committee is the most important procedure for nation-wide assessments. Standards for these assessments are also set using explicit procedures for consistency and elaborate prescriptions. These prescriptions are usually combined with pre-testing, while consensus in assessment committees is frequently accompanied by procedures for consistency.

Ministries of Education are responsible for collecting, marking, grading and reporting results in two-thirds of all countries. Schools also play an important role in the grading, marking and reporting of results. When school-based assessments are part of national-exams, the schools play an active role in grading, marking and reporting. Unsurprisingly, schools play an important role in instances where only school-based exams are held: Austria, the Czech-republic, Finland and Slovakia. Test development centres are put in charge of collecting, marking, grading and reporting in Estonia, Latvia and the Netherlands.

6.6 Use of Results

Results from nation-wide assessments are generally given as basic grades or a pass / fail decision. Added feedback can include basic grades and a pass / fail per subtest or a basic grade plus a profile across subtests. In nearly two-thirds of the European countries, school-based assessments also provide a basic grade or a pass / fail decision. Various other forms of feedback given to students include a basic grade or pass / fail per subtest (six countries), a profile of performance (five countries), a basic grade plus a profile across subtests and information about their relative position in the population (both in four countries). Nearly always, a single type of feedback is provided. Two types of feedback were indicated in just a few countries.

There are a number of stark differences between countries regarding the amount of information that is communicated to the public before and after assessments. Three countries do not release information beforehand and two countries release none afterwards. Test-content and the way in which the test is presented are addressed in information to the public beforehand in just four countries. Information on pass marks is released beforehand in Ireland, Poland, Portugal and Slovakia, while information on presentation is released in Estonia, Poland, Portugal and Sweden.

After nation-wide assessments take place, the public generally receives information about grades, the range of scores, pass marks and scoring. Taken at face value, this would mean that all information pertaining to the final scores is released. Of the various types of information distinguished in this study two or more are released beforehand in just five countries. In the rest of the EU, this is either one type or no information at all. A few countries release substantial amounts of information after assessments take place, with Sweden being the most elucidatory (eight types of information). Five to six types of information are released in Slovenia, The Netherlands, Hungary and Latvia.

6.7 Areas for future study

Three possible key areas for future study have been identified.

1. Efficiency: Questions to be addressed are: Are the testing systems efficient and are they used efficiently? Are the testing systems capable of adequately showing the progress pupils have made? Is it possible to make valid statements about an educational system on the basis of that country's assessment culture?

2. Teachers and pupils: How do teachers and pupils perceive the extent to which exams address the knowledge that has to be obtained, cover the curriculum and to what extent do the scores actually represent what pupils can do or what knowledge they have gained? How can those who have to take a test affect the different types of assessment? How do other stakeholders, such as (future) employers in the various EU-countries deal with test results?

3. Backwash effects on teaching: To what extent are backwash effects on teaching comparable between countries?

References

- Alderson, C.J., Figueras, N., Kuijper, H., Nold, N., Takala, S., & C. Tardieu. (2004). The Development of Specifications for Item Development and Classification within the Common European Framework of Reference for Languages: Learning, Teaching, and Assessment. Reading and Listening. Final Report of The Dutch CEF Construct Project. <http://www.ealta.eu.org/dutch/grid.htm>
- Blondin, C., Candelier, M., Edelenbos, P. Johnstone, R., Kubanek-German A. & Taeschner, T. (1998). *Foreign Languages in Primary and Pre-School Education: Context and Outcomes. A Review of Recent Research within the European Union*. CILT: London.
- Byram, M., Brumfit, C. et al. (2000). *Routledge Encyclopedia of Language Teaching and Learning*. Routledge: London/New York.
- Council of Europe (2001) *Common European Framework of Reference teaching, assessment*. Cambridge: Cambridge University Press.
- Cronbach, L.J. & K. Shapiro (Eds.) (1987). *Designing Evaluations of Educational and Social Programs*. Jossey-Bass: San Francisco
- Dickson, P. & A. Cumming (Eds.) (1996). *Profiles of Language Education in 25 Countries*. National Foundation for Education Research: Berkshire.
- Edelenbos, P. & A. Kubanek-German (2004). Teacher assessment: the concept of 'diagnostic competence'. *Language Testing*, 21 (3), 259-283.
- Philips J.K. (ed.) (1999). *Foreign Language Standards: Linking Research Theories and Practices*, Lincolnwood: National Textbook Company.
- Rea-Dickins, P. 2001: Mirror, mirror on the wall: identifying processes of classroom assessment. *Language Testing* 18 4), 429-462.
- West, A., Edge, A. & E. Stokes (2000). Examination and Assessment of Data on Foreign Language Learning: Final Report. Report commissioned by DG XXII under the Leonardo da Vinci Programme Strand III.2.b. (2000).

