



# First European Survey on Language Competences

*Technical Report*



## Technical Report authors

The table below presents the key authors of the Technical Report.

Key authors	Institutions	Position
Neil Jones	ESOL	Project Director
Karen Ashton	ESOL	Project Manager, Field Operations Lead
Gunter Maris	Cito	Data Analysis Lead
Sanneke Schouwstra	Cito	Questionnaires development Lead including framework and indices
Norman Verhelst	Cito	Standard Setting Lead
Ivailo Partchev	Cito	Weighing Lead, Data Management Lead
Jesse Koops	Cito	Data management Co-Lead
Martin Robinson	ESOL	Language Testing Team Lead
Manas Chattopadhyay	Gallup	Sampling and base weights Lead
Gergely Hideg	Gallup	Sampling Co-Lead
Jostein Ryssevik	Gallup	Software systems Lead

The following people were instrumental in the reviewing, proofreading and formatting chapters of this report: Erna Gille, Johanna Kordes, Robert Manchin, Agnes Ilyes, Peter Husztik, Anna Chan, Michaela Perlman-Balme, Julia Guess, Danilo Rini, Guiliana Bolli, Sylvie Lepage, Roselyne Marty, Heidi Endres, Inma Borrego, Joost Schotten, Remco Feskens, Rebecca Stevens and John Savage.

## Abbreviations and codes used in this report

The following educational system and language codes are used throughout this report.

Participating educational system	Educational system code	Questionnaire language(s)	Language code
Flemish Community of Belgium	BE nl	Dutch	Nl
French Community of Belgium	BE fr	French	Fr
German Community of Belgium	BE de	German/French	de, fr
Bulgaria	BG	Bulgarian	Bg
Croatia	HR	Croatian	Hr
England	UK-ENG	English	En
Estonia	EE	Estonian; Russian	et, er
France	FR	French	Fr
Greece	EL	Greek	El
Malta	MT	English	En
Netherlands	NL	Dutch	Nl
Poland	PL	Polish	Pl
Portugal	PT	Portuguese	Pt
Slovenia	SI	Slovene	Sl
Spain	ES	Spanish, Basque, Catalan, Galician, Valencian	es, Spanish-Basque Spanish-Catalan, Spanish-Galician, Spanish-Valencian
Sweden	SE	Swedish	Sv

The following abbreviations are used in this report.

Abbreviations	In full
<b>BoW</b>	Body of Work method
<b>CB</b>	Computer-based
<b>CD</b>	Compact Disc
<b>CEFR</b>	Common European Framework of Reference
<b>CFI</b>	Comparative Fit Index

Abbreviations	In full
CLIL	Content and Language Integrated Learning
COGN	Cognitive
CML	Conditional Maximum Likelihood
CMOS	Cumulative Measure of Size
DIF	Differential Item Functioning
DVD	Digital Versatile Disc
EC	European Commission
EILC	European Indicator of Language Competences
ENR	Enrolment
ESCS	Economic, social and cultural status
ESLC	European Survey on Language Competences
FL	Foreign Language
Gb	Gigabyte
HISEI	Parental Occupation
HOMEPOS	Home possessions
ICT	Information and Communication Technologies
ID	Identification
ILO	International Labour Organisation
INES	OECD Indicators of Education Systems
INT	International
IRT	Item Response Theory
ISO	International Organization for Standardization
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISCO_F	International Standard Classification of Occupation Father
ISCO_M	International Standard Classification of Occupation Mother
ISEI	International Socioeconomic Index
MM	Multiple Marking
MOS	Measure of Size
NFI	Normed Fit Index
NNFI	Non-Normed Fit Index
NRC	National Research Coordinator



<b>Abbreviations</b>	<b>In full</b>
<b>OECD</b>	Organisation for Economic Co-operation and Development
<b>OPLM</b>	One Parameter Logistic Model
<b>PARED</b>	Higher parental education expressed as years of schooling
<b>PB</b>	Paper-based
<b>PDF</b>	Portable Document Format
<b>PIRLS</b>	Progress in International Reading Literacy Study
<b>PISA</b>	Programme for International Student Assessment
<b>PCM</b>	Partial Credit Model
<b>PPS</b>	Probability Proportional to Size
<b>QC</b>	Quality Control
<b>RMR</b>	Root Mean Residual
<b>RMSEA</b>	Root-Mean Square Error of Approximation
<b>SC</b>	School Co-ordinator
<b>SCH</b>	School
<b>SCO</b>	Scored responses
<b>SE</b>	Standard Error
<b>SES</b>	Socio-Economic Status
<b>SRS</b>	Simple Random Sampling
<b>TA</b>	Test Administrator
<b>TALIS</b>	Teaching and Learning International Survey
<b>TCS</b>	Target Cluster Size
<b>TIMSS</b>	Third International Mathematics and Science Study
<b>TL</b>	Target Language/Test Language
<b>USB</b>	Universal Serial Bus



# TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION .....</b>	<b>2</b>
1.1	KEY ELEMENTS OF THE ESLC .....	4
1.2	THIS TECHNICAL REPORT .....	7
1.3	REFERENCES .....	8
<b>2</b>	<b>INSTRUMENT DEVELOPMENT - LANGUAGE TESTS .....</b>	<b>10</b>
2.1	DEVELOPMENT OF THE LANGUAGE TESTING FRAMEWORK .....	11
2.2	DEVELOPMENT OF THE LANGUAGE TESTS .....	27
2.3	TEST DEVELOPMENT PROCESS .....	34
2.4	MARKING .....	41
2.5	FINAL TEST DESIGN .....	44
2.6	REFERENCES .....	49
<b>3</b>	<b>INSTRUMENT DEVELOPMENT - QUESTIONNAIRES .....</b>	<b>51</b>
3.1	CONCEPTUALISATION .....	51
3.2	OPERATIONALISATION .....	74
3.3	REFERENCES .....	98
<b>4</b>	<b>OPERATIONS - SAMPLING .....</b>	<b>105</b>
4.1	TARGET POPULATION AND OVERVIEW OF THE SAMPLING DESIGN .....	105
4.2	POPULATION COVERAGE AND SCHOOL AND STUDENT PARTICIPATION RATE STANDARDS .....	107
4.3	COVERAGE OF THE INTERNATIONAL TARGET POPULATION .....	107
4.4	ACCURACY AND PRECISION .....	109
4.5	RESPONSE RATES .....	110
4.6	ESTABLISHING THE NATIONAL TARGET POPULATION .....	111
4.7	SAMPLING IMPLEMENTATION – TEST LANGUAGES .....	111
4.8	TESTING GRADES .....	112
4.9	SCHOOL SAMPLING FRAME .....	116
4.10	STRATIFICATION .....	117
4.11	ASSIGNING A MEASURE OF SIZE TO EACH SCHOOL .....	122
4.12	SORTING THE SAMPLING FRAME .....	122
4.13	SCHOOL SAMPLE ALLOCATION ACROSS EXPLICIT STRATA .....	123



4.14	PROBABILITY PROPORTIONAL TO SIZE SAMPLING .....	123
4.15	IDENTIFYING REPLACEMENT SCHOOLS.....	125
4.16	STUDENT SAMPLING .....	126
4.17	SELECTING THE SCHOOL SAMPLE PERSONNEL.....	129
4.18	SAMPLING FORMS .....	130
<b>5</b>	<b>OPERATIONS - TRANSLATION.....</b>	<b>133</b>
5.1	INTRODUCTION.....	133
5.2	OVERVIEW OF TRANSLATION SYSTEM, SUPPORT AND TRAINING .....	134
5.3	DOCUMENTATION NEEDING TRANSLATION AND THE TRANSLATION PROCESS .....	135
5.4	SURVEYLANG TRANSLATION GUIDELINES .....	140
5.5	QUESTIONNAIRE LANGUAGE, LOCALISATIONS AND AMENDMENTS TO STANDARD PROCESS .....	141
5.6	DEVELOPMENT OF SOURCE VERSIONS.....	143
5.7	FIELD TRIAL AND MAIN STUDY TRANSLATION PROCESSES.....	144
5.8	RECRUITMENT GUIDELINES FOR TRANSLATORS .....	145
5.9	REFERENCES.....	145
<b>6</b>	<b>OPERATIONS - THE SURVEYLANG SOFTWARE PLATFORM .....</b>	<b>148</b>
6.1	INTRODUCTION.....	148
6.2	REQUIREMENTS .....	148
6.3	ARCHITECTURE .....	150
6.4	TEST-ITEM AUTHORIZING TOOL .....	151
6.5	TEST-ITEM DATABANK .....	154
6.6	TRANSLATION MANAGEMENT .....	156
6.7	TEST ASSEMBLY .....	157
6.8	A: TEST ASSEMBLY .....	157
6.9	B: ALLOCATION.....	160
6.10	TEST MATERIALS PRODUCTION.....	162
6.11	THE USB MEMORY STICK PRODUCTION UNIT.....	163
6.12	TEST RENDERING .....	164
6.13	THE USB-BASED TEST RENDERING OPERATING ENVIRONMENT .....	166
6.14	DATA UPLOAD SERVICE .....	166
6.15	ADDITIONAL UTILITIES .....	167
6.16	SOFTWARE QUALITY AND TESTING .....	167



6.17	PERFORMANCE.....	168
<b>7</b>	<b>FIELD OPERATIONS .....</b>	<b>170</b>
7.1	OVERVIEW OF ROLES AND RESPONSIBILITIES .....	170
7.2	KEY NATIONAL RESEARCH COORDINATOR TASKS.....	170
7.3	COMMUNICATIONS BETWEEN SURVEYLANG AND NRCs .....	173
7.4	STAFF SELECTION AND STAFF TRAINING.....	174
7.5	NRC SAMPLING TASKS .....	175
7.6	NRC PRE-ADMINISTRATION TESTING TASKS.....	176
7.7	NRC TEST MATERIALS MANAGEMENT TASKS .....	176
7.8	KEY SCHOOL COORDINATOR TASKS .....	178
7.9	KEY TEST ADMINISTRATOR TASKS .....	181
7.10	KEY TECHNICAL SUPPORT PERSON TASKS (IF CB TESTING).....	183
7.11	RECEIPT OF MATERIALS AT THE NRC AFTER TESTING .....	183
7.12	DATA ENTRY TASKS.....	183
7.13	MARKING OF WRITING .....	185
7.14	DATA SUBMISSION .....	188
7.15	DATA CHECKING.....	189
7.16	CODING OF STUDENT QUESTIONNAIRES .....	189
7.17	MAIN STUDY: A REVIEW .....	191
7.18	REFERENCES .....	191
<b>8</b>	<b>OPERATIONS - QUALITY MONITORING .....</b>	<b>193</b>
8.1	AN INTRODUCTION .....	193
8.2	SUPPORT FOR NRCs IN QUALITY MONITORING.....	194
8.3	IMPLEMENTATION OF QUALITY MONITORING PROCEDURES .....	196
8.4	ESLC QUALITY MONITORS .....	197
8.5	QUALITY MONITORING DATA .....	200
<b>9</b>	<b>DATA PROCESSING - WEIGHTING.....</b>	<b>203</b>
9.1	MOTIVATION AND OVERVIEW .....	203
9.2	BASE WEIGHTS .....	205
9.3	ADJUSTING WEIGHTS FOR NON-RESPONSE .....	207
9.4	VARIANCE ESTIMATION.....	214
9.5	REFERENCES.....	215



<b>10</b>	<b>DATA PROCESSING - QUESTIONNAIRE INDICES .....</b>	<b>217</b>
10.1	TYPE OF INDICES.....	217
10.2	TESTING THE STRUCTURE OF LATENT VARIABLES .....	217
10.3	DATA PREPARATION .....	218
10.4	STUDENT QUESTIONNAIRE .....	219
10.5	TEACHER QUESTIONNAIRE .....	247
10.6	PRINCIPAL QUESTIONNAIRE .....	261
10.7	REFERENCES .....	273
<b>11</b>	<b>SETTING STANDARDS IN RELATION TO THE CEFR .....</b>	<b>275</b>
11.1	OUTLINE OF THE STANDARD SETTING CONFERENCE (SEPT 26-30 2011).....	276
11.2	STANDARD SETTING METHODOLOGY .....	277
11.3	THE STANDARD SETTING CONFERENCE - RESULTS .....	283
11.4	THE WRITING ALIGNMENT STUDY (AUGUST 2011) .....	289
11.5	WRITING ALIGNMENT STUDY OUTCOMES.....	291
11.6	THE STATUS OF THE STANDARDS.....	294
11.7	REFERENCES .....	295
<b>12</b>	<b>ANALYSES .....</b>	<b>298</b>
12.1	INTRODUCTION .....	298
12.2	ITEM RESPONSE THEORY.....	300
12.3	LINKING LANGUAGE PROFICIENCY TO KEY POLICY INDICATORS: LATENT REGRESSION .....	308
12.4	APPENDIX: GENERATING PLAUSIBLE VALUES.....	310
12.5	REFERENCES .....	320
<b>13</b>	<b>DATA PROCESSING - DATA SETS.....</b>	<b>332</b>
13.1	THE STUDENT QUESTIONNAIRE AND PERFORMANCE DATA FILE .....	332
13.2	LANGUAGE ASSESSMENT ITEMS DATA FILES.....	332
13.3	TEACHER QUESTIONNAIRE DATA FILE .....	333
13.4	SCHOOL QUESTIONNAIRE DATA FILES.....	333
13.5	RECORDS IN THE DATA SETS.....	334
13.6	RECORDS EXCLUDED FROM THE DATASETS .....	334
13.7	WEIGHTS IN THE DATASETS.....	334
13.8	REPRESENTING MISSING DATA .....	335



13.9	IDENTIFICATION OF RESPONDENTS, SCHOOLS AND MARKERS.....	335
<b>14</b>	<b>APPENDICES.....</b>	<b>338</b>
14.1	EXAMPLE LANGUAGE TEST TASK TYPES.....	338
14.2	QUESTIONNAIRES .....	373
14.3	SAMPLING FORMS .....	553
14.4	ESLC TECHNICAL STANDARDS.....	630
14.5	DEFINITIONS .....	641
14.6	MULTIPLE MARKING.....	630
14.7	MANAGING AND IMPLEMENTING THE ESLC.....	630

# ***Chapter 1:***

# ***Introduction***



# 1 Introduction

The European Survey on Language Competences (ESLC), the first survey of its kind, is designed to collect information about the foreign language proficiency of students in the last year of lower secondary education (ISCED2) or the second year of upper secondary education (ISCED3) in participating countries or country communities (referred to herein as 'educational systems', with the same meaning as 'adjudicated entities' used in other surveys). The intention was 'not only to undertake a survey of language competences but a survey that should be able to provide information about language learning, teaching methods and curricula.' (European Commission 2007a). As the European Commission (2005) states, 'it is important for Member States to be able to contextualise the data', and thus the language tests should 'be complemented by questionnaires to teachers and pupils to gather contextual information'.

The ESLC is a collaborative effort among the 16 participating educational systems and SurveyLang partners to measure the language proficiency of approximately 53000 students across Europe, to assist the European Commission in establishing a European Indicator of Language Competence to monitor progress against the March 2002 Barcelona European Council conclusions. These conclusions called for 'action to improve the mastery of basic skills, in particular by teaching at least two foreign languages from a very early age' and also for the 'establishment of a linguistic competence indicator' (European Commission 2005). As the Commission (European Commission 2005) states, the decision to launch the ESLC 'arose from the current lack of data on actual language skills of people in the European Union and the need for a reliable system to measure the progress achieved'. The ESLC was therefore initiated by the Commission with the aim that: 'the results collected will enable the establishment of a European Indicator of Language Competence and will provide reliable information on language learning and on the language competences of young people' (European Commission 2007a) as well as providing 'strategic information to policy makers, teachers and learners in all surveyed countries' through the collection of contextual information in the background questionnaires (European Commission 2007b).

Each educational system tested students in two languages; the two most widely taught of the five most widely taught European languages: English, French, German, Italian and Spanish. This effectively meant that there were two separate samples within each educational system, one for the first test language, and one for the second. Each sampled student was therefore tested in one language only. Students' proficiency was assessed in two of the three skills of Listening, Reading and Writing.

The ESLC sets out to assess students' ability to use language purposefully, in order to understand spoken or written texts, or to express themselves in writing. Their observed language proficiency is described in terms of the levels of the Common European Framework of Reference (CEFR) (Council of Europe 2001), to enable comparison across participating educational systems. The data collected by the ESLC will allow





participating educational systems to be aware of their students' relative strengths and weaknesses across the tested language skills, and to share good practice with other participating educational systems.

To 'facilitate a more productive comparison of language policies, and language teaching methods' (European Commission 2005:5) context questionnaires, covering the 13 policy issues detailed below, were administered to the students tested, their teachers of foreign languages, and their institution principals. In addition, system-wide information was collected through the National Research Coordinators.

Early language learning is explored through questions on the onset of foreign language learning, and the weekly amount of time for target and foreign language learning (lesson time and homework).

The diversity and order of foreign language teaching is explored through questions to principals and students on the number of foreign and ancient languages provided (schools) and learned (students).

The language friendly living environment explores the number of students' first languages, languages used at home, and parents' target language knowledge; also the ways in which students use the target language: at home, in the living environment, through visits abroad or through the media.

The concept of the language friendly school looks at the degree of language specialisation, for example, whether content and language integrated learning (CLIL) is practised.

A set of indices related to the use of ICT to enhance foreign language learning and teaching.

Intercultural exchanges arising from school trips, visits or language projects are explored from the perspective of students, teachers, principals and educational systems.

The impact of teachers from other language communities is explored.

Language learning for all looks at provision for immigrant students of the first and second generation.

Under approaches to foreign language teaching a large number of indices explore, for example, the relative emphasis teachers put on teaching the different skills, emphasis placed on similarities between the target language and other known languages, and use of the target language during lessons by teachers and students - all these from the perspective of teachers and students. Several questions probe students' attitudes to the target language: their perception of its usefulness, of how difficult it is to learn and of how they evaluate the lessons, teacher and textbooks.

Teacher initial and in-service training includes indices for teacher qualifications and competences. Questions to teachers and principals explore financial and other incentives for in-service training, how much training teachers attend, and whether the focus of training is on language teaching.

A period of work or study in another country addresses questions to teachers and principals on the number of such stays, financial incentives, and availability of funding for exchange visits or stays abroad.



Several questions on the use of existing European language assessment tools explore uptake of the CEFR and a language portfolio: is use of the CEFR compulsory, teachers' received training in the use of the CEFR and a language portfolio, and how do they use it?

Teachers' practical experience is explored through questions for example on years' experience in teaching the target language and other languages and the number of languages taught over the past five years.

The ESLC data adds significantly to the knowledge base that was previously available at European level or from official national statistics. The data should prove a valuable resource for researchers, policy makers, educators, parents and students and will enable them to review progress towards achieving the March 2002 Barcelona European Council conclusions of learning two foreign languages from an early age.

SurveyLang recognises the contribution of all of its partners and National Research Coordinators (NRCs) in the delivery of the survey. The ESLC is methodologically complex and its implementation has required a considerable collaborative effort by the participating educational systems with SurveyLang. The in-country administration of the survey was the responsibility of the representatives of each educational system (NRCs). Implementing the ESLC depended not only on this collaboration but also on pooling the expertise of SurveyLang partners to develop and exploit innovative methodologies, test instruments and technologies. This Technical Report describes these methodologies, together with other aspects of the methodology that have enabled the ESLC to provide data to support the European Commission in this area of policy. The descriptions are provided at a level that will enable review of the implemented procedures and the solutions adopted for the challenges faced.

This report contains a description of the theoretical underpinning of the complex techniques used for the ESLC and to create the ESLC data sets, which contain data on approximately 50000 students from 15<sup>1</sup> educational systems. The data sets include not only information on student performance in two of the three language skill areas of Listening, Reading and Writing, but also their responses to the Student Questionnaire that they completed as part of the administration. Data from the school principals and language teachers of participating schools teaching at the eligible ISCED level are also included in the data sets.

## **1.1 Key elements of the ESLC**

Elements central to the design of the ESLC are outlined in brief below. The remainder of this report describes these elements, and the associated procedures and methodology, in more detail.

---

<sup>1</sup> As England participated in the Main Study later than other adjudicated entities, at this stage data from England is not included in the data sets.



**Sample size:** Approximately 53000 students enrolled in schools in 16 participating educational systems were assessed in the ESLC Main Study 2011.

**Tested education level:** Students were tested at the last year of lower secondary education (ISCED2) or the second year of upper secondary education (ISCED3) in participating educational systems.

**Tests:** The tests covered three language skills: Listening, Reading and Writing in five test languages: English, French, German, Italian and Spanish. Each student was assessed in two out of these three skills in one test language and also completed a contextual questionnaire. The language tests measure achievement of levels A1 to B2 of the Common European Framework of Reference (CEFR) (Council of Europe, 2001). The pre-A1 level which is also reported indicates failure to achieve A1. Language teachers and school principals at sampled schools also completed a contextual questionnaire.

**Testing mode:** The ESLC was administered in both paper and computer-based formats. The Teacher and Principal Questionnaires were administered through an internet-based system.

**Testing duration:** Students had either 30 minutes or 45 minutes to complete each test. All Listening and Reading tests were set at 30 minutes. The low and intermediate Writing tests were set at 30 minutes, while the high level Writing test and Student Questionnaires (including a CEFR self-assessment) were set at 45 minutes. The total testing time for a student, including the questionnaire, was thus 105 or 120 minutes.

**Summary of tested languages, levels and testing mode across participating educational systems:** The tables below provide a summary of the tested languages, levels and testing mode of each educational system. Further details on the tested languages and levels can be found in Chapter 4 on sampling.

**Table 1 Educational system testing design summary**

Educational system	First most widely taught <sup>2</sup> foreign language	Testing grade for 'First' language	Second most widely taught foreign language	Testing grade for 'Second' language	Testing mode
Flemish Community of Belgium <sup>3</sup>	French	ISCED2	English	ISCED3	CB
French Community of Belgium	English	ISCED3	German	ISCED3	CB
German Community of Belgium	French	ISCED2	English	ISCED3	PB
Bulgaria	English	ISCED3	German	ISCED3	PB
Croatia	English	ISCED2	German	ISCED2	CB, PB
England	French	ISCED3	German	ISCED3	PB
Estonia	English	ISCED2	German	ISCED2	CB, PB
France	English	ISCED2	Spanish	ISCED2	PB
Greece	English	ISCED2	French	ISCED2	PB
Malta	English	ISCED2	Italian	ISCED2	PB
Netherlands	English	ISCED2	German	ISCED2	CB
Poland	English	ISCED2	German	ISCED2	PB
Portugal	English	ISCED2	French	ISCED2	CB
Slovenia	English	ISCED2	German	ISCED2	PB
Spain	English	ISCED2	French	ISCED2	PB
Sweden	English	ISCED2	Spanish	ISCED2	CB, PB

<sup>2</sup> Note, this refers only to the first and second most widely taught languages out of English, French, German, Italian and Spanish. For several adjudicated entities, their first or second most widely taught language is not one of these languages.

<sup>3</sup> The ESLC was carried out independently in the three constituent communities of Belgium

**Table 2 Tested languages summary**

Language	Number of countries testing language as first most widely taught language	Number of countries testing language as second most widely taught language
English	13	2
French	3	3
German	0	8
Italian	0	1
Spanish	0	2

**Table 3 Tested levels summary**

	Number of countries testing ISCED 2	Number of countries testing ISCED 3
First most widely taught language	13	3
Second most widely taught language	11	5

**Outcomes** – the ESLC delivers the following outcomes:

- A profile of the language proficiency of sampled students. Contextual indicators providing a broad range of information on the context of foreign language teaching policies and foreign language learning at student, teacher and school level.
- Information on the relationship between language proficiency and the contextual indicators.
- A resource and knowledge base for policy analysis and research.

## **1.2 This technical report**

This technical report is concerned with the technical aspects of the ESLC whereas the Final Report is concerned with the results of the ESLC. Policy recommendations are outlined in the conclusions of the Final Report and not discussed in this report.

This technical report describes the methodologies and procedures adopted to enable the ESLC to provide high quality data to support the European Commission in this area of policy. The descriptions are provided at a level that will enable review of the implemented procedures and solutions to the challenges faced.

The report covers the following areas:



- *Instrument design*: Chapters 2 and 3 describe the development of the language tests to produce measures comparable across languages and interpretable in relation to the CEFR, and the questionnaires, to address a range of European language policy issues.
- *Operations*: Chapter 4 describes the sampling procedures, Chapter 5 the translation of the questionnaires, Chapter 6 the innovative software platform developed for the ESLC to support both paper-based and computer-based administration, Chapter 7 the field operations and Chapter 8 the approach taken to quality monitoring.
- *Data processing, scale construction and data products*: Chapter 9 describes the handling of sampling weights, Chapter 10 design of the questionnaire indices, Chapter 11 the approach to setting CEFR-related standards for the five languages, Chapter 12 the analyses, and Chapter 13 the development and the contents of the data sets.
- *Appendices*: Examples of the language test task types, the complete set of Main Study Questionnaires, the sampling forms, the Technical Standards and a comprehensive report on multiple marking of Writing.

### 1.3 References

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

European Commission (2005) *Commission Communication of 1 August 2005 - The European Indicator of Language Competence [COM(2005) 356 final - Not published in the Official Journal]*, retrieved 18 January 2012, from [http://europa.eu/legislation\\_summaries/education\\_training\\_youth/lifelong\\_learning/c11083\\_en.htm](http://europa.eu/legislation_summaries/education_training_youth/lifelong_learning/c11083_en.htm)

European Commission (2007a) *Communication from the Commission to the Council of 13 April 2007 entitled 'Framework for the European survey on language competences'* [COM (2007) 184 final – Not published in the Official Journal]

European Commission (2007b) *Terms of Reference: Tender no. 21 'European Survey on Language Competences'*, Contracting Authority: European Commission.

# ***Chapter 2:***

# ***Instrument***

# ***development –***

# ***Language Tests***

## 2 Instrument development - Language tests

The ELSC is a collaborative effort by the participating countries, the European Commission and SurveyLang, guided by shared policy-driven interests. Each partner is responsible for particular areas of the survey and although these work areas vary in size, each is vital in ensuring the project's success. The aim is to deliver an indicator of language competences to provide information on the general level of foreign language knowledge of the pupils in the Member States in order to help policy makers, teachers and practitioners to take decisions how to improve the foreign language teaching methods and thus the performance of pupils. The aim of the SurveyLang language testing group has been to develop language tests the results of which are comparable across the five languages and all participating countries.

Developing the language tests was methodologically complex, requiring intensive collaboration among the members of the language testing group: University of Cambridge ESOL Examinations (Cambridge ESOL), Centre international d'études pédagogiques (CIEP), Goethe Institut, Università per Stranieri di Perugia and Universidad de Salamanca.

The successful delivery of the language test instruments depended on the use, and further development, of state of the art methodologies and technologies. This chapter describes the processes adopted to develop the language tests, and support the development of a European Indicator of Language Competences.

The approach adopted by SurveyLang in designing the language test instruments is summarised as follows:

- (i) define a language testing framework that incorporates the aims and objectives of the ESLC
- (ii) out of this framework, develop initial specifications, a set of draft task types and a draft test development process
- (iii) pilot the initial specifications and draft task types
- (iv) gather feedback from all relevant stakeholders including the Advisory Board, the participating countries, teachers and students. Review this feedback together with the analysis of the pilot results
- (v) further develop the initial specifications into final item writer guidelines and agree on a collaborative test development process to be shared across the five languages
- (vi) undertake a rigorous item development programme in order to develop language tests for the Main Study, the results of which would be comparable across the five languages and all participating countries.



To ensure that the items used in the Main Study were fit for purpose and of the required level of quality, the language testing team produced and trialled a large number of items over the course of the development programme. Over 100 tasks were piloted in 2008 in order to finalise the test specifications and agree on the most appropriate task types to be used in the ESLC. The team then produced over 500 tasks (2200+ items) which were then exhaustively trialled through the Pretesting and Field Trial stages before the best-performing items were selected. For the Main Study, 143 tasks (635 items) were used across the five languages.

The first part of this chapter describes the language testing framework that incorporates the aims and objectives of the ESLC and provides the basis for the development of the language testing instruments. Section 2.2 describes the item development process that was designed to allow the language partners to work together in a highly collaborative and intensive way. From section 2.2 the text goes on to describe the different work areas within the detailed, multi-stage development cycle designed to deliver high-quality, fit for purpose language tests. Section 2.5 describes the final test design implemented in the Main Study.

## **2.1 *Development of the language testing framework***

The Commission specified The Common European Framework of Reference for Languages: Learning, Teaching, Assessment as the framework against which to measure language learning outcomes for the ESLC, reflecting the widespread impact which this document has had since its publication in 2001. The language tests developed for the ESLC set out to reflect the CEFR's action-oriented, functional model of language use, while ensuring relevance for 15–17 year-olds in a school setting. The socio-cognitive model adopted is based on the CEFR's model of language use and learning, and identifies two dimensions – the social dimension of language in use, and the cognitive dimension of language as a developing set of competences, skills and knowledge. Applying these allowed the definition of testable abilities at each proficiency level. To enable the resulting test construct to be implemented comparably across languages, these abilities were mapped to specific task types, drawing chiefly on task types which had been used successfully by SurveyLang's language partners in their operational exams.

The approach to developing the language testing framework by SurveyLang is summarised as follows:

- identify the relevant aims and objectives of the ESLC, including the language skills to be tested
- for each skill, identify the test content and a set of testable subskills or abilities derived from a socio-cognitive model of language proficiency and a listing of language functions or competences found to be salient at each level from A1 to B2 in the descriptor scales of the CEFR
- identify the most appropriate task types to test these subskills

- create a test design that presents combinations of tasks to students in such a way as to maximise the quality of interpretable response data collected while not overburdening the sampled students
- adopt a targeted approach to testing where pupils are given a test at an appropriate level of challenge
- develop specifications, item writer guidelines and a collaborative test development process that are shared across languages in order to produce language tests that are comparable.

These steps are described in this chapter.

### 2.1.1 Requirements of the language tests

A number of key aims and objectives of the ESLC impacted on the design of the language testing instruments:

- for each country, the ESLC should cover tests in the first and second most commonly taught official European languages in the European Union from English, French, German, Italian and Spanish
- test performance should be interpreted with reference to the scale of the Common European Framework of Reference for languages (CEFR)
- the tests should assess performance at levels A1-B2 of the CEFR
- performance should be reported at the level of the group, not the individual
- the ESLC should assess competence in the 3 language skills which may be assessed most readily, i.e. Listening comprehension, Reading comprehension and Writing
- instruments for testing in these 3 competences should be developed, taking into account the previous experience and knowledge in the field at international, Union and national level
- results must be comparable across 5 languages and all participating countries
- tests must be available in both paper-based and computer-based formats.

Previous international surveys had translated tests across languages but it was a key aim of this survey to create parallel but not identical tests across the five languages, thereby making the issue of cross language comparability a crucial one.

### 2.1.2 Defining test content in terms of the CEFR

Test content was approached using the categories proposed by the CEFR (Council of Europe 2001 chapter 4). As the CEFR stresses, these categories are illustrative and suggestive, rather than exhaustive. However, the listed elements provide a useful starting point for selecting appropriate content.

The CEFR identifies four basic *domains* of language use:

- personal



- public
- educational
- professional

The first three are most relevant to the ESLC. The CEFR illustrates each domain in terms of *situations* described in terms of:

- the locations in which they occur
- relevant institutions or organisations
- the persons involved
- the objects (animate and inanimate) in the environment
- the events that take place
- the operations performed by the persons involved
- the texts encountered within the situation.

*Communication themes* are the topics which are the subjects of discourse, conversation, reflection or composition. The CEFR refers to the categories provided in Threshold {Van Ek, 1998 #722}, which appear in very similar form in the Waystage and Vantage levels {Van Ek, 1998 #723}, {Van Ek, 2000 #4292}. These too provide a useful starting point for selecting appropriate content. Example headings of these are:

- personal identification
- house and home, environment
- daily life
- free time, entertainment
- travel
- relations with other people.

Below these major thematic headings are sub-themes, each of which defines a range of *topic-specific notions*. For example, area 4, 'free time and entertainment', is subcategorised in the following way:

- leisure
- hobbies and interests
- radio and TV
- cinema, theatre, concert, etc.
- exhibitions, museums, etc.
- sports.

Topic-specific notions contrast with *general notions*, which are the meanings and concepts expressed through language whatever the specific situation. The lexicogrammatical means through which such general notions are expressed are an important aspect of selection and sequencing content in a communicatively-oriented syllabus.

Similarly the list of *language functions* provided in the Waystage-Threshold-Vantage levels, and discussed in the CEFR as an aspect of pragmatic competence, provide a general rather than setting-specific taxonomy of language in social use. The major headings relevant to the tested skills are:

- imparting and seeking information
- expressing and finding out attitudes
- deciding and managing course of actions (Suasion)
- socialising
- structuring discourse.

Together these *communication themes*, *notions* and *functions* provided the basis for categorising and selecting texts for use in the ESLC. The final choice of test content was made by considering the approach proposed by the CEFR in conjunction with the characteristics of the target language users, i.e. the 15–17 year old students participating in this survey.

Consideration of which domains of language use are most relevant to target language learners at different proficiency levels informed a decision as to the proportion of tasks relating to each of the *domains* mentioned above across the four levels of the ESLC.

**Table 4 Domain distribution across levels**

	A1	A2	B1	B2
personal	60%	50%	40%	25%
public	30%	40%	40%	50%
educational	10%	10%	20%	20%
professional	0%	0%	0%	5%

Each domain was then divided into topics and sub-topics as specified below:

Personal:

- family: family celebrations and events, relationships (parent-child, brothers-sisters, grandchildren-grandparents)
- friends: groups versus individuals, relationships between boys and girls, peer group identity, personal character, shared problems, shared tastes, hobbies
- leisure: sport, music, cinema, internet, reading, going out
- home: family, at friends', ideal home environment
- objects: those related to new technology (telephone, game consoles, computers, etc.), those related to fashion and brands
- pets: presence/absence, relations with animals.

Public:

- people: sports stars, musicians, actors, etc.

- official: representatives of the law (justice, police, etc.), administration, associations
- going out: cinema, restaurant, discotheques, stadiums, swimming pool, theatre, concerts, shopping
- holidays: beach, mountain, town, country, foreign travel
- objects: favourite food, relationships with money, modes of transport (bicycle, motorbike, learning to drive with parents, etc.)
- events: accidents, illness, health.

Educational:

- people: students, teachers, school staff
- school trips: exchanges with penpals, discovering a country, sociocultural experiences, studying languages abroad
- objects: books, other purchases for school, classroom equipment
- places: primary and secondary school, classrooms, school environment
- events: school festivals, open days, exam results, shows, etc.

Professional:

- people: careers advisors, representatives of the world of work
- professions: choice of future profession, favourite and least favourite jobs
- accessing the job market: workshops for students, documents outlining jobs and careers
- events: work placements and sandwich courses, summer jobs, etc.

As the above list suggests, domains overlap, and some tasks might be classified under more than one domain. To ensure adequate coverage across the ESLC, domains and topics were assigned to tasks at the commissioning stage.

It was important that test materials did not contain anything that might offend or upset candidates, thereby potentially affecting their performance or distracting them during the examination. Thus, certain topics such as war, politics, serious family problems, etc, were considered unsuitable. A detailed list of unsuitable topics was provided in the Item Writer Guidelines.

### **2.1.3 The constructs of Reading, Listening and Writing**

The socio-cognitive validation framework proposed by {Weir, 2005 #726}, an approach coherent with other recent discussions of theories of test design, was adopted as the means to identify the subskills to be tested. This complements the CEFR's treatment of the cognitive dimension and provides useful practical models of language skills as cognitive processes and ways of refining a description of progression.

## 2.1.4 The construct of Reading

Over the last century, reading research has moved from viewing the reading process as a bottom-up process to a top-down process and finally to an interactive one. Bottom-up models of reading comprehension pictured proficient readers as those who process a written text by working their way up the scale of linguistic units starting with identification of letters, then words, then sentences and finally text meaning. In top-down models, comprehension takes place when readers integrate incoming information with their existing 'schemata' (i.e. their knowledge structures); meaning is constructed as the readers integrate what is in the text and what they already have. Interactive models of reading comprehension expect both directions of processing (i.e. top-down and bottom-up) to proceed simultaneously as well as to interact and influence each other: 'reading involves the simultaneous application of elements such as context and purpose along with knowledge of grammar, content, vocabulary, discourse conventions, graphemic knowledge, and metacognitive awareness in order to develop an appropriate meaning' {Hudson, 1991 #5854}.

The process of reading can thus be regarded as an interaction of the reader's conceptual abilities and process strategies, language knowledge and content knowledge. This cognitive view of reading is currently shared by researchers in the fields of psycholinguistics, cognitive psychology, and language assessment and it applies to both L1 and L2 reading ability. A parallel sociolinguistic and discourse analytic view considers how textual products function within a given context, e.g. educational, socio-political, etc.

Weir (2005) brings together these two perspectives in a socio-cognitive framework for test validation. It allows us to describe progression across the CEFR levels to be surveyed in a way which practically informs test design and item writing. The cognitive validity of a reading task is a measure of how closely it elicits the cognitive processing involved in contexts beyond the test itself, i.e. in performing reading task(s) in real life.

Different types or purposes for reading are identified which employ different strategies and processing. A distinction is made between *expeditious* and *careful* reading, and between *local* and *global* reading (i.e. understanding at the sentence level or the text as a whole). In terms of cognitive demand a general progression is posited as follows:

- scanning – reading selectively to achieve very specific goals such as finding a name or date
- careful local reading – establishing the basic meaning of a proposition
- skimming for gist – quick superficial reading – “what is this text about”
- careful global reading for comprehending main idea(s). Global reading activates all components of the model, building a mental model that relates the text to the reader's knowledge of the world
- search reading for main idea(s). Search reading is sampling the text to extract information on a predetermined topic, when the reader is not sure what form the information may appear in

- careful global reading to comprehend a single text
- careful global reading to comprehend several texts.

## Reading at A1

The CEFR illustrative scales stress the very limited nature of reading competence at A1. Learners at this level can ‘recognise familiar names, words and very basic phrases’, ‘understand very short, simple texts a single phrase at a time’, ‘understand short, simple messages on postcards’, ‘follow short, simple written directions’ and ‘get an idea of the content of simpler informational material and short simple descriptions, especially if there is visual support’.

Decoding text and accessing lexical meaning represents a major cognitive load at this level. This limits capacity to apply syntactic knowledge to parse sentences and establish propositional meanings at clause or sentence level. Capacity to infer meaning is very limited, hence the importance of non-linguistic (e.g. graphic) support.

Appropriate communication themes relate to the personal and familiar, e.g. personal identification, house and home, environment, daily life, free time and entertainment. Appropriate macrofunctions for continuous texts are narration and description. Non-continuous texts (notices, advertisements etc) are appropriate for testing the ability to find specific information. Texts used in test tasks at A1 will be semi-authentic, i.e. controlled for lexicogrammatical difficulty.

## Reading abilities tested at A1

1	Reading a simple postcard or email, identifying factual information relating to personal and familiar themes
2	Understanding word-level topic-specific notions from personal and familiar domains
3	Understanding general notions (existential, spatial, relational) as used to describe pictures or graphically displayed information
4	Finding predictable factual information in texts such as notices, announcements, timetables, menus, with some visual support
5	Understanding signs, notices and announcements

## Reading at A2

Reading at A2 is described as still quite limited: learners can understand ‘short, simple texts containing the highest frequency vocabulary’, or ‘short simple personal letters’. However there is reference to a wider range of text types: ‘everyday signs and notices: in public places, such as streets, restaurants, railway stations’, or ‘letters, brochures and short newspaper articles describing events’. There is also a suggestion of some functional competence: ‘e.g. use the Yellow Pages to find a service or tradesman’, or ‘understand basic types of standard routine letters and faxes (enquiries, orders, letters of confirmation etc.) on familiar topics’.



More automated decoding enables the learner to deal with longer texts and make more use of syntactic knowledge to parse sentences and establish propositional meanings at clause or sentence level. The learner can begin to infer meanings of unknown words from context.

Themes are as A1, plus routine everyday transactions, e.g. free time and entertainment, travel, services, shopping, food and drink. Appropriate macrofunctions for continuous texts are narration, description and instruction. Non-continuous texts (notices, advertisements etc) are appropriate for testing the ability to find specific information. Texts used in test tasks at A2 will be semi-authentic, i.e. controlled for lexicogrammatical difficulty.

### Reading abilities tested at A2

4	Finding predictable factual information in texts such as notices, announcements, timetables, menus, with some visual support
5	Understanding signs, notices and announcements
6	Understanding the main ideas and some details of longer texts (up to c. 230 words)
7	Understanding routine functional exchanges, as occur in emails or conversation
8	Understanding personal letters
9	Understanding lexicostuctural patterns in a short text
10	Reading several short texts for specific information and detailed comprehension

### Reading at B1

The illustrative scales describe a useful functional competence with respect to texts which are 'short, simple', 'everyday', 'straightforward' concerning 'familiar matters of a concrete type'. The B1 reader can 'understand the description of events, feelings and wishes in personal letters well enough to correspond regularly with a pen friend'. Moreover, s/he can 'identify the main conclusions in clearly signalled argumentative texts', and 'recognise significant points in straightforward newspaper articles on familiar subjects'. Other text types referred to include 'letters, brochures and short official documents', 'advertisements, prospectuses, menus, reference lists and timetables'.

Better able to establish meanings at clause or sentence level, the B1 reader can begin to use inference and apply topical or general knowledge to building a mental model of the text as a whole. This corresponds to the notion of careful global reading to comprehend main ideas, in Weir's model, and relates to the PISA process of forming a broad understanding.

Range at B1 is still mainly limited to familiar, concrete themes, but there is more scope to introduce topics of general interest, including argumentative texts. Comprehension



extends beyond the retrieval of specific factual information to understanding the main points of longer texts, including identifying opinions and points of view. Texts used in test tasks at B1 will mostly be semi-authentic, i.e. controlled for lexicogrammatical difficulty.

### Reading abilities tested at B1

7	Understanding signs, notices and announcements
8	Understanding personal letters
9	Understanding lexicostructural patterns in a short text
10	Reading several short texts for specific information and detailed comprehension
11	Scanning a factual text for specific information
12	Reading for detailed comprehension and global meaning, understanding attitude, opinion and writer purpose
13	Using understanding of text structure, cohesion and coherence

### Reading at B2

The B2 reader 'can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. S/he can 'quickly identify the content and relevance of news items, articles and reports on a wide range of professional topics'. A wide range of more challenging text types is referred to: 'articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints', 'contemporary literary prose', 'specialised articles outside his/her field'. The B2 reader has a broad active reading vocabulary, though will still need to refer to a dictionary.

Already confident in the process of *careful global reading* to comprehend main ideas, the B2 reader can apply knowledge of text structure (genre, rhetorical tasks) to construct a text level understanding. This relates to the PISA process of *developing an interpretation*.

B2 readers can deal with a range of themes beyond the entirely familiar; however, it is important that topics selected for the ESLC should be relevant and interesting for the population tested. Informative, argumentative and expository texts will be appropriate, and may be taken from authentic sources.

### Reading abilities tested at B2

10	Reading several short texts for specific information and detailed comprehension
11	Scanning a factual text for specific information
12	Reading for detailed comprehension and global meaning, understanding attitude, opinion and writer purpose
13	Using understanding of text structure, cohesion and coherence

### 2.1.5 The construct of Listening

While reading has to be taught, listening ability in one's own language occurs naturally. In this sense, listening is the more basic form of language comprehension.

Nonetheless, many of the processes of language comprehension are assumed to be common to listening and reading. What is specific to listening is how speech is perceived, and the core process in this is word recognition. Moreover, once words have been recognised the prosodic and intonational structure of speech plays a key role in subsequent syntactic and discourse processing.

Traditionally, there have been two approaches to defining the listening construct – competence-based and task-based. The competence-based approach assumes that consistencies in listening performance are due to the characteristics of the test-taker and that test scores indicate the level of underlying competence that manifests itself across a variety of settings and tasks. Models of communicative competence set out to describe as comprehensively as possible the knowledge and skills L2 learners need in order to use the language (i.e. the listening skill) effectively. In the assessment context, however, a major disadvantage of the competence-based approach is that it can be very difficult to determine which test items actually assess the (sub)competencies (or subskills) of interest.

An alternative approach to defining the listening construct assumes that consistencies in listening performance are due to the characteristics of the context in which the listening takes place. In this more task-focused approach the interest is in what the test-takers can do under specific circumstances. The main problem, though, is how to define the target-language use (TLU) situation in an appropriate way for testing purposes. Do we need to cover all possible situations? If so, how can we realistically achieve this? And if not, which situations should we select? These issues have practical implications for available resources and pose significant challenges for establishing task comparability across test versions.

Buck (2001:108) proposes a construct definition for listening based on the interaction between competence and task: 'when making test tasks, the important thing is not that the test task is similar to the target-language use task, but the interaction between the test-taker and the test task is similar to the interaction between the language user and the task in the target-language use situation.'

In this approach both traits and tasks are used as the basis for construct definition and test tasks are regarded as requiring similar competencies. This interactive approach is consistent with the premise that use of language skills such as listening, reading, etc. is both psycholinguistically driven (i.e. competency-focused) and contextually driven (i.e. task-in-situation-focused).

Careful definition of the contextual parameters of the target-language use context can help determine which type of speakers, accents, level of phonological modification, speed, vocabulary, syntax, discourse structures, rhetorical functions and types of inferencing we need to include in the test. Additionally, we can determine the language functions, communicative load, the pragmatic implications (i.e. indirect meanings) and the appropriacy of linguistic forms. With these contextual features in mind, we can then select appropriate listening texts; we can also determine the cognitive skills and metalinguistic strategies that are of interest and construct appropriate tasks.

## **Listening at A1**

The illustrative descriptors stress that A1 represents a very low level of listening competence. The A1 listener can ‘follow speech which is very slow and carefully articulated, with long pauses for him/her to assimilate meaning’. Comprehension is limited to ‘familiar words and very basic phrases’, concerning immediate concrete topics such as personal identification and family.

Decoding speech to identify words and access lexical meaning represents a major cognitive load at this level. This severely limits capacity to apply syntactic knowledge to parse sentences and establish propositional meanings at clause or sentence level. A1 listeners operate in the here-and-now. They extract meanings at word and phrase level, heavily dependent on cues provided by the immediate context.

Appropriate communication themes relate to the immediate and personal, e.g. personal identification, house and home, the immediate environment, daily life. There should be coverage of general notions such as numbers, days, letter-sounds, and basic existential, spatial, temporal or quantitative notions. Basic social language functions may be tested.

Texts are very short dialogues and monologues, often with visual support. Texts used in test tasks at A1 will be semi-authentic, i.e. controlled for lexicogrammatical difficulty and delivered slowly, though with natural pronunciation, intonation and stress.

## **Listening abilities tested at A1**

1	Recording specific information in announcements or messages
2	Understanding general or topic-specific notions describing pictures or graphically-displayed information
3	Identifying communicative function
4	Identifying the situation and/or the main idea in announcements, messages or conversations (short)

## Listening at A2

The A2 listener can deal with a limited range of topics: 'e.g. very basic personal and family information, shopping, local geography, employment'. S/he can deal with speech 'delivered slowly and clearly'. S/he can 'catch the main point in short, clear, simple messages and announcements', and 'understand simple directions relating to how to get from X to Y'.

More automated decoding, particularly with respect to familiar word sequences, enables the learner to deal with slightly longer texts and make more use of syntactic knowledge to parse sentences and establish propositional meanings. A2 listeners are still dependent on sympathetic interlocutors and on contextual cues for understanding.

Communication themes as A1, plus routine everyday transactions, e.g. free time and entertainment, travel, services, shopping, food and drink.

Texts are short dialogues and monologues, often with visual support. Texts used in test tasks at A2 will be semi-authentic, i.e. controlled for lexicogrammatical difficulty and delivered slowly, though with natural pronunciation, intonation and stress.

### Listening abilities tested at A2

1	Recording specific information in announcements or messages
2	Understanding general or topic-specific notions describing pictures or graphically-displayed information
3	Identifying communicative function
4	Identifying the situation and/or the main idea in announcements, messages or conversations (short)
5	Understanding a longer dialogue (conversation, interview) True/False

## Listening at B1

The illustrative descriptors for B1 identify a useful functional competence, though still limited in range. The B1 listener can understand 'straightforward factual information' about 'familiar matters regularly encountered in work, school, leisure etc', 'short narratives', 'simple technical information' and 'the main points of radio news bulletins and simpler recorded material'. S/he can understand 'both general messages and specific details', always provided that speech is 'clearly and slowly articulated'.

The B1 listener can process clearly-spoken texts sufficiently automatically to begin using inference and topical or general knowledge to build a mental model of the text as a whole. S/he has sufficient autonomy to use listening to learn new language.

Range covers the same familiar, concrete themes as A2, with some scope to introduce topics of general interest. Comprehension concerns understanding main points as well

as details, including identifying opinions and points of view. Listening texts used in test tasks at B1 will be semi-authentic, i.e. controlled for lexicogrammatical difficulty.

### Listening abilities tested at B1

4	Identifying the situation and/or the main idea in announcements, messages or conversations (short)
5	Understanding a longer dialogue (conversation, interview) True/False
6	Understanding a longer dialogue (conversation, interview) MCQ
7	Understanding monologue (presentation, report) and interpreting information

### Listening at B2

The illustrative descriptors identify a wide-ranging functional competence, covering 'propositionally and linguistically complex speech on both concrete and abstract topics', 'technical discussions', 'extended speech and complex lines of argument', 'lectures, talks and reports and other forms of academic/professional presentation'. This assumes familiarity with the topic, speech 'delivered in a standard dialect', and presentation which is 'straightforward and clearly structured'. The B2 listener can 'identify speaker viewpoints and attitudes as well as the information content', and 'identify the speaker's mood, tone etc.' when listening to recorded or broadcast audio material.

Although listening to more complex texts requires conscious effort, the B2 listener has sufficiently automated decoding skills to focus on constructing text level understanding.

B2 listeners can deal with a range of themes beyond the entirely familiar; however, it is important that topics selected for the ESLC should be relevant and interesting for the population tested. Informative, argumentative and expository texts will be appropriate, and may be taken from authentic sources.

### Listening abilities tested at B2

4	Identifying the situation and/or the main idea in announcements, messages or conversations (short)
5	Understanding a longer dialogue (conversation, interview) True/False
6	Understanding a longer dialogue (conversation, interview) MCQ
7	Understanding monologue (presentation, report) and interpreting information

## 2.1.6 The construct of Writing

For many years the notion of writing was decontextualised and regarded primarily as product-oriented, where the various elements are coherently and accurately put

together according to a rule-governed system; the text product was seen as an autonomous object and writing was considered independent of particular writers or readers {Hyland, 2002 #663}. Written products were largely viewed as ideal forms capable of being analysed independently of any real-life uses.

More recently, writing has come to be viewed as a strongly contextualised phenomenon which should not be disconnected from the writer and the audience/purpose for whom/which the writer is writing. According to {Hayes, 1996 #5173}, writing is fundamentally a communicative act: 'We write mainly to communicate with other humans'. {Hamp-Lyons, 1997 #2987} offer a similar broad, conceptual view of writing: 'an act that takes place within a context, that accomplishes a particular purpose, and that is appropriately shaped for its intended audience'.

According to this view, the linguistic patterns employed in a piece of writing are influenced by contexts beyond the page which bring with them a variety of social constraints and choices. The writer's goals, relationship with readers and the content knowledge s/he wants to impart are accomplished by the text forms appropriate to that social context. This constitutes a socio-cognitive model of writing as Communicative Language Use which takes into account both internal processing (i.e. cognitive or psycholinguistic) and external, contextual factors in writing. Writing is considered a social act taking place in a specifiable context so particular attention needs to be paid to:

- the writer's understanding of the knowledge, interests and expectations of a potential audience and the conventions of the appropriate discourse community as far as this can be specified
- the purpose of the writing
- the writer taking the responsibility for making explicit the connections between the propositions and ideas they are conveying and structuring their writing
- the importance of the demands the task makes in terms of language knowledge: linguistic, discoursal and sociolinguistic, and content knowledge.

Research indicates that categories of L2 learners can be differentiated from each other by their age, standard of education, L1 literacy and by their ability and opportunity to write in a second language. These differences are especially important when constructing or developing appropriate tests of writing. A definition of writing ability for a specific context therefore needs to take account of the group of L2 writers identified and the kinds of writing they would typically produce.

In line with current views on the nature of writing, the model adopted for this survey looks beyond the surface structure manifested by the text alone; it regards the text as an attempt to engage the reader communicatively. The socio-cognitive approach is adopted where attention is paid to both context-based validity and to cognitive validity. Context-based validity addresses the particular performance conditions or the setting under which it is to be performed (such as purpose of the task, time available, length, specified addressee, known marking criteria as well as the linguistic demands inherent in the successful performance of the task) together with the actual examination

conditions resulting from the administrative setting. Cognitive processing in a writing test never occurs in a vacuum but is activated in response to the specific contextual parameters set out in the test task rubric. These parameters relate to the linguistic and content demands that must be met for successful task completion as well as to features of the task setting that serve to delineate the performance required.

## Writing at A1

A1 is identified as a very low level of competence, limited to 'simple isolated phrases and sentences'. Topics are the most immediate and personal: A1 learners can write about 'themselves and imaginary people, where they live and what they do', 'a short, simple postcard', and personal details, numbers and dates such as on a hotel registration form.

The A1 learner can produce very short texts based on a few learned phrases. S/he will rely heavily on models and can only adapt these in limited, simple ways.

As indicated by the above CEFR descriptors, A1 writing themes are immediate, personal and stereotypical. Postcards, notes and emails are appropriate text types. Forms have apparent authenticity at this level; however, they tend to test reading as much as writing at this level.

## Writing abilities tested at A1

1	Expressing general or topic-specific notions describing pictures or graphically-displayed information
2	Writing an email/postcard
3	Completing a form

## Writing at A2

Writing at A2 is limited to 'a series of simple phrases and sentences linked with simple connectors like "and", "but" and "because"'. Topics referred to include 'family, living conditions, educational background, present or most recent job', 'imaginary biographies and simple poems about people', 'matters in areas of immediate need', 'very simple personal letters expressing thanks and apology'. Letters and notes will tend to be 'short, simple' and 'formulaic'.

The A2 learner can begin to use writing as a genuine communicative act and thus form a conception of purpose and target reader. S/he can begin to use and adapt syntactic patterns to generate new propositions. Appropriate tasks relate to routine, everyday themes; basic personal and family information, school, free time, holidays, familiar events. Forms of writing include short letters and notes, possibly based on transforming information provided in text or graphic form.



## Writing abilities tested at A2

2	Writing an email/postcard
3	Completing a form
4	Completing a text, showing understanding of lexicogrammatical relations
5	Writing a referential text (intended to inform)

## Writing at B1

The illustrative descriptors at B1 identify a limited functional competence. The B1 writer can produce 'straightforward connected texts', 'by linking a series of shorter discrete elements into a linear sequence'. Text types referred to include: 'very brief reports to a standard conventionalised format', 'personal letters and notes', story narration and 'very short, basic descriptions of events, past activities and personal experiences'.

Topics include 'experiences, describing feelings and reactions', 'everyday aspects of his/her environment, e.g. people, places, a job or study experience', and 'messages communicating enquiries, explaining problems'.

The B1 learner still finds it difficult to plan, but can compose a simple referential text particularly given a clear set of content points to work from. S/he has a greater awareness of lexicogrammatical dependencies and may be able to self-correct.

## Writing abilities tested at B1

4	Completing a text, showing understanding of lexicogrammatical relations
5	Writing a referential text (intended to inform)
6	Writing a conative text (intended to persuade or convince)
7	Editing a piece of writing

## Writing at B2

The illustrative descriptors at B2 identify a good functional competence over a range of topic areas. The B2 writer can produce 'clear, detailed texts on a variety of subjects related to his/her field of interest, synthesising and evaluating information and arguments from a number of sources'. S/he can write 'an essay or report which develops an argument, giving reasons in support of or against a particular point of view and explaining the advantages and disadvantages of various options', and 'can synthesise information and arguments from a number of sources'. S/he can 'convey information and ideas on abstract as well as concrete topics', 'write letters conveying degrees of emotion' and 'express news and views effectively'.



The B2 learner can plan a piece of writing with a given audience in mind, and organize arguments. S/he can engage in 'knowledge transforming', rather than simply 'knowledge telling'. More extensive written stimuli provide a basis for constructing an argument or expressing opinions; reacting to an issue, etc. Letters, essays, reports are appropriate texts.

### **Writing abilities tested at B2**

5	Writing a referential text (intended to inform)
6	Writing a conative text (intended to persuade or convince)
7	Editing a piece of Writing

How the language testing framework presented above was implemented in the language tests is the subject of the next section.

## **2.2 Development of the language tests**

### **2.2.1 Major stages in the development process**

There were five main stages in the development of the language testing instruments, which can be summarised as follows:

- 2008 Development of the language testing framework
- 2008 The Pilot Study
- 2009 Pretesting
- 2010 The Field Trial
- 2011 The Main Study

The Pilot Study constituted a small-scale trial of proposed task types, and an exploration of collaborative working methods that would favour consistency of approach. A total of 106 tasks were developed across the skills and languages, with each language partner focusing on a different part of the ability range.

Pretesting was a large-scale trial of all the test material developed for potential use in the Main Study. Over 2000 items were pretested. Across languages 50 Reading tests, 35 Listening tests and 60 Writing tests were produced. Tasks were administered in schools made available by the NRCs and the language partners' centre networks. Most of these schools were in Europe. A total of 8283 students participated.

Table 5 below shows the countries which participated in pretesting and the numbers of students per country. For all languages the students were well distributed over the tested ability levels. This sample was wholly adequate for the purposes of pretesting.

Tasks for the Field Trial were selected on the basis of the pretest review, during which a third or more of tasks were dropped. This still meant that twice as much material could be used in the Field Trial as would be needed for the Main Study,

The Field Trial had the important aim of testing out the major technical and human systems upon which successful delivery of the Main Study depended:

- test construction, printing and despatch procedures
- delivery and administration of the language tests in both paper-based and computer-based formats.

Additionally it provided a final opportunity to identify any poorly performing tasks, to ensure that the test content (topic, cognitive demand, etc) was fully appropriate for the target population, and to revise features of the test design and administration procedures.

**Table 5 Pretesting: countries participating and numbers of students**

<b>English</b>		<b>French</b>	
Belgium	112	Bulgaria	150
Bosnia & Herzegovina	120	Estonia	16
Croatia	30	Ireland	100
Italy	375	Italy	295
Poland	200	Netherlands	695
Portugal	30	Scotland	274
Russia	150	Spain	195
Spain	355	Sweden	163
Ukraine	225	Turkey	105
<b>Grand Total</b>	<b>1597</b>	<b>Grand Total</b>	<b>1993</b>
<b>German</b>		<b>Italian</b>	
Belarus	280	Ireland	96
Brazil	49	Spain	244
Burkina Faso	148	Switzerland	406
Croatia	30	<b>Grand Total</b>	<b>746</b>
Denmark	84	<b>Spanish</b>	
Egypt	0	Belgium	73
Finland	36	Bulgaria	285
Germany	7	Czech republic	105
Ireland	73	France	655
Kazakhstan	41	Holland	96
Latvia	60	Hungary	184
Mexico	46	Italy	423
Portugal	15	Poland	199
Senegal	34	Portugal	139
Slovakia	30	Romania	54
Spain	17		



Turkey	442
UK	146
<b>Grand Total</b>	<b>1538</b>

Slovakia	66
Slovenia	47
Switzerland	18
Turkey	65
<b>Grand Total</b>	<b>2409</b>

Tasks were selected for the Main Study on the basis of the Field Trial review. In the Field Trial two tasks of each task type at each level were trialled in each language for all skills. The best-performing task in each pair was selected for the Main Study.

Each of these development stages contributed to the specification of the tests, in terms of content and task types, to the construction of a large body of test tasks, and to their progressive refinement through a series of empirical trials and the collection of qualitative feedback.

An important conditioning factor was the collaborative working methodology itself, developed by the language partners in order to maximize the quality and the comparability of the final tests.

### 2.2.2 General test design considerations

As in most complex surveys, each sampled student was to see only a proportion of the total test material. The total amount of test material was determined by the need to achieve adequate coverage of the construct; that is, to test all aspects of a skill considered important at a given level. In order to avoid fatigue or boredom effects for individual students it was necessary to utilise an incomplete but linked design where each student would receive only a proportion of the total test material.

A design constraint was adopted that the total language test time for a student should not exceed 60 minutes. A test for one skill would comprise 30 minutes of material. A student would only be tested in two of the three skills. Individual students would therefore receive Reading and Listening, Reading and Writing, or Listening and Writing. Students would be assigned randomly to one of these three groups.

The design needed to be implemented in the same way in each of the five languages, as consistency of approach would maximise the comparability of outcomes.

### 2.2.3 Targeted testing

An additional complexity followed from the early decision by SurveyLang to offer students a test targeted at their general level. This was important because the range of ability tested was very wide, and a single test covering this range would have been not only inefficient in terms of the information it provided about students' level, but also demotivating for most students, because parts would be far too easy or difficult.

Had it been possible to administer the entire survey via computer this requirement might have been more elegantly addressed. As it was, a simple approach common to computer- and paper-based tests was called for. This was to devise tests at three levels, with overlapping tasks to ensure a link in the response data across levels.

To assign students to a particular level it would be necessary to administer a short routing test to all participating students in advance of the ESLC. Section 2.3.7 provides details of how this routing test was designed and used.

#### **2.2.4 Test delivery channel**

The test delivery channel also impacted on the general design. The preferred option of the European Commission, as stated in the Terms of Reference, was to introduce computer-based testing where national and regional technical capabilities allowed but provide a paper-based testing alternative where participating countries had inadequate levels of readiness concerning testing with computers.

To enhance comparability of test results, the same test material and the same design principles were used for both modes.

#### **2.2.5 Task types**

Section 2.1 above describes the process of identifying the test content and the set of testable subskills or abilities to be assessed. The next step was to map each ability to a specific task type. A rigorous design was proposed which could be replicated across languages, thus maximising coherence and consistency in the implementation of the construct.

For Reading and Listening it was preferred to use selected response types, for ease and consistency of marking:

- multiple choice (graphic options, text options, true/false)
- multiple choice gap-fill (gapped texts, e.g. to test lexicogrammatical relations)
- matching texts to graphics (e.g. paraphrases to notices)
- matching texts to texts (e.g. descriptions of people to a set of leisure activities/holidays/films/books that would suit each of them)
- matching text elements to gaps in a larger text (e.g. extracted sentences) to test discourse relations, understanding at text level.

For writing a range of open, extended response task types was proposed, e.g., writing an email, postcard or letter, writing a referential or conative text (intended to inform, persuade or convince).

Eight task types were initially selected for Reading, five for Listening and four for Writing. Some task types were used across more than one level.

The 2008 Pilot Study informed the final selection of task types, and the construction of detailed test specifications and item writer guidelines for each of them. From the provisional set of task types each partner produced exemplar tasks. From these a smaller set of task types was selected and a draft specification written for each one that was identical across all languages. Existing test materials were also adapted for use in the Pilot Study. This ensured that the quantity of pilot material required for the pilot study could be created in the short space of time available. These tasks were amended and edited to fit the specifications. Any texts or tasks that were found to be inappropriate for the target population were excluded. A total of 106 tasks were developed across the skills and languages, with each language partner focusing on a different part of the ability range. A plan was agreed which achieved overall coverage of the construct and some linking across levels.

As national structures (i.e. NRCs) were not yet in place, SurveyLang's language partners made arrangements to administer the pilot tests through their own networks of test centres or other contacts in different countries. The majority of these were private language schools or other institutions outside the state sector. A few schools in the state sector were included. Administration of the pilot tests took place in October 2008. Over 2220 students in 7 countries completed tests in up to 3 skills, plus the routing test. Care was taken to target learners of an appropriate age group. Age ranged from 12 to 18 with the majority being between 15 and 17. All students were studying one of the 5 languages as a foreign or second language. In total 34 trial tests were created in Reading, Listening and Writing across the 5 languages. Tests followed the design proposed for the ESLC, being 30 minutes in length.

Feedback was elicited from teachers on their impressions of the tests, as well as from a range of stakeholders, including the Advisory Board and the participating countries, the Advisory Board's panel of language testing experts, and NRCs where these were in place. A booklet of tasks and feedback form were created for this purpose.

All analysis for the five languages was undertaken centrally. The purpose of analysis was to contribute to a decision on which task types to retain for the survey, and thus define the item writing requirements. Selection of actual tasks for use in the Main Study would follow subsequent stages (pretesting and the Field Trial). Approaches included classical analysis (facility, discrimination, reliability, and distractor analysis), Rasch analysis, and subjective cross-language comparison of the performance characteristics of the items.

The pilot test review thus focused on statistical evidence and feedback from different stakeholders. Feedback indicated general satisfaction with the task types. The feedback from the teachers of the students who took the trial tests was generally very positive. The review led to a reduction in the number of task types. Given the relatively small sample size agreed for the Main Study (1000 respondents per skill per country), it was important to avoid spreading responses too thinly over task types. Partners were satisfied that this reduction did not entail a substantive change to the construct. Some task types were retained but moved to another level, where this was seen to improve the articulation of the construct.

Further feedback on content was collected during the subsequent pretesting phase. This was strongly positive. Students agreed that topics were varied and suitable, there was enough time, and the instructions and layout were clear. The levels of difficulty were finely graded from easier tasks to more difficult ones and the clarity and speed of the recording was said to be excellent. Table 6 to Table 8 below detail the final selection of task types across the four levels for each of the three skills.

**Table 6 Main Study reading tasks**

Task	Test focus	Text type	Task type	Levels
R1	Identifying factual information relating to personal and familiar themes.	Short personal text (email, postcard, note).	3-option multiple choice with graphic options. Candidates choose the correct option.	A1
R2	Finding predictable factual information in texts such as notices, announcements, timetables, menus, with some visual support.	Notice, announcement etc. on everyday topic, with graphic support.	3-option multiple choice with short text-based options focusing on information. Candidates choose the correct option.	A1 A2
R3	Understanding signs, notices, announcements and/or labels.	A set of notices or signs etc. and a set of statements or graphics paraphrasing the message.	Candidates match the statements or graphics to the correct notices /announcements.	A1 A2
R4	Understanding the main ideas and some details of a text.	A newspaper/magazine article on familiar everyday topic.	Candidates answer 3-option multiple-choice questions.	A2
R5	Understanding information, feelings and wishes in personal texts.	A personal text (email, letter, note).	Candidates answer 3-option multiple-choice questions.	A2 B1
R6	Reading 3 (B1) or 4 (B2) short texts for specific information, detailed comprehension and (at B2) opinion and attitude.	A set of 3 (at B1) or 4 (at B2) short texts (e.g. ads for holidays, films, books), and a list of information/attitudes that can be found in the texts.	Candidates match the information to the text it is in.	B1 B2
R7	Reading for detailed comprehension and global meaning, understanding attitude, opinion and writer purpose.  B2: deducing meaning from context, text organisation features.	A text on familiar everyday topic.	Candidates answer 3-option multiple-choice questions.	B1 B2
R8	Understanding text structure, cohesion and coherence.	Text from which sentences are removed and placed in a jumbled order after text.	Candidates match the sentences to the gaps.	B2

**Table 7 Main Study listening tasks**

Task	Test focus	Text type	Task type	Levels
L1	Identifying key vocabulary/information (e.g. times, prices, days of weeks, numbers, locations, activities).	A simple dialogue.	Candidates match the name of a person to the relevant graphical illustration.	A1 A2
L2	Identifying the situation and/or the main idea (A1/A2) or communicative function (B1/B2).	Series of five short independent monologues or dialogues, e.g. announcements, messages, short conversations, etc.	Candidates choose the correct graphic (A1/ A2) or text (B1/B2) option from a choice of three.	A1 A2 B1 B2
L3	Understanding and interpreting detailed meaning.	A conversation or interview.	True/False.	A2
L4	Understanding and interpreting the main points, attitudes and opinions of the principal speaker or speakers.	Dialogue.	3-option multiple-choice.	B1 B2
L5	Understanding and interpreting gist, main points and detail, plus the attitudes and opinions of the speaker.	A longer monologue (presentation, report).	3-option multiple-choice.	B1 B2

**Table 8 Main Study writing tasks**

Task	Test focus	Text type	Task type	Levels
W1	Expressing general or topic-specific notions describing pictures or graphically-displayed information.	Short personal text (email).	Candidates write a short personal text making reference to the picture/graphically-displayed information.	A1
W2	Expressing general or topic-specific notions in response to input text and content points.	Short personal text (email, postcard).	Candidates write a short personal text explaining, describing etc.	A1 A2 B1
W3	Writing a referential text (intended to inform).	Personal text (email).  At B2 an article, essay, letter, report, review.	Candidates write a personal text explaining, describing etc.  At B2 candidates write an article etc explaining, describing, comparing etc.	A2 B1 B2
W4	Writing a conative text (intended to persuade or convince).	An essay, letter.	Candidates write an essay/letter describing, explaining, comparing, justifying, giving opinion etc.	B2

## 2.3 Test development process

The key aim was to produce language tests, the results of which would be comparable across all languages and in all countries. To this end many items of high quality had to be produced in a short space of time.

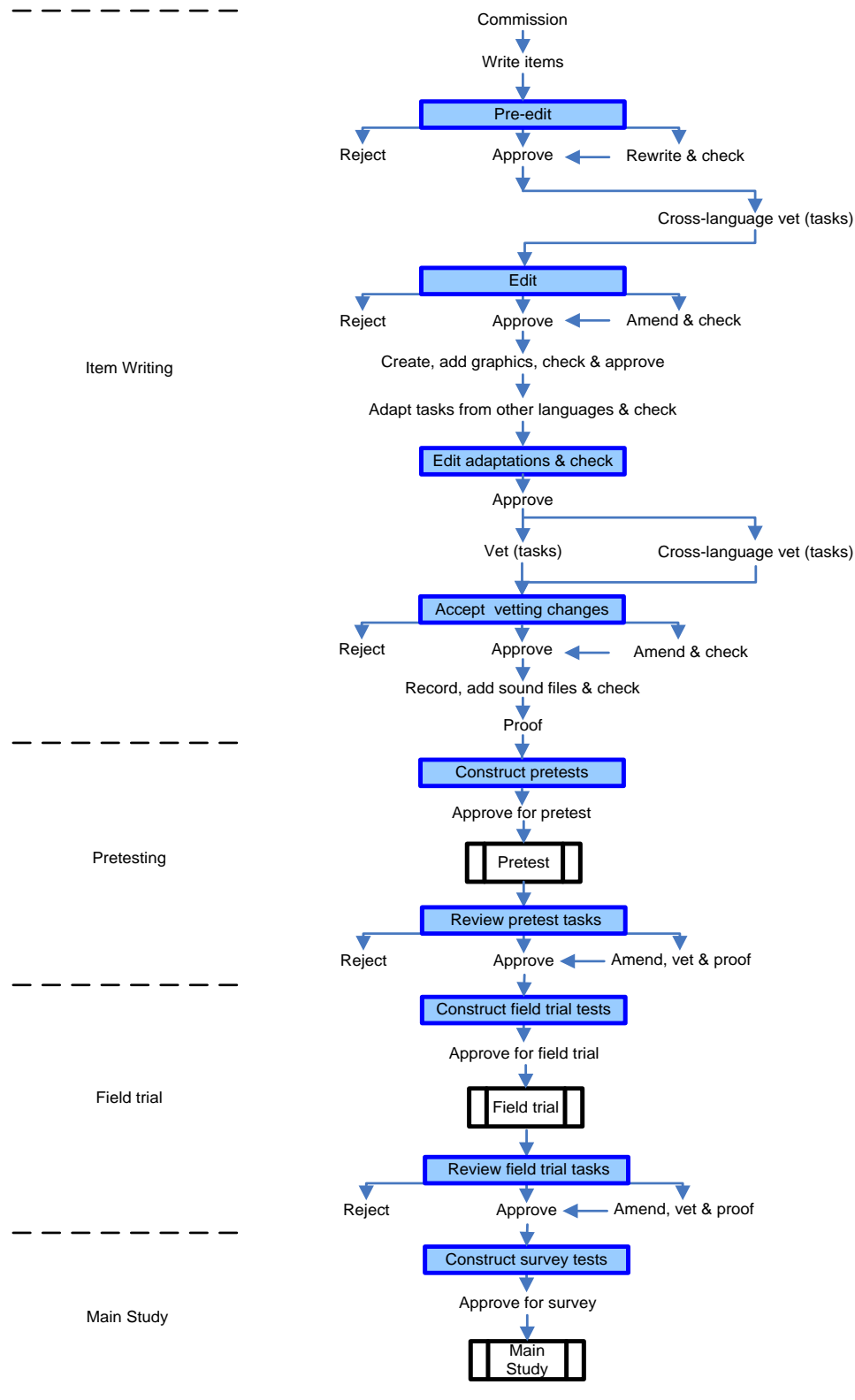
This comparability and quality required the close collaboration of the language partners, based on adoption of the same:

- test development cycle (pilot, pretesting, Field Trial, Main Study)
- test specifications and item writer guidelines
- test production process
- item authoring tool and item banking system
- quality control process
- standard setting process.

The steps in the test development process are shown in detail in Figure 1 below.



**Figure 1 Test development process**



### **2.3.1 Test specifications and item writer guidelines**

Following the Pilot Study, the test specifications were reviewed and finalised. Common test specifications across the 5 languages ensured that tasks across languages were almost identical in terms of number of items, number of options, text length, etc.

Detailed item writer guidelines were developed for each of the three skills. These guidelines specify the requirements of each task type at each level in terms of overall testing aim, testing focus, level of distraction in the options, input text length, etc. They also provide explicit guidance on the selection and manipulation of text types and topics, and the production of artwork and recordings. Quality criteria relevant to each task type are listed and these criteria provide the basis for the acceptance, rejection and editing of tasks as they proceed through the item production process.

### **2.3.2 Commissioning**

Before item writing began, the number of items required for the Main Study was calculated. As the pretesting and Field Trial stages were intended to enable selection of the best performing items for the Main Study, a much greater number of items than required for the Main Study were therefore commissioned. In total, over 500 tasks (2200+ items) were commissioned across the five languages. Given the large number of item writers commissioned it was imperative to plan for adequate coverage of construct, domains and topics for all tasks at each level across the five languages. Each item writer therefore received a detailed commissioning brief specifying the task types, levels and topics for to ensure adequate and consistent coverage of the CEFR domains as specified in Test Content in section 2.1.2 above.

Concerning the use of adapted tasks across languages it was agreed that all Writing tasks would be adapted as would all Reading and Listening tasks at levels A1 and A2. The work of creating and adapting these tasks was divided among the language partners (see 2.3.5 below).

Over 40 specialist item writers were commissioned across the five languages. For some languages, item writers specialised in certain skills, levels or task types. Item writers were organised into teams and managed by team leaders and specialist language testing product managers.

### **2.3.3 Recordings and artwork**

Professional recording studios employing native-speaker actors were used to record all Listening sound files. Listening test rubrics were standardised across the languages. A common style for producing artwork was agreed and the production of the graphics for all tasks was shared out among the five language partners. All artwork was commissioned from professional graphic artists.

### **2.3.4 Quality control**

Quality control procedures were included in each stage of the test production process which was developed for this survey.

The multi-stage, detailed test production process illustrated in Figure 1 above ensured that tasks were trialled several times before they appeared in the Main Study, to ensure they were fit for purpose. Figure 1 also illustrates how each task was thoroughly and repeatedly checked and proofread by external professional proof readers and signed off by internal team leaders and test production managers before being used in test construction.

### **2.3.5 Collaborative working methodology**

The common approach to item development described above was considered essential if the resulting tests in five languages were to be comparable in the way they related performance to the CEFR. Two specific aspects of this process are worth noting: cross-language vetting, and the use of task adaptation.

Cross-language vetting worked as follows:

- tasks from each language were vetted by at least 2 other language partners
- multi-lingual, experienced item writers vetted tasks from other languages to ensure that tasks, items and options would operate correctly
- a vetting form was created to ensure that vetting comments could be recorded consistently and electronically
- vetting comments were then passed back to the original language partner who could then compare comments from both their own vetters and the vetters from other language partners.

A review conducted at the end of the Pilot Study confirmed the value of cross-language vetting as an additional stage to the standard test production process. It not only provided an additional quality control, it also enabled the sharing of knowledge and experience among the language partners.

Task adaptation worked as follows. A proportion of the Reading and Listening tasks were adapted across languages. Each language partner was asked to adapt some tasks from two of the other four languages. There were several purposes for adapting tasks and including them in the pilot:

- it was seen as a valuable context for developing collaborative working methods between the language partners: studying each others' tasks in detail stimulated much critical reflection and interaction
- it might be a possible way of enhancing consistency and comparability across languages
- it might offer a straightforward, if not a quicker, way of generating new tasks

The Pilot Study review also confirmed the value of adapting tasks across languages. It appeared that most task types used in the Pilot Study could be successfully adapted from one language into another if the aim was to adapt but not translate. However, the process needed skilled item writers who were competent in two or more of the languages. Item writers needed to be aware of lexicogrammatical differences between the languages and how these differences might affect the perceived difficulty of the items. The only task type that appeared difficult to adapt was the multiple-choice cloze task where the testing focus was largely lexicogrammatical.

For the skill of writing, it was deemed practical and desirable to adapt the same set of writing tasks into all languages.

### 2.3.6 Selection of tasks for the Main Study

In the Field Trial two examples of each task type per level per skill per language were trialled. At Field Trial review one task from each pair was selected for the main survey.

All tasks were subject to expert review by each language team, taking into account feedback from administrators, coordinators, teachers and students, collected in the NRC and Quality Monitor reports. With analysis completed all tasks were again reviewed, this time combining expert judgement with the statistical analysis. In almost all cases, the judgement agreed with the analysis.

Each language team selected one task from each pair for the main survey, recording this in a spreadsheet with a justification. One spreadsheet was then created detailing the selection and the justification for all the common tasks across the five languages, i.e. all the Writing tasks and the A1 and A2 Reading and Listening tasks.

All five language teams then discussed the tasks common across languages, in separate meetings for each skill. Each of the common tasks was again reviewed, taking into account each team's selections, the statistical analysis and the feedback from NRCs and QM reports. In this way one task from each pair, common across the five languages, was selected for the main survey.

There were relatively few task pairs where the selection could be motivated by statistical evidence alone. In two cases the judgment of the teams went against statistical evidence.

Table 9 illustrates the statistics used in selection for English Listening in a summarised form. In this table:

- **Selected** indicates the selected task.
- **N responses**: the combined number of CB and PB responses. Smaller numbers mean that less confidence can be placed in the statistics.
- **Facility**: the mean score on a task as a proportion of the maximum score.

- **N score categories estimated:** where fewer categories are estimated than the maximum score this indicates that too few responses were available over the whole score range.
- **Difficulty:** an IRT estimated difficulty.
- **Fit** (average over score categories): an approximate indicator of fit, where values less than 5% indicate significant misfit. When summarised over score categories only 5 significant cases are found.

**Table 9 Illustration of statistics used selecting tasks at Field Trial review  
(English Listening)**

Task Pair	Specific ID	Selected (1=Yes)	N responses	Facility	N score categories estimated	Difficulty (average over score categories)	Fit (average over score categories)	Significant misfit
A1-L1	EL111	1	1596	0.66	5	-0.95	0.03	TRUE
	EL114	2	83	0.80	1	-1.47	0.86	FALSE
A1-L2	EL213	1	799	0.65	4	-0.97	0.29	FALSE
	EL212	2	880	0.78	4	-1.75	0.25	FALSE
A2-L1	EL123	1	1818	0.71	5	-0.72	0.19	FALSE
	EL121	2	1929	0.82	5	-1.20	0.20	FALSE
A2-L2	EL221	1	1913	0.73	4	-1.02	0.22	FALSE
	EL222	2	1834	0.70	4	-0.76	0.39	FALSE
A2-L3	EL321	1	1909	0.68	6	-0.51	0.35	FALSE
	EL323	2	1838	0.71	6	-1.19	0.13	FALSE
B1-L2	EL231	1	379	0.89	3	0.23	0.31	FALSE
	EL233	2	2217	0.73	5	-0.02	0.23	FALSE

One of the considerations in selecting the set of tasks for the Main Study was to preserve as far as practical the proportion of tasks addressing each domain. This was achieved reasonably well, as Table 10 shows. The figures in brackets are the original target proportion, as shown in Table 4.

**Table 10 Distribution across domains – Main Study tasks**

Domain	A1	A2	B1	B2	Grand Total
personal	43% (60%)	38% (50%)	34% (40%)	11% (25%)	31%
public	57% (30%)	50% (40%)	47% (40%)	50% (50%)	51%
educational	0% (10%)	12% (10%)	16% (20%)	33% (20%)	16%
professional	0% (0%)	0% (0%)	3% (0%)	6% (5%)	2%
Grand Total	100%	100%	100%	100%	100%

### 2.3.7 The routing test

As explained in 2.2.3 above, the decision to adopt a targeted testing approach necessitated the administration of a routing test for each language, which would be used to place students into one of three level groups.

The routing tests were developed and trialled in the Pilot Study, and further revised for the Field Trial. Each test was 15 minutes long, and for simplicity consisted of 20 Reading-focused items, ordered to be progressive in difficulty. This was considered adequate to the purpose of the test: to make a very broad classification into three levels. Items were taken from the language partners' existing item banks, already calibrated on a scale related to the CEFR, as conceived and implemented by each language partner. Thus they could also be used in the pilot to anchor the Reading and Listening tests to existing proficiency scales. Each candidate completing the Reading and/or Listening test would also take the routing test, so that, to the extent that the routing test was linked to the CEFR, all tasks could be linked.

It is worth stating that the reference to partners' existing CEFR-related proficiency scales had no direct impact on the final standard setting process (see Chapter 11); however, there is no doubt as to the great practical utility for the development of having such points of reference.

The score on the routing test did not count as part of the language test performance of any student, but was used to allocate the student to an appropriate level. Nor did the score on the routing test influence the sampling probability of any student.

The requirement to administer a routing test added to the administrative complexity of the ESLC. NRCs ensured that schools administered the routing test. It was administered to all eligible students or to the sampled students only, depending on the participating country. In a few cases countries proposed an alternative procedure to the routing test: a teacher-rated can-do questionnaire, or a comparison with exam results. SurveyLang accommodated these requests. NRCs ensured that the scores from the routing test were returned to SurveyLang so that students could be allocated to a low, medium or high level test accordingly.

The final allocation determined what proportion of students saw the low, middle or high-level tests. It was considered important, other things being equal, that a sufficient number of responses were collected for each level for the purpose of analysis. Thus the cut-offs for the routing tests were modified where thought fitting, with reference to the consequences in terms of allocation.

## **2.4 Marking**

As noted in 2.2.5 above, an early design decision was to use objectively-marked task types for Reading and Listening, and subjective marking for Writing.

### **2.4.1 Marking of Reading and Listening**

For computer-based tests, responses for Reading and Listening were captured and automatically marked against an answer key. Paper-based tests had to be manually marked in-country and the marks uploaded to a central point.

For the Field Trial and Main Study an electronic data-entry tool was provided to countries, fully customised to contain the IDs of all sampled students. The tool allowed double mark entry, and countries were recommended to use a proportion of double keying as a check on quality. However, this was not a required procedure.

### **2.4.2 Marking of Writing**

The approach to marking went through several revisions between the 2008 Pilot Study and the Main Study.

The mark scheme originally used for the pilot was somewhat complex:

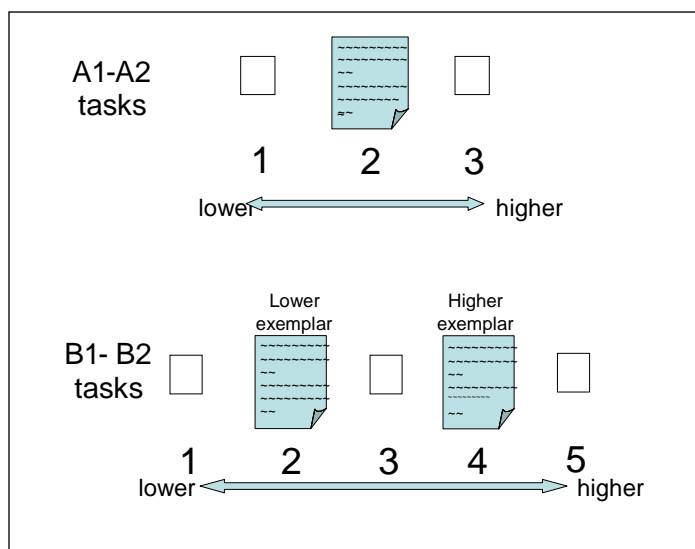
- it contained four analytic scales: Task fulfilment, Communicative command, Discourse and Linguistic accuracy. The first two of these focused on functional communication, the second two on formal linguistic features
- the scales had five score categories (0-4)

- marks were to be awarded in the order that the scales are listed above
- there were two slightly different versions of the scheme, one for A1-A2 and one for B1-B2
- the scheme could also incorporate task-specific elements.

At the Field Trial stage a different and quite innovative approach was introduced. Rather than ask markers to make absolute judgments about a student's CEFR level, it was decided to require a comparative judgment, where the marker's task was to say whether a student's performance was lower than, equal to or higher than an exemplar text. For levels A1-A2 one exemplar was provided, defining a 3-point scale: 1, 2 or 3. For the B1-B2 levels two exemplars (a higher and a lower one) defined a 5-point scale, see Figure 2 below.

Exemplars were chosen at a level to elicit the widest possible range of marks, and were informed to an extent by Field Trial experience of the general level of the student population for each language. As explained in training, exemplars were not intended to represent a specific performance level in CEFR terms, but rather a level where a roughly equal number of worse and better performances might be expected to be produced. Choice and use of the exemplars did not pre-judge the subsequent standard-setting.

**Figure 2 Marking of Writing against exemplars**



Four criteria were retained at B1-B2, but just two at A1-A2.

In preparation for the Main Study further revisions and additions were made to the design of the Writing tests, the marking criteria, training, and quality assurance procedures. The number of tasks a student responded to was reduced to 3 at Level 1, 2 at the higher levels, aiming at quicker marking and fewer missing or partial responses. The same two criteria – *Communication* and *Language* – were used for all 4 test levels, to make marking quicker and easier.



Training procedures were improved, with more stress on practice and standardisation of marking, with provision of detailed, automatically generated feedback on performance, aiming at improving the accuracy of marking. All multiple-marked scripts, rather than a proportion, were to be returned to SurveyLang for central marking, aiming at more reliable comparison across countries.

### **2.4.3 Systems support: the testing tool**

Close collaboration between partners in the development of the tests, and consistent implementation and presentation of test tasks, were supported by the item authoring, banking and test assembly functionality of the testing tool specifically developed for the ESLC. The item authoring tool is web-based and allowed item writers across Europe to create items with task templates created directly from the test specifications. Once created, the items could be uploaded directly into the shared item bank. This item bank also allowed the language partners to describe their tasks using exactly the same system of metadata. At the time of test production, the test design (see section 2.5) was implemented in the testing tool so that full tests could be produced in both computer and paper-based formats. See Chapter 6 for further details of the testing tool functionality.

Pretesting used test material authored on and generated out of the testing tool, though not administered through it. The Field Trial enabled a full-scale trial of every aspect of the testing tool.

The reviews following the Pretest and Field Trial phases led to a series of amendments to tasks. Changes to the tasks, commissioning of new graphics, or re-recording of audio files all led to updating of the test material on the system.

Tasks were checked and signed off in both paper-based and computer-based format. In paper-based format this was done in booklets created by the test assembly tool and in computer-based format the tasks were signed off in CB tests created by the rendering tool.

### **2.4.4 Ensuring familiarity with the form of the tests**

Much consideration was given in the language testing group to how to ensure that students would be sufficiently familiar with the form of the tests for them to demonstrate their ability.

Evidence from trialling and pretesting suggested that students had no real problems in understanding how to respond to the test tasks in their paper-based form. The instructions included in the paper-based and computer-based tests were also rendered in the students' questionnaire language, i.e. in most cases their first language. The provision of additional on-screen help in the CB mode was thus felt to be unnecessary (and would have been very expensive).

It was decided to make familiarisation material available to students or teachers who wished to make use of it, but not to impose a compulsory familiarisation activity as part of the test administration in schools.

Familiarisation materials were created with descriptions of the task types and sample materials. They were reviewed following the Field Trial. Additional clarification was added to the School Coordinator Guidelines to stress that the tasks were intended only for familiarisation and only if teachers judged this necessary. These materials were made available by the NRCs to all participating teachers and were available on the SurveyLang website.

For the Main Study both paper-based and computer-based familiarisation materials were available. The sample computer-based tests on the SurveyLang website enabled the student to choose a language to be tested in, as well as the language for the on-screen instructions

## 2.5 Final test design

Table 11, Table 12 and Table 13 below illustrate the test design for each skill. The paragraphs below describe how to read the test design tables.

Each table consists of a number of columns:

- The leftmost column contains the test task ID in a generic way. More explanation on this is given below.
- The second to sixth columns contains the specific task labels across languages (E=English, F=French, G=German, I=Italian and S=Spanish, R=Reading, L=Listening, W=Writing).
- The seventh column is the time load for each task: 5; 7.5; 15 or 30 minutes.
- On the right are a number of columns filled with coloured blocks. The columns represent test form (test booklet), labeled 'b1', 'b2', etc. For Writing 12 different test forms have been defined; for Reading 18 and for Listening 7.

**Testing time:** The bottom row of each table is the total testing time for the test form (booklet). For example, in the test design for Reading, Booklet 1 consists of 4 tasks, each of 7.5 minutes making a total testing time of 30 minutes. All test booklets are 30 minutes except for Listening level 1 and Writing level 3.

- For Listening at Level 1, the total testing time for Booklet 1 is 25 minutes. The original design was for 30 minutes but it was agreed by the Advisory Board after the Field Trial that 6 tasks was too many for Level 1 students and the number of tasks was reduced to 5.
- For Writing at Level 3, the total testing time for each booklet is 45 minutes. Two tasks were required for a linked design and since the B2 level tasks were 30 minutes each and the B1 level tasks were 15 minutes the total testing time was greater than the 30 minutes as specified in the original design. However, only a small proportion of the total number of students received these



booklets: those at level 3 and receiving one of the combinations of skills that included Writing.

The generic task ID which is in the leftmost column is constructed in the following way:

- The first two positions indicate the CEFR level of the task. This can be A1, A2, B1 or B2.
- The next two positions (between the dashes) indicate the task type as illustrated in Table 6 to Table 8 above.

**The coloured cells:** A coloured cell (with a number written in it) indicates that the task (row) is part of the test form (column). To help in structuring the perception of the tables, four different colours have been used: yellow for the A1 tasks, green for the A2 tasks, dark blue for the B1 tasks and light blue for the B2 tasks.

In each column of a table (each test form) two, three, four or five tasks are coloured. The numbers written in the coloured cells indicate the sequence of the tasks in the test form (read vertically). In constructing the design the following principle has been used throughout:

- All tasks at a lower CEFR level precede all tasks at a higher CEFR level. All yellow cells precede all green cells; all green cells precede all dark blue cells and these always precede all light blue cells.

**Italian:** Note, there is a different design for Italian which takes into account the smaller number of students taking these tests. Note though that the Italian designs are mapped to the design for the other languages and therefore although a smaller number of booklets are used, the booklets used match the design used for the booklets for the other four languages.

**Table 11 Main Study test design for Reading**

Level 1												
Tasks	E	F	G	I	S	Time	Booklet 1	Booklet 2	Booklet 3	Booklet 4	Booklet 5	Booklet 6
A1-R1	ER112	FR112	GR111	IR113	SR112	7,5	1	2		2	1	
A1-R2	ER211	FR211	GR213	IR211	SR211	7,5	2		1	1		2
A1-R3	ER312	FR311	GR312	IR313	SR312	7,5		1	2		2	1
A2-R2	ER223	FR223	GR221	IR223	SR223	7,5	3		4		3	
A2-R3	ER321	FR322	GR321	IR323	SR322	7,5	4	3				4
A2-R4	ER423	FR423	GR421	IR421	SR423	7,5		4	3	4		
A2-R5	ER523	FR523	GR522	IR521	SR523	7,5				3	4	3
					Testing time		30	30	30	30	30	30
Level 2												
Tasks	E	F	G	I	S	Time	Booklet 7	Booklet 8	Booklet 9	Booklet 10	Booklet 11	Booklet 12
A2-R2	ER223	FR223	GR221	IR223	SR223	7,5	2		1		2	
A2-R3	ER321	FR322	GR321	IR323	SR322	7,5	1	2				1
A2-R4	ER423	FR423	GR421	IR421	SR423	7,5		1	2	1		
A2-R5	ER523	FR523	GR522	IR521	SR523	7,5				2	1	2
B1-R5	ER532	FR531	GR533	IR531	SR531	7,5	3	4		4	3	
B1-R6	ER631	FR631	GR633	IR632	SR631	7,5	4	3	3			4
B1-R7	ER731	FR733	GR731	IR733	SR733	7,5			4	3	4	3
					Testing time		30	30	30	30	30	30
Level 3												
Tasks	E	F	G	I	S	Time	Booklet 13	Booklet 14	Booklet 15	Booklet 16	Booklet 17	Booklet 18
B1-R5	ER532	FR531	GR533	IR531	SR531	7,5	1	2				
B1-R6	ER631	FR631	GR633	IR632	SR631	7,5		1	1			
B1-R7	ER731	FR733	GR731	IR733	SR733	7,5	2		2			
B2-R6	ER642	FR642	GR642	IR642	SR641	15			3	1		2
B2-R7	ER741	FR743	GR741	IR743	SR741	15	3			2	1	
B2-R8	ER841	FR843	GR842	IR842	SR841	15		3			2	1

Italian

2  
1

4  
3

(b1)\*

Italian

2  
1

3

					<b>Testing time</b>	<b>30</b>	<b>30</b>	<b>30</b>	<b>30</b>	<b>30</b>	<b>30</b>
--	--	--	--	--	---------------------	-----------	-----------	-----------	-----------	-----------	-----------

\* Same design as booklet 8 in other languages but called Booklet 1

\*\* Same design as booklet 14 in other languages but called booklet 2

**Table 12 Main Study test design for Listening**

							<b>Level 1</b>		
<b>Tasks</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>I</b>	<b>S</b>	<b>Time</b>	<b>Booklet 1</b>		
<b>A1-L1</b>	EL111	FL112	GL112	n.a.	SL112	5	1		
<b>A1-L2</b>	EL213	FL211	GL214	n.a.	SL214	5	2		
<b>A2-L1</b>	EL123	FL123	GL123	IL123	SL121	5	3		
<b>A2-L2</b>	EL221	FL222	GL222	IL222	SL222	5	4		
<b>A2-L3</b>	EL321	FL321	GL321	IL322	SL322	5	5		
<b>Testing time</b>							<b>25</b>		

							<b>Level 2</b>		
<b>Tasks</b>	<b>Time</b>						<b>Booklet 2</b>	<b>Booklet 3</b>	<b>Booklet 4</b>
<b>A2-L1</b>	EL123	FL123	GL123	IL123	SL121	5	3	2	1
<b>A2-L2</b>	EL221	FL222	GL222	IL222	SL222	5	2	1	3
<b>A2-L3</b>	EL321	FL321	GL321	IL322	SL322	5	1	3	2
<b>B1-L2</b>	EL231	FL232	GL233	IL233	SL232	7,5	4	5	
<b>B1-L4</b>	EL432	FL433	GL433	IL432	SL433	7,5	5		4
<b>B1-L5</b>	EL531	FL531	GL531	IL531	SL533	7,5		4	5
<b>Testing time</b>							<b>30</b>	<b>30</b>	<b>30</b>

							<b>Level 3</b>		
<b>Tasks</b>	<b>Time</b>						<b>Booklet 5</b>	<b>Booklet 6</b>	<b>Booklet 7</b>
<b>B1-L2</b>	EL231	FL232	GL233	IL233	SL232	7,5	2	1	
<b>B1-L4</b>	EL432	FL433	GL433	IL432	SL433	7,5	1		2
<b>B1-L5</b>	EL531	FL531	GL531	IL531	SL533	7,5		2	1
<b>B2-L2</b>	EL242	FL241	GL241	IL242	SL241	7,5	3		4
<b>B2-L4</b>	EL442	FL442	GL443	IL443	SL442	7,5	4	4	
<b>B2-L5</b>	EL543	FL541	GL541	IL541	SL541	7,5		3	3

		<b>Italian</b>
		1
		3
		2
		4
		5
		(Booklet 1)*

		<b>Italian</b>
		2
		1
		4
		3

Testing time	30	30	30
--------------	----	----	----

(Booklet 2)\*\*

\* Same design as booklet 4 in other languages but called booklet 1

\*\* Same design as booklet 7 in other languages but called booklet 2

**Table 13 Main Study test design for Writing**

Level 1										
Tasks	E	F	G	I	S	Time	Booklet 1	Booklet 2	Booklet 3	Booklet 4
A1-W1	EW113	FW111	GW113	n.a.	SW113	7,5	1	2	1	
A1-W2	EW212	FW212	GW212	n.a.	SW213	7,5	2	1		1
A2-W2	EW222	FW222	GW221	IW221	SW223	7,5	3		3	2
A2-W3	EW322	FW322	GW321	IW322	SW323	7,5		3	2	3
Testing time							30	30	30	30

Level 2										
Tasks	E	F	G	I	S	Time	Booklet 5	Booklet 6	Booklet 7	Booklet 8
A2-W2	EW222	FW222	GW221	IW221	SW223	7,5	1	1		
A2-W3	EW322	FW322	GW321	IW322	SW323	7,5			1	1
B1-W2	EW233	FW233	GW234	IW234	SW231	15	2		2	
B1-W3	EW331	FW334	GW332	IW333	SW332	15		2		2
Testing time							30	30	30	30

Italian

1
2

(Booklet 1)\*

Level 3										
Tasks	E	F	G	I	S	Time	Booklet 9	Booklet 10	Booklet 11	Booklet 12
B1-W2	EW233	FW233	GW234	IW234	SW231	15	1	1		
B1-W3	EW331	FW334	GW332	IW333	SW332	15			1	1
B2-W3	EW342	FW343	GW341	IW342	SW343	30	2		2	
B2-W4	EW443	FW443	GW441	IW441	SW444	30		2		2
Testing time							45	45	45	45

Italian

1
2

(Booklet 2)\*\*

\* Same as booklet 8 in other languages but called booklet 1

\*\* Same as booklet 12 in other languages but called booklet 2

## 2.6 References

- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Buck, G (2001) *Assessing Listening*, Cambridge: Cambridge University Press.
- Hamp-Lyons, L and B. Kroll (1997) *TOEFL 2000 Writing: Composition, community and assessment*, Princeton, NJ: Educational Testing Service.
- Hayes, J R (1996) *A new framework for understanding cognition and affect in writing*, in Levy, C M and Randsdell, S, *The Science of Writing: Theories, Methods, Individual Differences, and Applications*, Mahwah, New Jersey: Lawrence Erlbaum Associates, 1-28.
- Hudson, R (1991) *English Word Grammar*, Cambridge, MA: Blackwell.
- Hyland, K (2002) *Teaching and Researching Writing*, London: Longman.
- Van Ek, J and Trim, J (1998) *Threshold 1990*, Cambridge: Cambridge University Press.
- Van Ek, J and Trim, J (1998) *Waystage 1990*, Cambridge: Cambridge University Press.
- Van Ek, J and Trim, J (2000) *Vantage*, Cambridge: Cambridge University Press.
- Weir, C J (2005) *Language Testing and Validation: An Evidence-based Approach*, Basingstoke: Palgrave Macmillan.

# ***Chapter 3:***

# ***Instrument***

# ***development –***

# ***Questionnaires***



### 3 Instrument development - Questionnaires

The ESLC seeks to provide policy-relevant information about students' foreign language competence. The main goal of the contextual information is to 'facilitate a more productive comparison of language policies, and language teaching methods between Member States, with a view to identifying and sharing good practice' (Communication from the Commission to the European Parliament and the Council 2005:5). A lot of the factors contributing to foreign language competences are largely beyond the control of the countries, such as their general demographic, social, economic and linguistic contexts. Other contextual factors are more readily amenable to intervention through targeted educational policies, such as the age at which foreign language education starts, the intensity of the foreign language courses and the initial and in-service training of teachers. For a fuller appreciation of what the ESLC results mean and how they may be used to improve student learning in foreign languages, it is crucial to map and monitor the supranational and national contexts in which foreign language learning takes place. Contextual information allows the detection of factors that are related to foreign language competences and which, therefore, might be relevant for their improvement.

This mapping of the foreign language learning context was to be achieved by means of context questionnaires to the students tested, their teachers of foreign languages and their institution principals. In addition, system-wide information was to be collected through the NRCs. The context questionnaires aimed to provide a broad range of information on the foreign language teaching policies, foreign language teaching and learning policies and to provide a sound comparison between Member States.

Two broad stages can be identified in the development of context questionnaires: the conceptualisation stage - during which it is determined what concepts should be measured - as described in section 3.1 of this chapter, and the operational stage, during which an empirical indicator for each of the concepts is developed (described in the second part of this chapter.)

#### **3.1 Conceptualisation**

Before questions can be formulated, a decision has to be made as to what concepts should be measured given the research objectives. The first step, therefore, in the development of context questionnaires is to determine the purpose, specific research objectives and conditions (which we intend to study, when and where) of the ESLC and the procedure for selecting concepts, described in section 3.1.1. On the basis of the purpose, objectives and conditions we can specify what concepts should be measured, which is described in section 3.1.2.

### 3.1.1 Development of the conceptual framework

This section details the development of the conceptual framework.

#### **Purpose**

As written previously, the main goal of the contextual information is to 'facilitate a more productive comparison of language policies, and language teaching methods between Member States, with a view to identifying and sharing good practice' (Communication from the Commission to the European Parliament and the Council 2005:5). The provision of internationally comparable data on the policies regarding the teaching and learning of foreign languages constitutes relevant information for national policy makers, school leaders, teachers and parents. Contextual information can also reveal interesting disparities in the distribution of educational resources and opportunities among different groups of students, teachers, schools and countries (Willms 2006). Furthermore, the context questionnaires should allow an in-depth analysis which may provide insight in how the foreign language teaching policies are related to developing language competences (Communication from the Commission to the Council 2007). The contextual data may contribute to explaining why countries have different results, why some teachers or schools are more effective than others or why some students are better foreign language learners than others. Apart from a description of the foreign language teaching policies and how these policies are related to foreign language competences, the contextual data needs to serve two other main functions.

The second function of the context information is detecting and reporting group differences in foreign language achievement. The data should facilitate the definition of subgroups of the populations of students, teachers, schools and principals. The context questionnaires provide the information needed for reporting the foreign language competences of the students by subgroup. For example, it enables documenting the differences in foreign language competences between privileged and non-privileged students, schools, regions and countries.

A third function of the context questionnaires is of a more technical nature, which is the enhancement of data quality and usability of the data. Non-cognitive variables may play an important role in the sampling, stratification and weighting procedures, and sometimes in checking the validity of results. Some of the context questions will be used to assess the potential bias resulting from non-participation of students and schools. Another type of technical use is to estimate plausible values (see chapter 12). Furthermore, the Commission required that 'existing concepts and classifications should be used and links to similar international surveys should be explored' allowing secondary analyses and facilitating international comparison (Communication from the Commission to the Council 2007:5).

## Identification of specific objectives

In the ESLC the focus of the context information is on the European language policies and language teaching methods that help the identification of good practice within foreign language teaching in secondary education<sup>4</sup> and that might be relevant for improving foreign language competence in the European Union.

The key policy documents of the European Commission regarding multilingualism (European Commission 2008) were studied to establish the major European policy issues and related actions that have direct bearing on foreign language teaching and learning in lower secondary education or impact on the outcome of this process, i.e. foreign language competence.

The procedure of studying all key policy documents ensures that all core educational issues are European rather than country specific and that they are consistent with the primary goals of the survey. This procedure also ensures that we take account of previous work in the field at Union level, as the Council required (Council of Europe 2006:2), because these policy documents are the result of extensive preparation, studies within the European Union and of consultation processes (with teachers, the public, policy makers and scientists from various fields) in which all Member States have had their say.

The overview yielded several general policy issues that are aimed at improving foreign language teaching and learning in secondary education. To cross-validate the importance of these European educational policy issues we asked the Advisory Board Members for feedback using a feedback form. We approached the Advisory Board Members in order to obtain feedback from all Member States and because they are experts in language teaching policy, language teaching and/or international studies. The feedback was analysed and presented to the Advisory Board of the European Commission. Thirteen specific European foreign language teaching policy issues were selected as research objectives (see section 3.1.2).

## Conditions

As we have to arrive at a productive comparison (Communication from the Commission to the European Parliament and the Council 2005:5), we also studied the various educational systems in the European Union using the information available in the database of the Eurydice Information Network on Education in Europe (Eurydice) and the database of the statistical office of the European Union (Eurostat 2008). Our

---

<sup>4</sup> 'The "total population" of the survey, in statistical terms, should be the total number of pupils enrolled in the final year of lower secondary education (ISCED2), or the second year of upper secondary education (ISCED3), if a second foreign language is not taught at lower secondary education.' (Communication from the Commission to the Council 2007: 5)



aim was to identify the differences between the structures of the various educational systems that impact on:

- (i) the comparability of the data  
the level(s) at which concepts have to be measured (in other words in which of the four questionnaires)  
the localisation of the questionnaires (see section 3.2.3.2).

The main differences found in the structure are that the age at which compulsory ISCED1 education starts and ends, and the duration of ISCED1 and ISCED2 education differs between Member States. As a consequence, participating students from the Member States will have different ages, might still be receiving compulsory education or not, and will have received a different number of years of education. Furthermore during their educational career, students in one Member State may have had to change from one institute to another or may have had to choose between different areas of study, while students of other Member States are all enrolled in the same study programme.

### **Selection of the concepts**

The specification of the concepts started with the analysis of the conceptual frameworks of similar international surveys, such as:

- the IEA foreign languages studies (Carroll 1963), (Lewis and Massad 1975)
- PISA (Adams and Wu 2002), (OECD 2005), (Kuhlemeier 2007a), (Kuhlemeier 2007b), (OECD 2007)
- PIRLS (Campbell, Kelly, Mullis, Martin and Sainsbury 2001), (Mullis, Kennedy, Martin and Sainsbury 2004), TIMSS (Mullis I. et al. 2003), (Mullis I., Martin, Ruddock, O'Sullivan, Arora and Erberber 2005)
- The European study of English as a foreign language (Bonnet 1998) (Bonnet 2002).

Taking these conceptual frameworks as a starting point ensures that existing and comparable concepts are chosen and conform with the requirements of the Commission (Communication from the Commission to the Council 2007:5). Furthermore, this analysis ensures that we optimally employ the knowledge gathered and used before, as the conceptual frameworks of these international surveys are based upon combined knowledge from the many different scientific fields that deal with educational achievement and specifically foreign language achievement.

On the basis of this analysis, an overview was created of all concepts that could be considered for inclusion in the conceptual framework and of criteria for the selection of relevant concepts of which a reliable and valid measurement is feasible (see Table 14). The various criteria have to be carefully balanced, as they are sometimes in conflict with each other. For example, teaching time might be a very relevant concept. An accurate measurement of teaching time would need many detailed questions increasing the burden on the respondents. The increased burden is likely to result in less valid data due to non-response and recall problems.

The overview contained over 150 concepts reflecting the characteristics and malleable aspects of each level of the educational system, being the national educational system, the educational institutions, the instructional setting (teacher and classroom) and the individual participants (students).

**Table 14 Criteria for selecting concepts**

<p><b>Relevance</b></p> <ul style="list-style-type: none"> <li>• The constructs and variables chosen should be consistent with the primary goals of ESLC and its major policy priorities. The concepts should be relevant for foreign language teaching and learning policies.</li> <li>• The choice of possible concepts should be guided by empirical evidence of their relationship with foreign language competence. If empirical evidence is lacking, a relation with foreign language competence should at least be conceivable.</li> <li>• The concepts should provide relevant information for all Member States participating in the ESLC. Country-specific interests can be pursued through additional country-specific questions.</li> <li>• The concepts should support cross-country comparisons, have a comparable meaning and interpretation across countries and cultures, be culturally appropriate and be easily translated.</li> <li>•</li> </ul> <p><b>Reliable &amp; valid measurement feasible</b></p> <ul style="list-style-type: none"> <li>• The gathering of the contextual data should not overburden students, teachers, principals, or National Research Centres. In particular, completing the student questionnaire should be feasible in the testing time of half an hour.</li> <li>• Concepts should not arouse controversy nor be too sensitive.</li> <li>• The choice of the concepts should be in line with the possibilities and restrictions of the sampling design and the data collection methods.</li> <li>• The proposed questionnaire logistics should be feasible in terms of time, costs, personnel, administration, coding, data analysis, reporting and so on. The questionnaire should not endanger the timeliness of the reporting. The implementation of the questionnaire should not be too expensive for participating Member States.</li> </ul>
--

Based upon the description of the policy issues in the EC key documents of the Commission on Multilingualism and the European studies referred to in those documents, those concepts were selected from the overview that are related to the identified and cross-validated policy issues. The procedure followed of identifying and cross-validating exclusively European policy issues that are consistent with the primary goals of the survey and that are based upon up-to-date, relevant and comparable information in the Member States about the factors that impact the outcome of foreign



language teaching in secondary education (foreign language competence) ensures that the directly related concepts are relevant.

Because a reliable, valid measurement of the concepts should be feasible within the given testing time of half an hour, we particularly focused on concrete concepts. In contrast to concrete concepts, complex and abstract concepts are likely to have different meanings making interpretations across countries and a valid measurement more difficult to attain. The more complex and abstract the concepts, the larger the number of questions required to adequately represent the entire concept and the fewer additional concepts that can be measured within the given testing time. As a consequence we would risk threatening the validity due to construct underrepresentation, which occurs when the concept is not adequately represented in the measurement (specification error). Furthermore, we would risk that some of the identified policy issues would be inadequately addressed within the context questionnaires.

### **3.1.2 Conceptual framework for the context questionnaires**

The EC has developed a range of policies and actions regarding multilingualism based upon extensive consultation processes and studies (European Commission 2008). Parts of these policies and actions are aimed at improving the outcome of foreign language teaching and learning in secondary education which is tested in this survey. Three general and strongly related objectives can be distinguished in these European policy issues that are consistent with the primary goal of the survey. The first objective is to stimulate Member States to provide a sound basis for the life-long learning of foreign languages through the teaching of at least two foreign languages from an early age. This first objective will be discussed in section 0. The second objective is to stimulate the creation of a language-friendly environment, both in school (see section 0) and at home (section 0) where different languages are heard and seen, where speakers of all languages feel welcome and language learning is encouraged. Because the quality of teacher training is a key factor in ensuring the quality of school education (Commission of the European Communities 2007b), the third general objective is to improve teacher training which will be discussed in the section 0.

For each of these general objectives the identified European policies are described in the conceptual framework as well as the concepts that are directly related to the identified policies. These concrete and feasible concepts are organised in tables displaying at what levels of the educational systems these concepts have to be measured (e.g. a concept can be measured at various levels for quality control). In addition to the malleable aspects related to the identified policies, the antecedent conditions are also displayed in the tables. Antecedent conditions might put constraints on the impact of the malleable concepts upon foreign language competence. Those antecedent conditions might also be needed for the description of subpopulations and quality control (see section 0). An antecedent condition that is of particular importance is the organisational structure of European educational systems (discussed in the





section 0), because we have to arrive at a productive comparison (Communication from the Commission to the European Parliament and the Council 2005:5).

### **Basis for life-long learning of foreign languages**

All policy documents studied stress the importance of promoting language learning and linguistic diversity. Communication in foreign languages is one of the key competences for life-long learning (European Parliament and the Council 2006). The Barcelona European Council of 15 and 16 March 2002 called for further action to improve the mastery of basic skills, in particular by teaching two foreign languages to all from a very early age (Council of the European Union 2002b:19). In 2008, the council considered that 'the importance attached to multilingualism and other language policy issues in the context of common EU policies imposes the need to pay these matters the attention they deserve, as well as the need for the European institutions to re-emphasise their long-standing commitment to the promotion of language learning and linguistic diversity' (Council of the Europe 2008).

### **Early language learning: foreign language teaching time and onset**

Early language learning is one of the issues highlighted in recent policy documents which the EU is planning to work on in the immediate future (European Commission 2008). The Eurydice Key data report (2005) on teaching languages at school states that countries have gradually increased the total period during which languages are taught, in particular through the provision for learning at an increasingly early age. In 2006, in most countries, more than half of the ISCED1 pupils studied a foreign language, but the percentages varied widely (Eurostat 2008). The Council affirmed in 2008 that early language learning (among others) is an effective means of improving language learning provision (Council of the Europe 2008). However the High Level Group on Multilingualism (Final report 2007) advises to study the effect of early language learning.

Starting foreign language education at an earlier age (at ISCED1 level) usually coincides with an increased duration of foreign language education and an increased total teaching time for foreign language education. Therefore, we have to assess the recommendations in the national curriculum regarding the onset (starting age), duration and teaching time of foreign language education.

Because in many countries educational institutions have a considerable curricular autonomy (Eurydice 2008) or a new starting age is slowly phased in, the foreign language learning time and onset depends on the curriculum of the particular school(s) the student attends/has attended. We should be aware, however, that in many countries we can only assess the teaching time during ISCED2, because different institutions provide ISCED1 and ISCED2 education. In these cases, we can only assess the minimum amount of teaching time allocated to foreign language education



and allocated to the specific language tested in the ESLC (from here on called target language<sup>5</sup>, because the questionnaires are targeted at this specific language).

Foreign language teaching time and onset may also vary between individual students because the target language may be a curricular option, changes of school and/or programmes may have occurred and the national curriculum may have changed during the educational career of students. Therefore, at student-level we should measure the onset of foreign language and target language learning and the time spent weekly on foreign language and target language learning (lessons and homework). The time spent on language learning does not solely depend upon the length of periods and number of periods per week, but also on the time spent on homework. The European study of pupils skills in English (Bonnet 2002) showed that the time spent on homework differs markedly between countries.

### Issue 1: Concepts related to early language learning

Level	Antecedents	Malleable aspects
Individual participant (Student Questionnaire)		FL and TL teaching onset FL and TL learning time a week (lessons and homework)
Instructional setting (Teacher Questionnaire)		
Educational institutions (Principal Questionnaire)		FL and TL teaching onset FL and TL teaching time
National educational system (National Questionnaire)		FL and TL teaching onset FL and TL teaching time

Note: FL = foreign language; TL = target language

### Diversity and order of foreign languages supply

A prominent issue within all policy documents is the diversity of languages on offer. In the Action Plan 2004–2006 (2003:8) it is stated that 'Member states agree that pupils should master at least two foreign languages' and that 'the range on offer should include the smaller European languages as well as all the larger ones, regional, minority and migrant languages as well as those with 'national' status, and the languages of our major trading partners throughout the world' (2003:9). However, the current Eurostat data (reference year 2006) show that both the different languages on offer and the number of languages students learn seems to fall short of this aim. Furthermore, the diversity of foreign languages offered seems to be limited in most countries with English being the most widely taught language in most countries. In

<sup>5</sup> The two target languages for each country are the 1<sup>st</sup> and 2<sup>nd</sup> most widely taught official European languages of the European Union, from among English, French, German, Spanish and Italian.





2008, the Council affirmed that the broadest possible range of languages should be available to learners (Council of the Europe 2008). The council invited the Member States to increase the diversity of languages offered and encourage the learning of less widely used EU languages and non-European languages.

Even though the ESLC is not studying the competence in less widely used EU languages or in non-European languages, the diversity of languages on offer and the linguistic repertoire of students is very important for another reason. Research (Cenoz, Hufeisen and Jessner 2001) has shown that the existing knowledge of other languages can affect the learning of a new language. Pupils will use the skills and knowledge of known languages that are most similar to the language to be learned (Cenoz, Hufeisen and Jessner 2001). Within education, teachers can also build on this existing repertoire of learners (see section 0). For this reason, we have to measure which languages are taught and the order in which they are taught. As was the case with the first issue (foreign language learning time and onset), the diversity of foreign language supply depends to a varying extent on the national curriculum, the school curriculum and the choice of the individual student.

## Issue 2: Concepts related to diversity and order of foreign languages supply

Level	Antecedents	Malleable aspects
Individual participant (Student Questionnaire)		Learned foreign languages Learning order of foreign
Instructional setting (Teacher Questionnaire)		Offered foreign languages Teaching order of foreign languages
Educational institutions (Principal Questionnaire)		
National educational system (National Questionnaire)		Recommended/ allowed foreign languages Teaching order of foreign languages

## Language-friendly living environment

Another highlighted issue on which the EU is planning work in the immediate future is a language-friendly living, learning and working environment. A language-friendly environment is an environment where different languages are heard and seen, where speakers of all languages feel welcome and language learning is encouraged (European Commission 2008). The Action Plan 2004–2006 (2003) states that 'every community in Europe can become more language-friendly by making better use of opportunities to hear and see other languages and cultures'.



## **Informal language learning opportunities**

Living in a language-friendly environment where different languages are heard and seen creates opportunities for informal language learning. 'Non-formal and informal learning are important elements in the learning process and are effective instruments for making learning attractive, developing readiness for life-long learning and promoting the social integration of young people' (Resolution of the Council on non-formal and informal learning 2006:2). The High Level Group on Multilingualism (Final report 2007) considers research into the long-term effects of bilingual upbringing and of out-of-school contacts with speakers of other languages – in combination with educational measures – of particular interest.

The languages that are used in the home environment are particularly important as the home environment can provide very frequent exposure and use of other languages. Students can also be exposed informally to foreign languages through direct contact with native speakers in their living environment (e.g. relatives, friends, neighbours and tourists) and through visits to countries where the foreign language is spoken. The potential for students to come into contact with foreign languages is of course influenced by the linguistic heterogeneity of the population in their home town. European countries differ in the linguistic heterogeneity given the different number of official national and indigenous languages (Eurydice 2008) and the size and languages of the immigrant populations (Eurostat 2008).

As this kind of direct exposure to foreign languages is difficult to influence, EC policies focus particularly on the role of the media. In the Action Plan 2004–2006 (2003) and in the communication from the Commission on multilingualism (2008), emphasis is placed on the use of subtitles in film and television because research has shown that subtitles can encourage and facilitate language learning. The internet and so-called “edutainment” programmes may also influence and motivate informal language learning. The new media do not only offer exposure but also the possibility of using a foreign language, for example through MSN, blogs and online gaming.

### Issue 3: Concepts related to informal language learning opportunities

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>	Languages in the home-environment Target language exposure and use through home environment Target language exposure and use through visits abroad Target language exposure and use through traditional and Home location new	
<b>Instructional setting (Teacher Questionnaire)</b>		
<b>Educational institutions (Principal Questionnaire)</b>		
<b>National educational system (National Questionnaire)</b>	National & indigenous Size and languages of language Immigrant population	Use of subtitles on television and film

### Language-friendly schools

Several issues and actions are mentioned in the policy documents that are helpful in creating a language-friendly school. A language-friendly school is a school where different languages are heard and seen, where speakers of all languages feel welcome and language learning is encouraged.

### School's foreign language specialisation

Schools can offer a type of provision in which pupils are taught subjects in more than one language, called Content and Language Integrated Learning (CLIL). While the schools offering this provision are often referred to as bilingual or immersion schools, CLIL pupils learn a subject through the medium of a foreign language. This is considered an effective means of improving language learning provision (Council of the Europe 2008).

In the report on the implementation of the Action Plan (2007c) the following conclusion is provided: 'In 2006, the Eurydice network published a survey on "Content and Language Integrated Learning (CLIL) in schools in Europe", setting out the main features of CLIL teaching in European countries. While interest in CLIL provision is growing, only a minority of pupils and students are currently involved, with the situation varying greatly from country to country. The survey showed that if CLIL provision is to be generalised, it has to be supported in most countries by a significant effort in teacher training. Another area demanding further work is evaluation: because CLIL is still in its early stages in most countries, evaluation of CLIL practices is not widespread.' CLIL is, therefore, highlighted in recent policy documents as an area in



which the EU is planning immediate future work (European Commission 2008) and we should assess the extent to which foreign languages, and specifically the target languages, are used in schools for instruction in other subjects.

Schools that do not offer CLIL can also profile themselves as specialized in foreign languages. Because in many countries schools have some curricular autonomy, schools can 'introduce some subjects of their own choice – and in particular foreign languages – as part of the minimum level of educational provision' (Key Data on Teaching Languages at School in Europe - 2008 Edition 2008:32) or dedicate more teaching time to foreign languages than other schools. Furthermore, schools can offer enrichment lessons in foreign languages.

#### **Issue 4: Concepts related to the school's foreign language specialisation**

<b>Level</b>	<b>Antecedents</b>	<b>Malleable aspects</b>
<b>Individual participant (Student Questionnaire)</b>		Participation in FL and TL enrichment and remedial
<b>Instructional setting (Teacher Questionnaire)</b>		
<b>Educational institutions (Principal Questionnaire)</b>		Use of FL and TL for the instruction in other subjects Specialist language profile FL and TL enrichment and remedial lessons
<b>National educational system (National Questionnaire)</b>		

#### **Information and communication technology to enhance FL learning and teaching**

Another highlighted area for EU work is Information and Communication Technologies (Communication from the Commission about Multilingualism 2008). 'Information and communication technologies (ICT) offer more opportunities than ever before for learners and teachers to be in direct contact with the target language and target language communities' (European Commission 2008), for example through pedagogical use of ICT for learning (eLearning) and through Internet-facilitated school 'twinings' (Action Plan 2004–2006 2003). ICT offers flexibility in terms of time and place for accessing language learning opportunities and therefore can make language learning more widely available, accessible and attractive to all. ICT can also be used to increase the diversity of languages offered, to maintain links between teachers, and for independent learning and distance learning.

To address this policy issue the frequency with which teachers and pupils use ICT in the context of foreign language education and the purpose of the use (e.g. direct contact with the target language, lesson preparation, contacts with other FL teachers, school twinning, homework, making exercises) should be assessed. The use of ICT

might vary due to different ICT facilities in school and in the home environment of teachers and pupils. The European study of pupils' skills in English (Bonnet 2002) showed that resources like computer programs or the internet were very rarely used, but it remained unclear whether the finding reflected the constraints and availability of such media. Therefore, the ICT facilities in school and in the home environment of teachers and pupils also need to be assessed.

### **Issue 5: Concepts related to information and communication technology to enhance FL learning and teaching**

<b>Level</b>	<b>Antecedents</b>	<b>Malleable aspects</b>
<b>Individual participant (Student Questionnaire)</b>	ICT-facilities at home  Frequency and purpose of using ICT for personal	Frequency and purpose of using ICT in FL learning
<b>Instructional setting (Teacher Questionnaire)</b>	ICT-facilities at teachers home  Frequency and purpose of using ICT for personal ICT-facilities in school	Frequency and purpose of using ICT in FL teaching
<b>Educational institutions (Principal Questionnaire)</b>	ICT-facilities in school	
<b>National educational system (National Questionnaire)</b>		

### **Intercultural exchanges**

The EU has very actively promoted intercultural exchanges through the mobility schemes of several educational programmes (Comenius, Leonardo and Erasmus). Exchanges provide direct experience with the target language and target culture, which is considered to be helpful for increasing communicative and intercultural competence and awareness. Language and culture awareness is another highlighted area in which the EU intends further work (European Commission 2008). According to the Action Plan 2004–2006 (2003) all pupils should have the experience of taking part in Comenius school language projects, in which a class works together on a project with a class abroad, and in a related language exchange visit.

The extent to which schools create opportunities for intercultural exchanges is subject to a certain extent to financial constraints. Exchanges may be funded nationally, locally, or by parents. When schools create opportunities for exchange visits or school language projects, these opportunities are not necessarily provided for all foreign languages and participation may be optional. Therefore, we should assess whether pupils received these opportunities, specifically for the target language.

## Issue 6: Concepts related to intercultural exchanges

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>		Received opportunities regarding the target language for exchange visits and school language projects
<b>Instructional setting (Teacher Questionnaire)</b>		Created opportunities for exchange visits and school language projects
<b>Educational institutions (Principal Questionnaire)</b>		Created opportunities for exchange visits and school language projects Funding of intercultural exchanges
<b>National educational system (National Questionnaire)</b>		Funding of intercultural exchanges

### Staff from other language communities

According to the Action Plan 2004–2006 (2003) all secondary schools should be encouraged to host staff from other language communities, such as language assistants or guest teachers, because such exchanges ‘can improve the skills of young language teachers whilst at the same time helping to revitalise language lessons and have an impact upon the whole school, in particular by introducing schools to the value of teaching less widely used and less taught languages’.

At school level we should assess whether and how often they host language assistants and guest teachers from other language communities. Furthermore, the number of foreign language teachers that are native speakers of the target language should be assessed. As teaching a language to native speakers is quite different from teaching the language as a foreign language, we should also assess whether the native-speaking teachers have received training to teach their native language as a foreign language.

## Issue 7: Concepts related to staff from other language communities

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>		
<b>Instructional setting (Teacher Questionnaire)</b>	Teachers 1st language(s)  Teachers from other language communities	Training of teachers from other language communities to teach the target language as a foreign language
<b>Educational institutions (Principal Questionnaire)</b>		Language assistant and guest teachers from other language communities
<b>National educational system (National Questionnaire)</b>		

### Language learning for all

A language-friendly school is also a school where speakers of all languages feel welcome. Language learning should be for everybody. Improving equity in education and training is one of the eight key policy domains of the Education and Training 2010 strategy (Communication from the Commission: 'A coherent framework of indicators and benchmarks for monitoring progress towards the Lisbon objectives in education and training' (COM (2007) 61 final 2007b)). In 2008 the Council invited Member States to 'take appropriate steps to improve effective language teaching and continuity for language learning in a life-long learning perspective, including by making existing resources and infrastructure more widely available, accessible and attractive to all' (Council of the Europe 2008).

The equity dimension is usually studied through breaking down data by the sex, age and socio-economic background of learners. As agreed at the Advisory Board meeting of 19-20 June 2008, the measurement of socio-economic status will be, where possible, consistent with the measurement in PISA surveys, although the extent to which this is possible may be limited by the difference in populations of PISA and of the ESLC. In PISA (OECD 2007) assessing the socio-economic status is made operational through assessing the parental occupational status (six questions), parental educational status (four questions) and household possessions (three questions).

Another group of students, specifically mentioned, are immigrants. In 2008 the Council affirms that 'to help them integrate successfully, sufficient support should be provided to migrants to enable them to learn the language(s) of the host country, while members of the host communities should be encouraged to show an interest in the cultures of newcomers' (Council of the Europe 2008). The 2005 Eurydice Key Data report on teaching languages at school states that certain schools enrol large numbers



of pupils whose mother tongue is not the language of instruction (Eurydice 2005). Furthermore, there is evidence that ability grouping/tracking places a disproportionately high share of migrant pupils into lower-ability streams (Green Paper on Migration and Mobility 2008b).

At school-level, several approaches to helping immigrant children acquiring the host language can be discerned (Eurydice report on integrating immigrant children 2004), such as extra-curricular or pre-school language lessons in the host language and extra homework or attention during lessons. In order to address this issue, not only the immigrant status of pupils should be assessed, but also the provided/received help to master the host language. This approach can be combined with another support measure for immigrant pupils, which is the teaching of the 1<sup>st</sup> language(s) of immigrant children.

### Issue 8: Concepts related to language learning for all

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>	Immigrant status  Gender  Age  Socio-economic status (parental occupational status, parental educational status, household possessions)	Received help in mastering host language Received formal education in language(s) of origin
<b>Instructional setting (Teacher Questionnaire)</b>		
<b>Educational institutions (Principal Questionnaire)</b>	Percentage of immigrant students	Provisions for help in mastering host language Teaching of language(s) of origin
<b>National educational system (National Questionnaire)</b>		

### Foreign Language Teaching Approach

In 2002 the council invited Member States 'to promote the application of innovative pedagogical methods, in particular also through teacher training' (Council resolution on linguistic diversity and language learning 2002). The EU does not promote a particular teaching method with a clear defined set of activities, but rather a broad holistic approach to teaching in which emphasis is placed upon communicative ability and multilingual comprehension. According to the Action Plan 2004–2006 (2003:8) 'the emphasis should be on effective communicative ability: active skills rather than passive knowledge' during secondary education. Furthermore, the potential value of a multilingual comprehension approaches are emphasised (European Commission



2008). 'It is important that schools and training institutions adopt a holistic approach to the teaching of language, which makes appropriate connections between the teaching of "mother tongue", "foreign" languages, the language of instruction and the languages of migrant communities; such policies will help children to develop the full range of their communicative abilities. In this context, multilingual comprehension approaches can be of particular value because they encourage learners to become aware of similarities between languages, which is the basis for developing receptive multilingualism' (Action Plan 2004–2006 2003:9). In a multilingual comprehension approach the linguistic similarities between languages of the same language group are exploited to make the first steps of foreign language learning easier. Acknowledging and building on the existing linguistic repertoire of learners, is an aspect much emphasised in the guide for the development of language education policies of the Council of Europe (2007).

In contrast to the multilingual approach, the implementation of the communicative approach has been evaluated in several European studies. Within the national curricula of lower secondary education few differences in emphasis are found: 'the great majority of countries issue recommendations to attach equal emphasis to all four communication skills' (Eurydice 2008). The emphasis on other aspects though, such as grammar, vocabulary and pronunciation, is not reported. As for the actual implementation of teaching methods, a European study of pupils' skills in English (Bonnet 2002) showed marked differences in the use of the target language during lessons, whereas few differences were found between other aspects of the teaching method employed by foreign language teachers. We should note that information about teaching methods was reported by teachers themselves, not by their students, and a combination of student and teacher viewpoints might have proved invaluable.

To summarise, we should assess the emphasis on the four communicative skills compared to the emphasis on language content (grammar, lexis and pronunciation) within the national curriculum and within the teaching activities (instruction, classroom activities, homework and assessment) and resources used (books, video tapes, etc). Furthermore, the emphasis on similarities between known languages and the use of the target language during foreign language lessons should be measured. The viewpoints of the teacher should be triangulated with the viewpoints of the students.

In addition to the perception of students regarding teaching activities, their perception regarding foreign language learning and foreign language lessons may provide important insights. The European study of pupils' skills in English (Bonnet 2002) mentioned previously shows marked differences between the pupils of various countries in the perceived importance and appreciation of English. Like those pupils, adults of different European countries differed in the perceived usefulness of foreign languages and in the perceived impediments to foreign language learning (Eurobarometer 2006). An impediment very frequently mentioned, one that might also apply to students studying foreign languages, was 'not being good at languages'.

## Issue 9: Concepts related to teaching approach

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>		Perceived emphasis on the four communicative skills and language content within the teaching activities and resources used Perceived emphasis on similarities between known languages Use of the target language during foreign language lessons Perception (attitude) of foreign language, foreign language learning and foreign language lessons
<b>Instructional setting (Teacher Questionnaire)</b>		Emphasis on the four communicative skills and language content within the teaching activities and resources used Emphasis on similarities between known languages Use of the target language during foreign language lessons
<b>Educational institutions (Principal Questionnaire)</b>		
<b>National educational system (National Questionnaire)</b>		Emphasis on the four communicative skills and language content within the national curriculum

### Teacher initial and in-service training

Better language teaching is not only associated with a language-friendly school, but also with language teacher training. Improving the quality of initial teacher education and ensuring that all practicing teachers take part in continuous professional development has been identified as key factors in securing the quality of school education (Commission of the European Communities 2007b). The European policies and action have to a great extent been aimed at the language teacher. The Council affirmed in 2008 that 'Quality teaching is essential for successful learning at any age and efforts should therefore be made to ensure that language teachers have a solid command of the language they teach, have access to high quality initial and continuous training and possess the necessary intercultural skills. As part of language



teacher training, exchange programmes between Member States should be actively encouraged and supported' (Council of the Europe 2008). As directly testing the language skills of teachers is beyond the scope and purpose of the ESLC, the focus is on the efforts made to ensure the competence of teachers.

### **Access to high quality initial and continuous training**

According to the Eurydice Key Data report on teaching languages at school in Europe (2008), the level of initial teacher training tends to be ISCED5, but the duration of training can vary. The foreign language teachers in secondary education generally have to be specialists, but not in every country. Furthermore, the teachers can be specialised to teach one foreign language, several foreign languages or two subjects, one of which is a foreign language (Eurydice 2008). Even though the national recommendations are quite similar, there may be a difference between the recommendations and the implementation as some Member States face shortages of adequately-qualified language teachers. Furthermore, the national recommendations may have changed resulting in older teachers having different qualifications. Therefore, both at a national level and teacher level the duration, level, and specialisation of initial teacher training and the teacher qualifications should be assessed.

Like for students in secondary education, life-long learning for foreign language teachers is actively promoted. According to the Action Plan 2004–2006 (2003) all teachers should have regular opportunities to update their training and to keep their language and teaching skills up-to-date through e-learning and distance learning *inter alia*. The European Profile for Language Teacher Education (Kelly, Grenfell, Allan, Kriza and McEvoy 2004) also emphasizes the continuous improvement of teaching skills through in-service education.

We should assess the extent to which and how (e.g. via e-learning), teachers have participated in in-service training, as well as the focus of the training (e.g., language skills, ICT skills, language teaching methods). At school-level we could assess the incentives for participation in in-service training (e.g. rise in income, position, promotion).

## Issue 10: Concepts related to access to high quality initial and continuous training

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>		
<b>Instructional setting (Teacher Questionnaire)</b>	Age of teacher  Gender of teachers	Level and duration of initial training Qualifications and specialisation of teachers Participation in in-service training Mode and focus of inservice-training Incentives for inservice training
<b>Educational institutions (Principal Questionnaire)</b>		Incentives for inservice training
<b>National educational system (National Questionnaire)</b>		Required level and duration of initial teacher training Specialisation and qualifications of teachers

### A period of work or study in another country

Intercultural exchanges for teachers obviously benefit teachers in the same way that they benefit pupils in secondary education: increasing communicative and intercultural competence and awareness through direct experience with the target language and target culture; see among others, Lace (2007). Exchanges for teachers have the additional benefit of helping Member States with the introduction of Content and Language Integrated Teaching. and of helping Member States that face shortages of adequately-qualified language teachers (Action Plan 2004–2006, 2003). Furthermore, an exchange of teachers facilitates contacts and networking among teachers and between educational providers.

In the Action Plan 2004–2006 (2003:34-35) it is recommended that (future) teachers stay for an extended period in the country where the language to be taught is spoken. A period of work or study in a country or countries where the trainee's foreign language is spoken as a native language and the opportunity to observe or participate in teaching in more than one country are also included in the European Profile for Language Teacher Education (Kelly, Grenfell, Allan, Kriza and McEvoy 2004).

The report on the implementation of the Action Plan (2007c) however concludes that 'in many Member States language teachers are not obliged to spend a period abroad in the country whose language they teach', but 'the need is widely recognised among practitioners and teacher trainers, who make use of the mobility schemes offered by European educational programmes (Erasmus, Comenius, Leonardo) to improve their



language skills in many Member States'. As teacher mobility is still rather low (Council of the Europe 2008), the Council affirms that 'as part of language teacher training, exchange programmes between Member States should be actively encouraged and supported' and invites Member States to 'promote mobility among language teachers to enhance their language and intercultural skills'.

The extent to which foreign language teachers stay abroad for an extended period depends to a certain degree upon financial possibilities. The funding of such stays can be obtained through mobility schemes offered by European educational programmes (Erasmus, Comenius, Leonardo), national schemes or by opportunities found or created by the teachers themselves.

### Issue 11: Concepts related to a period of work or study in another country

Level	Antecedents	Malleable aspects
Individual participant (Student Questionnaire)		
Instructional setting (Teacher Questionnaire)		Stay in target culture and reason (study, work, other)
Educational institutions (Principal Questionnaire)		Incentives for stays abroad Funding of stays abroad
National educational system (National Questionnaire)		Requirements regarding stay abroad during initial training Funding of stays abroad

### Use of existing European language assessment tools

Both high quality initial and continued training and studying/working abroad are efforts to ensure that language teachers have a solid command of the language they teach. Another effort to increase foreign language competence and motivation for foreign language learning of both teachers and their pupils is the use of the European Language Portfolio (Council of Europe 2008a), which is based upon the CEFR (Council of Europe 2008b). In 2008, the council invited Member States to 'use existing tools to confirm language knowledge, such as the Council of Europe's European Language Portfolio and the Europass Language Portfolio' (Council of the Europe 2008). According to the European Profile Language Teacher Education (Kelly, Grenfell, Allan, Kriza and McEvoy 2004), (future) teachers should be trained in the use of the European Language Portfolio for self-evaluation.

Over half the Member States have formulated recommendations for 'the use of the CEFR as an assessment tool' (Eurydice 2008:108). A survey of the Council of Europe showed that the CEFR is quite widely used and used mostly by teachers, teacher trainers, test writers and material writers (Council of Europe 2005:3). We should assess the purpose and context in which foreign language teachers use the CEFR. Furthermore, we should assess whether teachers use the European Language Portfolio and whether they have been trained in the use of the Portfolio.

## Issue 12: Concepts related to the use of existing European language assessment of tools

Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>		
<b>Instructional setting (Teacher Questionnaire)</b>		Use of CEFR and received training in use Use of European Language Portfolio and received training in use
<b>Educational institutions (Principal Questionnaire)</b>		
<b>National educational system (National Questionnaire)</b>		Recommendations for the use of the CEFR and the European Language Portfolio

### Practical experience

Foreign language teaching also requires considerable practical skills. According to the Action Plan 2004–2006 (2003) 'Initial training should equip language teachers with a basic 'toolkit' of practical skills and techniques, through training in the classroom'. The importance of an internship is also stressed in the European Profile for Language Teacher Education (Kelly, Grenfell, Allan, Kriza and McEvoy 2004). According to the European Profile for Language Teacher Education teacher training should have an explicit framework for teaching practice (stage/practicum) and a curriculum that integrates academic study and the practical experience of teaching. Trainees should be trained in skills to incorporate research into teaching and in the practical application of curricula, syllabuses, teaching materials and resources.

Not only the practical experience acquired during initial training can differ between Member States, but the teaching experience acquired as a qualified teacher can also differ significantly. Partially due to different national recommendations regarding teacher training (Eurydice 2008), some teachers only have experience in teaching the target language, while others may also have experience in teaching other foreign languages or other subjects. Furthermore, to counter teacher shortages, sometimes teachers are re-trained to teach a different foreign language to the one for which they were originally trained.

### Issue 13: Concepts related to practical experience

Level	Antecedents	Malleable aspects
Individual participant (Student Questionnaire)		
Instructional setting (Teacher Questionnaire)		Stage during initial training Teaching experience in FL, TL and other subjects
Educational institutions (Principal Questionnaire)		
National educational system (National Questionnaire)		Stage required during initial training

### Organisational structure of the educational systems

When studying the relationship between policy actions and foreign language competences, we need to take the organisational structure of European education systems into account. The onset and duration of compulsory education differs between Member States, as well as the number of institutional distinctions and streaming within the education provided. Streaming can occur at the institutional level based upon exams and/or teacher assessment, but can also occur within institutions. Within institutions, students can be grouped in classes according to general ability or grouped within a class according to ability in particular subjects (a practice known as 'setting').

The effect of those different ways of grouping students on educational outcomes depends upon the size of the groups. Class sizes vary considerably from one country to the next and from one school to the next (Eurydice 2005). Whereas sometimes the class size is subject dependent, few countries establish class size norms specifically for foreign language teaching (Eurydice 2008). How education is organised and can be organised is of course restrained by the general affluence of the country and the investment in education.

## Issue 14: Concepts related to the organisational structure of the educational systems

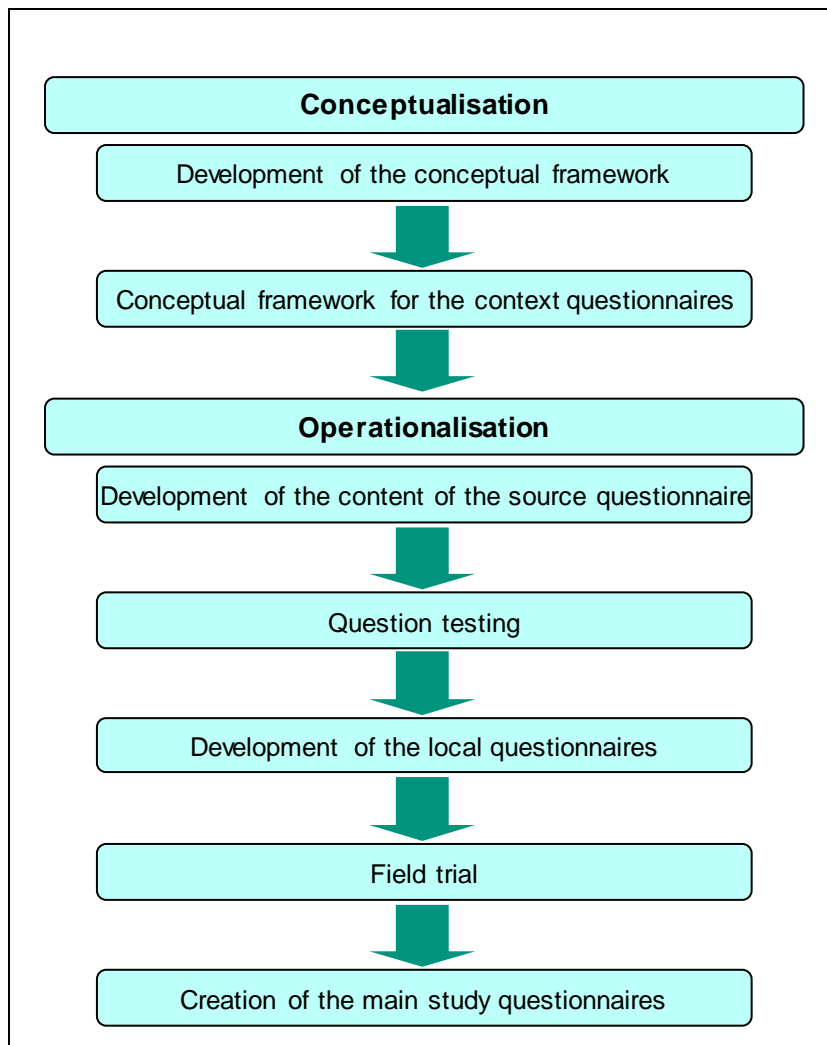
Level	Antecedents	Malleable aspects
<b>Individual participant (Student Questionnaire)</b>	Class size Study program Grade	
<b>Instructional setting (Teacher Questionnaire)</b>	Within class ability grouping (setting) TL class size	
<b>Educational institutions (Principal Questionnaire)</b>	Within school streaming based on general ability Admission criteria Class size	TL compulsory in curriculum
<b>National educational system (National Questionnaire)</b>	General affluence Investment in education Onset and duration of compulsory education Institutional distinctions Streamed educational systems Class size norms	TL & FL learning compulsory in curriculum

### 3.2 Operationalisation

The process that leads from concepts to survey questions (the operationalisation) consisted of five phases (see Figure 3 Phases in the development of the context questionnaires). First, the source questionnaires were developed (section 3.2.1). The content of the source questionnaires went through thorough question pretesting (section 3.2.2). Once the source questionnaires were agreed upon, the local questionnaires were created for administration in each Member State (section 3.2.3). On the basis of the outcomes of the Field Trial (section 3.2.4), the final questionnaires were created for the Main Study (section 3.2.5).



**Figure 3 Phases in the development of the context questionnaires**



### **3.2.1 Development of the content of the source questionnaire**

To guide the item writing process (see section 3.2.1.3) and the development of the testing tool (see chapter 6), first the general design of the questionnaires had to be specified. The general design consisted of a description of the general structure of the questionnaires, the question types and the question elements.

#### **General design of the questionnaires**

To be able to transform the concepts into actual question content, the question-types one intends to use needs to be decided upon. Which question-types are feasible and likely to generate a valid measurement depends to a certain extent upon the mode in which the questionnaire is administered (De Leeuw 2008). Each administration mode



has certain advantages and creates particular possibilities in terms of question types. Therefore when constructing a questionnaire we should take the mode of administration into account (Dillman 2008).

In the ESLC we had a dual administration mode: computer-based and paper-based (Communication from the Commission to the Council 2007:6). In the case of a dual administration mode, two approaches can be discerned in designing a questionnaire: a mode-specific construction and a unified mode construction. Within the mode-specific design the questionnaire is constructed separately for each mode, independent of what might be done in the other mode. In the unified mode the aim is to provide the same stimulus across modes in order to prevent unnecessary divergence across modes (Dillman 2000).

As the mode was expected to vary between countries, with some countries using only the computer-based mode, others only the paper-based mode and some using both, we were aware that differences between the preferred modes might cause systematic differences in responses between countries. Therefore, a unified mode of construction was preferred whereby the aim was to create the same question types, questions, question order, questionnaire lay-out and situation of questionnaire administration for both modes.

Taking into account that the paper-based and computer-based had to be equivalent and logistically feasible, the general structure was based upon the structure of similar international context questionnaires, such as PISA (OECD 2008), TIMSS and PIRLS (TIMSS & PIRLS International Study Center 2008), TALIS (OECD 2008). The questionnaire also had to be efficient and easy to use for respondents, coders and data analysts, because a questionnaire that is too time-consuming or complex for respondents, coders, and data-analysts, is very likely to produce unreliable outcomes.

### **General structure of the questionnaires**

The questionnaires consist of several parts. The front page of the questionnaires displays the name of study and the name of the questionnaire, the version (Field Trial, Main Study etc.), the date, the author (SurveyLang and EC) and an identification label. Following this front page a general introduction is presented in which the purpose and content of the questionnaire is explained, how long it will take to complete and what will be done with the answers.

The questionnaire is divided in several sections, grouping questions within the same general subject area together, for example 'about you, about your family, about your school environment, about foreign languages, about your foreign language lessons'. Some sections start with a short explanation of the kind of questions contained in that particular section.

The last section of the questionnaire may contain up to five country-specific questions. In addition to the European policy issues mentioned in the conceptual framework, other important issues can apply within each participating country that are deemed



less relevant to other participating countries. For this reason each participating country (from here on called educational system<sup>6</sup>) could add up to five questions to the questionnaires to pursue such issues.

After the last section of the Student Questionnaire a self-assessment section follows. The sixteen short self-assessment questions consist of “Can Do” statements, similar to the ones used as performance descriptors in the CEFR. This self-assessment is important for a cross-language linkage of language skills to the CEFR (see chapter 2). After the self-assessment the respondent is thanked for his/her cooperation. Note, due to the inclusion of the Can Do self-assessment statements, it was agreed by the European Commission on advice of the Advisory Board (18-19 March 2009) that the length of the Student Questionnaire be lengthened from 30 minutes to 45 minutes.

### **3.2.1.1 Question types**

In total nine different question types are used. The main consideration in selecting question types was that each question type should comprise very concrete and easy tasks for the respondents, in particular for the students. The question types used differ in three aspects: the question format (multiple items or no items), the response format (closed or open) and the number of responses required (one or several).

Two question formats are used: questions with several items (grid questions, see Figure 4) and simple questions without items (see Figure 5). Both question formats can be combined with a closed response format or an open response format. The most frequently occurring response format is the closed response format in which a limited number of pre-defined response options are presented from which the respondent has to choose. In single choice questions the respondent has to choose one single response option (see Figure 5) and in the free-choice questions the respondent is free to choose any number of options (see Figure 6).

---

<sup>6</sup> The educational system is a country, geographic region, or similarly defined population, for which the Consortium fully implements quality assurance and quality control mechanisms and endorses, or otherwise, the publication of separate ESLC results

**Figure 4 Example of a question with several items (grid question)**

---

24 How often do you use a computer outside school time for the following?

*(Please select one answer from each row)*

	<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>A few times a week</i>	<i>(Almost) every day</i>
1) For homework or school assignments -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) For homework or assignments for the subject of [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) For finding information-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) For games -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) For entertainment (e.g. music, movies, video clips) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) For contact with others (e.g. email, chatting, blogging, {MySpace, Skype}) ---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

---

**Figure 5 Example of a simple question without items (single choice)**

---

60 How much time do you usually study for a [target language] test?

*(Please select only one answer)*

No time at all ----- ☐ 0

Less than one hour ----- ☐ 1

About one to two hours ----- ☐ 2

About two to three hours ----- ☐ 3

More than three hours ----- ☐ 4

---

**Figure 6 Example of a simple free choice question**

---

30 Is participation in in-service training an obligation, a right or an option for you?

*(Please select the answer(s) that describes your situation best)*

Participation in in-service training is an obligation for teachers ----- ☐ 0

Participation in in-service training is a right for teachers----- ☐ 1

Participation in in-service training is required for promotion ----- ☐ 2

Participation in in-service training is optional ----- ☐ 3

---

The open response format is less frequently used, as this format tends to yield more invalid responses and outliers. Furthermore, open responses are more cognitively demanding for respondents and more costly in terms of data-analysis. Especially, open-ended questions that require a text response are very difficult to standardise and costly in terms of coding and data-analysis. However, four open-ended questions that require a text response have been used, because the Commission decided that the index of socio-economic status had to be comparable to the index used in PISA. The coding of these textual open-ended questions was a task for the NRCs (see section 7.16). All other open questions required one or two numerical answers.

In some Field Trial questions the last response category (“Other, namely ...”) was open-ended, for example in the questions about the language(s) spoken at home. These were a safeguard in case an important response category (e.g. a widely spoken language) in the explicit list had been overlooked, and would give information on how to change the explicit list in the questionnaire for the Main Study. These open response categories do not occur in the Main Study questionnaires.

**Figure 7 Question types in the Main Study questionnaires**

Response format	Question format	Number of responses	Question type	SQ	TQ	PQ
<b>Closed</b>	<b>Simple</b>	<b>one</b>	Closed single choice question	16	12	3
		<b>several</b>	Closed free choice question	6	9	5
	<b>Grid</b>	<b>one</b>	Closed single choice grid question	32	28	22
		<b>several</b>	Closed free choice grid question			1
<b>Open</b>	<b>Simple</b>	<b>one</b>	Open (numerical) question	3	3	3
			Open (tekst) question	4		
		<b>several</b>	Date	1		
	<b>Grid</b>	<b>one</b>	Open (numerical) grid question	1	7	6
		<b>several</b>	Open (numerical) grid requiring two responses			5

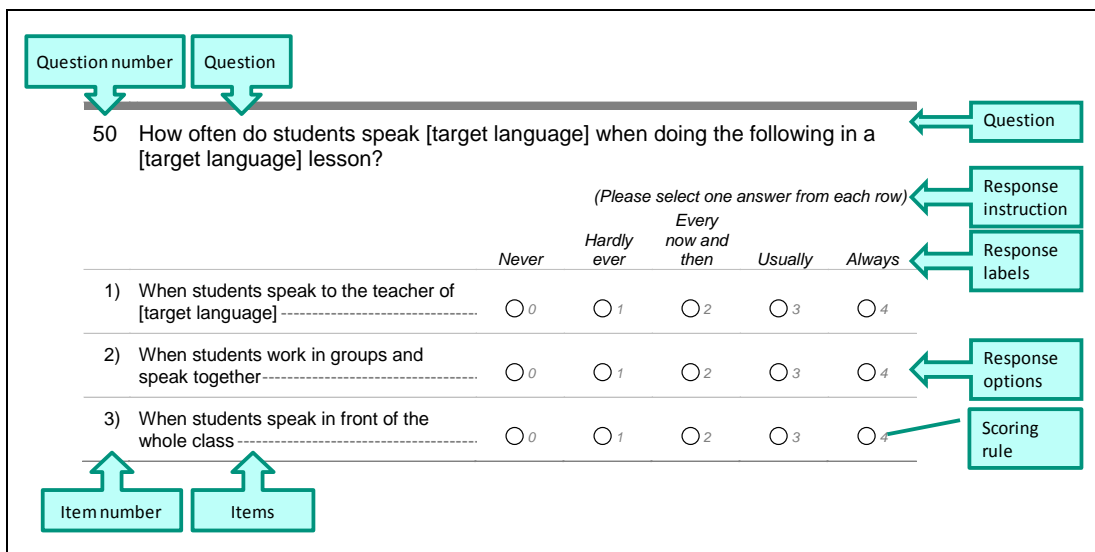
### 3.2.1.2 Lay-out of the questions

All questions consist of at least four elements and a maximum of eight elements (see Figure 8). All questions consist of a numbered question with the option of clarification of the question or intended response, a response instruction and response option(s). To distinguish optimally between the different response formats, all closed single choice questions are presented with an option button (see Figure 5). All closed free choice questions are presented with check boxes (see Figure 6) and all open questions are presented with a text box.

All closed questions also have response labels, which in the paper-based version are accompanied with a scoring rule for data entry. All grid questions have items, which are numbered in the paper-based version.

To allow for the display of the questions on a small computer screen the maximum number of items in a grid question was set at 10. The maximum number of response options displayed vertically was set at fifteen. The maximum number of response options displayed horizontally was set at four or five (see Figure 8) depending on the type of response scale. In Likert-type scales (most often used in attitude and personality measurement) an even number (four) of categories is used to avoid central tendencies. In intensity and frequency scales an uneven number of options are used as well (with a maximum of five). For the intensity and the frequency scales, there is no danger for central tendency, as the labels express something in an increasing degree, such that there is no neutral point.

**Figure 8 Question elements**



The diagram illustrates the components of a grid question. It shows a sample question: "50 How often do students speak [target language] when doing the following in a [target language] lesson?". Below the question is a response instruction: "(Please select one answer from each row)". The response labels are "Never", "Hardly ever", "Every now and then", "Usually", and "Always". The response options are "0", "1", "2", "3", and "4". The scoring rule is "1 point per correct answer". The item number is "50" and the items are "1) When students speak to the teacher of [target language]...", "2) When students work in groups and speak together...", and "3) When students speak in front of the whole class...".

	Never	Hardly ever	Every now and then	Usually	Always
1) When students speak to the teacher of [target language].....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) When students work in groups and speak together.....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) When students speak in front of the whole class.....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### 3.2.1.3 Question writing

The main concern in developing the questions was to obtain a valid measurement. Validity is built-in from the outset of questionnaire development. The rationales underlying questionnaire development form an integral part of validity evidence, see among others Anastasi (1986), Messick (1995), Kane (1992), Schouwstra (2000).

We needed to ensure that the questions captured the concepts and different situations adequately and that respondents could and would respond to the questions as intended. In question writing the conceptual framework was strictly followed and the questions from other international surveys were used as examples. In order to prevent unnecessary divergence across the two administration modes (paper-based and computer-based), the five target languages and the different Questionnaire Languages (see section 3.2.3.1) for this first cycle of the ESLC, one common set of questions was developed.

To ensure that respondents could answer the questions as intended the recommendations from survey methodologists regarding the question wording were

followed, see for example, Fowler & Cosenza (2008), Schouwstra (2000), Heuvelmans (2006). The aim was to prevent misunderstandings of the questions and to prevent putting too high demands on the cognitive skills required for responding to questions. Tourangeau (Tourangeau 1984), (Tourangeau, Rips and Rasinski 2000) gives a global description of the response process involved. In responding, the respondent has to understand the literal and pragmatic meaning of the question, retrieve relevant information from memory, formulate a judgment based upon the retrieved information and give a response using the response scale offered.

- (i) To aid question comprehension we formulated short, concrete questions with familiar, non-technical words and grammar.

To allow the retrieval of relevant information from memory we tried to prevent recall problems due to questions referring to very detailed information about the past whenever such detail was unnecessary for constructing the indices that would be used in the analysis.

Furthermore, as respondents should be able to formulate a judgment and give a response using the response scale provided, we were particularly concerned with avoiding ambiguity. Ambiguity can arise due to the use of negations, hidden premises, double-barrelled questions and ambiguous words or due to using response alternatives that do not match the question or that are not exhaustive and mutually exhaustive.

### **3.2.2 Question pre-testing**

Even well designed questionnaires can lead to unintended question interpretations; therefore the questions have to be (pre-) tested. Rather than choosing only one of the procedures developed for question testing, 'it is best to combine methods and take advantage of the strong points of each method' (Campanelli 2008:197). During the questionnaire development process informal methods of question (pre-) testing were employed to find errors early in the questionnaire development period. Colleagues were asked to provide feedback throughout the question writing process. Most importantly, we used a thorough question testing approach that consisted of cognitive labs and an extensive expert review.

The main purpose of cognitive labs is to identify problems respondents might have during the cognitive process of question answering (question comprehension, recall, judgment and response) and to gain insight into the source of the problems. Cognitive labs have proven useful in pinpointing problems in less time, with less effort and at lower costs than a field-trial (Campanelli 2008), like for example in PISA. The cornerstone of the cognitive labs method is the think-aloud procedure in which a small number of respondents (10 to 12) are instructed to verbalise their thoughts while answering a question followed by a short interview with an interviewer after each question is completed (Paulsen and Levine 1999); (Levine, Huberman, and Buckner 2002); (Campanelli 2008).

The draft versions of the source questionnaires were translated and localised for Dutch students and teachers. This pretesting of the translation and localisation was important

for developing notes for the upcoming translation and localisation. In sessions of about one hour the respondent was asked to think aloud during question-answering or to retrospectively explain how he/she came up with the answer for questions. After each question was completed a short interview was held in which the interviewer asked special questions (probes) to explore the response provided. The presence of an interviewer also enabled the use of observation. During their training the interviewers were shown how to pick up on verbal and non-verbal cues that could indicate guessing or problems in understanding or answering a question (for example, when the respondent looks surprised for an instant when answering the question or makes an annoyed sound). For registering the reactions of the students, a custom-made registration form was used.

Based on the outcome of the cognitive labs the question wording was refined and the terms that might need adapting to the situation in each country (localisation) were marked.

### **3.2.2.1 Expert feedback**

All draft questionnaires went through an intensive expert review process. An expert review is important for preventing unintended question interpretations and to allow the educational systems involved to check whether the concepts are adequately represented in the questions. Furthermore, the review is especially important for getting a cross-cultural input for question formulation.

The consortium members, NRCs and the Advisory Board Members received a form containing the draft source questionnaires in order to review the drafts. On the form above each question it was indicated what was the intended concept and policy issue. This allowed the reviewers to check the adequacy of the concept coverage. Below each question two fields were placed (see Figure 9 into which the reviewers could type their comments. In the first field they could indicate if they expected (some) students in their country not to be able to answer the question as intended. In the second field they could indicate if they foresaw that response alternatives or terms should need adapting to the situation in their country (localisation).



**Figure 9 Example of a draft question in the review form**

*Q1 Gender; Issue: Language learning for all (8)*

1	Are you female or male?	<i>Female</i>	<i>Male</i>
		<input type="checkbox"/> 0	<input type="checkbox"/> 1

*Question comprehension*

*Localisation*

All feedback was added to a database and evaluated. In reviewing the feedback we always considered whether it was best to implement the suggestions in the source questionnaire or to address the suggestions in the localisation (see section 3.2.3.2). The suggestions received were very useful for improving the question wording and to better anticipate where localisation might be needed. Furthermore, the queries and questions raised were very helpful for writing notes to aid the upcoming translation and localisation process.

On the basis of on the expert review the source questionnaires were finalised. By incorporating suggestions from the reviewers we have tried to keep the questions and response options as short and concrete as possible, wanting to ensure that the questionnaires would be easy to fill out. We also took care not to ask for highly detailed information whenever such detail was deemed unnecessary for constructing the indices.

Another key aspect in the questionnaire development was to consider the workload of the NRCs. We therefore did not add additional open-ended questions requiring a written text response. This is because the coding of open-ended questions requires a lot of time and work for the NRCs.

Of course, we also needed to keep in mind the conceptual framework and question formats agreed upon. In general we did not add items (or questions) whenever additionally proposed questions fell outside the scope of the agreed upon framework. Similarly, we did not remove questions whenever removal would yield an inadequate operationalisation of an important concept from the conceptual framework.

In this latter respect we need to mention the concept of the measurement of economic, social and cultural status (ESCS hereafter). Quite a few reviewers commented on the questions intended to measure SES. In June 2008, the European Commission, on advice of the Advisory Board, decided that SurveyLang should deliver a measure of

ECSC comparable to the one used in PISA. The PISA measure of ESCS is fairly complicated and requires responses from students regarding the occupation and education of both parents and regarding the presence and number of particular possessions in the home (e.g. books). To ensure comparability with PISA we needed to obtain responses from students to the same set of questions (most of the questions in the section “About your family and your home”). Therefore, this set of questions was kept despite the comments of reviewers.

To aid the upcoming translation process full note versions of the source questionnaires were created for the NRCs. The full note versions contained four types of notes below each question (see Appendix 2):

- (i) Notes for WebTrans (the system used for translation, see chapter 5) indicating the recurring question elements that were linked in WebTrans, so these elements had to be translated only once.
- Notes for the NRC clarifying terms and options, noting where localisations should be made, and providing a rationale for the question’s inclusion.
- Notes for the translator clarifying terms and options, noting where response categories and/or terms should not be translated, because they had to be localised.
- Notes for the Test Administrators giving some guidance in how to answer questions that students might ask during the administration.

### **3.2.3 Development of the local questionnaires**

For creating the local student, teacher and principal questionnaires, the source questionnaires had to be translated<sup>7</sup>. The main purpose of translating the questionnaires was that all (or almost all) respondents could comprehend the intended meaning of the questions and would feel at ease when Reading and responding to the questions. It was also crucial that the language in which the questionnaires would be administered fitted the legal, political and social situation of each Member State.

The complete process of developing the local questionnaires consisted of several distinct steps (see Figure 10). Before the translation could start, the languages into which the questionnaires (SQ, TQ and PQ) would be translated needed to be agreed upon (see section 3.2.3.1).

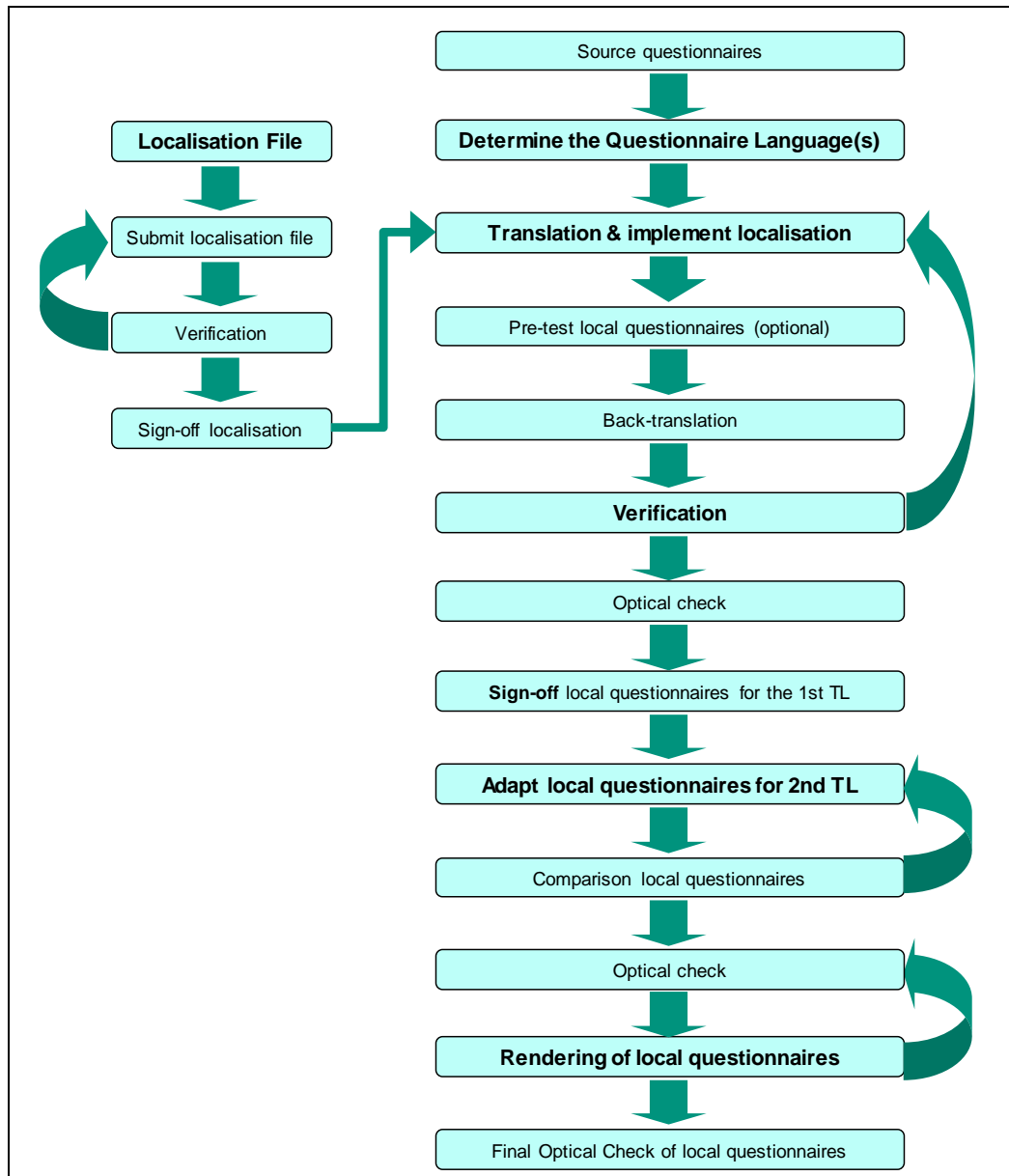
Furthermore, some terms and lists of response options needed no translation or adaptation, but needed to be replaced with a term or a list of response options that covered the concept of interest adequately in the educational system. This replacement of terms or lists of response options is called localisation, described in section 3.2.3.2 “Localisation”.

---

<sup>7</sup> The National Questionnaire did not have to be translated, but was administered to all NRCs in English.

After agreement was reached about the localisation, the actual translation (section 3.2.3.3) was started, followed by the verification of the translation (section 3.2.3.4). After finalising the translation, the second version of the questionnaires was made (section 3.2.3.5) and all local questionnaires were rendered (section 3.2.3.6).

**Figure 10 Development process of the local questionnaires**



### 3.2.3.1 Determining the questionnaire languages

To establish the questionnaire language<sup>8</sup> for each educational system a Questionnaire Language Form was sent to the Advisory Board Members of all potentially participating Member States. The form contained three lists for each potential Member State (see Figure 11). The lists of the official state language(s) and the official regional and minority language(s)<sup>9</sup> were based upon the information from Key Data on Teaching Languages at School in Europe (Eurydice 2008)

- The (official) state language(s)
- The official regional and minority language(s)
- Other language(s) that are used within the country as instructional language, for most communicative situations and/or that will be used as questionnaire language

**Figure 11 Questionnaire Language Form**

Adjudicated entity:		Please select your Adjudicated Entity					
Progress		0 %					
Language(s)	Generally used as instructional language		Used for most communicative situations		Questionnaire language		
	Answer	Comment	Answer	Comment	Answer	Comment	
<b>Official state language(s)</b>							
1							
2							
3							
4							
5							
<b>Regional or minority languages with official status</b>							
1							
2							
3							
4							
5							
6							

<sup>8</sup> "The 'Questionnaire Language' is the language that the questionnaires, sampling forms, testing tool navigation details, guidelines and manuals will be administered and available in. This language must be one of the official languages within the Member State which is used in most or most important communicative situations (for work, life in society, etc.) in the region where the school is located and that is a language of instruction in the school's region."

<sup>9</sup> "Many EU Member States use the definition of regional or minority languages contained in the "European Charter or Regional or Minority Languages" an international treaty supervised by the Council of Europe. This defines regional or minority languages as "those traditionally used by part of the population in a state, but which are not official state language dialects, migrant languages or artificially created languages" (European Commission 2008)



Advisory Board Members were requested to fill out for each of the listed languages (and each language added):

- (i) whether the language is used as instruction language in schools in the entire country, in particular regions or in particular communities in the country
- whether the language is used for common communicative situations (everyday life, shops, work, etc.) in the entire country, in particular regions or in particular communities of the country
- whether their country proposed to translate the questionnaires in those particular languages

On the basis of the responses to the Questionnaire Language Form the language(s) into which the source questionnaires had to be translated was agreed upon with SurveyLang (see chapter 5).

### 3.2.3.2 Localisation

In the questionnaires, several response options and terms occurred that had to be localised. Localisation is needed to ensure that the questions and response options adequately cover the concepts in each educational system. For example, the most widely spoken language is different within each Member State and questions or response options referring to the most widely spoken language had to be localised (in each educational system the appropriate language should be mentioned).

A Localisation File was sent to the National Research Coordinator to help with localising the lists of response options that occurred several times in the Student Questionnaire, Teacher Questionnaire and Principal Questionnaire. The information provided was part of the context information on the national level (see 3.2.4.2).

In the Localisation File, six tables had to be filled out:

- (i) In the Study programme Table the study programmes at ISCED2 and ISCED3 level<sup>10</sup> had to be listed. In many countries students can follow alternative or different study programmes at ISCED2 and/or ISCED3 level. Often these different study programmes (with a somewhat different curriculum and/or aimed at another level of ability) are offered at different types of institutions, but these study programmes can also be offered in the same school. In the countries where no administrative or structural boundary between (some) successive ISCED levels exists (e.g. between ISCED2 and 3), the grades of the study programmes that represent ISCED2 level and those that represent ISCED3 level had to be listed as separate study programmes.

---

<sup>10</sup> See the Manual for ISCED-97 Implementation in OECD Countries, 1999 Edition (OECD 1999)

- (ii) In the Language Table the most widely spoken languages in the educational system had to be listed. The five most widely spoken “indigenous” languages had to be listed. In the source questionnaires the term “indigenous” language is used to denote all the state and/or national languages, the regional and minority languages<sup>11</sup> and non-territorial languages spoken by part of the population in the educational system. The “indigenous” languages included did not need to be official languages. Furthermore, the five most widely spoken “non-indigenous” languages had to be listed. In the source questionnaires the term “non-indigenous” language is used to denote all the languages spoken by part of the population that are neither a state and/or national language, are not a regional or minority language and are not a non-territorial language. The native languages of the largest immigrant groups in the educational system had to be included. The languages had to be listed in descending order, with the most widely spoken language mentioned first and the least widely spoken language mentioned last.
- (iii) In the Taught Languages Table the ten most widely taught foreign languages and ancient languages in the educational system had to be listed. The languages that are most widely taught in primary and secondary education (ISCED1, ISCED2 and ISCED3) had to be included. If ancient languages, like Latin, ancient Greek and ancient Hebrew, can be studied in the educational system, these languages had to be included as well. The languages taught had to be listed in descending order, with the most widely taught language mentioned first and the least widely taught language mentioned last.
- (iv) In the Country Table the seven most frequent countries of origin (excluding the educational system) of immigrants living in the educational system had to be listed. The countries listed had to include the countries of origin of the largest immigrant groups in the educational system. The countries of origin had to be listed in descending order, with the most frequent occurring country of origin mentioned first and the least frequent occurring country of origin mentioned last.
- (v) In the ISCED-levels Table the different educational levels had to be listed. Most EU countries have officially classified their educational system using the ISCED classification of educational levels, see the manual of the OECD (1999).
- (vi) In the Country-specific questions sheets the NRC could indicate which country specific questions the educational system wished to include.

---

<sup>11</sup> “Many EU Member States use the definition of regional or minority languages contained in the “European Charter or Regional or Minority Languages” an international treaty supervised by the Council of Europe. This defines regional or minority languages as “those traditionally used by part of the population in a state, but which are not official state language dialects, migrant languages or artificially created languages” (European Commission 2008).

Each Localisation File was verified by SurveyLang. The internal consistency of the information provided was checked (e.g. the correspondence between the information in the Study Programme Table and the information in the ISCED Table) and the information was cross-checked with information from PISA, Eurydice and Eurostat.

As for the country-specific questions, we wanted to give educational systems as much freedom as possible in formulating the country-specific questions. However, the country-specific questions had to fit within the existing constraints of the ESLC. We checked, therefore, whether:

- (i) The question format of the country-specific questions were in a format (see section 3.2.1.1 “Question types”) already used within the rest of the questionnaires

Each question would fit on a small computer screen.

Answering the country-specific questions would not take too much time. In general, how much time it takes respondents to answer questions depends on the length of the question, the amount of information the respondent is asked to remember or reflect upon, the complexity of the judgment the respondent has to make, the number of judgments and responses a respondent has to make, and the length of the response that is asked for.

In the verification process each entry was signed-off separately. In case of queries the Localisation File was resent to the educational system with queries to be addressed by the NRC. This process was repeated until all entries in the Localisation file had been signed off.

### **3.2.3.3 Translation of the source questionnaires**

The source questionnaire was double translated into the questionnaire language(s) (see chapter 5). Even though the translation of the questionnaires had to match the source questionnaire as closely as possible, a complete literal translation was not looked for. Many terms and expressions needed a form of adaptation, see also Harkness (2008:73-74). The terms and expressions used needed to be adapted to the questionnaire language and to the cultural norms of communication and expression. Most importantly, the terms needed to be adapted such that they would be easily understood by students aged 14 to 16 in each educational system. So, in educational systems with English as a questionnaire language (see chapter 5) the translation process was in fact an adaptation process.

After the double translation the reconciler had to reconcile both translated versions and to implement the localisation agreed on. The NRCs were invited to pre-test the local version with a few students (similar to the cognitive labs) and, if necessary, to improve the translation and localisation. The reconciled version was then back translated into English (see chapter 5).



### 3.2.3.4 Verification of the translations

In the third phase the local questionnaire was verified. The back translation of the local questionnaires was checked against the source version by a verification team. The main concern in verifying the local questionnaires was comparability across educational systems; see also Harkness (2008).

During the verification of the translations attention was paid to three broad issues: comparability of the meaning conveyed, comparability of the scope of the questions, and whether the translation was consistent with the general question wording guidelines.

- (i) For comparability, the intended meaning should be conveyed accurately. For example, the phrase “For learning to write in [target language]” should not have changed into the phrase “To be able to write in [target language]”, because the last phrase conveys a slightly different meaning than the first phrase.

The scope of the questions had to remain comparable as well. After translation the generality of the terms used would have to have remained comparable and the situations, places, frequency, intensity and affective nuances referred to in the questions would have to have remained comparable. In particular the use of a plural form rather than a singular form (or vice versa) and omissions of clauses, adverbs and adjectives would be sources of changes in the scope of questions. For example, the phrase “For learning to pronounce [target language] correctly” is not comparable with the phrase “For learning to pronounce [target language]”.

Furthermore, the translated questions had to be consistent with the general question wording guidelines (see section 3.2.1.3 “Question writing”). The wording of the questions had to be easy, neutral and unambiguous. For example, an ambiguous phrase that occurred after translation was the phrase “French teacher” which could either be the teacher of the subject of French or a teacher from France.

In addition to verifying the translation (or adaptation), the implemented localisation and the lay-out were checked as well. The WebTrans system (see chapter 5) prevented any change in the order of question elements or general changes in lay-out. The only aspect NRCs had to implement was the underlining of words or phrases. It was carefully checked whether all words that had to be underlined in a question element, were also underlined in the translation.

Within WebTrans (see chapter 5) the translation of each question element had to be accepted separately. The verifier wrote a note with every question element in which the meaning conveyed or in which the scope appeared to have changed, in which the localisation was implemented differently than agreed or underline was missing. The NRC would then receive a list of question elements of which the translation was still pending and therefore, had to be attended to. The process of verification and correction continued until the translations of all question elements were accepted (in





total 788 question elements in each Student Questionnaire, 686 question elements in each Teacher Questionnaire and 542 question elements in each Principal Questionnaire).

Before the local questionnaires were signed-off an Optical Check was performed. In the Optical Check it was ascertained that no translations or localisations were accidentally omitted.

The process of localisation and translation yielded 21 different local questionnaires for each source questionnaire (Student Questionnaire, Teacher Questionnaire and Principal Questionnaire), see Table 15).

### **3.2.3.5 Creating the questionnaires for the second target language**

The local questionnaires had to be adapted to the target language. For example, when students' skills in English were to be tested, the questionnaire would ask about students' experiences with English and the school lessons in English. When students' skills in French were to be tested, the questionnaire would ask about students' experiences with French and the school lessons in French. After sign-off of the local questionnaire for the first target language, the local version for the 2<sup>nd</sup> target language had to be produced. This was a relatively easy step. For example, if the first target language was English and the second target language was German, all that needed to be done for the second version was to replace the (translated word for) English by (the translated word for) German. The NRCs received a list of all question elements in which the first target language had to be replaced with the second target language.

After the local questionnaire for the second target language was created both versions were compared highlighting all differences between the two versions. In case the two versions differed in other respects than the target language the NRC was notified to make both versions equal. Furthermore, an optical check was done checking whether any translation or localisation<sup>12</sup> was accidentally omitted.

---

<sup>12</sup> In some the localisation had to be adapted as well for the second target language version.

**Table 15 Different local questionnaires for each source questionnaire**

Localisation (for each Adjudicated Entity)			Translation into Questionnaire Language		Version 1 (First target language)		Version 2 (Second target language)	
1	BE de	Belgium - German-speaking community	de	German	FR	French	EN	English
2	BE fr	Belgium - French community	fr	French	EN	English	DE	German
3	BE nl	Belgium - Flemish community	nl	Dutch	FR	French	EN	English
4	BG	Bulgaria	bg	Bulgarian	EN	English	DE	German
5	EE	Estonia	et	Estonian	EN	English	DE	German
6			ru	Russian	EN	English	DE	German
7	EL	Greece	el	Greek	EN	English	FR	French
8	ES	Spain	es	Spanish	EN	English	FR	French
9			Spanish-Basque	Basque	EN	English	FR	French
10			Spanish-Catalan	Catalan	EN	English	FR	French
11			Spanish-Galician	Galician	EN	English	FR	French
12			Spanish-Valencian	Valencian	EN	English	FR	French
13	FR	France	fr	French	EN	English	ES	Spanish
14	HR	Croatia	hr	Croatian	EN	English	DE	German
15	MT	Malta	en	English	EN	English	IT	Italian
16	NL	Netherlands	nl	Dutch	EN	English	DE	German
17	PL	Poland	pl	Polish	EN	English	DE	German
18	PT	Portugal	pt	Portuguese	EN	English	FR	French
19	SE	Sweden	sv	Swedish	EN	English	ES	Spanish
20	SI	Slovenia	sl	Slovene	EN	English	DE	German
21	UK-ENG	England	en	English	FR	French	DE	German

### 3.2.3.6 Questionnaire rendering

After a complete sign-off, the local questionnaires were rendered for each administration mode. Depending on the educational system the administration mode of the Student Questionnaire, like the language tests, was paper-based (nine educational systems), computer-based (four educational systems<sup>13</sup>) or both (three educational systems, see chapter 6). All Teacher and Principal Questionnaires were administered through the Internet (a Web survey). A Final Optical Check was done for all 112 rendered questionnaires (see Table 16). In the Final Optical Check attention was paid to:

- (i) the occurrence of characters that should not be there, like square brackets, curly brackets ([, ], <, >, {, }), double apostrophes (e.g. student's) or double question marks
- missing question elements
- incorrect hyphenation or missing hyphenation
- incorrect question or item numbering
- incorrect lay-out (paper-based version).

---

<sup>13</sup> Those numbers are based on the Main Study. Portugal administered the Student Questionnaire in the paper-based only for the Field Trial in computer-based format only for the Main Study.

**Table 16 Rendered questionnaires**

Local questionnaires	Administration mode	
Local Student Questionnaires (two versions)	Paper-based	32
	Computer-based	16
Local Teacher Questionnaires (two versions)	Websurvey	32
Local Principal Questionnaires (two versions)	Websurvey	32
Total		112

Note: Also for educational systems who administered the Student Questionnaire completely computer-based, a paper-based version was rendered as back-up.

### 3.2.4 Evaluation of the Field Trial results

#### 3.2.4.1 Local questionnaires for students, teachers and school principals

The goal of the Field Trial was to test all local questionnaires with real respondents from all educational systems under real survey conditions. The results of the Field Trial (including the observations made by the Test Administrators and National Research Coordinators) were intended to amend the questionnaires when necessary.

After the Field Trial all data were merged. The data of the country specific questions were extracted from the database and sent to the countries for analysis. Furthermore, the students' responses to the four open-ended text questions, meant to provide information on the parental occupational status, were sent to the countries for coding (see section 7.16). After the completion of the coding the codes were sent back to SurveyLang and added to the database.

After data preparation (coding and recoding), the Field Trial data were analysed to detect items or questions that malfunctioned internationally or locally (in a particular educational system). For each educational system three reports were prepared about the items of the questionnaires: about the Student Questionnaire, the Teacher Questionnaire and the Principal Questionnaire. The purpose(s) of the reports were the following:

- (i) provide information on the responses in each educational system to each of the questionnaire items
- provide information that might help with evaluating whether the translation or localisation of particular items needed to be corrected for the Main Study.

Each report consisted of two parts, a description of the item responses in all participating educational systems and a description of the item responses in each educational system. For each item of the questionnaire the following information was provided:

- a description of the item content

- total and valid number of responses in the sample
- proportions of missing responses, a distinction being made between:
  - (i) the proportion of respondents that gave an invalid response to the item
  - the proportion of respondents that did not respond to any item of the question
  - the proportion of respondents that did respond to other items of the question, but not to this item
- proportions and frequency of valid responses to the categories.
- descriptive statistics of each item (measures of central tendency and measures of dispersion)
- flags indicating when an item behaved differently in the educational system and the reason for the flag.

Items were flagged when the proportion of missing responses were high and/or much higher than internationally (a high proportion of invalid responses, a high item non-response or a high question non-response). A high proportion of invalid responses or missing responses may indicate that several respondents did not understand the item well or that the item was not applicable for many respondents. As it is important for the quality of a survey that as many respondents (students, teachers and principals) give a valid response that can be used in the data analysis, we asked the NRC to carefully evaluate the translation and/or localisation of items that were flagged for invalid responses.

Items were also flagged, when they showed a lack of variation denoted by an extremely high or low proportion of responses to certain response categories. One of the reasons that a response category is used relatively often or little (or not at all) may be that the translation and/or localisation of the response category are not optimal. When a localised response option(s) is chosen very infrequently and/or many respondents choose the "Other" response option, this might indicate that the localisation could be improved by offering some other options. In these instances, the NRCs were recommended to analyse the open responses (written in the "Please specify"-boxes of the questionnaires).

After an NRC training about the Field Trial analyses and evaluation, NRCs received the reports and a file with the open text responses for inspection.

#### **3.2.4.2 National information**

Because the national information was to be collected through the NRCs and was, therefore, comparatively small scale, the pre-testing phase and Field Trial phase were combined. A lot of the national information we obtained through the Localisation File (see section 3.2.3.2 "Localisation"). The remaining national information we aimed to obtain through the National Questionnaire. In the National Questionnaire questions were asked about three issues:



- (i) the official regulations or recommendations regarding the foreign language curriculum, for example the age at which foreign language education is recommended to start

the official regulations or recommendations regarding the qualifications and specialisation of teachers, for example, the extent to which teachers should be specialised to teach foreign languages

some additional questions about the educational system in each educational system, for example class size norms.

For collecting the national information SurveyLang has sought a collaboration modality with the Eurydice network, at the request of the European Commission. The Eurydice Network has a long standing experience in collecting and analysing national information on the education systems and policies. It consists of 35 national units based in all 31 countries participating in the EU's Life-long Learning programme (EU Member States, EEA countries and Turkey) and is coordinated and managed by the EU Education, Audio-visual and Culture Executive Agency in Brussels, which drafts its publications and databases (for more information see Eurydice's website).

One of Eurydice's publications is the Key Data on Teaching Languages at School in Europe (the latest edition appeared in 2008; the next edition will be published in 2012). This report gives a picture of the language teaching systems in place in the schools of the 31 countries covered by the Eurydice Network.

Collaborating with Eurydice served several purposes. It allowed us to benefit from Eurydice's expertise in this field, but it also avoided doubling questionnaires and effort and ensured that the collected data would be coherent across the publications of Eurydice and the ESLC. Each National Eurydice Unit has been requested to collaborate with the NRC to ensure high quality national information. The National Eurydice Units were requested to check the Field Trial information provided through the localisation file and to check the additional information provided through the National Questionnaire during the Field Trial.

Below each question in the Field Trial National Questionnaire two fields were placed (see Figure 12). In the first Field the NRC could type an additional explanation or clarification of the situation in their educational system, as well as suggestions or comments regarding the question posed. In the second field below the questions the Eurydice National Unit could type their comments or suggestions and the agreement status.

**Figure 12 Example of a Field Trial question of the National Questionnaire**

G7. In your Adjudicated Entity, is it custom to subtitle or to dub in the following media?		
Television programs from abroad? .....	<input type="radio"/> Subtitle	<input type="radio"/> Dub <input type="radio"/> Neither
Cinema movies from abroad? .....	<input type="radio"/> Subtitle	<input type="radio"/> Dub <input type="radio"/> Neither
Additional note NRC	Suggestions of Eurydice's National Unit	Agreement status

The National Questionnaire was also sent to the Advisory Board Members for an expert review. All filled-out National Questionnaires and those containing feedback were added to a database for evaluation of the questions.

### 3.2.5 Creation of the Main Study Questionnaires

#### 3.2.5.1 Revisions of the source questionnaires

In principle, any change in the question formulation would require a new thorough question testing and translation process. For this reason the intention was to remove malfunctioning questions only. A proposal was made for the Advisory Board regarding the removal of questions or items. The proposal was based on the Field Trial results, the Field Trial expert reviews (see section 3.2.2.1), the translation comments of NRCs (see chapter 5) and the Educational system Feedback Reports from the Field Trial (see chapter 7).

In all three questionnaires, we saw that open-ended questions had a higher non-response, probably because answering those questions is cognitively demanding. In the Student Questionnaire the questions for assessing SES seemed particularly problematic. From some educational systems we received objections stating that those questions were too problematic or difficult to answer. Furthermore, the questions about home possessions showed a lack of variation across and within countries and the questions about parents' occupations showed a high level of non-response. The European Commission, on advice of the Advisory Board, decided, however, to maintain those questions in order to obtain a measure of SES comparable to the one used in PISA.

Four types of improvements were implemented in the source questionnaire:

- (i) a few malfunctioning questions (or items) were removed. Care was taken that the conceptual framework was still adequately covered
- for a few questions the open question format was changed into a closed question format. This was only possible when the range of answers was limited



the open-ended response category occurring as the last option in some questions (see section 3.2.1.1 “Question types”) was removed  
small inconsistencies in wording between similar items within and between the questionnaires were resolved and some additional notes for the Test Administrator were written

The Main Study source questionnaires can be found in Appendix 2.

### **3.2.5.2 Revisions of the localisation**

The NRCs first received the Localisation File again (see section 3.2.3.2), containing all information from the Field Trial. The NRC was requested to check all information and to correct any information if necessary. Reasons for corrections were:

- (i) during the Field Trial the Localisation File was reviewed by the Eurydice Head of Unit in each educational system. Their review might have shown the need to correct, for example, the information about ISCED-levels or Study Programmes  
some educational systems might have wanted to correct or add new country-specific questions  
the Field Trial results and the analysis of the open answers written in the “Please specify” boxes of the Student Questionnaire in the educational system might have shown the need to correct, for example, the Language Table or Country Table

The proposed corrections were verified by SurveyLang using the same procedure as prior to the Field Trial (see section 3.2.3.2). Once full agreement was reached between the NRC and SurveyLang the Localisation File was signed-off.

### **3.2.5.3 Revisions of the local questionnaires**

All NRCs received an Excel-based form for proposing corrections to the local questionnaires. Within each file, the complete source questionnaire was displayed, the corrections to the source questionnaire and the complete Field Trial local questionnaires (both versions). Next to each element of the Field Trial questionnaires the NRC could indicate the intended correction (and a back-translation). All proposed corrections went through a verification process similar to the procedure prior to Field Trial (see section 3.2.3.4 “Verifications of the translations”).

After the corrections were agreed, the NRC was asked to implement the corrections to both the translation and localisation in WebTrans. All changes were tracked and verified. Furthermore, the NRC was asked to compare the two versions of each questionnaire and perform an optical check. Once the NRC had signed-off the questionnaires, the verification team performed an additional verification, comparison and optical check similar to the one in the Field Trial. The NRC was notified when the verifier had detected any more changes to the questionnaires than agreed upon or had detected differences between the two versions.





After sign-off and agreement between the NRC and the verification-team on the local questionnaires, all questionnaires were rendered and were given a Final Optical Check like in the Field Trial (see section 3.2.3.6 "Questionnaire rendering").

#### 3.2.5.4 Revisions of the National Questionnaire

On the basis of the answers, notes and feedback received during the Field Trial, we have made a glossary containing a description of all the terms for which the intended meaning might be unclear. In some questions terms were added clarifying the intended response. For example, in the question about the number of languages students have to study we have more clearly indicated that we would like to know the minimum number of languages and the maximum number of languages.

Furthermore, in three questions (about teaching time, ancient languages and the end of compulsory education) we asked for more exact information to enable a productive comparison between educational systems. For example, in the question about the end of compulsory education we made a distinction between the end of full-time compulsory education and the end of part-time compulsory education.

For the Main Study we asked the NRCs to evaluate whether the pre-filled answers were appropriate given the reference year 2010/2011 and given the description of the terms in the glossary. Because each educational system has its unique characteristics, below each question a field was placed into which the NRCs could type an additional explanation or clarification of the situation in their educational system. The Main Study National Questionnaire can be found in Appendix 2.

### 3.3 References

- Adams, R and Wu, M (Eds) (2002) *PISA 2000 technical report*, Paris: OECD.
- Anastasi, A. (1986) Evolving concepts of test validation, *Annual Review of Psychology* 37, 1-15.
- Bonnet, G (Ed.) (1998) *The Effectiveness of the Teaching of English in the European Union: Report and Background Documents of the Colloquium Held in Paris on October 20th and 21st 1997*, Paris: Ministère de l'éducation nationale.
- Bonnet, G. (Ed.) (2002) *The Assessment of Pupil's Skills in English in Eight European Countries: An European project*, Paris: Ministère de l'éducation nationale.
- Campanelli, P (2008) Testing survey questions, in de Leeuw, E, Hox, J and Dillman, D (Eds), *International Handbook of Survey Methodology*, New York: Lawrence Erlbaum Associates, 176-200.
- Campbell, J, Kelly, D, Mullis, I, Martin, M and Sainsbury, M (2001) *Framework and specifications for PIRLS assessment 2001*, Chestnut Hill, MA: Boston College.





- Carroll, J (1963) A model for school learning, *Teachers College Record* 64 (8), 723-733.
- Cenoz, J, Hufeisen, B and Jessner, U (2001). *Cross-linguistic Influence in Third Language Acquisition*, Multilingual Matters Ltd.
- Commission of the European Communities (2003) Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions - Promoting Language Learning and Linguistic Diversity: an Action Plan 2004 - 2006. COM(2003) 449 final. Brussels.
- Commission of the European Communities (2007b) Communication from the Commission - A coherent framework of indicators and benchmarks for monitoring progress towards the Lisbon objectives in education and training. COM(2007) 61 final. Brussels.
- Commission of the European Communities (2007c) Report on the implementation of the Action Plan 'Promoting language learning and linguistic diversity'. COM(2007) 554 final/2. Brussels.
- Commission of the European Communities. (2008). Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions - Multilingualism: an asset for Europe and a shared commitment.COM(2008) 566 final. Brussels.
- Communication from the Commission to the Council. (2007). Framework for the European survey on language competences. COM(2007) 184 final. Brussels.
- Communication from the Commission to the European Parliament and the Council. (2005). Communication from the Commission to the European Parliament and the Council: The European Indicator of Language Competence. COM(2005) 356 final. 5. Brussels.
- Communication of the European Communities. (2008b). Green Paper- Migration & Mobility: challenges and opportunities for EU education systems. COM(2008) 423 final. Brussels.
- Council of Europe (2005) *Survey on the use of the Common European Framework of Reference for Languages: Synthesis of Results*, Strasbourg: Language Policy Division.
- Council of Europe. (2006) Council conclusions on the European Indicator of Language Competence.2006/C 172/01.



- Council of Europe (2008a) *European Language Portfolio*, retrieved from [http://www.coe.int/T/DG4/Portfolio/?L=E&M=/main\\_pages/introduction.html](http://www.coe.int/T/DG4/Portfolio/?L=E&M=/main_pages/introduction.html)
- Council of Europe (2008b) *Common European Framework of Reference for Languages*, retrieved from [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf)
- Council of Europe (2007) *From linguistic diversity to plurilingual education: Guide for the development of language education policies in Europe*. Strasbourg.: Language Policy Division.
- Council of the Europe (2008) Council conclusions of 22 May 2008 on multilingualism. *Official Journal of the European Union*, C 140 , 06/06/2008, 14-15.
- Council of the European Union. (2002, 2 23). Council resolution of 14 February 2002: On the promotion of linguistic diversity and language learning in the framework of the implementation of the objectives of the European Year of Languages 2001. *Official Journal of the European Communities*(C 50).
- Council of the European Union. (2002b, March 15-16). Barcelona European Council 15 and 16 March 2002: Presidency conclusions. Barcelona.
- Council of the European Union. (2006, 7 20). Resolution of the Council and of the Representatives of the Governments of the Member States, meeting within the Council: On the recognition of the value of non-formal and informal learning within the European youth field. *Official Journal of the European Union*(C 168).
- De Leeuw, E (2008) Choosing the method of data collection, in de Leeuw, E, Hox, J and Dillman, D (Eds), *International handbook of survey methodology*, New York: Lawrence Erlbaum Associates, 113-135.
- Dillman, D A (2000) *Mail and Internet Surveys: The Tailored Design Method (2nd ed.)*, New York: Wiley.
- Dillman, D A (2008). The logic and psychology of constructing questionnaires, in de Leeuw, E, Hox, J and Dillman, D (Eds), *International handbook of survey methodology*, New York: Lawrence Erlbaum Associates, 161-175.
- Eurobarometer (2006). *Special Eurobarometer: Europeans and their Languages*. EB 64.3.
- European Commission. (2008). *EU Language Policy: Policy documents*. Retrieved from European Commission: [http://ec.europa.eu/education/languages/eu-language-policy/doc124\\_en.htm](http://ec.europa.eu/education/languages/eu-language-policy/doc124_en.htm)



- European Commission. (2008). *Language Teaching: In the spotlight*. Retrieved from [http://ec.europa.eu/education/languages/language-teaching/doc24\\_en.htm](http://ec.europa.eu/education/languages/language-teaching/doc24_en.htm)
- European Commission. (2008). *Languages: Facts*. Retrieved from European Commission: [http://ec.europa.eu/languages/languages-of-europe/facts\\_en.htm](http://ec.europa.eu/languages/languages-of-europe/facts_en.htm)
- European Parliament and the Council. (2006). Recommendations of the European Parliament and of the Council of 18 December 2006 on key competences for life-long learning. *Official Journal of the European Union*, L 394/10 of 30.12.2006.
- Eurostat. (2008). *Eurostat*. Retrieved from [http://epp.eurostat.ec.europa.eu/portal/page?\\_pageid=1090,30070682,1090\\_33076576&\\_dad=portal&\\_schema=PORTAL](http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33076576&_dad=portal&_schema=PORTAL)
- Eurydice. (2004). *Integrating immigrant children into schools in Europe*.
- Eurydice. (2005). *Key Data on Education in Europe 2005*.
- Eurydice. (2008). *Key Data on Teaching Languages at School in Europe - 2008 Edition*.
- Eurydice. (n.d.). *Eurypedia - The European Encyclopedia on National Education Systems*. Retrieved from Eurydice: [http://eacea.ec.europa.eu/education/eurydice/eurybase\\_en.php](http://eacea.ec.europa.eu/education/eurydice/eurybase_en.php)
- Fowler, F and Cosenza, C (2008) Writing effective questions, in de Leeuw, E, Hox, J and Dillman, D (Eds), *International handbook of survey methodology*, New York: Lawrence Erlbaum Associates, 136-160.
- Harkness, J A (2008) Comparative survey research: goals and challenges, in de Leeuw, E, Hox, J and Dillman, D (Eds), *International handbook of survey methodology*, New York: Lawrence Erlbaum Associates, 56-77.
- Heuvelmans, A M (2006). *Constructie en verwerking van vragenlijsten. POK Memorandum*. Arnhem: CITO.
- High Level Group on Multilingualism (2007). *Final report*, Luxembourg: Commission of the European Communities: Office for Official Publications of the European Communities.
- Kane, M T (1992) An argument based approach to validity, *Psychological Bulletin* 112, 527-535.
- Kelly, M, Grenfell, M, Allan, R, Kriza, C and McEvoy, W (2004) *European Profile for Language Teacher Education – A Frame of Reference: Final Report. A Report to the European Commission Directorate General for Education and Culture*.



- Kuhlemeier, H (2007a) *Development of the Field Trial Questionnaires for PISA 2009 [EDU/PISA/GB(2007)32]*, Paris: OECD.
- Kuhlemeier, H (2007b) *Proposal for the international questionnaire options for PISA 2009 [EDU/PISA/GB(2007)40]*, Paris: OECD.
- LACE (2007) *The Intercultural Competences Developed in Compulsory Foreign Languages Education in the European Union*.
- Levine, R, Huberman, M and Buckner, K (2002). *The measurement of instructional background indicators: Cognitive laboratory investigations of the responses of fourth and eighth grade students and teachers to questionnaire items*, Washington, DC: U.S. Department of Education.
- Lewis, E and Massad, C (1975) *The Teaching of English as a Foreign Language in Ten Countries*, New York: Wiley.
- Messick, S. (1995) Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry to score meaning, *American Psychologist* 50, 741-749.
- Mullis, I, Kennedy, A, Martin, M and Sainsbury, M (2004) *PIRLS 2006 assessment framework and specifications*, Chestnut Hill, MA: Lynch School of Education Boston College.
- Mullis, I, Martin, M, Ruddock, G, O'Sullivan, Y, Arora, A and Erberber, E (2005) *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I, Martin, M, Smith, T, Garden, R, Gregory, K, Gonzalez, E, Chrostowski, S and O'Connor, K (2003) *TIMSS assessment frameworks and specifications 2003*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education Boston College.
- OECD . (2005). *PISA 2003 technical report*. Paris: OECD.
- OECD. (1999). *Classifying Educational Programmes — Manual for ISCED-97 Implementation in OECD Countries*, 1999 Edition. Paris: OECD.
- OECD. (2007). *PISA 2006 Technical report*. Paris: OECD.
- OECD. (2008). Retrieved from Programme for International Student Assessment (PISA):  
[http://www.oecd.org/pages/0,3417,en\\_32252351\\_32235731\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/pages/0,3417,en_32252351_32235731_1_1_1_1_1,00.html)



- OECD. (2008). *OECD Teaching and Learning International Survey (TALIS) Home*. Retrieved from OECD: [http://www.oecd.org/document/0/0,3746,en\\_2649\\_39263231\\_38052160\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/0/0,3746,en_2649_39263231_38052160_1_1_1_1,00.html)
- Paulsen, C, and Levine, R (1999). *The applicability of the cognitive laboratory method to the development of achievement test items*, paper presented at the annual meeting of the American Educational Research Association, Montreal. Montreal.
- Schouwstra, S (2000) *On testing plausible threats to construct validity*, doctoral dissertation, the Netherlands: University of Amsterdam.
- TIMSS & PIRLS International Study Center (2008), retrieved from TIMSS & PIRLS International Study Center: <http://timss.bc.edu/>
- Tourangeau, R (1984) Cognitive sciences and survey methods, in Jabine, T, Straf, M, Tanur, J and Tourangeau, R (Eds), *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*, Washington, D.C.: National Academy Press, 73-100.
- Tourangeau, R, Rips, L and Rasinski, K (2000) *The Psychology of Survey Responses*, Cambridge: Cambridge University Press.
- Willms, J (2006) *Learning Divides: Ten Policy Questions About The Performance and Equity of Schools and Schooling Systems*, Montreal: Unesco.

# ***Chapter 4:***

# ***Operations -***

# ***Sampling***

## 4 Operations - Sampling

This chapter describes the sampling procedures and implementation for the ESLC.

### **4.1 *Target population and overview of the sampling design***

The purpose of the ESLC is to measure the foreign language ability of students in participating educational systems. The languages tested in the survey were those among the official languages of the European Union that were the first and second most commonly taught foreign languages in the participating educational systems. Based on the latest available Eurostat data at the time of selection of the languages, the following five languages were chosen: English, French, German, Italian and Spanish. The choice of the first and second most taught languages was made by each participating educational system on the basis of the latest available documented data from Eurostat. For each of the five selected languages, the survey included testing of three skills: (i) Reading comprehension, (ii) Listening comprehension, and (iii) Writing.

The target population for each language in an educational system consisted of students enrolled in ISCED2 level (final year) or after the first completed year of ISCED3 level. Hence, the international target population corresponds to the total number of students in eligible grades (ISCED2 or ISCED3) that were 1) attending educational institutions located within the educational system and 2) studying the language to be tested for a minimum period of one academic year prior to testing.

As mentioned above, there were two possible test populations in the ESLC: those at the end of lower secondary education (ISCED2) and those at the second year of upper secondary education (ISCED3). As a result, there was variation in the age and duration at which students have been learning a foreign language in the survey. Also, some of the students in certain educational systems learn foreign languages voluntarily and for some educational systems it is obligatory for the students to learn foreign languages. Hence, to ensure comparability of results across educational systems, the primary testing grade ESLC aimed for in each educational system was the last grade in the ISCED2 education for both languages. Participating educational systems were strongly encouraged to aim for this level, and ESLC standards allowed exceptions only in special situations where the use of ISCED3 level (instead of ISCED2 level) could be really justified. This was allowed, for example, in situations where the language was not taught at ISCED2 level or the number of eligible students who were taught the language at ISCED2 level was insufficient for generating estimates of acceptable precision.

It was clear for SurveyLang that, given the wide variation in how ISCED2 and ISCED3 levels are defined and operated in different educational systems, no single approach would be effective in producing the desired result in all educational systems. It was acknowledged that there might be situations requiring special treatment. While the primary goal was to select ISCED2 level of education, SurveyLang's plan was to



review, with the help of the NRCs, the school system/actual structure in terms of the ISCED2 and ISCED3 levels as well as the typical age/grade at which the first and second foreign languages were introduced within each level. After completion of that review, SurveyLang came up with appropriate rules to address issues that were unique to specific educational systems. Table 17 presents, for each participating educational system, the list of testing grades for both languages. These levels were formally agreed with the each participating educational system and the European Commission. It can be seen that ISCED3 level was used for testing Language 1 in three out of the sixteen participating educational systems while for language 2, ISCED3 level was used for five educational systems.

A two-stage stratified sample design was used for the ESLC. For the purpose of testing students in the first and second foreign language in each participating educational system, two separate independent samples were chosen in each educational system: one sample for the first foreign language and one for the second foreign language. The two samples, therefore, could overlap with common schools and students within a school sampled for both languages could be eligible for student sampling for both languages. However, no pupil was sampled (and therefore tested) in both foreign languages.

The sample was designed to satisfy all the general and technical requirements for testing of this kind. The design was consistent with international scientific standards of sampling methods for such a survey (for example, PISA and TIMSS). Following the two-stage sample design, schools were sampled, at the first stage, using a stratified sample design. Within each stratum, schools were selected using PPS (probability proportional to size) method of selection where the measure of size was a function of the number of eligible students enrolled for the language to be tested (due to significant primary data collection need and the limited time available in the testing year, the figure from the previous academic year was used for this purpose). The second stage sampling units were students within sampled schools.

Once schools were selected to be in the sample, a list of eligible students was prepared. Depending on whether the school was sampled for one or both languages and whether the students learned the two languages in the same or different grades within the sampled school, several scenarios could occur. From schools that were selected for one language only (or selected for both languages with no overlapping students since the two languages were taught in different grades), the goal was to sample 25 students (a figure that varied somewhat according to the availability of eligible students locally, as described in later sections), with equal probability (using simple random sampling). If the total number of eligible students fell below 25, all students were selected in the student sample from such schools. For schools that were selected for both languages and had students eligible for both languages, the students within those schools were stratified based on whether they were eligible for the first language (language 1) only, the second language (language 2) only or both. Sample allocations across these three strata were determined by taking into consideration the number of students in each stratum and the overall sample size requirements for each language. At the final stage, student samples of appropriate



sizes were drawn with equal probability from each stratum independently. The goal of sampling for any language was to select 25 students with representation from students learning that language only and also students learning both languages. Once the student sample was selected for any language, each student was randomly assigned for testing of two of the three skills: (i) Reading comprehension, (ii) Listening comprehension, and (iii) Writing. Additional details of student sampling are given in section on student sampling (section 4.16).

## ***4.2 Population coverage and school and student participation rate standards***

In order to generate valid survey-based estimates of student ability, it was important to employ a sample design that would produce a representative sample of the target population. As outlined in the previous section, the sample design was developed based on international scientific standards for sampling methods for such surveys. Quality standards were developed and maintained with respect to (i) the coverage of the international target population, (ii) accuracy and precision, and (iii) the school and student response rates.

## ***4.3 Coverage of the international target population***

In order to ensure data quality, highest priority was given to the task of minimising the coverage error, i.e. for minimising the difference between the national desired target population and the international desired target population. SurveyLang made all possible efforts to limit exclusions from the national target population.

The excluded population refers to all schools and students from the national target population that are not incorporated into the sampling frame. The reasons that were generally put forward for excluding part of the school and student population are usually of a practical nature, for instance higher survey costs, challenging test setup or possibly due to other political and/or operational reasons. These exclusions could take place at two levels: (i) at school level, i.e. entire schools are left out of the sample design, or (ii) within schools, i.e. some students within sampled schools are not included in the sample.

SurveyLang's goal, as previously mentioned, was to limit exclusions from the national target population as far as possible. The challenges related to small schools, (i.e. few eligible students enrolled), or other schools where it is logistically challenging or costly to conduct assessments, were addressed whenever possible by other modifications to reduce the number of such schools in the sample rather than by excluding them altogether. The same strategy was used with special education students and students with limited proficiency in the agreed questionnaire language(s) (see section 3 for a definition of this term), thereby limiting their exclusion to a minimum.

School level exclusions mostly involved the following situations:

- the size of the school is particularly small
- special need schools
- physical access to the school is difficult.

Special need schools included students with special education needs and those that provided instruction only to students in the excluded categories such as schools for the blind.

In general, schools containing fewer than 10 eligible students were considered “extremely small” and were possible candidates for exclusion. However, the total exclusion (exclusions due to other reasons plus exclusions from ‘extremely small schools’) was not allowed to exceed 2 percent of the total population of all enrolled eligible students across all schools. If necessary, the definition of ‘extremely small schools’ was modified i.e. the cut-off value of enrolment was lowered (from 10) to keep the total exclusions to less than 2 percent. However, all schools containing fewer than 6 eligible students were considered ‘extremely small’ and were systematically excluded.

All special situations were reviewed on a case by case basis to minimise exclusions and thereby to avoid sample bias. SurveyLang received detailed information from the NRCs of all cases of school-level exclusions and of their rationale. Minimising school level exclusion was the central element in the quality strategy and one of the quality indicators suggested by SurveyLang.

Besides the school level exclusions, student level exclusions constituted the basis for another quality indicator of the national survey samples. It was foreseen that definitions of within-school exclusions would be different from one educational system to another and that is why SurveyLang requested NRCs to adapt specific rules so that they could be applied in their respective educational systems.

Within-school exclusion rules applied to the following groups:

- Functionally disabled students – students suffering from a permanent disability that prevented them from taking part in the ESLC test. The exclusion did not apply to functionally disabled students who actually had the physical ability to participate.
- Intellectually disabled students – this intellectual disability should have been previously diagnosed by professionals such as the school principal, qualified staff members or psychologists. Students who were emotionally or mentally not capable to follow even the general instructions of the test were included in this group. However, students who did not do well academically or had standard discipline problems did not fall under this category. Severely dyslexic children were excluded in countries where it was a legal requirement to exempt such children from written tests in general (such exclusion was used in very small numbers in France, Greece, Poland and Portugal).
- Students with insufficient command of the questionnaire language of the educational system.

- Any other reason of within-school exclusion had to be documented in detail on the sampling form.

NRCs were requested to provide a list of all eligible students within the sampled schools, that is, a list of all the students in the target grades who were learning the respective target language (or languages if the school was sampled for both languages).

Students whom the NRCs considered for exclusion from the sample were retained, and a variable maintained to briefly outline the reason for exclusion. By proceeding this way, SurveyLang was able to assess the extent of the within-school exclusions from the sample data.

It is important to stress the difference between within-school exclusions and non-response. Exclusions are about the incapacity to take part in the test mainly due to a permanent functional or intellectual condition. Non-response is about a temporary condition or circumstance at the time of testing that prevents the student from taking the test. The objective was to limit the overall school-level and within-school exclusions to at most 5 percent of the national target population.

#### **4.4 Accuracy and precision**

In the school sample, a minimum of 71 schools for each of the two designated languages were selected in most of the participating educational systems. Within each participating school sampled for one language (or for schools sampled for both languages but with no students eligible for both languages), 25 students were selected (on average) with equal probability per language if there were at least 25 eligible students available. If the total number of eligible students for such schools was less than 25, then all available eligible students were included in the sample with certainty. For other schools (sampled for both languages and containing students eligible for both languages), the goal was to sample 25 students (or the maximum number available) per language. In addition, representatives (teachers, principals) of each sampled school also supplied information about the school itself and the practices implemented.

Following the rules outlined above, roughly 1775 ( $71 \times 25$ ) students were sampled, in general, for each language. This was the standard sample size requirement at the national level for any educational system to participate in the ESLC. Based on an overall response rate of 85%, about 1500 students per educational system per language were expected to be tested. The precision or accuracy of an estimate depends on the effective sample size which in turn depends on the underlying design effect. For each language, the measurement model implied that there were 3 tests (Reading, Listening, Writing), with one student taking only two of them. Therefore, for any single test, an average sample size (or cluster size) of 14 ( $=25 \times (2/3) \times 0.85$ ) per school was expected to be achieved. Given this cluster size, and anticipating an intra-class correlation coefficient of 0.1, the design effect could be roughly estimated to be

about  $(1 + 13 \cdot 0.1) = 2.3$ . This was an approximation and was obviously expected to vary depending on the exact value of intra-class correlation coefficient in specific educational systems and estimates. However, based on this simplifying assumption, the effective sample size corresponding to 1000 completed cases was expected to be around 437 ( $1000/2.3=437$ ). This was expected to result, at the educational system level, in a minimum precision (or maximum sampling error) of +4.7 percent for estimation of an unknown population proportion. The precision associated with any estimator for any other subgroup (region, demographic groups etc.) was of course dependent on the corresponding sample size and also on the nature of the estimator. Stratification was employed in the sample design with the goal to further reduce the variance of the survey-based estimators.

#### **4.5 Response rates**

As in the case of similar international education surveys like PISA, SurveyLang did set, in each participating educational system, the eligibility bar for response rates both at the school and student level. In terms of data quality standards, it was important to determine minimum participation rates for schools as well as for students. The purpose of these standards is to limit the risk of response bias. For both schools and students, there was one participation rate for each tested language in each participating educational system.

In the ESLC, we set the bar at a minimum participation rate of 85% of originally sampled schools. We accept in principle that sampled schools choosing to opt out of the test be substituted with “replacement schools” (from the same explicit stratum) to meet sample size and response rate requirements. The educational systems were expected to maximise the number of responding schools by (i) ensuring maximum co-operation from the originally sampled schools, and then (ii) gaining co-operation from replacement schools in case the originally sampled school did not respond.

Along the same lines, the bar for students was set at a minimum participation rate of 80% within participating schools (sampled and replacement). It was acknowledged that follow-up sessions might be necessary in some schools where too few students took part in the tests originally conducted. It was left to the School Coordinators and Test Administrators to decide together with the NRCs whether additional sessions were needed. The recommendation was that a follow-up administration had to be held if 15% or more of the sampled students (from all students on the student tracking form excluding exclusions) were absent at the original test administration. For example, in cases where there were 25 students sampled, this means that if more than 4 students were missing, a follow-up administration should have been held.

National student participation rates consisted of an average of student participation rates in all participating schools, be they originally sampled or replacement schools, and in all sessions, whether originally scheduled or additional. The goal was to reach the target student participation rate that was set at national level, but not necessarily at the school level.

The ESLC school personnel sample was self-selecting – each of the participating school's principals and language teachers teaching the test language at the testing level were invited to fill in the School and Teacher Questionnaire, respectively. Where a school was selected for both test languages, the school principal was randomly allocated to complete the School Questionnaire for one test language only rather than having to complete the two questionnaires, one for each test language. Similarly, for teachers teaching both test languages at the eligible level in a school selected for both test languages, the teacher was randomly allocated to complete the teacher questionnaire for one test language only.

There was no official participation criterion for the teachers and principals. Educational system samples were eligible to be included in the international sample, even if the response rate for questionnaires among teachers remained low. However response was monitored by the NRC and SurveyLang and all possible efforts were made by the NRC to obtain as high a response rate as possible.

#### **4.6 *Establishing the national target population***

It was every NRC's role to define and describe the educational system's target population. The national target population definition addressed the requirements of the international target population outlined above. The goal for defining the national target population was to provide as exhaustive a national coverage of eligible students as possible. Any difficulties in accomplishing that goal were specified, documented and approved in advance.

NRCs were strongly encouraged to provide complete national coverage in their national target population. In fact, according to the data submitted, NRCs did not exclude specific regions on the basis of problematic access in any of the educational systems covered. Hence, the national target population matched the international target population in each entity surveyed in terms of geographic coverage.

#### **4.7 *Sampling implementation – test languages***

One of the important early objectives was to clarify the two most commonly taught languages among those eligible for testing in the ESLC in each participating educational system. According to the reports of NRCs and on the basis of Eurostat data and the Eurydice report, the following languages were eligible for testing in each educational system. The below table provides the two most commonly taught languages for each participating educational system.

**Table 17 List of tested languages in each participating educational system**

Educational system	Most commonly taught (eligible) foreign language ('first' language)	Second most commonly taught (eligible) foreign language ('second' language)
Flemish Community of Belgium <sup>14</sup>	French	English
French Community of Belgium	English	German
German Community of Belgium	French	English
Bulgaria	English	German
Croatia	English	German
England	French	German
Estonia	English	German
France	English	Spanish
Greece	English	French
Malta	English	Italian
Netherlands	English	German
Poland	English	German
Portugal	English	French
Slovenia	English	German
Spain	English	French
Sweden	English	Spanish

#### **4.8 Testing grades**

After identification of the languages to be tested, the next important step in defining the national target population was to determine the appropriate test population (ISCED2 or ISCED3) for each participating educational system and for each language to be tested. As mentioned before, there were two possible test populations in the ESLC: those at the end of lower secondary education (ISCED2) and those at the second year of upper secondary education (ISCED3). Hence the age and time during which students have been learning a foreign language could be different for the students tested in the survey. Also some of the students in certain educational systems learn foreign languages voluntarily and for some educational systems it is obligatory for the students to learn foreign languages. The above mentioned rule (for choosing ISCED2 or ISCED3) was dealt with at national level. A number of simplifications were used in the ESLC to ensure that the determination of the appropriate level was clear and easy to

<sup>14</sup> The ESLC was carried out independently in the three constituent regions of Belgium



execute for participating educational systems. One important simplifying assumption was that the language most frequently used nationally tended to be the first foreign language taught for students, and second most frequently used foreign language was generally taken up as a second language by the students in a particular educational system.

The process of identification of the appropriate level (ISCED2 or ISCED3) considered a single parameter, which determined the strategy applied for a particular educational system: the compulsory introduction age of the two foreign languages (in general education). The first language was, in almost all educational systems, introduced at an age that made all ISCED2 final year students eligible for taking the test (having completed at least one academic year of training in the particular language prior to testing). There was, on the other hand, a huge variation as to when the second language was introduced, and consequently, at what level the second language could be tested.

The table below summarises possible scenarios and indicates the levels at which the first and second tested language proficiency were to be assessed.



**Table 18 Testing Grade Allocation Scheme (considering the whole territory of the educational systems<sup>15</sup>)**

	ISCED2 level testing		ISCED3 level testing	
	Language 1 <sup>16</sup>	Language 2 <sup>17</sup>	Language 1	Language 2
i) Educational systems where both languages were introduced in “due time”, e.g. at or before the penultimate year of ISCED2 education	X	X	N/A	N/A
ii) Educational systems where the first language was introduced in “due time” but not the second language	X	N/A	N/A	X
iii) Educational systems where there was NO requirement for a second foreign language	X	X	N/A	N/A
iv) Educational systems where none of the two languages was introduced in “due time”	N/A	N/A	X	X

As shown above, educational systems were classified by the single parameter of the compulsory introduction age for the two languages, with typically 13 years of age as the threshold. This was, however, just a general approximation of the introduction grade (that had to be verified for each educational system with the help of NRCs), as the typical grade preceding the final grade in ISCED2 education.

Another important issue to note was that a language might be taught at ISCED2 level but it might be just for a very short period for some students who would be eligible for sampling. The data obtained by testing students with very limited exposure to the language did not provide useful results – not even reaching a minimal proficiency that

<sup>15</sup> In the French Community of Belgium, most students learn a first modern language (Dutch, English or German) from the fifth grade up, which would make ISCED2 level testing possible. However, according to a linguistic law, some specific areas are subjected to different rules. In the “Région de Bruxelles-Capitale”, notably, the modern language courses begin earlier (3rd grade) and the first language taught must be Dutch: thus in this area, neither German nor English can be taught before the 9th grade. Hence, the testing grade was shifted to ISCED3 in the French Community of Belgium.

<sup>16</sup> Most commonly taught (eligible) foreign language (‘first’ language)

<sup>17</sup> Second most commonly taught (eligible) foreign language (‘second’ language)



could be tested. The approach described above helped to comply with the eligibility criteria for students of at least one full school year of tuition before testing.

The four possible classes of educational systems and the matching approach were as follows:

- (i) Educational systems where both languages were introduced in “due time” and this was the most frequent scenario: here all testing was carried out in ISCED2 level, for both languages, among students who received respective language training for at least one full year/grade prior to testing.

Educational systems where the second language was not introduced in “due time”: here all testing for the nationally first language (Language 1) will be carried out in ISCED2 level, among students who received respective language training for at least one full year/grade prior to testing. In the sampling frame for Language 1, the ISCED3 level education was not involved. For Language 2, the situation was the opposite, all testing was carried out at ISCED3 level, and the sampling frame for Language 2 did not include ISCED2 level education.

Educational systems where there was NO requirement for a second foreign language: the rule for type (i) applied, as students might take the first and second most commonly taught language as their only language, but with a relative early start. However, if one (or both) of the two most commonly taught languages was not introduced in “due time” in any of these educational systems, the target population for that language(s) within that educational system consisted of eligible students at the ISCED3 level.

Educational systems where none of the two languages was introduced in “due time”: here all testing for the two most commonly taught languages was carried out at ISCED3 level.

It was clear for SurveyLang that, given the wide variation in how ISCED2 and ISCED3 levels are defined and operated in different educational systems, no single approach would be effective in producing the desired result in all educational systems. The proposed approach described above was, in general, effective and applicable for the participating educational systems. It was, however, understood that there might be other situations requiring special treatment.

In general, participating educational systems were strongly encouraged to aim for ISCED2 level, and ESLC standards allowed exceptions in situations where really justified. However, the number of students who satisfied the eligibility criterion of learning the test language for at least one full academic year prior to testing could pose logistical challenges in achieving this goal. If only a small proportion of the students in the last grade of ISCED2 education were eligible to be tested for a particular language, the number of schools needed to achieve the desired sample size could be significantly higher than the number that was set by ESLC standard (71, with a provision that it might be increased to some extent to make up for the missing sample size that results from inclusion of schools that are generally small or have

fewer eligible students than the standard sample size per school). In order to make testing feasible in those educational systems where resources were not sufficient or the eligible student population was too small to test students in the last grade of ISCED2 education, ESLC allowed testing at ISCED3 level in those special situations.

With the help of the NRCs, the school system/actual structure in terms of the ISCED2 and ISCED3 levels as well as typical age/grade of introducing the first and second foreign languages within each educational system was reviewed, and SurveyLang came up with specific rules to address issues that were unique to specific educational systems. Table 19 below presents the testing grades for each participating educational system for both languages. These levels were formally agreed and approved by the European Commission with a warning of the impact on comparability of data.

**Table 19 Testing Grades for participating educational systems for both languages**

Educational system	Most commonly taught (eligible) foreign language ('first' language)	Second most commonly taught (eligible) foreign language ('second' language)
Flemish Community of Belgium	ISCED2	ISCED3
French Community of Belgium	ISCED3	ISCED3
German Community of Belgium	ISCED2	ISCED3
Bulgaria	ISCED3	ISCED3
Croatia	ISCED2	ISCED2
England	ISCED3	ISCED3
Estonia	ISCED2	ISCED2
France	ISCED2	ISCED2
Greece	ISCED2	ISCED2
Malta	ISCED2	ISCED2
Netherlands	ISCED2	ISCED2
Poland	ISCED2	ISCED2
Portugal	ISCED2	ISCED2
Slovenia	ISCED2	ISCED2
Spain	ISCED2	ISCED2
Sweden	ISCED2	ISCED2

#### **4.9 School sampling frame**

On the basis of their national target population framework, NRCs constructed their school sampling frame. As in the case of national target population versus international

target population, discrepancies between the national target populations and the corresponding frames were also related to practical issues. As in the previous cases, SurveyLang did endeavour to limit these discrepancies to a minimum.

It was the responsibility of the NRCs to generate the school sampling frame based on the approved definition of their national target population. This was to be done for both languages and the same school could have enrolment in one or both languages depending on whether one or both languages to be tested were taught in that particular school. The importance of the quality of the sampling frame in terms of its impact on sampling, weighting, estimation and hence on the final survey results was emphasised. NRCs were therefore advised to generate the sampling frame with utmost care and to make sure that the frames were free of any incorrect or duplicate entries and included all schools that were part of the national target population. It may be noted that the school frames did include schools that were marked for exclusion together with the reasons for exclusion.

The most important information to be included in the school frames was ENR (i.e. the enrolment or the number of eligible students learning the language in the selected level – ISCED2 or ISCED3) for that school. At the time of frame construction (for use in school sampling), however, the exact information on enrolment was not available, as the construction of the sampling frame often required primary data collection from the schools, which was logistically only possible a year prior to the ESLC administration. That required using alternative methods of coming up with the best available estimates of enrolment. For the ESLC, NRCs provided enrolment estimates for several categories of students in the sampling forms (see section 4.18 on Sampling Forms for further details) including the following: (i) students eligible (those who had prior instruction in the language for at least one year) for testing in the eligible grade, (ii) students (number of all students) enrolled in the grade below the eligible grade, and (iii) students learning the language in the grade below the eligible grade.

For the purpose of estimating the ENR to be used for school sampling, the number of students learning the language in the grade below the eligible grade (as reported in (iii) above) was used as the best estimate. It should be noted that those students were expected to be eligible for testing in the next academic year during which the data collection was planned. All educational systems could provide these estimates and so it was possible to carry out the PPS sampling scheme for schools by deriving the 'size' of schools based on these enrolment estimates. Besides the enrolment numbers, NRCs also supplied other useful information at the school level including (i) school identification information (national school id) and name/address of the school, (ii) educational level (ISCED2 or ISCED3), and (iii) information on exclusions along with reasons and (iv) information on suggested explicit and implicit stratification variables.

#### **4.10 Stratification**

Before the beginning of the actual sampling exercise, schools were stratified in the sampling frame. Stratification is about dividing schools up into homogenous groups (of

schools) according to relevant variables, called the stratification variables. Use of stratification in the ESLC had several advantages including (i) maximising the efficiency of the sample design, and thereby improving the reliability of survey estimates, (ii) using different sample designs such as using disproportional sample allocation across different groups (strata) of schools at stake (see the next paragraph for examples), (iii) ensuring adequate (or minimum) representation of schools from different school groups and guaranteeing that all population segments are incorporated in the sample, and (iv) obtaining reliable estimates for specific strata if necessary.

Several stratification variables were used in the ESLC. Examples of stratification variables used include, but were not limited to the following:

- regions (for example, states/provinces)
- school size
- school types (for example, public/ private)
- school programmes (for example, academic/vocational)
- urbanisation (rural areas, urban areas)
- socio-economic status (for example, low/ medium/high income).

Two types of stratification variables (explicit and implicit) were used. Explicit stratification implies constructing sampling frames based on the explicit stratification variables identified. Using stratification, it is possible to employ different sample designs (for example, disproportional sample allocation) across different explicit strata. It is possible to sample the same number of schools from each explicit stratum, irrespective of the relative size of each stratum. In that case, the idea would be to produce equally reliable estimates for each stratum. In a proportional allocation however, large strata would cover more sampled schools than small strata. The challenge with a proportional allocation is that the sample size can be often too small in small strata to generate reliable estimates. For the ESLC, the major advantage of stratification was to have the flexibility to implement disproportional allocation of the sample across explicit strata whenever it was found necessary to ensure adequate representation of certain types of schools (size, public/private etc.) or geographic regions.

Implicit stratification involves sorting the schools within each explicit stratum by a set of implicit stratification variables before randomly sampling them with a specified sampling interval. Implicit stratification is, therefore, essentially about categorising the school sampling frame via a set of implicit stratification variables. It is within the explicit strata that this categorisation takes place. It basically offers a very simple and effective way of guaranteeing a strictly proportional sample allocation of schools across all implicit strata. Another advantage is that it is likely to increase the reliability of survey estimates, as long as the implicit stratification variables considered are correlated with ESLC ability at school level. Implicit stratification, therefore, uses proportionately allocated classes to ensure systematic coverage in various relevant aspects.

Some general guidelines were followed when selecting stratification variables:

- every school on the frame needed to be coupled with a potential stratification variable
- it is essential that each school in the sampling frame be only allocated to one level of each stratification variable
- the link between the stratification variables and the variables of interest to be measured in the survey, e.g. education performance, should be plausible
- the size of the explicit strata, namely both the number of schools and the number of eligible students for each stratum, should be known
- defining very small strata, especially explicit strata, was avoided to the extent possible
- the goal was to select at least two schools from each explicit stratum to be able to compute the sampling error of estimates. In general, efforts were also made to limit the number of explicit strata. In some special situations, selection of one school from a few small strata was allowed although, as mentioned before, at least two schools were allocated to almost all explicit strata
- NRCs were requested to suggest the stratification variables (explicit and implicit) taking into consideration the special requirements of the corresponding educational systems. SurveyLang then reviewed those suggestions and finalised the stratification variables. Table 20 below provides the details of the stratification scheme used in each educational system for both languages.

**Table 20 Stratification Scheme in each participating educational system**

Educational system	Language	Explicit Stratification	Number of Explicit Strata	Implicit
Flemish Community of Belgium	1 & 2	Size (<25/25-34/>34); Area (2)	4	Net (2); Onderwijstype (4)
French Community of Belgium	1	Size (<25/25-34/>34); Type of School/SES (9)	11	Region (1-6)
French Community of Belgium	2	CENSUS		
German Community of Belgium	1 & 2	CENSUS		
Bulgaria	1 & 2	Size (<25/25-34/>34); Type of School (3)	5	Location (1-3)
Croatia	1	Size (<25/25-34/>34); Region (6)	8	NONE
Croatia	2	Size (<13/13-24/25-34/>34); Region (6)	9	NONE
England	1	Size (<25/25-34/>34)	3	Region (1 -4); School Type (1-3); Achievement (1-6)
England	2	Size (<13/13-24/25-34/>34)	4	Region (1 -4); School Type (1-3); Achievement (1-6)
Estonia	1	Size (<13/13-24/25-34/>34); Location (2)	7	Region (1-2)
Estonia	2	CENSUS		
France	1 & 2	Size (<25/25-34/>34); Type of School (3)	5	NONE
Greece	1	Size (<25/25-34/>34)	3	Region (1-7)
Greece	2	Size (<13/13-24/25-34/>34)	4	Region (1-7)
Malta	1 & 2	CENSUS		
Netherlands	1 & 2	Size (<25/25-34/>34); Type of Education (2)	4	Study Programme (1-6)
Poland	1	Size (<25/25-34/>34); Type of School (2)	6	Locality Size (3); School Size
Poland	2	Size (<13/13-24/25-34/>34)	4	Locality Size (3); School Size
Portugal	1 & 2	Size (<25/25-34/>34); Location (7)	9	School Nature (1-2)

Educational system	Language	Explicit Stratification	Number of Explicit Strata	Implicit
Slovenia	1	Size (<25/25-34/>34)	3	Region (1-8); School Size;
Slovenia	2	Size (<13/13-24/25-34/>34)	4	Region (1-8); School Size;
Spain	1	Size (<25/25-34/>34); Region	20	Region (1-16); School Type (2)
Spain	2	Size (<13/13-24/25-34/>34); Region	21	Region (1-16); School Type (2)
Sweden	1 & 2	Size (<25/25-34/>34)	3	Merits in English & Spanish (1)

As seen in Table 20 above, the size (number of eligible students enrolled) was always chosen as a stratification variable. For most educational systems, these strata were defined as follows: Large (>34), medium (25-34) and small (<25). For some educational systems with relatively larger numbers of schools with enrolment less than 25, four strata based on size were defined as follows: Large (>34), medium (25-34), small (13-24) and very small (<13). The primary goal for stratification by size was to ensure minimum representation of smaller schools in the sample within the overall constraints on the number of schools and students in the sample. In most cases, additional stratification variables (as shown in Table 20 above) were used. One or more of the size-based strata were further stratified into additional explicit strata. Within each educational system and for each language, schools containing fewer than 10 eligible students (extremely small schools with 9 eligible students or fewer) were excluded from the sampling frame as long as the total exclusion (exclusions due to other reasons plus exclusions from 'extremely small schools') was less than 2 percent of the total population of all enrolled eligible students across all schools. If necessary, the definition of 'extremely small schools' were modified i.e. the cut-off value of enrolment was lowered (from 10) to keep the total exclusions to less than 2 percent. However, all schools containing less than 6 eligible students were considered 'extremely small' and were excluded.

As mentioned before, the school sample size for all participating educational systems was at least 71 and the minimum target sample size for students was 1775 (71\*25). In small educational systems with fewer schools for any language, a census was undertaken by including all schools teaching that language. For some educational systems (Flemish Community of Belgium and Spain), additional schools beyond the number of schools required for the ESLC standard design were included in the school sample to meet additional sample size requirements for specific regions. For these two entities, therefore, two different samples (the national/entity sample drawn for ESLC and the full sample to accommodate additional sample size requirements) were created and the two samples were also weighted separately.



For the Flemish Community of Belgium, the full sample, as compared to the national sample, included 10 additional schools chosen from the Brussels region for each of the two languages (French and English). The national sample size for Language 1 and 2 for the Flemish Community of Belgium was 75 and 74 respectively whereas the corresponding full sample size was 85 and 84 for the two languages. For Spain, the additional sample size requirement was for specific regions. For Language 1 (English), a sample size of about 50 schools was allocated to each of the three regions (Andalusia, Canary Islands and Navarra) whereas for Language 2 (French), a sample size of about 50 schools was allocated for regions Andalusia and Canary Islands. The full sample size for Spain in Language 1 and Language 2 was 206 and 154 respectively whereas the corresponding national sample size was 76 and 82. The national samples were created by sub-sampling from the full samples.

#### ***4.11 Assigning a measure of size to each school***

A measure of size (MOS) had to be assigned to each school on the sampling frame for carrying out the PPS method of sampling schools. For the ESLC, MOS was derived from ENR and was defined as  $MOS = \max(ENR, 25)$ . The MOS was therefore equal to the enrolment number (ENR), unless ENR was less than 25 (the target cluster size, TCS, per school). When ENR was less than the cluster size, the MOS was set equal to the cluster size (25).

As mentioned above, “size” was always used as a variable for explicit stratification. For Large (>34) and Medium (25-34) strata, the MOS, as defined above, was equal to ENR. For schools with enrolment less than the TCS, the MOS for all schools was equal within the same stratum. As a result, the sampling of schools from such strata (where size was the same for all schools) was effectively based on equal probability of selection through simple random sampling.

#### ***4.12 Sorting the sampling frame***

For the purpose of sampling schools, the school frame was sorted within each explicit stratum by implicit stratification variables and then by ENR within each implicit stratum. The schools were first sorted by the first implicit stratification variable and then, within the levels of the first implicit stratification variable, by the second implicit stratification variable, and so on, until all implicit stratification variables were used. At the lowest level (i.e., for cells defined by different levels of the implicit stratification variables), the schools were sorted by ENR within each cell (or implicit strata). Within each explicit stratum, the sort order by ENR within each cell was changed from one implicit stratum to the next by using a high to low sort order in one followed by a low to high sort order in the next.



#### **4.13 School sample allocation across explicit strata**

The sample allocation of schools across all explicit strata was done such that the proportion of students sampled in any explicit stratum was roughly the same as the population proportions of eligible students in the corresponding explicit stratum.

However, exceptions had to be made to meet other constraints. First of all, the goal was to sample a minimum of 71 schools and, more importantly, at least 1775 students per language to meet the precision requirements. It was also important to have a minimum representation of relatively smaller schools (small schools (<25) and medium schools (25-34)). Sampling schools from the small and medium-sized schools would result in lower overall sample size in terms of students. As a result, disproportionate allocation of samples across strata based on size had to be allowed to meet the dual objective of representing schools of all sizes and at the same time keep the overall sample size of schools within reasonable limits. In some instances, the cluster size (number of students to be selected from sampled schools) had to be increased to make up for the loss in student sample size due to selection of relatively smaller schools in the school sample. This was mostly done in schools from large size stratum (>34) and sometimes also from medium size stratum (25-34).

The other constraint in the ESLC was that no student could be sampled for both languages although it was possible to select the same school for both languages. In some of these “overlapping” schools (sampled for both languages), there could be students who were eligible for both languages although they could be sampled for only one of the two languages, resulting in some loss in student sample size for one or both languages. The two school samples for the two languages were drawn independently and no specific steps were taken at the stratification stage (or at any other stage of sampling) to control or minimise the overlap of schools. In general, the goal was to increase the school sample size to the extent possible to account for the potential loss in student sample size. In some of these overlapping schools, the total number of eligible students for both languages in a few cases was not enough to meet the sample size requirement for each language. In the majority of these situations, relatively more students were allocated to the second (nationally less frequently taught) language because the availability of students for the second language was in general lower than that for the first language. In most cases involving overlapping schools, however, it was possible to sample enough students for each language (meeting the TCS requirement), without selecting the same student for both languages.

#### **4.14 Probability proportional to size sampling**

The schools within each explicit stratum, as mentioned before, were selected with PPS sampling. The procedures used to implement the PPS selection within each explicit stratum consisted of the following steps:

- (i) deriving the total measure of size ( $M$ ) for all schools in any explicit stratum. So, if there were  $N$  schools in any particular stratum, then  $M$  is the sum of the size measure of all  $N$  schools in that stratum  
recording the number of schools ( $n$ ) to be sampled from the specified explicit stratum. This sample size ( $n$ ) was determined based on sample allocation across all explicit strata  
calculating the sampling interval,  $I$ , as follows:  $I = M/n$   
within each explicit stratum, selection of a random number  $R$ , drawn from a uniform distribution between 0 and 1.

At the next step, the selection numbers for each of the ' $n$ ' schools to be selected from that explicit stratum were generated following the procedure described below:

- (i) obtaining the first selection number  $U$ , by multiplying the sampling interval,  $I$ , by the random number,  $R$ . This first selection number ( $U = RI$ ) was used to select the first sampled school in the specified explicit stratum  
obtaining the second selection number by adding the sampling interval,  $I$ , to the first selection number. The second selection number was used to identify the second sampled school  
continuing the same process i.e., adding the sampling interval,  $I$ , to the previous selection number to obtain the next selection number. This was continued until all ' $n$ ' selection numbers (one for each of the ' $n$ ' schools to be sampled) were generated.

Following this process described above, the ' $n$ ' selection numbers for the ' $n$ ' schools to be selected were as follows:  $U$ ,  $U + I$ ,  $U + 2I$ , .....,  $U + (n-1)I$ .

The process described above was carried out in each explicit stratum independently. For any specific explicit stratum, the sample size ' $n$ ' was based sample allocation for that stratum whereas the random number  $R$  was chosen independently for each stratum.

It should also be noted that for some explicit stratum there were some units that had to be chosen with certainty (i.e. with probability 1) because of their relatively large size. Specifically, schools, if any, with size ( $S$ ) equal to or greater than (Total Size for that explicit stratum/sample size for that explicit stratum) were selected with certainty and were set aside. The 'Total Size for an explicit stratum' was the sum of size measures of all schools belonging to that stratum whereas the sample size was the original sample size allocated to that explicit stratum. Once all the certainty selections were identified, that explicit stratum consisted of all schools not already selected with certainty and the total measure of size ( $M$ ) and sample size ( $n$ ) were computed based on those schools in that stratum. At that point, the selection numbers for the modified sample size were generated using the process described above. At the end of the process, the schools selected with certainty, if any, and those selected with probabilities less than 1 (non-certainty selections) were all included in the school sample.

The next task was to identify the schools to be sampled corresponding to the selection numbers already generated following the procedure described above. The first task was to compile a cumulative measure of size (CMOS) in each explicit stratum of the school sampling frame that determined which schools were to be sampled. Sampled schools were determined as follows.

Let  $U$  denote the first selection number for a particular explicit stratum. It was necessary to find the first school in the sampling frame where the cumulative measure of size (CMOS) equalled or exceeded  $U$ . This was the first sampled school. So, if  $C_s$  was the CMOS of a particular school  $S$  in the sampling frame and  $C_{(s-1)}$  was the CMOS of the school immediately preceding it on the sorted list, then the school in question (with CMOS equal to  $C_s$ ) was selected if:

- $C_s$  was greater than or equal to  $U$ , and
- $C_{(s-1)}$  was strictly less than  $U$ .

For a given explicit stratum, this rule was applied to all selection numbers and the corresponding selected schools generated the original sample of schools for that stratum. As mentioned before, the certainty selections, if any, were also added to the sample.

#### **4.15 Identifying replacement schools**

For each sampled school in the Main Study, up to two replacement schools were assigned from the sampling frame at the time of the selection of the main sample. Replacement schools were identified as follows: for each sampled school, the schools immediately preceding and following it on the sorted list (frame) in the same explicit stratum were designated as its replacement schools. The strata were considered as a continuous list, the last entry of which was “followed” by the first, and the first “proceeded” by the last one. The school immediately following the sampled school was identified as the first replacement, while the school immediately preceding the sampled school was identified as the second replacement. The within-stratum ordering of the school sampling frame by ENR ensured that any sampled school’s replacements were expected to have similar size characteristics.

Sometimes problems could be encountered when trying to identify two replacement schools for each sampled school. A sampled school could never be designated as the replacement school for another sampled school. It was also difficult to assign replacement schools to some very large sampled schools because such schools appeared very close to each other in the sampling frame. At times, it could only be possible to assign a single replacement school and perhaps none when two consecutive schools in the sampling frame are sampled. No replacement schools could obviously be assigned to any school from educational systems where a census of all schools was conducted. NRCs were encouraged to make every effort to confirm the participation of as many originally sampled schools as possible to minimise the potential for non-response bias. They contacted replacement schools after all attempts to obtain co-operation from the originally sampled schools were made. Each sampled

school that did not participate was replaced by replacement schools whenever possible.

#### **4.16 Student sampling**

Once schools were selected in the sample, the next step was to compile a list of students in the target grade(s) (ISCED2 or ISCED3) who were studying the language relevant to the school sampled. Student lists contained names, assigned IDs, and all relevant information defined by the student sampling form provided by SurveyLang; notably, the results of the routing test which provided a rough indication of each student's language proficiency in order to assign a test at the appropriate level and the foreign language. Student lists had to be exhaustive within the sampled schools and exclusions had to be documented.

The student sample for each language was chosen at the second stage from the sampled schools that responded to the survey (including replacement schools that were used as replacements for non-responding schools). The selection of the students took place via Simple Random Sampling (SRS) from the list of eligible students.

The two school samples for the two languages in any educational system were selected independently and so there could be some overlap between the two samples. In other words, there could be sampled schools within an educational system where two separate student samples (one for each language) had to be selected from the same school. Since no student could be tested in more than one language, the two student samples from such schools needed to be disjointed. Hence, the student samples were drawn in such a way that no student was included in both student samples.

For other schools that were sampled for one language only (or had no common eligible student even if they were sampled for both languages), the sampling of students was straightforward SRS selection from the lists provided.

As previously explained, the goal was to have a minimum sample of 25 students for each language tested in a school. In the student sample, use of replacement students was not allowed and a minimum response rate of 85% was expected<sup>18</sup>, i.e. an actual minimum sample size of about 21 participating students was targeted. An additional requirement for the selection of student sample was that a single student could undergo at most two of the three skill tests for the assessed language (Reading comprehension, Listening comprehension and Writing). Hence, for each of the three tests for a language, approximately 14 of these 21 students were selected so that no single student was chosen for more than two tests.

The following scenarios occurred for student sampling:

---

<sup>18</sup> although formal participation response rate had to be 80% at national level

- (i) scenario 1: if only one language had to be tested in the sampled school, a simple random sample of students (of size 25 as default) was drawn from among eligible students for that language. All eligible students were selected if fewer than 25 were enrolled for a particular language. It was found that the vast majority of schools in both samples actually required sampling for one language only.

scenario 2: if both languages had to be tested in the school selected, the two student samples needed to be mutually exclusive. If the testing grades for the two languages were different, i.e. if there were no common students who were eligible for both languages, the situation was similar to scenario 1 described above (for each language, a simple random sample of students was drawn from the corresponding population of all eligible students for that language). In situations where there were common students, the list of students was then divided up on the basis of the languages learned (this information was derived from the school sampling frame information):

- stratum 1: students that exclusively learned Language 1<sup>19</sup> in the school sampled (at the eligible grade)
- stratum 2: students that exclusively learned Language 2<sup>20</sup> in the school sampled
- stratum 3: students who learned both tested languages

Based on the numbers in the three categories, the two student samples were drawn so that the sample for each language could be considered a representative sample for that language and, at the same time, the sample size requirements for the two samples could be met. In most cases, the number of students for the most taught language (Language 1) in that educational system had more eligible students in a school as compared to the second most commonly taught language (Language 2). However, in a particular school, the situation could be reversed. Based on the number of students in each of the three strata (described above) and the sample size requirements for each language, sample allocation across the three strata were done to meet, as far as possible, the following objectives: (i) for the most commonly taught language (language 1), the sample size allocation to stratum 1 and stratum 3 would be proportional to the total number of students learning language 1 in those strata, and (ii) for the second most commonly taught language (language 2), the sample size allocation to stratum 2 and stratum 3 would be proportional to the total number of students learning language 2 in those strata.

For example, let's assume there were 80 students in stratum 1, 40 students in stratum 2 and 20 students in stratum 3 and a sample of 25 students were to be drawn for each language. Then, the 25 students to be sampled for language 1 would consist of 20 from stratum 1 and 5 from stratum 3, i.e. they would be in proportion to the total number of students in those strata (80:20 vs. 20:5). Similarly, the 25 students sampled

---

<sup>19</sup> Nationally most commonly taught (eligible) foreign language

<sup>20</sup> Nationally second most commonly taught (eligible) foreign language

for language 2 would consist of 17 from stratum 2 and 8 from stratum 3 and that would be roughly in proportion to of the total number of students in those strata (40:20 vs. 17:8). The eight students chosen from stratum 3 (containing students learning both languages) for language 2 were chosen from those students that were not already selected for language 1.

However, there could be several scenarios involving the number of students in each of these three strata and the two samples also had to be mutually disjointed. So, a strictly proportional sample allocation scheme was not always feasible under these constraints but the goal was to meet those objectives to the extent possible in each situation.

Finally, all sampled students within a school for a given language were sampled for the three skill tests: Reading comprehension, Listening comprehension and Writing. Each sampled student was randomly assigned to two of the three tests such that each student was assigned to exactly two tests. In other words, the sample size for each of these three tests was roughly two-thirds of the student sample size for that language. For example, if 18 students were sampled, then the sample size for each test was 12 and each student was assigned to two of the three tests. If the number of sampled students was 25 (not an exact multiple of 3), then two of the three tests had a sample size of 17 whereas that for the third test was 16. Again, each student was assigned to exactly two of the three tests.

The achieved sample size of students per educational system is shown in Table 21.

**Table 21 Student sample size for participating educational systems for both languages**

Educational system	Student Sample Size ('first' language)	Student Sample Size ('second' language)
Belgium (Flemish Community <sup>21</sup> )	1824	1813
French Community of Belgium	1805	1297
German Community of Belgium	1006	761
Bulgaria	1806	1808
Croatia	1796	1803
England	1778	1747
Estonia	1779	1489
France	1811	1799
Greece	1761	1488
Malta	1366	1381

<sup>21</sup> With the booster sample; first language: 2069, second language: 2048

Educational system	Student Sample Size ('first' language)	Student Sample Size ('second' language)
Netherlands	1633	1607
Poland	2132	1787
Portugal	1781	1838
Slovenia	1775	1775
Spain <sup>22</sup>	1905	1856
Sweden	1849	1785
Total	27807	26034

**Table 22 Student sample summary across languages**

Language	Number of sampled students for 'first' language	Number of sampled students for 'second' language	Total number of students sampled per language	Total percentage of students sampled per language
English	23199	4321	27520	51%
French	4608	5182	9790	18%
German		11566	11566	21%
Italian		1381	1381	3%
Spanish		3584	3584	7%
<b>Total</b>	<b>27807</b>	<b>26034</b>	<b>53841</b>	<b>100%</b>

#### **4.17 Selecting the school sample personnel**

School personnel: the ESLC school personnel sample was self-selecting – each participating school's principal and all respective language teachers teaching the test language at the testing level were invited to fill in the School and Teacher Questionnaire, respectively.

The goal was to administer questionnaires to all language teachers for each of the tested languages in each sampled school at the testing grade. No sampling among the language teachers was implemented; all listed teachers were invited to fill in the survey.

<sup>22</sup> With the booster sample; first language: 5046, second language: 3332



Where a school was selected for both test languages, the school principal was randomly allocated to complete the School Questionnaire for one test language only rather than having to complete the two questionnaires, one for each test language. Similarly, for teachers teaching both test languages in a school selected for both test languages, the teacher was randomly allocated to complete the Teacher questionnaire for one test language only.

There was no participation criterion or minimum response rate set for the teacher sample. NRCs made all efforts to decrease teacher non-response and have as many respondents as possible in each country.

#### **4.18 Sampling forms**

For the purpose of sampling schools and students, all educational systems were required to provide information using suitably designed sampling forms. Example sampling forms can be seen in Appendix 3. Using these forms, the NRCs submitted necessary information on languages to be tested, testing grades, target populations, exclusions, stratification variables, school and student sampling frames and all other relevant details for SurveyLang to be able to carry out the sampling task. Once these forms were received, they were checked and reviewed by SurveyLang for accuracy and consistency and, if necessary, educational systems were asked to make necessary revisions. Final decisions on all issues were made in consultation with the NRCs of the corresponding educational systems. A brief description of the main scope and purpose of the different sampling forms used is given below. Similar forms were used for both languages.

- Sampling Form 1: Organisation, Logistics: information on participation in the ESLC, and on NRCs and experts responsible for sampling information
- Sampling Form 2: Language and Grade Definition: confirmation of the two languages and the corresponding testing grades
- Sampling Form 3: School Level Exclusions: Types of exclusions, reasons for each exclusion type, estimated percentage of students to be excluded for each exclusion type
- Sampling Form 4: Student Level Exclusions: Types of exclusions, reasons for each exclusion type, estimated percentage of students to be excluded for each exclusion type
- Sampling Form 5: Explicit Stratification: suggested explicit stratification variables, if any, and their categories (levels); estimated percentage of students by strata and suggested sample allocation across strata (proportional or any specified disproportional allocation)
- Sampling Form 6: Implicit Stratification: up to three suggested implicit stratification variables and their categories
- Sampling Form 8: Unified School Master List: comprehensive listing of all schools to be included in the ESLC; includes information testing grades, exclusions, and stratification variables at the school level; most importantly,





contains information on enrolment of eligible students in relevant grades to be used for estimating enrolment (size) for schools for the purpose of PPS sampling

- Sampling Form 9: Student Listing: provides student level data (name, sex, date of birth, academic years of language instruction, level of language proficiency) for creating student sampling frames within sampled schools
- Sampling Form 10: Teacher Listing: provides teacher level information for each tested language
- Tracking Forms: Two tracking forms (T1 and T2) were used to record information on participation and test administrations for the school and student samples.

# ***Chapter 5:***

# ***Operations -***

# ***Translation***

## 5 Operations - Translation

This chapter provides an overview of the translation process for the ESLC. Note, the discussion focuses on the Main Study processes, Field Trial processes are not discussed unless relevant.

### 5.1 Introduction

A large number of the documents created by SurveyLang needed to be translated and localised to the questionnaire language(s) (see 3.2.3.2 above for a definition of this term) of the participating educational systems.

Good translation and ensuring the quality of all educational system questionnaires and documentation was essential to the overall success of a multilingual project like the ESLC, where international comparability is the key requirement. It was, therefore, crucial to ensure that the translation process did not introduce bias likely to distort these comparisons. ESLC, therefore, implemented strict translation procedures which are described in this chapter.

Translation work and costs were borne by the participating educational systems. Participating educational systems were responsible for conducting the translation and localisation work and for recruiting and training their national translation teams. Their tasks with respect to translation were as follows:

- to attend the central SurveyLang training session on translation
- to manage the process for all documents requiring translation and localisation
- to coordinate translation roles and schedules
- to recruit translators according to the criteria set by SurveyLang
- to ensure protocol and guidelines set by SurveyLang were followed by all translators
- to take overall responsibility for quality control and the signing-off of the final versions of all documents, including final optical checks of the questionnaires before they were produced.

SurveyLang's responsibility was to provide the source documents in English (the source language for the project) and to set translation standards, as well as providing the necessary guidelines and manuals, checklists and tools. SurveyLang also had a role in the quality control of the output.

For the ESLC, unlike in other international surveys, the questionnaires (for students, teachers and principals) were the only test instruments that required translation. This was because the language tests were created in the test languages (English, French, German, Italian and Spanish) by the language testing group (see chapter 2).

## **5.2 Overview of translation system, support and training**

To help national teams with the complicated task of translation and localisation, an internet-based translation system called WebTrans was provided to manage and facilitate the translation process for the project. Source documents were provided by SurveyLang in English.

The topic of translation and hands-on experience using WebTrans was part of a two-day training session provided by SurveyLang. Documentation was also made available to NRCs to assist them in their translation tasks. NRCs were required to cascade the training and provide the following documentation to the translators they appointed:

- translation guidelines
- a user manual for each role in WebTrans
- general quality control and process guidelines
- Do's and Don'ts of translation
- recruitment checklists
- Frequently Asked Questions.

WebTrans contained the translation guidelines, the instruction material on how to use WebTrans, as well as the relevant user manual for each profile of the translation procedure. SurveyLang team members were also available to respond to questions at any time during the translation phase.

WebTrans enabled NRCs to follow translation procedures step by step whilst also allowing SurveyLang to review and quality-control the documents submitted by NRCs. WebTrans is a secure online translation system, so as long as users had a stable internet connection they could access it anywhere. Users received their user name and password to access WebTrans from SurveyLang. Different members of the NRC translation team in charge of each phase of the translation process performed their tasks (i.e. translation, reconciliation/review, back translation, evaluation, and approval of documents) in different profiles in WebTrans; each specially tailored to the requirements of their respective tasks. Once a phase was complete, the respective profile was closed and that user was no longer able to make further changes in his/her profile. The next user coming into the translation process then started performing his/her task. A user may have had more than one role and after logging into WebTrans, they were able to see all of their roles listed in their profile.

The training session and the documentation outlined the importance of the impact of translators' contributions as well as the significant difference their commitment could make to the success of the ESLC. All translators recruited to work on the ESLC were asked to subscribe to the guidelines provided. Although the recruitment of translators was the responsibility of participating educational systems, SurveyLang provided recommended criteria for recruitment which can be seen at the end of this Chapter.

### 5.3 Documentation needing translation and the translation process

There were two different translation procedures depending on the type of document. For the sake of clarity, SurveyLang divided documents into two main categories: Type A and Type B. The translation procedures for Type A and Type B documents were different, and are outlined below. In brief, Type A documents related to the test instruments and required a more extensive translation process. Type B documents related to the operational documentation used in-country. The process for Type B documents was also rigorous while allowing more flexibility for the in-country team. There was a third category, Type C, which were provided as finalised versions by SurveyLang and did not need translating. The majority of these were in English and were intended for the NRC. However, there was also the marking of writing documentation which was provided in the five test languages.

The full list of documentation needing translation was as follows:

**Table 23 Translation process for each document needing translation**

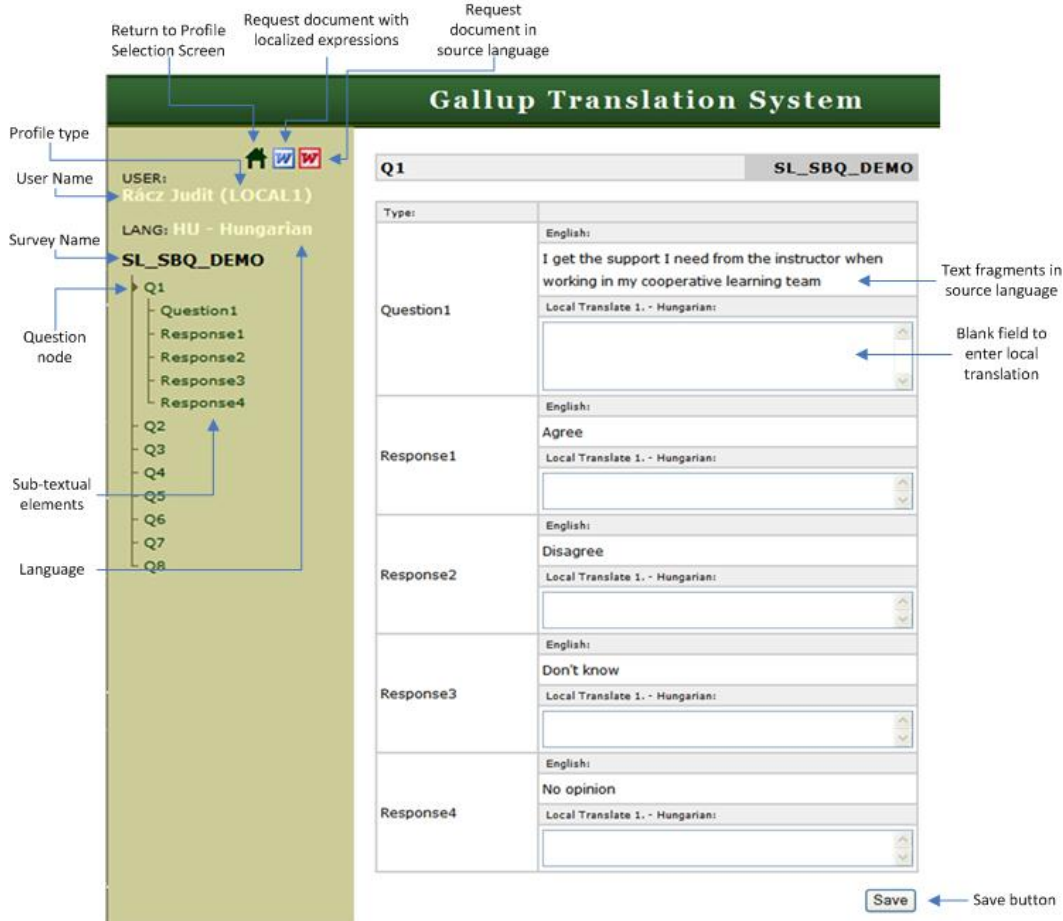
Document name	Translation process
Student Questionnaire	<i>A</i>
Teacher Questionnaire	<i>A</i>
Principal Questionnaire	<i>A</i>
Testing Tool navigation	<i>A</i>
Test Administration Manual (paper-based)	<i>B</i>
Test Administration Manual (computer-based)	<i>B</i>
School Coordinator Guidelines (paper-based)	<i>B</i>
School Coordinator Guidelines (computer-based)	<i>B</i>
Language Test Familiarisation Materials;	<i>B</i>
Testing Tool Guidelines for Students	<i>B</i>
sampling guidelines	<i>B</i>
sampling forms	<i>B</i>
Data Entry Guidelines (language tests)	<i>B</i>
Data Entry Guidelines (questionnaires)	<i>B</i>
Routing Test Instruction Sheet	<i>B</i>

Type A documents were translated online on WebTrans. Translators could save their work and log on and off the system when convenient. The Type B documents could be downloaded in Word document format from WebTrans and translators worked on these documents offline. Translators could upload their completed translation and WebTrans automatically managed version control.

For Type A documents, the process was as follows:

- Double forwards translation [LOCAL1, LOCAL 2]: two independent translators translated the source document producing two parallel translations in the questionnaire language. Translators did this work directly on the WebTrans system as the screen shot below illustrates. WebTrans opened up in a simple browser where all translation information was displayed clearly. The hierarchical structure of the document was visible on the left-hand side of the screen. All itemized text fragments (in the example below: survey questions) appeared as “main nodes” which branched off into sub-textual elements (in the example below: questions and responses). For ease of use, these fragments were labelled as they were in the source text, e.g. Q3 for question 3 etc. The navigation bar enabled users to go back and forth between the various parts of the itemised document. Once a fragment and its sub-elements had been translated, a tick appeared beside that fragment in the left-hand side menu. Translators were also able to add comments in the questionnaire language below each translated field. For this purpose, a notepad appeared in the left-hand side menu beside the fragment that had been commented on.

**Figure 13 WebTrans screen for Type A forward translation**

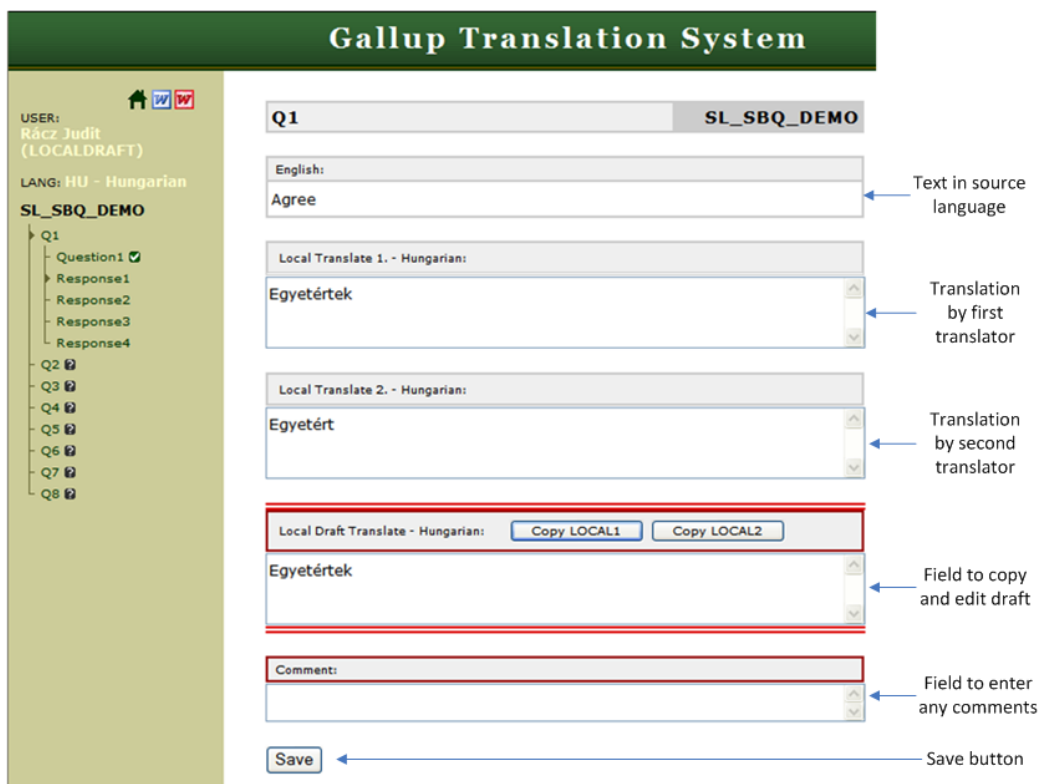


The screenshot displays the Gallup Translation System interface. On the left, a navigation pane shows a hierarchical tree structure with nodes for Profile type, User Name, Survey Name, Question node, Sub-textual elements, and Language. The main area on the right is titled 'Gallup Translation System' and shows a table of text fragments for translation. The table has columns for 'Type' and 'English' (source language). The first row shows 'Question1' with the English text 'I get the support I need from the instructor when working in my cooperative learning team'. Below this, there is a 'Local Translate 1. - Hungarian:' field with a blank space for the translator to enter the translation. The table continues with 'Response1' (Agree), 'Response2' (Disagree), 'Response3' (Don't know), and 'Response4' (No opinion), each with a corresponding 'Local Translate 1. - Hungarian:' field. A 'Save' button is located at the bottom right of the interface.

Type	English	Local Translate 1. - Hungarian
Question1	I get the support I need from the instructor when working in my cooperative learning team	
Response1	Agree	
Response2	Disagree	
Response3	Don't know	
Response4	No opinion	

- Reconciliation [LOCAL DRAFT]: a third person worked on the two parallel translations and created one unified version. This provided an opportunity to build on the respective strengths of the two parallel translations to produce an enhanced version. In order to do this, the person performing the reconciliation was required to regularly refer to the source document. The figure below shows the screen that the translator performing this work would see. The central part of the screen displays in tabulated format the sub-textual elements of the fragment currently being translated. The work of the first and second forward translators can be seen (e.g. Local 1 and Local 2). The reconciler, using the WebTrans role of LOCALDRAFT, could see both versions on the screen could copy the text from either translator in the “Local Draft Translate” box, and could then be fine-tune the reconciled version on that basis.

**Figure 14 WebTrans screen for Type A reconciler**



**Gallup Translation System**

USER: Rácz Judit (LOCALDRAFT)  
LANG: HU - Hungarian  
SL\_SBQ\_DEMO

Q1 SL\_SBQ\_DEMO

English: Agree

Local Translate 1. - Hungarian: Egyetérték

Local Translate 2. - Hungarian: Egyetérték

Local Draft Translate - Hungarian: Egyetérték

Copy LOCAL1 Copy LOCAL2

Comment:

Save

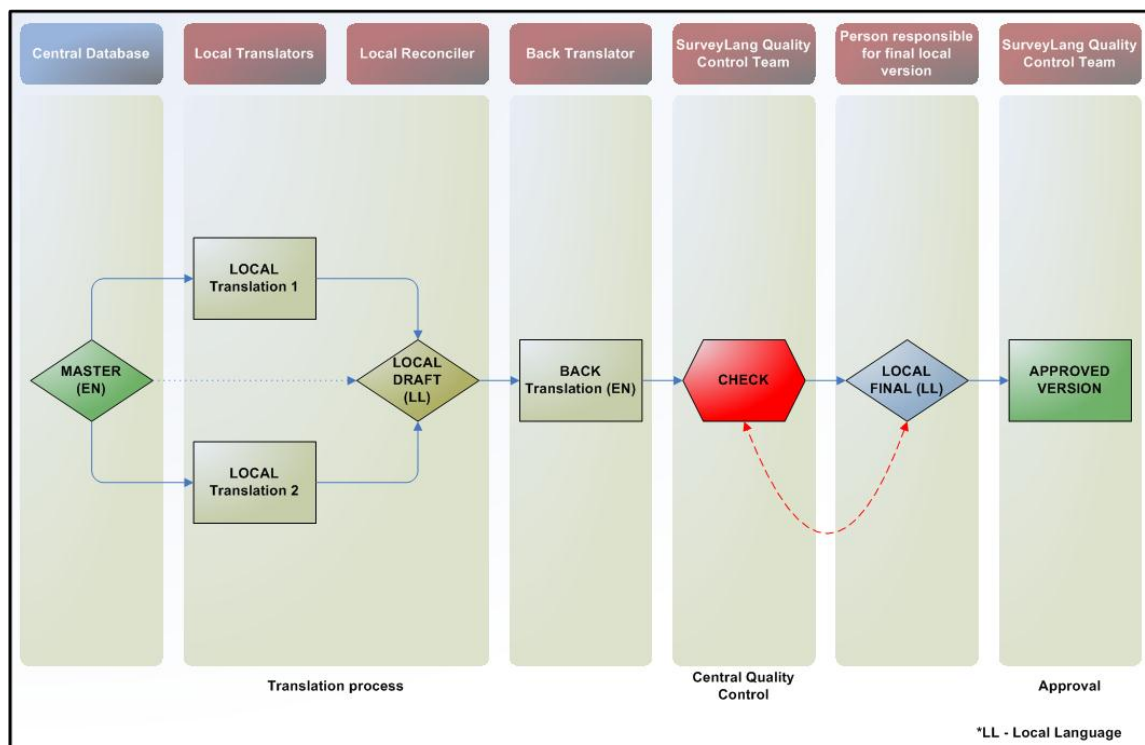
- Back translation [BACK]: a fourth person who had not seen the source version translated the reconciled version back to English.
- Verification process and sign-off [CHECK, LOCAL FINAL]: a SurveyLang team member with experience of the verification process from other international surveys performed the verification and sign-off of the questionnaires for the ESLC. This involved a close comparative analysis of the source text and the back translation. Where the SurveyLang team had comments or queries, they could flag items on WebTrans for the NRC’s attention. All of this communication was done in English. The entire process

was documented on WebTrans with all comments visible. Once all queries were resolved, the document could be signed-off. More details about the verification process are detailed in Chapter 3.

- Second test language questionnaire: after sign-off, each of the three questionnaires was prepared for the second test language, as each educational system was tested in two languages. This meant that the NRC had to make minor changes throughout the questionnaires, including changing the first test language to the second test language as well as potentially making changes to the localisations (see Chapter 3 for further details on the localisations). This stage also required sign-off by the verification team.
- Optical sign-off: the final task for NRCs was to sign-off their questionnaires in the final format that they were produced in. This was important as the earlier sign-off had been element by element. This was the final step before test production.

The process described above for the verification and sign-off of the first test language questionnaire is shown in the figure below where the role profiles within WebTrans are in capital letters.

**Figure 15 WebTrans process for Type A translation**



For Type B documents, the process was as follows:

- One forward translation [LOCAL]: two independent translators translated the source document producing two parallel translations of that document in the

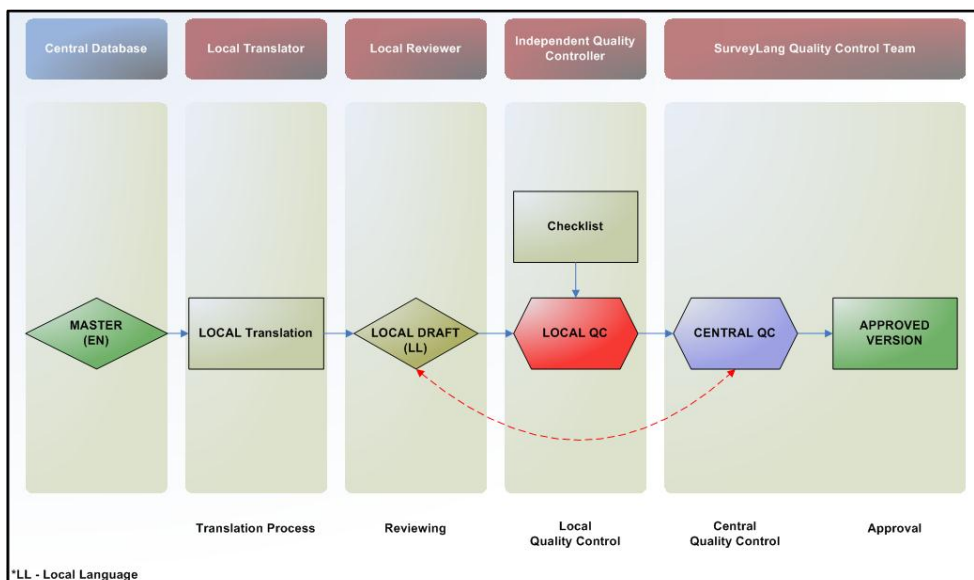


questionnaire language. *Note: SurveyLang had originally required two forward translators: however, after discussion with educational systems, this was felt to be too onerous and costly and it was agreed that there would only be one forward translator. This was endorsed by the European Commission at the Advisory Board meeting on 5 December 2008.*

- Review [LOCAL DRAFT]: this involved another translator going through the translation and producing an enhanced draft version.
- Quality of translation (Stage 1 Local quality control) [LOCAL QC]: an independent person performed a comparative analysis of the local version to the source text on the basis of a document-specific checklist created by SurveyLang. The checklist focused on content criteria essential for the document.
- Quality of translation (Stage 2 Central quality control) [CENTRAL QC]: SurveyLang reviewed the final document together with the source version. This was an optical check rather than a language check and was done paragraph by paragraph to ensure that the same number of paragraphs, bullet points etc. were used for both documents. Particular attention was paid to checking the points in the text relating to the checklist criteria. SurveyLang engaged in a comment/revision loop until all points had been clarified or amended.
- Approval of the translation, or formal sign-off, took place once all points had been clarified and all necessary adjustments made.

The process described above is shown in the figure below where the role profiles within WebTrans are in capital letters.

**Figure 16 WebTrans process for Type B translation**



To sum up, some of the main differences between the translation process of Type A and B documents are as follows:

**Table 24 Summary of differences between Type A and B translation**

	Type A documents	Type B documents
How/where is the translation carried out?	In WebTrans, translated item by item	In a Word file which is uploaded in WebTrans once each phase is complete
Translation	By two independent translators	By one translator only
Back translation	There is a back translation	There is no back translation
Quality control/checking of translation	Checked once by SurveyLang, based on linguistic/semantic criteria	Two stage checking (local and then central by SurveyLang) based on document specific checklist / content criteria

WebTrans automatically stored all interactions performed in its framework; therefore, all changes and each version's history were recorded on the system. For Type A documents, these audit trails systematically included:

- the two parallel translations
- the reconciled version
- the back translation
- the modifications agreed during the verification process
- all comments from each of the above stages
- the final version of the translated document.

For Type B documents, these audit trails systematically included::

- the comments based on the checklist criteria
- all versions together with unique name, the user profile where the document was uploaded and date.

## **5.4 SurveyLang translation guidelines**

There are no universally agreed standards or principles as to what constitutes a good translation. The main reason for this is that translation is generally perceived as an art as much as it is a science. Like any art, translation is a creative process and defining quality standards for a creative process is notoriously difficult.

Most translators acknowledge that absolute equivalence does not exist in translation. The most frequent dilemma that translators face revolves around the issue of 'under' versus 'over' translation: in other words, around the degree of freedom with which translators should perform translation tasks. While some believe that a translation should stay as close as possible to the syntactical and lexical features of the source language, others think that a translation should primarily remain faithful to the spirit, if

not to the letter, of the source text and essentially seek functional equivalence. SurveyLang recommended the latter approach, yet urged translators to also keep in mind the pan-European comparability context at all times. The aim was really to strike a good balance; the translation must not be literal to the point that it sounds awkward, but neither should it deviate too freely from the source version, which might affect the functioning of the measurement items in unexpected ways.

As in most disciplines, preparation in translation is crucial. Most of the preparation work in the run-up to the actual translation was about gaining a deep and thorough understanding of the source text; ‘understanding is already translating’, as translation theorist José Ortega y Gasset famously said.

SurveyLang had three benchmark criteria that we believed covered the notion of “functional equivalence”:

***Accurate:** the text should reproduce as accurately as possible the contextual meaning of the source text and the goal should be semantic equivalence between the two.*

***Natural:** the text should use natural forms of the questionnaire language in a way that is appropriate to the source text being translated; a good test is to check whether the text reads like a translation or like a document originally written in the questionnaire language.*

***Communicative:** the text should express all aspects of the contextual meaning in a way that is readily understandable to the intended audience; it should attempt to produce the same effect on the readers as the source text.*

In the ESLC, the quality of translations was assessed on the basis of these three criteria.

## **5.5 Questionnaire language, localisations and amendments to standard process**

Before the translation work commenced, each educational system agreed with SurveyLang the languages that they intended to translate their documentation into. This language was known as the ‘questionnaire language’. The term ‘Questionnaire Language’ was used in place of the terms ‘local language’, ‘national language’ and ‘language of instruction’ which had been criticised for their lack of clarity. The questionnaire language was defined as the language that the questionnaires, testing tool navigation details, sampling forms, guidelines and manuals were administered and available in. This language had to be agreed upon with SurveyLang and had to be one of the official languages within the educational system which is used in most or most important communicative situations (for work, life in society, etc.) in the region where the school is located and that is the language of instruction in the school’s region. These agreed languages can be seen in Table 25 below.



For some educational systems (e.g. Croatia, Slovenia, Sweden), it was considered whether it was necessary to translate the documentation into other languages in addition to the language below, however each NRC assured SurveyLang in writing that it was sufficient to translate only into the languages listed below.

As can be seen, Estonia and Spain translated their documentation into more than one questionnaire language. In Estonia, all documents were translated into both Estonian and Russian. In Spain, all documentation was translated into five languages: Spanish (Castilian), Basque, Catalan, Galician and Valencian. As each questionnaire had to be available for the two tested languages, this meant that there were 10 versions of each of the three questionnaires (student, teacher and parent) created for Spain.

France and the French Community of Belgium had an agreement to share the translation process. This meant that, in practice for each document, one educational system took the lead and the other acted as the 'donor'. The donor received the document in their own profile area of WebTrans after SurveyLang had signed off the document for the lead educational system. The donor then made any necessary changes to the localisations as already agreed with SurveyLang. Once they had made their changes, the quality control and review process was undertaken by SurveyLang before the document could be signed off.

In the German Community of Belgium, it was agreed that the documentation intended for students would be translated into German; however, the documentation intended for participants other than students, e.g. School Coordinators and Test Administrators could be in French as these personnel could all speak French fluently. This also means that the Teacher and Principal Questionnaires were administered in French. This lowered the translation burden on the German Community of Belgium as they agreed with the French Community of Belgium that they would use the French documentation created by them (or by France) and localise as agreed with SurveyLang for their own context.

For Malta, there was a discussion point as although English was the first most widely taught language and hence a language for testing, English is also an official language and was nominated by Malta as the questionnaire language rather than Maltese. Malta reasoned that this is what is done in other surveys and therefore was also acceptable for the ESLC. After discussion with the teams managing TIMSS and PIRLS and Malta's assurances that administering the questionnaires in English would not be detrimental to students, it was agreed that English rather than Maltese could be used. The team managing TIMSS and PIRLS stated that 'Malta is an example of a country that administers TIMSS in English as the language of instruction, even though Maltese is the mother tongue. We have no evidence that administering the TIMSS assessment in English caused undue problems for the Maltese students' (Michael Martin 2009, personal communication). Furthermore recent 'experience from PIRLS 2011, which was administered in both English and Maltese to the same students, suggests that students performed as well if not better on the English version' (Michael Martin 2012, personal communication).

**Table 25 Agreed questionnaire languages for each educational system**

Educational system	Educational system code	Questionnaire language(s)	Language code
Flemish Community of Belgium	BE nl	Dutch	nl
French Community of Belgium	BE fr	French	fr
German Community of Belgium	BE de	German/French	de, fr
Bulgaria	BG	Bulgarian	bg
Croatia	HR	Croatian	hr
England	UK-ENG	English	en
Estonia	EE	Estonian; Russian	et, er
France	FR	French	fr
Greece	EL	Greek	el
Malta	MT	English	en
Netherlands	NL	Dutch	nl
Poland	PL	Polish	pl
Portugal	PT	Portuguese	pt
Slovenia	SI	Slovene	sl
Spain	ES	Spanish, Basque, Catalan, Galician, Valencian	es, Spanish-Basque Spanish-Catalan, Spanish-Galician, Spanish-Valencian
Sweden	SE	Swedish	sv

Another essential task before the translation work could commence was for each educational system to standardise and agree their localisations with SurveyLang. SurveyLang created a localisation spreadsheet where educational systems needed to formally record aspects about their educational context and have these signed off by SurveyLang. Clear guidance was given by SurveyLang on each step. Each NRC completed this task with the assistance of their local Eurydice representative where available. Further details on the localisation process can be seen in Chapter 3.

## **5.6 Development of source versions**

All of the operational source documentation was developed by SurveyLang in close accordance with the standards of other international surveys but clearly tailored towards the specific needs of the ESLC. The language test materials were developed by the specialist Language Testing Group in the test language with feedback and input



at various stages from NRCs, the European Commission and their Advisory Board members, as well as students and teachers who participated in the pretesting and Field Trial phases of the survey (see chapter 2 for further details of the development process and use of feedback and test statistics for each stage of development). The questionnaires were developed by the specialist SurveyLang team, again with close input and feedback from NRCs, the European Commission and their Advisory Board members as well as students and teachers who participated in the Field Trial (see chapter 3 for further details).

After the Field Trial, through their feedback in the NRC Feedback Report and through the Quality Monitor report, NRCs contributed to the revision and modification of all source documents.

## **5.7 *Field Trial and Main Study translation processes***

The bulk of the translation work had to be done prior to the Field Trial. Before the Field Trial, the steps for NRCs in regard to translation were to:

- agree the questionnaire language
- agree the Localisation spreadsheet
- attend the central SurveyLang training
- recruit translators according to set criteria and send details to SurveyLang
- agree translations with SurveyLang
- carry out and sign-off all translations according to the schedule and criteria set by SurveyLang
- store finalised translations on the ESLC Basecamp website.

After the Field Trial, the source questionnaires and all operational documents were modified following a detailed SurveyLang review. SurveyLang also used information and feedback received from NRC teams in the Quality Monitor report, NRC feedback report and from the statistical analysis from the Field Trial.

After the Field Trial and before the Main Study, the steps for NRCs in regard to translation were to:

- amend the Localisation spreadsheet and sign-off with SurveyLang if necessary
- recruit translators according to set criteria and send their details to SurveyLang
- agree all translation modifications with SurveyLang
- carry out and sign-off all translations according to the schedule and criteria set by SurveyLang
- store finalised translations on the ESLC Basecamp website.

For all documents, SurveyLang clearly indicated which changes and modifications had to be made to the Main Study versions. NRCs could make additional amendments if

they felt that there was enough evidence from the Field Trial to justify the changes but these had to be agreed by SurveyLang. The translation process followed the standard process outlined in paragraphs 0 to 0 above, including the verification process for the questionnaires and in all cases SurveyLang had to sign-off the final versions.

## **5.8 Recruitment guidelines for translators**

NRCs needed to assess how many translators they needed based on the source materials that required translation and or localisation, whether those materials were Type A or Type B and the number of questionnaire languages they had.

SurveyLang set strict criteria for translators participating in the ESLC. All translators had to be fully trained and have a perfect command of the source language, English, which had to be documented. Translators also had to be native speakers of the questionnaire language and have a proven track record of undertaking high level translation for at least 3-5 years prior to the ESLC. They needed to be specialised, or at least well-versed, in educational issues and they should have had an extensive knowledge of the school system of their home educational system, and, preferably, also of various other school systems across Europe. Translators should also have been familiar with the challenges of translating from English to their mother tongue and ideally, they should not only have had bilingual ability but also bicultural vision. They should also all have been computer literate enough to be able to use an internet-based tool such as WebTrans. Applicants should have resided in the educational system and they should have been able to provide references or agree to do a test translation of a maximum of 400 words related to the topic. It was recommended that translators were individual contractors rather than translation agencies. The purpose of this was so that the task was not allocated or assigned to anyone other than the selected and approved translator.

SurveyLang imposed even stricter requirements for the reconciler. In view of the strategic role played by the reconciler in the translation process, it was essential that s/he should have had an in-depth understanding of the organic nature of both source and questionnaire languages, good familiarity with the terminology used and meticulous attention to detail. SurveyLang was happy for the NRC to act as reconciler.

SurveyLang urged NRCs to test translators before entrusting them with the job, unless, of course, NRCs had a positive track record of working with translators in the past.

## **5.9 References**

Ortega y Gasset, J (1993) The misery and the splendor of translation, in Schulte, R and Bigunenet, J (Eds), *Theories of Translation: an Anthology of Essays from*



*Dryden to Derrida* (translated by Elizabeth Gamble Miller), Chicago: University of Chicago Press, 93 -112.



# ***Chapter 6:***

# ***Operations - the***

# ***SurveyLang***

# ***software platform***

## 6 Operations - the SurveyLang software platform

This Chapter provides a detailed description of the requirements, architecture and functionality of the SurveyLang software platform.

### 6.1 *Introduction*

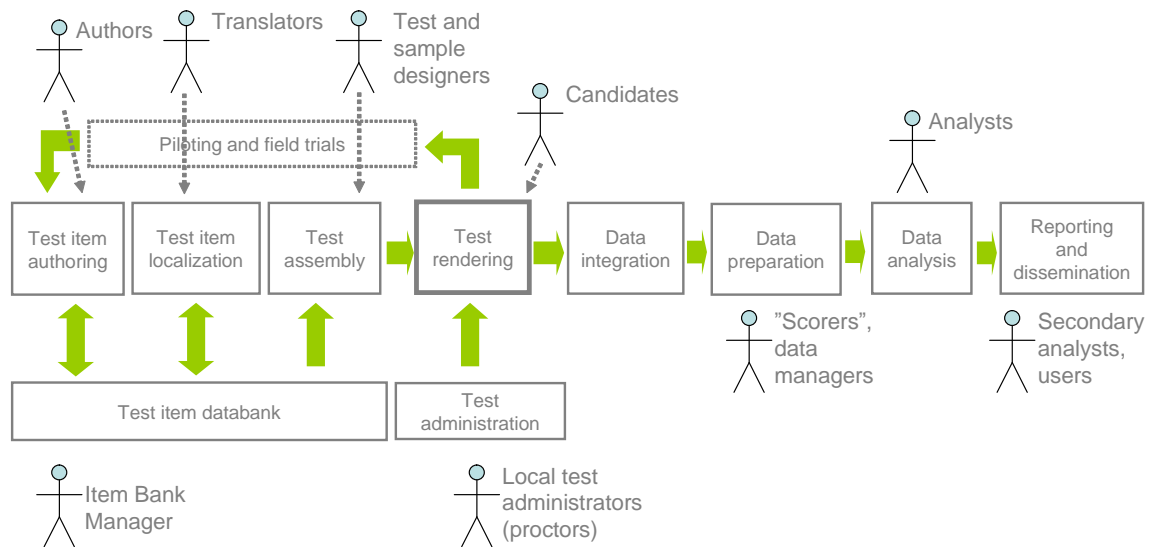
SurveyLang has developed an integrated, state-of-the-art, functionality-rich software system for the design, management and delivery of the language tests and accompanying questionnaires. The platform is fine-tuned to the specific set of requirements of the ESLC project and is designed to support the delivery of the paper-based and computer-based tests. The software platform also supports all major stages of the survey process.

### 6.2 *Requirements*

The technical and functional requirements of the software platform were developed in close cooperation between the SurveyLang partners. At a high level, the software platform should:

- Support all various stages in the development and implementation of the survey (see Figure 17)
- enable the automation of error-prone and expensive manual processes
- be flexible enough to handle the variety of task types used by the survey
- support the implementation of the complex test design used in the survey
- meet the high security requirements of international assessment surveys like the ESLC
- reduce the technical and administrative burden on the local administrators to a minimum
- run on existing hardware platforms in the schools
- be an open source platform (to be made available by the European Commission after the completion of the project), for free use by any interested party.

**Figure 17 Stages and roles in the development and delivery of the survey**



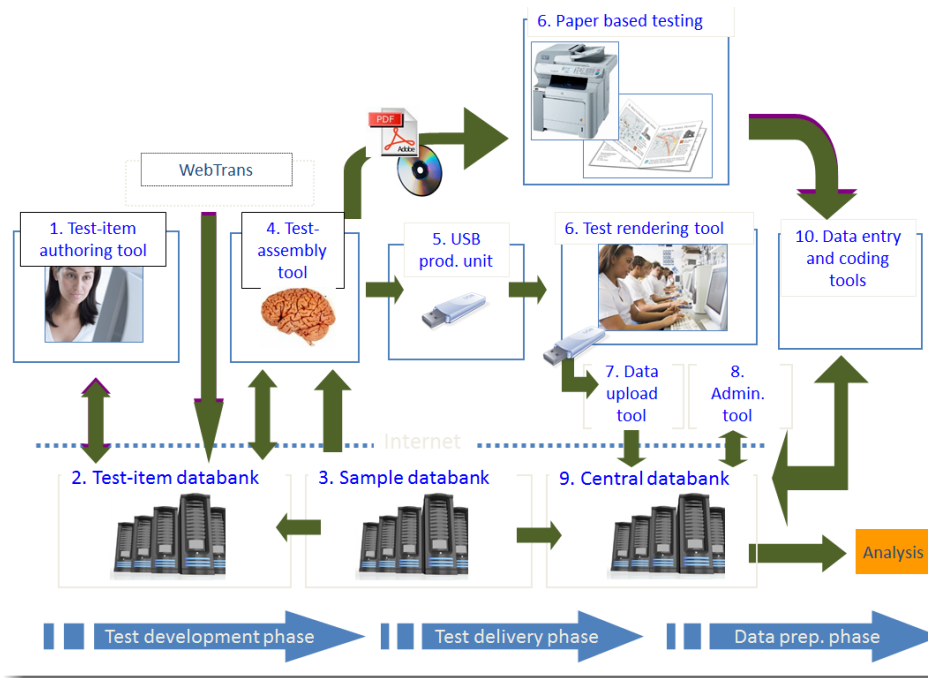
In terms of functionality, the following tools and components were needed:

- Test-item authoring, editing and preview functionality supporting an environment of distributed authors scattered around Europe.
- Test-item databank functionality providing efficient storage, management and version control of test-items. This tool should also encourage visibility and sharing of resources between the various roles associated with the stages of the test-item life-cycle.
- Test-item translation functionality, supporting the localization of test-items, instructions and accompanying questionnaires to national languages.
- Test construction functionality, supporting the assembly of individual test-items into complete test sessions as well as the allocation of students across tests at different levels.
- Test material production functionality for computer-based as well as paper-based testing.
- Test administration functionality supporting the management of respondents and test administrations at the school level.
- Test rendering functionality supporting efficient and user-friendly presentation of tests-items to respondents as well as the capturing of their responses (for computer-based-testing).
- Data integration functionality supporting efficient assembly of response data coming from the participating schools.
- Data preparation functionality supporting all tasks related to the preparation of data files ready for analysis, including data entry of paper-based responses and support for manual marking/scoring of open ended items.

## 6.3 Architecture

The high level architecture of the software platform that has been designed to provide this functionality can be seen in Figure 18.

**Figure 18 High level architecture**



The platform consists of a central Test-item databank interfacing two different tools over the Internet: the Test-item authoring tool and the Test assembly tool. In addition, an interface to a translation management system is also provided. As a whole, these distributed tools, plus the Test-item databank, are designed to support the central test development team in their efforts to develop and fine-tune the language tests.

The production of paper-based and computer-based test materials is handled by the Test assembly tool. The physical production of computer tests (delivered on USB memory sticks) is, however, done by a USB memory stick production unit containing specially built hardware as well as software components.

To support the test-delivery phase of the project, another set of tools are provided. These are i) a Test-rendering tool to be installed on the test computers in all the schools taking computer-based-testing and ii) a data upload service which allows the test administrator to upload student test data from the test USB memory sticks to the central databank.

The various tools and components are described in further details in the following paragraphs.

## 6.4 Test-item authoring tool

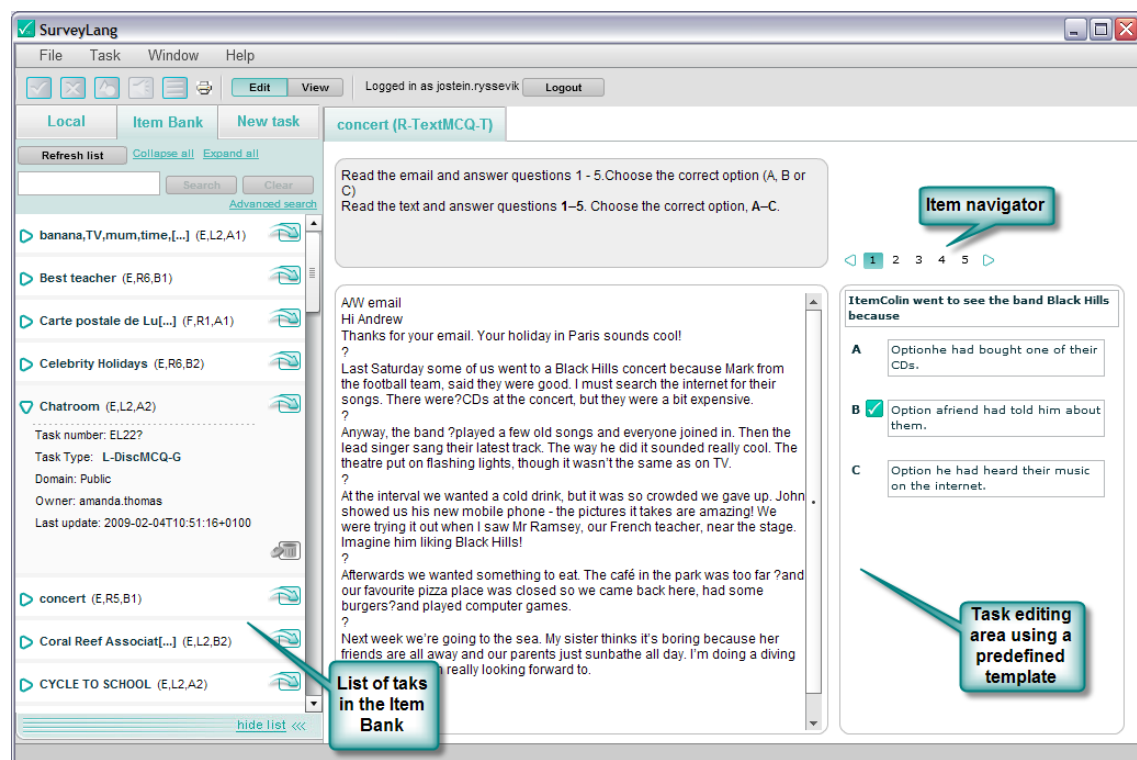
The test-items of the ESLC have been developed by an expert team of 40+ item writers distributed across Europe doing their work according to specifications and guidance provided by the central project team. Items have moved through various stages of a predefined life-cycle including authoring, editing, vetting, adding of graphics and audio, pilot-testing, Field Trial etc., each stage involving different tasks, roles and responsibilities.

The Test-item authoring tool was designed to support this distributed and fragmented development model. It was also designed to allow non-technical personnel to create tasks in an intuitive way by means of predefined templates for the various task-types that are used in the survey. At any stage in the development, a task can be previewed and tested to allow the author to see how it will look and behave when rendered in a test. The authoring tool also supports the capture and input of all the metadata elements associated with a task, including descriptions, classifications, versioning metadata, test statistics etc.

The tool is implemented as a rich client by means of technologies like Adobe Flex and Adobe Air. This provides a user-friendly and aesthetically pleasing environment for the various groups involved in the development of the tasks.

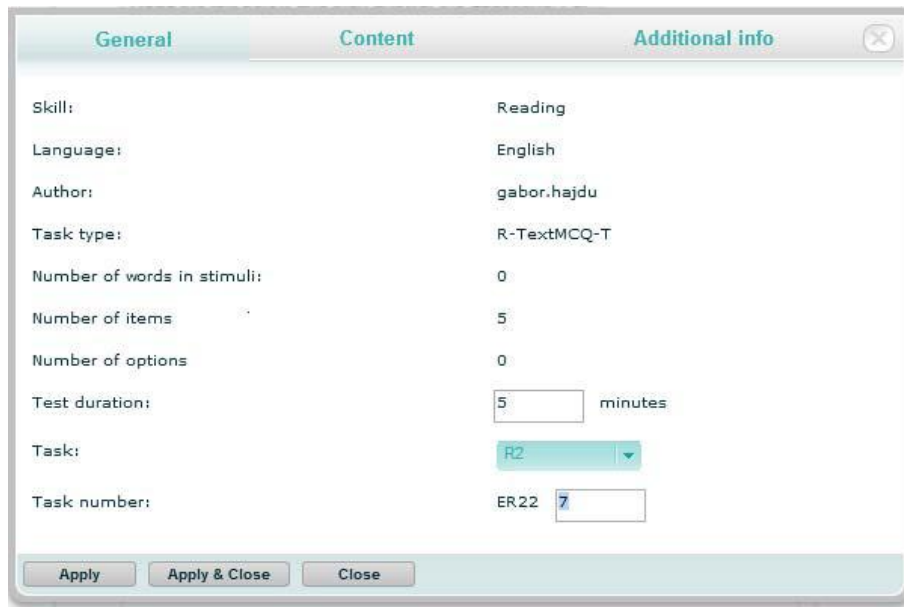
Below a few screenshots are presented:

**Figure 19 Item authoring using a predefined template**



The left navigation frame, shown in Figure 19 above, allows the user to browse the Item Bank, find or create tasks or upload tasks to the Item Bank etc. The various elements of the task are shown in the task display area to the right where the user can add or edit the content, upload multimedia resources like images and audio and define the properties of the task. The elements and functionality of the task display area are driven from a set of predefined task templates.

**Figure 20 Metadata input window**



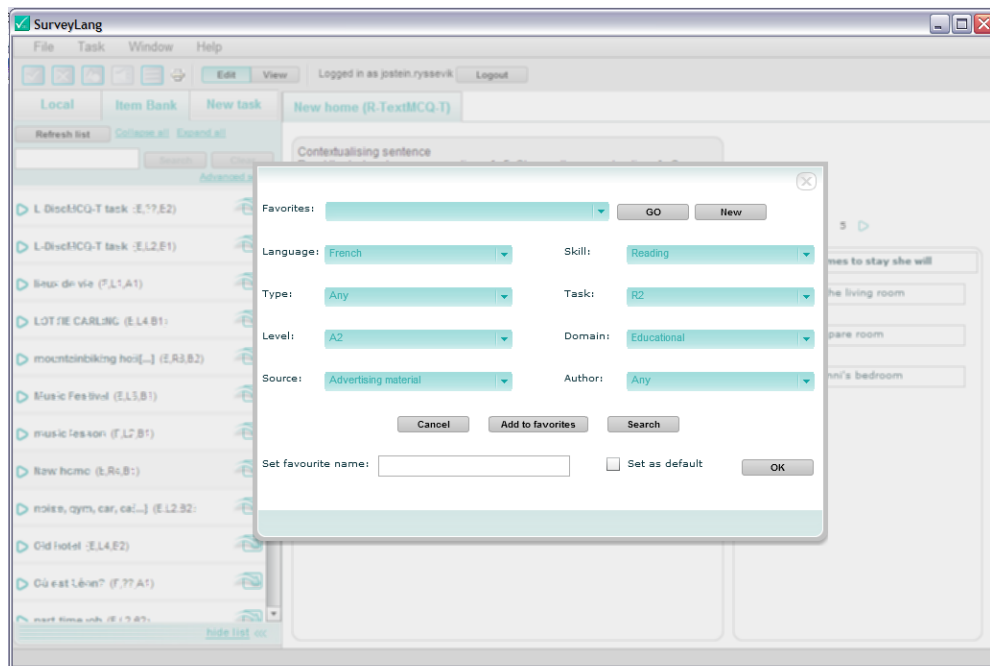
The screenshot shows a web-based form titled 'Metadata input window' with three tabs: 'General', 'Content', and 'Additional info'. The 'Content' tab is active. The form contains the following fields and values:

Field	Value
Skill:	Reading
Language:	English
Author:	gabor.hajdu
Task type:	R-TextMCQ-T
Number of words in stimuli:	0
Number of items:	5
Number of options:	0
Test duration:	5 minutes
Task:	R2
Task number:	ER22 7

At the bottom of the form are three buttons: 'Apply', 'Apply & Close', and 'Close'.

The content and structure of a task can be described by a series of metadata elements. Metadata elements are entered and edited in specially designed forms like the one displayed above in Figure 20.

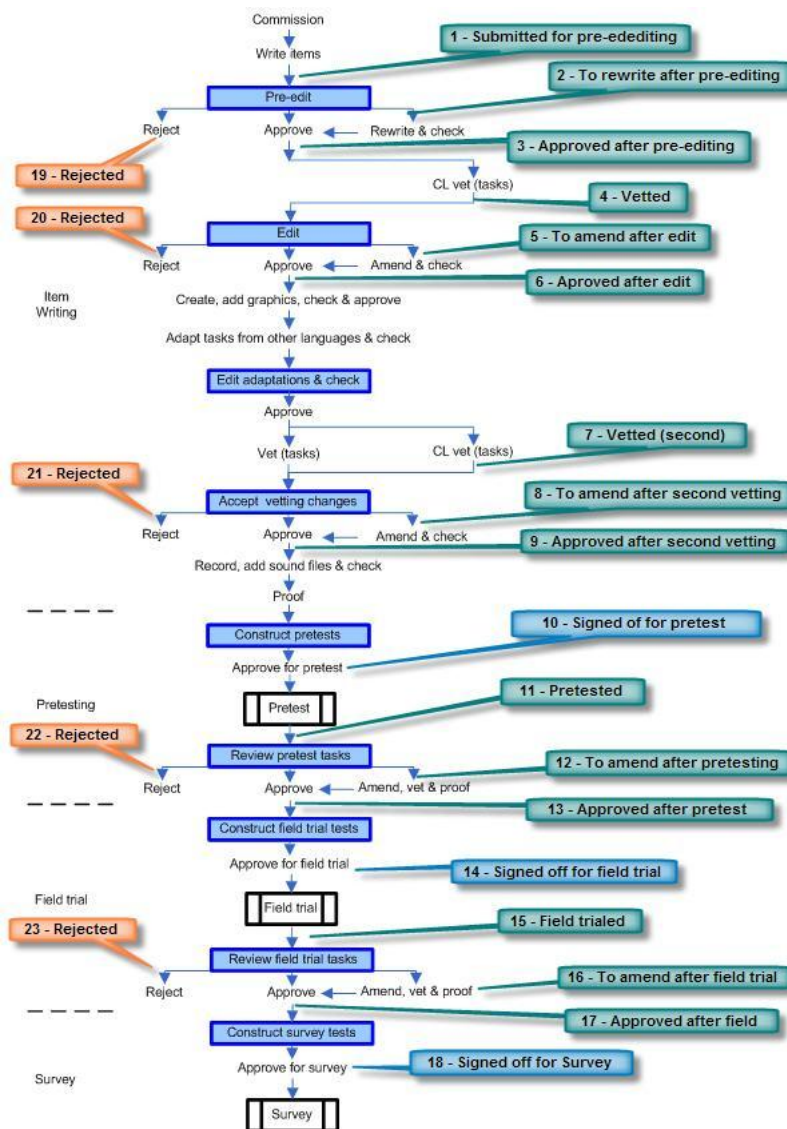
**Figure 21 Task search dialogue**



Tasks can be found in the Item Bank by searching their metadata. A free-field text search as well as an advanced structured search dialogue is provided. The example of the latter is displayed in Figure 21.

One of the metadata elements describes at what stage in the life-cycle a task is currently positioned. This defines who will have access to the tasks and what type of operations can be performed on the task. The structure of the implemented life-cycle is described in Figure 22.

**Figure 22 The structure of the SurveyLang task life-cycle**



As an integrated part of the life-cycle system, functionality-to-version and adapt-tasks have been implemented. When a task is versioned, any changes to this task will only affect the latest version of this task. Adaption is, on the other hand, a procedure that allows a task developed in one test language to be adapted to another language.

## 6.5 Test-item databank

The Test-item databank is the hub of the central system providing long-term storage, version control and management of test-items and their associated metadata and rich media resources. Test-items are uploaded to the item bank by the item writers to be seen and shared by others. When, as an example, a task has reached a stage in the development where an audio file should be added, the person responsible for this stage will download the task, read the audio transcript, create and attach the



soundtrack and load the task back up to the databank. The databank includes a version control mechanism keeping track of where the task is in the life-cycle as well as a secure role-based authentication system, making sure that only authorized personnel can see or change a task at the various stages in the life-cycle.

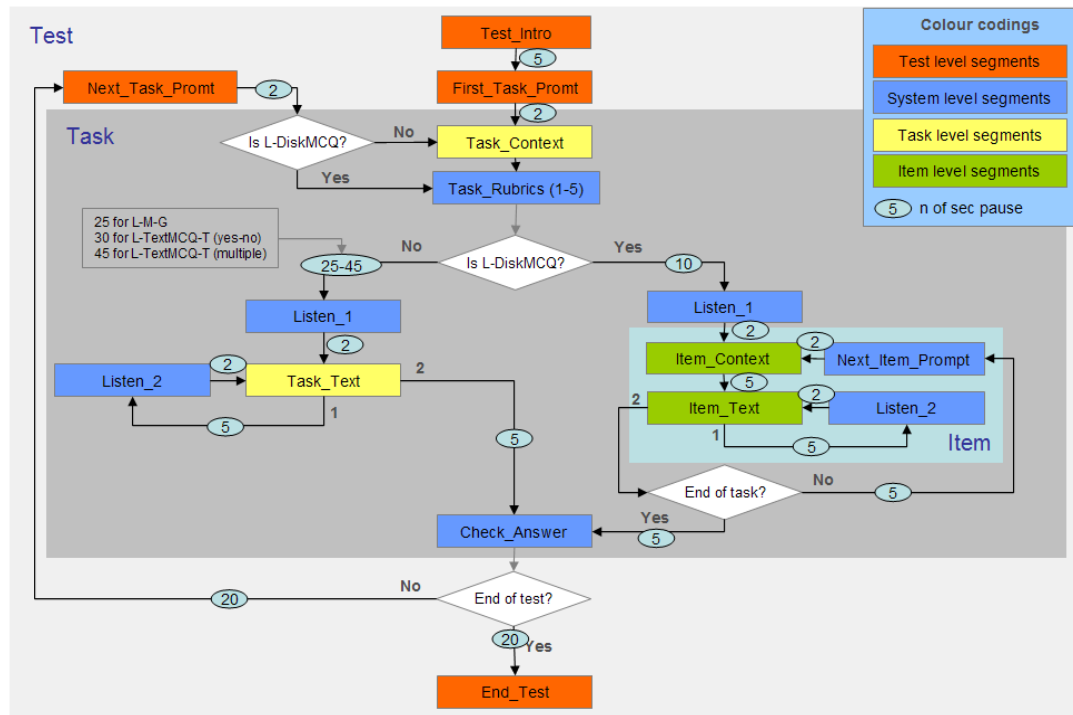
The Test-item databank is implemented in Java on top of Apache Tomcat and MySQL and communicates with the various remote clients through Adobe Blaze DS.

One of the most innovative features of the Item Bank is its ability to manage the audio tracks of the Listening tasks. Creating high quality audio is normally a time consuming and expensive operation. Traditionally the full length track of a task has been created in one go and stored as an audio-file. If a change is made to this task at a later stage, the audio-file is no longer usable and a completely new recording is thus required. To avoid this, an audio segmentation model has been developed whereby the audio files can be recorded as the shortest possible audio fragments. The various fragments are stored along with the other resources of the task and are assembled into full-length audio-tracks when the test materials are produced.

The basic principles of this audio segmentation model are shown below. The model consists of:

- Test level segments, which are reusable segments used to introduce and end the test as a whole, as well as to introduce the next task within the test.
- System level segments, which contain the fixed task rubrics as well as shorter prompts between items, like “Listen carefully” and “Please check your answers”. These are also reusable segments.
- Task level segments, which contain task specific audio.
- Item level segments, which contain audio specific to the various items within a task.

**Figure 23 SurveyLang audio segmentation model**



The model also specifies the number of seconds of silence between the various types of segments when these are assembled into full-length audio-tracks. The assembly of segments is handled by the system when the various test series are defined and as an early step of the test material production process.

## 6.6 Translation management

It goes without saying that a software platform developed for foreign language testing will need to be genuinely multilingual. Not only are language tests of a comparable level of difficulty needed for the five target languages but manuals and guidelines, navigation elements and the questionnaires are offered in all the questionnaire languages (see Chapter 5 for a definition of this term) of the educational systems where the tests are taken. Each concrete test presented to a respondent will thus have two different languages; a test language and the questionnaire language of the educational system where the test takes place. This requires efficient language versioning and text string substitution support. It also requires an efficient, robust and scientifically sound translation management system.

Gallup Europe had already developed a translation management system called WebTrans for their large scale international survey operations, amongst other the Commissions' Flash Eurobarometer project. This WebTrans system supports central management of translators scattered all over Europe (see Chapter 5 for further details

on translation). To allow for efficient use of WebTrans for the translation of questionnaires, an interface between that WebTrans and the Item Bank has been created.

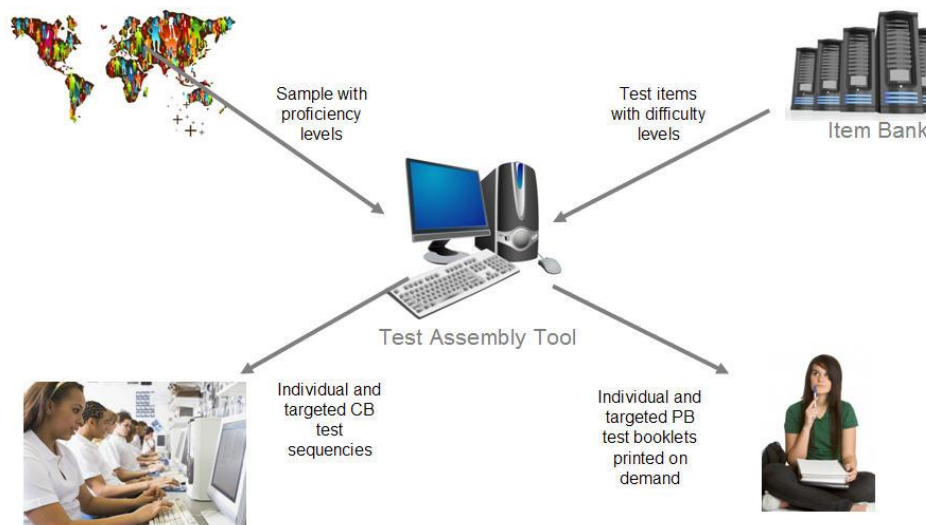
## 6.7 Test assembly

The Test assembly tool is without doubt the most sophisticated piece of software in the SurveyLang platform. The tool is designed to support four important functions:

- the assembly of individual test items into a high number of complete test sequences
- the allocation of students across these test sequences according to the principles and parameters of the predefined survey design
- the production of the digital input to the computer-based-test production unit
- the production of the digital documents to be used to print the paper-based test booklets.

This crucial role of the Test Assembly Tool is illustrated in Figure 24.

**Figure 24 The crucial roles of the Test Assembly Tool**



## 6.8 A: Test assembly

The assembly of test items into complete test sessions is driven by the test designs that are defined for the ESLC survey. This is a design that defines a series of individual test sequences (booklets) for each proficiency level.

An example of these principles applied to one skill section (Reading) is shown in the illustration below:

**Figure 25 Example of test design for a single skill section (Reading)**

tasks/booklets	time	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	b11	b12	b13	b14	b15	b16	b17	b18	b19	b20	b21	b22	b23	b24	b25	b26	b27	b28	b29	b30	b31	b32	b33	b34	b35	b36
A1-R1-1	5	1	1	2	3																																
A1-R1-2	5					1	2	2	3																												
A1-R4-1	5	2	3					1		2																											
A1-R4-2	5			3	2	3			1																												
A1-R5-1	5	3			1		2			3																											
A1-R5-2	5		2			1		3			1																										
A2-R4-1	7.5	4		5							1																										
A2-R4-2	7.5		5				4																														
A2-R5-1	7.5	5				4																															
A2-R5-2	7.5			5					4	2	1																										
A2-R6-1	7.5		4							5				1	2																						
A2-R6-2	7.5			4			5																														
A2-R8-1	7.5			4			5				1																										
A2-R8-2	7.5				5		4					3		1																							
B1-R8-1	7.5										1	4																									
B1-R8-2	7.5											4																									
B1-R10-1	7.5												3																								
B1-R10-2	7.5													4																							
B1-R12-1	15														3																						
B1-R12-2	15															5																					
B2-R10-1	15																																				
B2-R10-2	15																																				
B2-R12-1	15																																				
B2-R12-2	15																																				
B2-R13-1	15																																				
B2-R13-2	15																																				

Each line in this table is a task and each column is a test (or more accurately, the Reading section of a test administration). In this example a total of 25 individual tasks are used to construct a series of 36 unique tests. The 8 tests to the left are A1-A2, the middle group of tests are A2-B1 and the rightmost group are B1-B2. See section 2.5 for further details on the test design.

The illustration below is focusing on the A1-A2 group of tests:

**Figure 26 Example of a test design for a single difficulty level of a single skill section**

tasks/booklets	time	b1	b2	b3	b4	b5	b6	b7	b8	
A1-R1-1	5	1	1	2	3					4
A1-R1-2	5					1	2	2	3	4
A1-R4-1	5	2	3				1		2	4
A1-R4-2	5			3	2	3		1		4
A1-R5-1	5	3			1		2		3	4
A1-R5-2	5		2			1		3		4
A2-R4-1	7.5	4		5						2
A2-R4-2	7.5		5				4			2
A2-R5-1	7.5	5				4				2
A2-R5-2	7.5				5				4	2
A2-R6-1	7.5		4						5	2
A2-R6-2	7.5				4			5		2
A2-R8-1	7.5			4			5			2
A2-R8-2	7.5					5		4		2
testing time		30	30	30	30	30	30	30	30	

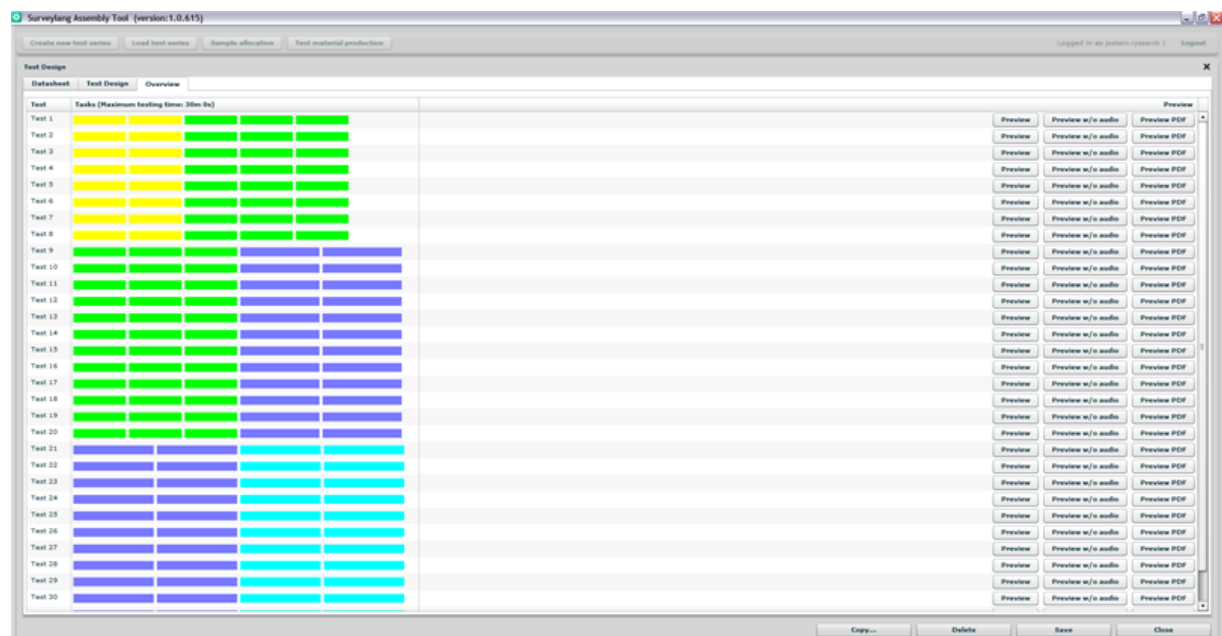
The first column in this table contains the task identifiers. The second column shows the testing time of each task and the bottom line the overall testing time for this skill section in each test. The column to the right shows the number of tests that each task

is included in. Coloured cells signal that a task is used in a test and the number shows the sequence in which the selected tasks will appear in the test. Similar designs are developed for Listening and Writing.

The Test Assembly Tool has a specialized graphical interface allowing the user to specify the test designs as illustrated above. Tasks in the Item Bank which are signed off and thus approved for inclusion in a test are dragged and dropped into the first column of the table. The content of each single booklet or test sequence is then defined clicking the cells of the table.

The Tool has another graphical interface which allows the user to inspect and preview the various test sequences which are created. An example of this interface is shown in Figure 27.

**Figure 27 Test preview interface**



The interface provides a graphical overview of the test series using colour-coded bars to indicate tasks of different difficulty levels. The length of a task bar is proportional to the length of the task (in seconds). Each line in this display is thus a booklet or test sequence. By clicking on the buttons to the right, the user will either be allowed to preview how a test will render in the test rendering tool for computer-based testing, or produce and inspect the test booklet which will be produced for paper-based-testing. This functionality proved to be very useful in the very last rounds of quality assurance of the test materials.

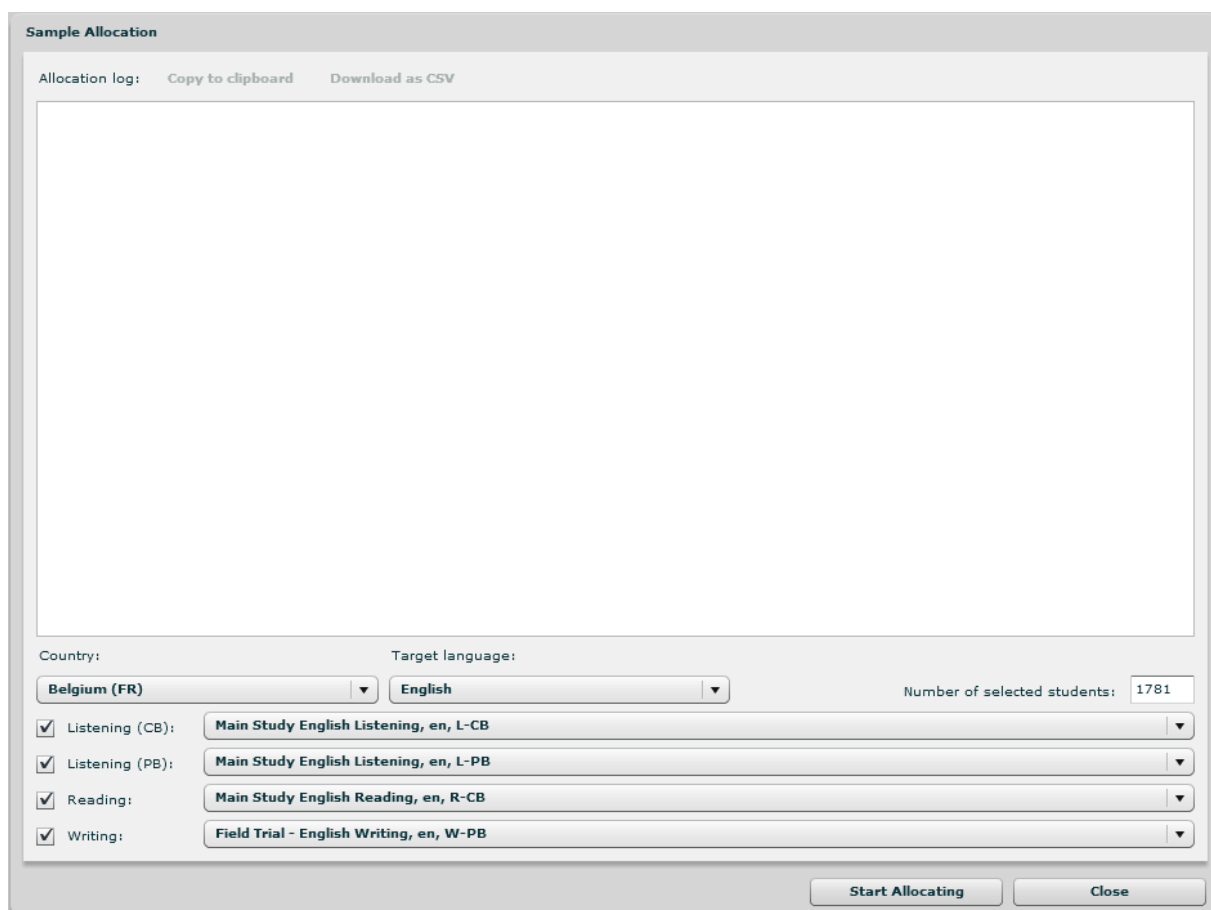
## 6.9 B: Allocation

The second role of the Test Assembly Tool is to decide which booklet or task sequence each single student will get. To accomplish this, the system is combining information from the previously described test design, with information about the student samples from the Sample Data Base. The latter is a database designed to enter and store the information about student samples (see chapter 4 for further information about the sampling process).

The allocation is partly targeted and partly random. The system makes sure that each single student receives a test sequence which corresponds to their proficiency level, as indicated by the routing test. The rest is random. For each single student the system first selects the two skills which the student will be tested in (from Listening, Reading and Writing) and then booklets for these skills are randomly selected from the set of available booklets corresponding to the student's proficiency.

The allocation process is managed through the interface displayed in Figure 28 below:

**Figure 28 Allocation interface**



The screenshot shows the 'Sample Allocation' window. At the top, there are links for 'Allocation log:', 'Copy to clipboard', and 'Download as CSV'. Below these is a large empty rectangular area. At the bottom, there are controls for 'Country:' (set to 'Belgium (FR)') and 'Target language:' (set to 'English'). To the right of these is a field for 'Number of selected students:' with the value '1781'. Below these controls are four rows of skill selection, each with a checked checkbox and a dropdown menu:

- ☒ Listening (CB): Main Study English Listening, en, L-CB
- ☒ Listening (PB): Main Study English Listening, en, L-PB
- ☒ Reading: Main Study English Reading, en, R-CB
- ☒ Writing: Field Trial - English Writing, en, W-PB

At the bottom right, there are two buttons: 'Start Allocating' and 'Close'.

The user will first decide which educational system (sample) and testing language the allocations should be run for. Secondly, the relevant test designs for the various skills are indicated. An allocation for an educational system/ language combinations takes about a minute and produces a log where the user can inspect the properties of the allocations. Due to the random nature of the allocation process, the resulting allocation will in many cases be biased in one direction or another. Normally it will therefore be necessary to run the allocation process a few times before a balanced allocation appears. Each allocation was reviewed and when satisfactory was signed off by the Project Director.

An example of the allocation log is displayed in Figure 29 below:

**Figure 29 Allocation log example for sample country**

Sample country is 100% paper-based
<b>Target language: English</b>
TOTAL NUMBER OF STUDENTS: 1000
NUMBER OF ALLOCATED STUDENTS: 1000
ALLOCATION TIME: 01:04
ERRORS DURING ALLOCATION: No errors
STUDENT DISTRIBUTION BY TEST TYPE: paper-based: 1000
STUDENT DISTRIBUTION BY SKILL: L: 656 R: 667 W: 677
STUDENT DISTRIBUTION BY DIFFICULTY: 1: 153 2: 322 3: 525
STUDENT DISTRIBUTION BY TEST: 01. EL1/1: 50 02. EL1/2: 50 03. EL2/3: 104 04. EL2/4: 118 05. EL3/5: 171 06. EL3/6: 163 07. ER1/1: 6 Etc.....

This is the log for a sample country consisting of 1000 students to be paper-based tested. We can see that the distribution by skill is balanced. By inspecting the

distribution by test, we can also decide to what extent the algorithm has produced a balanced allocation across the various test booklets that are available.

The following table shows the percentages of students for each educational system allocated to each testing mode (paper or computer-based).

**Table 26 Number of students allocated Main Study test by educational system and mode**

<b>Educational system</b>	<b>CB</b>	<b>PB</b>
French Community of Belgium	100%	
German Community of Belgium		100%
Flemish Community of Belgium	100%	
Bulgaria		100%
Croatia	21%	79%
England		100%
Estonia	17%	83%
France		100%
Greece		100%
Malta		100%
Netherlands	100%	
Poland		100%
Portugal	100%	
Slovenia		100%
Spain		100%
Sweden	28.5%	71.5%
<b>Grand Total</b>	<b>16390</b>	<b>42487</b>

## 6.10 Test materials production

The last role of the Test Assembly tool is to produce the test materials for computer-based as well as paper-based testing. The system implements a genuine two-channel solution where materials for the two modes can be produced from the same source.

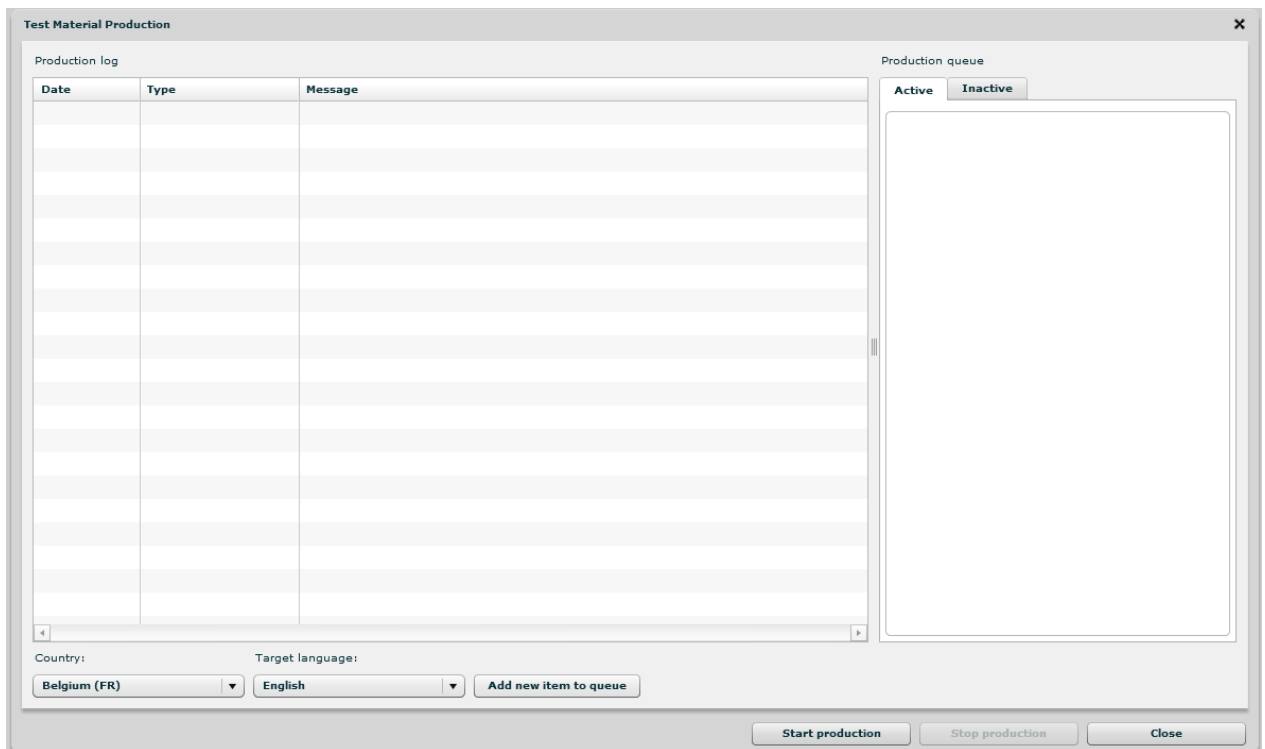
- For paper-based-testing the system produces test booklets in pdf format to be printed by the countries. It also produces the assembled audio-files to be used for the Listening tests in paper-based-mode.
- For computer-based-testing, the system produces the digital input to the test rendering tool including resources and audiotracks.

The materials production is managed through the interface displayed in Figure 30 below. The user decides which educational system and target language to produce for



and adds the request to the production queue. Several production requests can be started and run in parallel. Due to the fact that the system produces individualized test materials for each single student, the production for a single educational system/target language combination can take several hours.

**Figure 30 Test materials production interface**



When completed, the materials for paper-based testing are copied on DVDs (booklets) and CD (audio) and distributed to the countries. The materials for computer-based testing are transferred to the USB memory stick production unit for the physical production of test USB memory sticks. There was quality control over all of these steps with manual checks and sign-off of each stage in the process.

### **6.11 The USB memory stick production unit**

The test materials for computer-based testing were distributed on USB memory stick. The stick included the test material and the Test Rendering software, as well as an operating environment based on Linux (see more about this below). In order to produce the USBs in an efficient way, two specialized USB production units were built. Regarding hardware, the units are built from standard components, but specialized software was developed to manage the production process. A picture of the second USB production unit can be seen in Figure 31. The unit has slots to produce a kit of 28 USB memory sticks in one go.

**Figure 31 USB memory stick mass production unit**

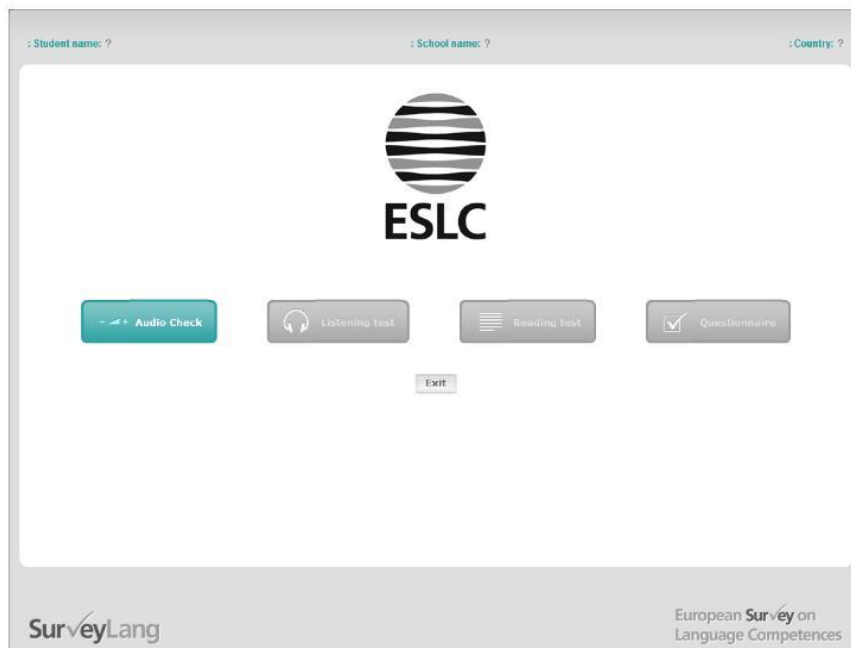


## **6.12 Test rendering**

The Test Rendering Tool is the software that delivers the test and the Student Questionnaire to the students and captures their responses. It is implemented to run on the test computers set up in each of the schools and are distributed on the USB memory sticks produced by the USB memory stick production units described in the previous section. The tool is implemented as a rich client by means of technologies like Adobe Flex and Adobe Air. It is designed to support the rich multimedia test format generated from the Test assembly tool.

Below is an example of the opening screen which allows the student to test the audio and to start the various skill sections. As skill sections should be taken in a predefined order, the next section cannot be opened before the previous is completed.

**Figure 32 Rendering tool opening screen**



An example of how tasks are rendered can be seen below:

**Figure 33 Rendering tool task display**



The navigation bar at the bottom of the screen is used to navigate through the test and also to inform the student about the length and structure of the test. Colour-codes indicate whether a task is completed or not.

Prior to testing, students were referred to the ESLC Testing Tool Guidelines for Students and demonstrations which were on the SurveyLang website (<http://www.surveylang.org/Project-news-and-resources/CB-Familiarisation-materials.html>). This ensured that students were familiar with both the task types and software system before taking the actual Main Study tests.

### **6.13 The USB-based test rendering operating environment**

One of the critical challenges related to computer-based delivery of assessment tests is, in general, security. On the one hand, it is crucial that the content of the tests are protected from disclosure before and during the testing period. On the other hand, it is of importance to create testing environments that are as equal as possible for everyone and where the students are protected from external influences of any sort (like access to the web, chatting channels, digital dictionaries etc.) while taking the tests.

If the tests could have been taken on dedicated test computers brought into the schools for that purpose, the problems would have been trivial. However, in a scenario where all tests are taken on the schools' existing hardware platforms, this is more of a challenge.

The solution developed by SurveyLang is literally taking full control of the test computers preventing any access to the computers hard disk, networks or any other external devices. This is done by booting the computers by a minimum level Linux operating system which only includes the components and drivers which are needed to run the Test Rendering software and to capture the students responses through the keyboard and the mouse.

The operating environment is distributed on the test USBs along with the test materials and the test renderer. All the USBs for a single educational system are in principle identical, as the USBs contains all the test materials for the educational system in both test languages. However, to increase the security, each kit of USBs (for a single Test Administrator) is encrypted with a different key. The materials on the USBs can only be unlocked by a Test Administrator's password in combination with a student password.

The USBs are also including a non-encrypted partition used to store the student's response data.

To store the necessary software and information, 4Gb USB memory sticks were required.

### **6.14 Data upload service**

One of the challenges of the described USB-based test delivery model is the fact that student response data from a single school will be distributed across a number of USB devices. To make it easy for the Test Administrators to consolidate and store these

despair data files, a Data Upload Service was provided. This is a web-based solution (implemented in Java), extracting relevant data from the test USBs one by one. The solution checks the integrity of the student data by opening files and comparing the incoming data with information from the sample data base. The system also provides a log where the Test Administrator can check whether all data files are uploaded or not.

### **6.15 Additional utilities**

Several other tools and utilities have been implemented to perform specific tasks, especially when it comes to data integration. Among others, these are:

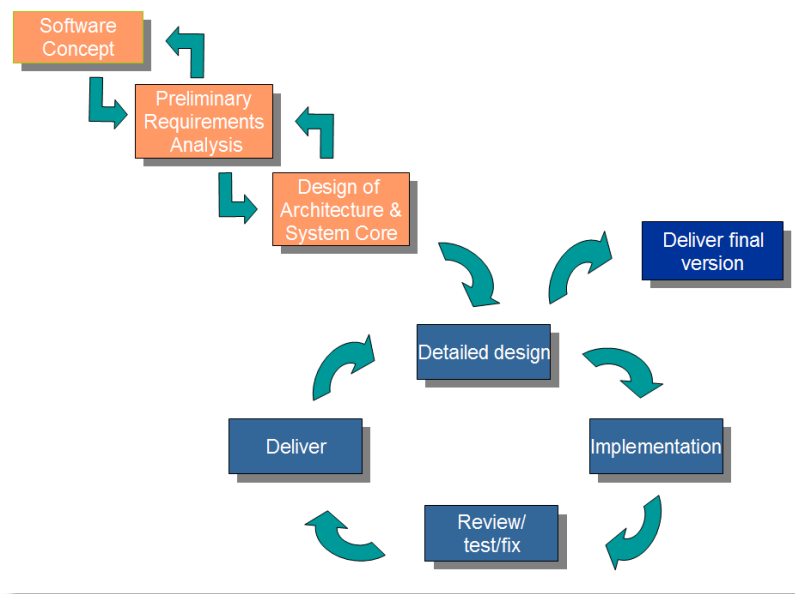
- A data entry tool for the paper-based test data (see section 7.12 for further details of this tool)
- A tool for coders to use for the open responses of the Student Questionnaire (see section 7.16 for further details of this tool)

### **6.16 Software quality and testing**

All parts of the software platform have been developed according to an iterative software development model called Staged Delivery (see Figure 34). In this model, an initial phase, including the requirements analysis and the technical specification of the architecture and the system core, is followed by an iterative process where the various components of the system are delivered in two or more stages. Each of these stages will normally include detailed design and implementation as well as testing and fixing. Towards the end of each stage the software is driven to a releasable state and made available for external testing by the representatives of the various user groups.

In addition to the testing done by external users, all software components are extensively tested by the software development team. This includes automated unit and build tests, as well tests of functionality and usability.

**Figure 34 Software development model**



All software code and documentation are stored in a state-of-the-art code management system (GForge) and a system and procedures for handling bugs and requests has been implemented. This code, together with components of the software developed specifically for the ESLC, have made available directly to the European Commission.

## **6.17 Performance**

The developed technology proved to be very efficient and robust. The system provided the necessary support for the test developers and was able to reduce the amount of manual errors to a minimum. The amount of mistakes and errors in the computer-based and paper-based test materials was negligible in the Field Trial as well as in the Main Study. It would have been impossible to implement a survey of this complexity without the support of a system like this.

The system also performed well during the test delivery phase. In the Field Trial, a data loss of 1.9 percent was caused by different types of issues related to the computer-based testing tools. To reduce this number even further, we systematically addressed all issues reported from the Field Trial and improved the system wherever possible. This was done by reviewing all issues reported to SurveyLang during the Field Trial and through the NRC feedback report and Quality Monitor reports (see Chapter 8 for further details). The amount of data loss in the Main Study was consequently reduced to 0.75 percent. These figures include all data loss occurrences and not only those relating to computer-based administration.

# ***Chapter 7: Field Operations***

## 7 Field Operations

This chapter provides an overview of the field operations for the European ESLC. Key in-country tasks and processes are discussed. *Note: the discussion focuses on the Main Study processes only. Field Trial processes are not discussed unless relevant.*

### 7.1 Overview of roles and responsibilities

The ESLC was implemented in each educational system by a National Research Coordinator (NRC). NRCs were typically assisted by a small team in a location referred to as the National Research Centre.

The NRC implemented procedures prepared by SurveyLang and agreed upon by participating educational systems. Their role was crucial in terms of managing the survey in-country and included quality control over every step.

NRCs appointed Test Administrators to administer the assessments in each school. School Coordinators were nominated by participating schools to liaise with the NRC and Test Administrator and also to undertake a number of preparatory tasks before the day of administration.

### 7.2 Key National Research Coordinator tasks

NRCs were responsible for implementing the survey within their own educational system. They:

- Acted as the key liaison person with SurveyLang.
- Signed a confidentiality agreement with SurveyLang and established procedures for the security and confidentiality of materials and data during all phases of the survey ensuring all NRC staff employed adhered to the confidentiality agreement signed with SurveyLang.
- Attended NRC training meetings led by SurveyLang concerning all aspects of the ESLC and passed on this training to relevant staff in-country.
- Recruited and trained staff in-country e.g. additional NRC support staff, sampling experts, translators, data entry staff, coding staff, Test Administrators, marking of writing staff, Quality Monitors.
- Negotiated specific aspects of the implementation of the ESLC with SurveyLang, such as administration of the routing test to all eligible students or sampled students only; amendments to standard procedures and national options for including country-specific questions in the questionnaire.
- Informed SurveyLang of any local legislation impacting upon SurveyLang procedures.





- Developed a communications plan for promoting school participation, effective implementation of the survey and dissemination of results amongst relevant national stakeholders.
- Ensured that technical standards were adhered to.
- Gave feedback on the development of the language tests, e.g. NRCs gave feedback on task types in 2008 during the pilot phase and they were involved in the sign-off process for the test specification and finalised task types in December 2009. They also had the opportunity to participate in pretesting in 2009 and give feedback on Field Trial tasks and tests.
- Gave feedback on the development of the questionnaires, e.g. NRCs gave feedback on the conceptual framework for the questionnaires as well as on the Field Trial versions of the questionnaires.
- Organised the translation, localisation and modification of all ESLC documentation necessary for administration, including the Student, Teacher and Principal Questionnaires and all operational documentation and manuals into the questionnaire language of the country (in some cases this was more than one language).
- Followed agreed procedures of document and version control.
- Prepared information for documenting issues related to sampling of the national educational system and the school sampling frame.
- Provided the list of eligible schools for SurveyLang to draw the school sample from.
- Provided the list of eligible students for each sampled school for SurveyLang to draw the student sample with.
- Ensured the required sample sizes and response rates were met.
- Organised the administration of the routing test with sampled schools prior to the survey administration.
- Organised the administration of the survey, including all logistical elements such as coordination with schools over dates and agreeing room plans and timetables according to SurveyLang specifications.
- Ensured each school appointed a School Coordinator to act as the key liaison between the NRC and the school. The NRC managed the School Coordinator in ensuring that a number of in-school preparatory tasks for the survey were completed prior to administration.
- Provided test administration dates to SurveyLang.
- Managed a help desk throughout the Main Study administration.
- Maintained a central Administration Issues Log (a document for recording any technical or administrative issues experienced) from Test Administrators which was forwarded to SurveyLang eight weeks after the end of the testing window.
- Provided a written report to SurveyLang on the operational processes following administration.
- Completed the NRC Questionnaire.



- Collated paper-based test papers and prepared them for marking of writing and data entry.
- Returned 150 multiply-marked Writing test booklets for each test language to SurveyLang for central marking.

The Field Operations Manual provided detailed information about the NRC's duties and responsibilities and was the NRC's main reference document for practical information about their role in administering the survey in-country. Supplementary documentation, with detailed information on particular aspects of the survey, was also provided; for example:

- Technical standards
- Translation manuals
- Sampling manual and guidelines
- Routing Test Instruction Sheet
- Routing tests and keys
- Test design tables
- Test Administration Manual (paper-based)
- Test Administration Manual (computer-based)
- School Coordinator Guidelines (paper-based)
- School Coordinator Guidelines (computer-based)
- Language Test Familiarisation Materials
- Testing Tool Guidelines for Students
- Room plans
- Frequently Asked Questions on the ESLC Basecamp website
- Data Entry Guidelines
- Coding Guidelines
- Quality Plan for Quality Monitors and the Quality Monitor Report
- Marking of Writing documentation
- ESLC Certificate of Participation
- SurveyLang Leaflet for Schools

The ESLC Basecamp website provided an additional and crucial source of support. More information on the ESLC Basecamp website is provided in the following section.

SurveyLang also provided a website <http://www.surveylang.org> which was kept up-to-date throughout the project to assist NRCs with their communications. Additionally, several brochures were made available; for example, a general leaflet about the survey and a brochure intended for schools, which was translated by NRCs and designed in print format by SurveyLang.

### **7.3 Communications between SurveyLang and NRCs**

The ESLC Basecamp website, a dedicated project management tool, was the main channel through which SurveyLang and NRCs communicated with each other during the course of the project.

Each educational system had their own private area on the ESLC Basecamp website to communicate securely with any member of the SurveyLang team. Messages could be sent by any member of SurveyLang or any member of the NRC team to query or comment on any task or aspect of the project. The central SurveyLang office received a copy of every message sent so that they could manage and track all queries, ensuring that NRCs were responded to as quickly as possible.

For the Main Study administration period, in addition to the support provided by the ESLC Basecamp website, a help desk was set up to provide immediate support for NRCs.

Task lists covering every aspect of the project were maintained in the form of 'To Do' lists and 'Milestone' tasks on each educational system's own area of the ESLC Basecamp website and were checked off after completion. In addition to the messaging and task management functionality, each educational system's area on the ESLC Basecamp website managed document control for all aspects of the project. Each educational system's final signed-off documentation was stored and categorised in their private area of the ESLC Basecamp website.

There were several shared areas for all NRCs on the ESLC Basecamp website. One area was called NRC Tasks which was where the general files and documentation provided by SurveyLang could be accessed by all NRCs. Messages of interest and relevance to all NRCs could also be found on this area. There was also an area called NRC Training Sessions where all information and documentation relating to each of the training sessions provided by SurveyLang were made available.

There were also 'Frequently Asked Questions' areas for specific aspects of the project, such as the marking of writing, data sets, sampling, Main Study general administration and Main Study computer-based testing administration. These areas were updated with key questions and answers about particular processes, allowing NRC teams to build up a store of practical knowledge.

A specific area for data analysts was also set up so that analysts could form a community of practice. In this area they could chat and communicate with each other on any aspects of interest in relation to the data sets and data analysis. SurveyLang also posted information of interest to all data analysts in this area.

## **7.4 Staff selection and staff training**

NRCs were responsible for recruiting, hiring and training additional staff necessary for the project based on the guidelines in the role specifications provided by SurveyLang. There were times when there was a greater need for further administrative support, for example, the months leading up to and during the Main Study administration.

NRCs were responsible for recruitment, training and ensuring the quality control of all work of the following staff:

- translators
- Test Administrators (external to school if at all possible)
- Quality Monitors
- data entry staff
- markers of writing
- coders

Role credential sheets for each of the above roles were provided by SurveyLang. NRCs were welcome to use additional criteria or, in cases where it was felt that the criteria were too strict, to discuss this with SurveyLang. The overarching responsibility for NRCs was to ensure that any changes made to the criteria specified by SurveyLang did not impact on the quality of the tasks carried out or on the data collected.

In terms of training, NRCs were required to:

- attend the centralised SurveyLang training sessions themselves or nominate an alternative NRC team member where appropriate
- pass on this training to staff they recruited through in-country training sessions
- provide all staff with the training documentation they needed, including any inserts or amendments that may have been released subsequent to the original manuals/guidelines
- ensure staff were clear on the survey and its aims
- ensure staff were clear on all tasks and deadlines
- be available to staff for questions that arose during the course of the project
- monitor tasks throughout and provide quality assurance to SurveyLang.

The following centralised training sessions were provided by SurveyLang

- Introduction to ESLC and Overview of Tasks and Work Areas (October 2009)
- Translation and Sampling (June 2009)
- Test Administration (January 2010, January 2011)
- Marking of Writing (for Team Leaders) (March 2010, March 2011)
- Analysis (September 2010, November 2011)

## **7.5 NRC sampling tasks**

SurveyLang was responsible for the sampling of both schools and students, which differed from the practice of other international surveys. This was done in order to minimise the potential for error and ensured uniformity in the outputs and more efficient data processing later on. It also relieved the burden of this task from NRCs. A web portal was provided for NRCs to manage the flow of data outlined below. This system was separate to the ESLC Basecamp website and was known as the 'Sampling Portal'.

NRCs were required to provide the following data to SurveyLang in order for SurveyLang to draw the school sample:

- available empirical information about foreign language training in ISCED2 and ISCED3 levels – the number and percentage of students taking the various languages, etc. so that SurveyLang could discuss and approve the two test languages and level of testing for each educational system
- explicit and implicit stratification criteria for each test language
- school level exclusion categories
- a full list of all schools that were eligible to participate in the survey

In providing the school sample, SurveyLang assigned two replacement schools to each sampled school so that where necessary, NRCs could use these schools to ensure their sample size and response rate requirements were met. For the specific response rate rules, see Chapter 4.

NRCs were required to provide the following data to SurveyLang in order to draw the student sample:

- student level exclusion categories so that SurveyLang could standardise a list of exclusion codes for all educational systems
- a full list of all eligible students for each sampled school.

NRCs were also required to list all teachers teaching the test language at the eligible level. Although no sampling was done at teacher level, this information was needed to provide teachers with access to the web-based Teacher Questionnaires.

In addition to the provision of the data above, specific sampling tasks included:

- contacting schools and engaging them to participate in the Main Study administration
- ensuring that the required school and student sample sizes and response rates were met.

After the Main Study test administration, NRCs were responsible for:

- uploading all school and student participation information
- resolving any discrepancies in the participation information with SurveyLang.



SurveyLang recommended maintaining a database of schools so that careful tracking of schools and their participation was possible.

Further details about sampling for the Main Study are provided in Chapter 4.

## **7.6 NRC pre-administration testing tasks**

NRCs were responsible for ensuring that schools administered the routing test, a short test which was designed to quickly elicit a student's proficiency. NRCs were also responsible for ensuring that the scores from the routing test were returned to SurveyLang so that students could be allocated a low, medium or high level test accordingly. Depending on the preferences of the educational system, the routing test was administered to all eligible students or to the sampled students only. The routing test scores were not used in the sampling process; they were only used for closely allocating students to tests (see section 2.5 for further information on the test design).

Prior to the actual Main Study administration, NRCs were responsible for ensuring that the School Coordinator received and used the Language Test Familiarisation Materials (for paper-based tests) and/or the Testing Tool Guidelines for Students (for computer-based tests). These Language Test Familiarisation Materials were sample materials representing a range of tasks and levels that were not necessarily the same level of difficulty that students saw in the actual Main Study tests. The materials were designed so that students could become familiar with the task types and task instructions which were in the test language for the Main Study tests. The Testing Tool Guidelines for Students, together with a demonstration, provided on the SurveyLang website, were designed to familiarise students with the computer-based testing screens and test format.

## **7.7 NRC test materials management tasks**

With regard to the Main Study tests (including Listening, Reading and Writing tests for paper-based administrations and Writing tests for computer-based administrations), SurveyLang was responsible for the creation of individualised test booklets. This measure was taken to ensure the standardisation of the survey materials and in order to help minimise the potential for error across such a complex test design. This decision also relieved the burden of this task from NRCs. Each language test booklet was pre-allocated to each sampled student and contained information such as school sampling ID, school name, unique ESLC ID (made up of codes for the country, questionnaire language<sup>23</sup>, test language, skill, booklet number and unique student data entry ID), and the student name. In some cases, educational systems chose to use their own IDs rather than names where regulations prevented this information being printed on the test booklets.

---

<sup>23</sup> See paragraph 3.2.3.1 in Chapter 3 for a definition of this term



Student language test booklets were provided on a DVD and sent to countries. The booklets were contained in the following file structure on the DVD: test language, language skill (i.e. Listening, Reading or Writing), school ID and then each individual student booklets which were identified by the unique ESLC ID.

For the Student Questionnaires, two non-personalised versions were provided, one for each test language. NRCs were responsible for 'over-printing' the above details (i.e. school sampling ID, school name, unique ESLC ID, student name) onto the Questionnaires. SurveyLang provided a spreadsheet with all the data needed for over-printing and NRCs were required to perform a test print and provide this to SurveyLang for review and sign-off.

There were no printing requirements for the Teacher and Principal Questionnaires as these were all administered in a web-based environment.

Other materials requiring high quality printing included:

- Routing test Instruction Sheet
- Language Test Familiarisation Materials
- routing tests and keys
- School Coordinator Guidelines (paper-based)
- School Coordinator Guidelines (computer-based)
- room plans
- Test Administration Manual (paper-based)
- Test Administration Manual (computer-based)
- Quality Plan for Quality Monitors / Quality Monitor Report
- Testing Tool Guidelines for Students
- marking of writing documentation
- ESLC Certificate of Participation
- SurveyLang Leaflet for Schools

NRCs were required to:

- Ensure the survey materials were printed to a high-quality by a professional printing company
  - SurveyLang provided the specifications for printing which included:
    - A3 paper
    - paper quality minimum 80gsm, preferred 100gsm
    - double-sided
    - centre stapled and folded
    - greyscale (SurveyLang created the PDFs in greyscale).
- contract a professional print company and send a test print run to SurveyLang for review at least one and a half months before the first administration date



- Manage the printing process by ensuring the quality of all test booklets and documents as specified above. They needed to check in particular that the layout and pagination had not been altered from the original and that the print and graphics were clear and legible.
- Ensure the process for over-printing on the questionnaires using the personalised spreadsheet provided by SurveyLang was done accurately and to a high-quality.
- Package the printed materials by school and dispatch to School Coordinators (this applied to paper-based administration and to the Writing tests only for computer-based administration). The following materials were also sent to School Coordinators at different intervals prior to testing: routing tests and keys, room plans, Language Test Familiarisation Materials, Testing Tool Guidelines for Students, the Student Tracking Form (a form which contained sampled student names, IDs, allocated test skills, level and booklets), Materials Reception Form (a form for schools to confirm that they have received test materials), log-in details for the web-based Principal Questionnaire and Teacher Questionnaires, student participation certificates and the SurveyLang Leaflet for Schools
- Package the Listening CDs and dispatch to Test Administrators (this applies primarily to paper-based administration but also as a back-up for computer-based administration should students need to change mode). Test administrators of computer-based tests were also sent test USBs containing both the computer-based system and individualised language tests and questionnaires, audio USBs for the Listening tests and back-up boot CDs should the USB not load the computer-based system and Test Administrator and student login and passwords for the USBs. The following materials were also sent to the Test Administrator at different intervals prior to testing: the Student Tracking Form, the timetable provided by school, Administration Report Forms (a form to record timing and conditions of each administration), Materials Return Form (a form detailing the number of completed and unused test booklets and USBs if computer-based administration) and the Administration Issues Log.

## **7.8 Key School Coordinator tasks**

School Coordinators were appointed by the school and acted as the key liaison with the NRC and Test Administrator over in-school preparation for the administration. Detailed School Coordinator Guidelines were provided as guidance for the role. Separate versions were available depending on whether the school had selected computer or paper-based administration.

The tasks that the School Coordinators were responsible for included:

- Student lists: providing the list of eligible students to the NRC who then forwarded the information to SurveyLang to draw the student sample



- Teacher lists: providing the list of language teachers who teach the test language at the tested educational institution
- Routing test: organising the administration of the routing test. Depending on the agreement the NRC had with SurveyLang, routing tests were administered to either a) all eligible students, or b) sampled students only. *Note: in no case did the routing test score impact on sampling. Sampling was performed randomly and was independent of any students' score on the routing test*
- Routing test: liaising with language teachers over the routing test administration and ensuring scores were recorded on student lists and returned to the NRC who then forwarded the lists to SurveyLang
- Selection of administration dates: School Coordinators gave their preferred dates to the NRC who confirmed the assigned date of administration. They discussed any necessary changes of dates with the NRC or Test Administrator as necessary
- Communications: informing school staff of the assigned administration date
- Communications: disseminating information about the survey to school colleagues, parents and students and obtaining consent if necessary
- Assessing the school's computers for computer-based testing: receiving the dummy USB stick, audio USB stick and back-up boot CD. Arranging for a computer technician to assess whether all computers proposed for the test administration met the technical specification outlined by SurveyLang and testing all Main Study administration computers with the dummy USB stick. Returning the 'Computer Facility Test Report Form' detailing the results of testing to the NRC. Informing NRC of a change to paper-based testing if necessary
- Student familiarisation in advance of administration (paper-based administration): providing teachers with the Language Test Familiarisation Materials
- Student familiarisation in advance of administration (computer-based administration): Ensuring teachers had access to the Testing Tool Guidelines for Students, the computer-based demonstrations as well as the Language Test Familiarisation Materials for Writing tests
- Student familiarisation: ensuring students had used the Language Test Familiarisation Materials and/or the Testing Tool Guidelines for Students so that students understood how they should respond to the different task types and were familiar with the computer-based testing system
- Room planning: receiving the student allocation which listed sampled students, allocated test booklets and levels to assist with room planning
- Room planning and timetabling: organising the assessment event itself (timetabling, arranging rooms, etc.) according to the rules set out by SurveyLang and finalising details with the NRC
- Room planning and timetabling: arranging an additional invigilator if necessary and ensuring that they had access to the documentation needed

- Room planning and timetabling: ensuring that each student knew where they had to be on the day of testing by creating individualised timetables if necessary
- Technical support: ensuring a technical support person was available for the day of administration and that they were available when the Test Administrator arrived if computer-based testing
- Test materials management: receiving test materials, confirming receipt with the NRC and storing in a secure location until the day of administration. Test materials included Language Tests and questionnaires for paper-based testing and Writing tests for computer-based testing. Other materials received included the Student Tracking Form. *Note: the Listening CDs were sent to the Test Administrator who brought them with them on the day of administration.* For computer-based testing, the Test Administrator brought all materials with them, for example, test USBs, audio USBs, back-up boot CDs and Test Administrator and student login and passwords for the USBs
- Communications: ensuring any language teacher observing the administration was aware they could observe only and could not participate in the administration
- Communications: notifying school staff and reception that there may be an unannounced visit by a Quality Monitor
- Student Tracking Form: identifying sampled students who could no longer participate
- Planning with the Test Administrator: talking to the Test Administrator by telephone and working through the preparations checklist provided ahead of the day of administration
- Teacher and Principal Questionnaires: providing teachers and the principal (or his/her nominee) with their web log-in details for their Questionnaires
- Teacher and Principal Questionnaires: ensuring the completion of the web-based Teacher and School Questionnaires, liaising with the NRC over response rates to ensure a high number of Teachers and Principals responded
- Day of administration: ensuring the test administration day goes smoothly by preparing the test room, ensuring that all students were present and assisting the Test Administrator and technical support person (if computer-based testing) as necessary
- Day of administration: being prepared to be observed and interviewed by the Quality Monitor
- Follow-up sessions: assessing the need for a follow-up administration with the Test Administrator and making arrangements if a follow-up administration is needed, keeping the NRC informed
- Student Tracking Form: completing the student tracking form with the Test Administrator after the administration
- Student Tracking Form: storing copies of the Student Tracking Form.

Note that in some cases, NRCs approached SurveyLang and requested the transferring of tasks that had been designated to the School Coordinator to the Test Administrator. These requests were assessed by SurveyLang on a case by case basis to ensure that the quality of the administration was not compromised by the changes.

## **7.9 Key Test Administrator tasks**

Test Administrators were appointed by the NRC to administer the survey in sampled schools. They also acted as the key liaison with the NRC and School Coordinator over in-school preparation for the administration. Detailed Test Administrator Manuals were provided. Separate versions were available depending on whether the school had selected to undertake computer or paper-based administration.

Test Administrators were primarily responsible for ensuring that the ESLC Language Tests and Questionnaires were administered the same way in all schools and in all participating educational systems. To maintain fairness, a Test Administrator could not be the language teacher of the students being assessed and it was preferred that they were not a staff member at any participating school.

The tasks that they were responsible for included:

- Training: attending Test Administrator training provided by the NRC. This training included a thorough review of the Test Administrator Manual and a walk-through of all tasks needed, including a close review of the individualised test booklet covers, the importance of administering test booklets to the correct students and the script to be delivered to students during the administration. Information to assist students with queries that they had regarding particular questions in the Student Questionnaire was also provided
- Test administration logistics: ensuring the NRC had details of availability and receiving information from the NRC about the schools to visit
- Documentation: reviewing documentation sent by the NRC including Test Administrator Manual, Language Test Familiarisation Materials, Testing Tool Guidelines for Students and the Demonstration version, section 2.4 of the School Coordinator Guidelines detailing the process for loading USBs (for computer-based administration), and a sample Student Questionnaire together with notes for assisting students with particular questions
- Test administration logistics: receiving School Coordinator contact details and the dates and times for the schools to visit
- Test materials management: receiving test materials and confirming receipt with the NRC and storing in a secure location. Test materials included CDs for paper-based Listening tests and testing kits of USBs for computer-based testing. These kits consisted of 1) test USBs containing both the test environment and individualised language tests and questionnaires, 2) audio USBs for computer-based Listening, 3) back-up boot CDs should the USB not be able to load the test environment 4) Test Administrator and student login and passwords for the USBs. Other materials received included the Student

Tracking Form containing the list of sampled students. *Note: the School Coordinator received the paper-based language tests and questionnaires directly*

- Prior to administration talking to the School Coordinator by telephone one to two weeks before the administration and working through the preparation checklist provided
- On the day of administration bringing all necessary documentation and materials as indicated on the materials preparation checklist in the Test Administrator Manual and anything additional as agreed with the NRC or School Coordinator
- Day of administration: meeting the School Coordinator, and additional invigilator if applicable, in advance of the administration and reviewing the Student Tracking Form as well as any other preparation details as necessary
- Day of administration: setting up the test room and test materials in advance of the administration
- Day of administration: setting up the test computers together with the technical support person for computer-based administration
- Day of administration: ensuring that notes on administration (security, attendance, observers, malpractice, student assistance) are adhered to
- Day of administration: being prepared to be observed and interviewed by the Quality Monitor
- Day of administration: distributing and collecting the test papers, administering and invigilating the test administration and ensuring that students received correct individualised tests. If computer-based testing, ensuring that the USB number each student used is correctly recorded on the Student Tracking Form against each student's unique ID
- Day of administration: managing any changes in administration from computer-based to paper-based testing if necessary using the back-up materials provided by SurveyLang
- Post administration: completing the Administration Report Form and Administration Issues Log if necessary
- Post administration: assessing the need for a follow-up administration and making arrangements with the School Coordinator if this is needed whilst keeping the NRC informed
- Post administration: completing the Student Tracking Form together with the School Coordinator
- Post administration: uploading of USBs to the central SurveyLang data server if computer-based administration
- Post administration: packaging up materials, checking all carefully and returning to NRC with the Materials Return Form
- Post administration: storing copies of the Student Tracking Form and Administration Report Form and responding to any queries from the NRC

Note that in some cases, NRCs approached SurveyLang and requested the transferring of tasks that had been designated to the Test Administrator to the School Coordinator. These requests were assessed by SurveyLang on a case by case basis to ensure that the quality of the administration was not compromised by the changes.

### **7.10 Key Technical Support Person tasks (if CB testing)**

A technical support person was nominated by the school coordinator in cases where the school opted for computer-based testing. The technical support person was responsible for the following tasks:

- testing all school computers prior to the day of administration with a dummy USB stick and reporting back findings to the NRC via the School Coordinator on the 'Computer Facility Test Report Form'
- helping to ensure that teachers had access to the Testing Tool Guidelines for students and the computer-based demonstration and assisting if any issues arose
- ensuring their availability on the day of administration to meet with the Test Administrator when they arrived
- preparing all test computers with the Test Administrator and being available if any issues arose
- ensuring test USBs were removed safely after testing so that they could be taken away by the Test Administrator and uploaded to the central data server.

### **7.11 Receipt of materials at the NRC after testing**

SurveyLang recommended that the NRCs keep a careful record of the assessment materials sent out so that they could carefully check back the materials as they were returned. NRCs received returned completed Language Tests and questionnaires, unused Language Tests and questionnaires, completed Student Tracking Forms, and completed Administration Issues Logs and Administration Report Forms.

NRCs prepared the test booklets for data entry and marking of writing. For computer-based testing, NRCs checked that all expected computer-based data had been uploaded to the central SurveyLang data server and accounted for any that had not.

### **7.12 Data entry tasks**

SurveyLang provided each NRC with customised data entry tools to enter their data from the paper-based testing booklets. There were separate tools for the data entry of the language test data and the questionnaires. These tools ran in Microsoft Excel and each educational system's data entry tools were designed by SurveyLang to accept all student data from that educational system. After entering elements from the ESLC ID, which were printed on each student's test booklet front cover, a customised form

popped up allowing data entry for that particular student. The student response data could then be entered directly into the data entry tools from the test booklets and questionnaires by specialist data entry staff recruited by the NRC.

The data entry process for the Reading and Listening tests was relatively simple as the data entry person entered each individual student's response (ranging from A to G) for each test question where a single test booklet contained a maximum of 27 single test questions. For Writing, the data entry person was required to enter two figures ranging between 1 and 5 (representing the mark awarded for each criteria, which were communication and language, see section 2.4.2 for more details) for each task that the student sat. A single Writing test booklet contained two or three tasks depending on the level of the test. As all Writing tests were administered in paper-based format, all NRCs had to arrange for the data entry of the Writing booklets to take place after marking.

For the Student Questionnaires, the process was similar to that of the language test data although. The students' questionnaires yielded three types of data, depending on the question type (see chapter 3). For closed single choice questions (with radio buttons), the data entry person had to type the number presented next to selected radio button. For closed free choice question (with check boxes) the data entry person had to type a 1 for each ticked check box and a zero for check boxes that were not ticked. For open (numerical or text) questions the data entry person had to type the text literally as it was written on the booklet of the student.

The number of questions and sub-questions varied slightly across educational systems. This is because some educational systems opted to include up to five country-specific questions and also because localisations of questions differed across educational systems. The response formats for the country-specific questions were the same as other questions in the questionnaire. There were only four open-response questions which related to parental occupation. These were to be entered as text exactly as the student had written them. The questionnaire tool adapted automatically to the appropriate number of questions and options in each educational system.

NRCs were strongly recommended to perform double data entry for all data.

Data Entry Guidelines for the language tests and questionnaires were provided separately and provided detailed information on the data entry process. They also provided guidance on how to review data that had been double entered and how to correct any entries if discrepancies had been found. Such discrepancies had to be resolved before the data was submitted to SurveyLang.

For the computer-based tests, the Test Administrator uploaded the Listening, Reading and Student Questionnaire data from the USBs directly to the central data server after each test administration.

As the Teacher and Principal Questionnaires were web-based, no entering or submitting of data was necessary after testing.



### **7.13 Marking of Writing**

The Writing tests had open responses and required marking by specialist staff recruited by the NRC before the data entry could be done. The Team Leaders for the marking of writing, appointed by the NRC, attended two training sessions (one before the Field Trial and one before the Main Study) provided by SurveyLang.

Following this centralised training, the Team Leaders passed on what they learnt from this training to the team of markers of writing in-country who were also appointed by the NRC. A standardisation exercise was included as part of this training so that Team Leaders could be confident that their team members were marking consistently and to the correct standard. It also meant that Team Leaders could identify any issues at an early stage and rectify them with the marker concerned. A spreadsheet was provided by SurveyLang which enabled Team Leaders, after entering the scores that markers of writing awarded for each task on a number of trial scripts, to assess both the consistency and level against the official scores awarded by SurveyLang. Consistency was defined as the ability to systematically distinguish superior and inferior performances, i.e. whether markers displayed a high correlation with the official scores awarded by SurveyLang. Level was defined as the comparison of the mean score of the marker and the reference marks, which monitored the ability to mark to the same average level as the official scores awarded by SurveyLang.

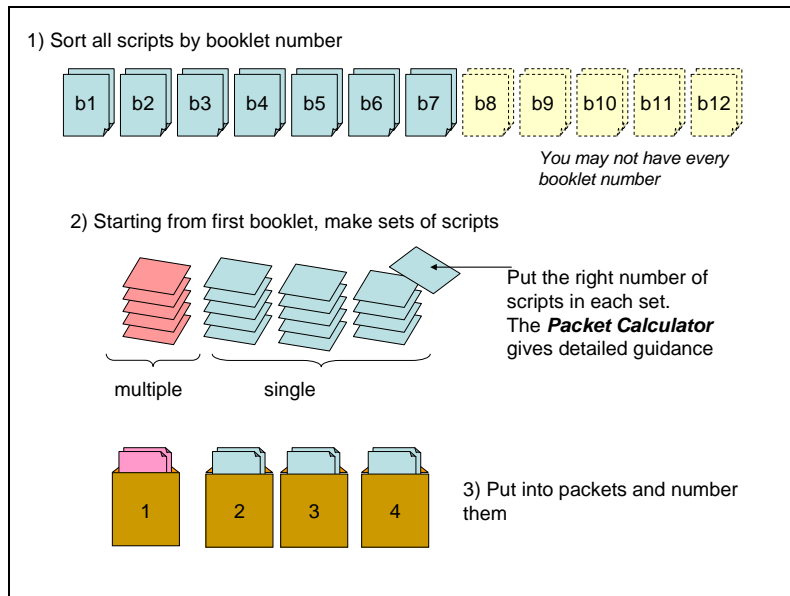
Extensive documentation was provided to support this process (see Chapter 2 for further details).

- Marking of Writing Administration Guide (for NRCs)
- Marking of Writing Guidelines for Markers (For Team Leaders and all markers)
- Exemplar Booklet (For Team Leaders and all markers)
- Training Powerpoint (For Team Leaders to use in their training)
- Accompanying Notes for Markers (For Team Leaders to use in their training)
- Standardisation Scripts 1 (For Team Leaders and all markers)
- Standardisation Scripts 2 (For Team Leaders and all markers)
- Standardisation Spreadsheet (For Team Leaders to use in their training)

Markers met at a central location in-country and marked the test papers over the course of several days. A proportion of the scripts for each test language (150) were multiply marked, that is, 150 scripts were marked by each member of the marking team while all other scripts were marked by a single marker only. A 'packet calculator' (see below) was provided in Microsoft Excel format to assist NRCs with the process of dividing the scripts up into 'packets' for marking. This included allocating the scripts for the multiple marking and ensuring a random allocation over a range of levels and test booklets. The 150 scripts which were multiply marked per test language were sent to SurveyLang at the end of the marking process where they were also centrally marked by the language testing group. Chapter 12 describes how this data was used in the analysis.


Before the day of marking, the NRC had to organise the scripts as shown in the figure below.

**Figure 35 Script sorting process by NRC**






The figure below shows how the packet calculator worked.

**Figure 36 Screen shot of packet calculator**


**SurveyLang**

### Packet Calculator for Marking Writing

Calculate the number of packets and the number of scripts in each packet

	Level 1	Level 2	Level 3		
What is the total number of students doing writing?	59	359	792	total students	1210
Number multiple-marked	14	42	93	total multiple-marked	150
Target script packet size	22	20	15		
Number of packets (multiple)	1	2	7	total packets (multiple)	10
Number of packets (single)	2	16	49	total packets (single)	67
total packets	3	18	56	Grand total packets	77
Construct packets in groups of	3	9	8		
	 1 multiple, 2 single	 1 multiple, 8 single	 1 multiple, 7 single		
Number of scripts in each multiple-marked packet	14	21	13		
Number of scripts in each single-marked packet	22	20	14		
<i>To the nearest whole number - may not be exact!</i>					
language	English				



Only the yellow boxes needed completing. NRCs entered the number of Writing scripts at each level, whereby:

- booklets 1-4 = level 1 (A1-A2)
- booklets 5-8 = level 2 (A2-B1)
- booklets 9-12 = level 3 (B1-B2)

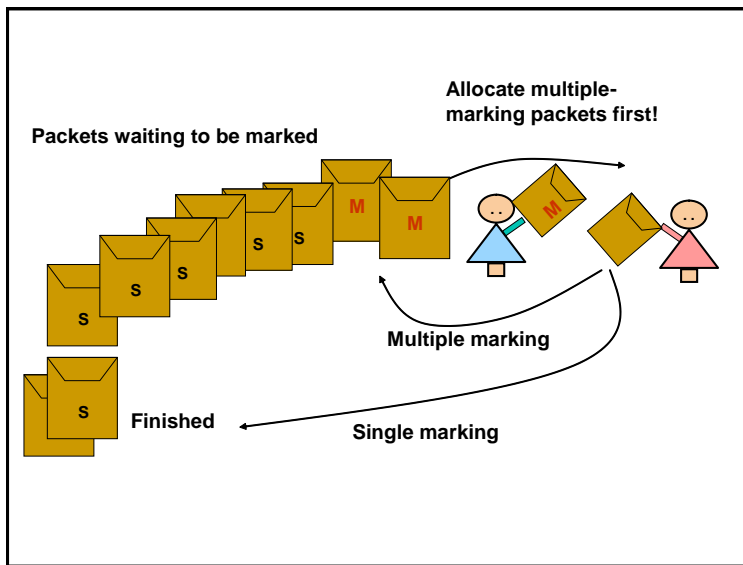
The calculator told NRCs:

- the number of packets for multiple- and single-marking
- the number of scripts to put in each multiple- and each single-marking packet
- how to select the scripts to put in each packet.

Different target packet sizes were identified per level, because high level scripts took longer to mark than low-level scripts.

A process for the Team Leader to allocate marking packets, prioritising the multiply-marked scripts was also defined as the figure below illustrates. This process is also described in the text below.

**Figure 37 Allocation process to prioritise multiple-marking packets**



A marker completed a packet and returned it to the Team Leader.

- When a single-marking packet was complete it went to the back of the queue of packets, i.e. it is finished.
- A multiple-marking packet was put back at the front of the queue of packets, and allocated to the next marker who has not yet marked it.

When all markers had marked a multiple-marking packet it was finished and goes to the back of the queue.

Team Leaders were asked to keep a record of who has marked which packet. A form was provided for SurveyLang for this purpose.

**Table 27 Sample form for tracking marking**

Packet	Single / Multiple	Marker ID					
		AB	NJ	KL	HK	PO	
1	M	✓✓	✓✓	✓✓		✓	
2	M		✓✓	✓✓	✓		
3	S	✓✓					
4	S		✓✓				
Etc .....							

One tick ✓ against the marker shows that the packet has been allocated. Two ticks ✓✓ show that the packet has been marked and returned.

In the example above, multiple-marked packet 1 has been completed by three markers and is currently with a fourth.

Single-marked scripts are only allocated to a marker once.

Once all the scripts were marked, they marks could then be entered into the data entry tool mentioned above. NRCs were strongly recommended to perform double data entry.

The Data Entry Manual for the language tests provided instructions on how to enter the data for the Writing tests and how to check for and correct any discrepancies between the first and second data entry. Any discrepancies found had to be corrected before the data was submitted to SurveyLang.

After the marking was completed and the data entered, NRCs were required to send the 150 multiply-marked scripts for each test language to SurveyLang for central marking. Chapter 12 describes how this data was used in the analysis.

## **7.14 Data submission**

Within eight weeks of the last testing date, NRCs were required to upload the following:

- all school and student tracking information showing final participation and any further exclusions that NRCs were not aware of prior to student sampling



- data for the language test booklets and questionnaires (via the data entry tools for paper-based tests or via the data upload portal for computer-based tests)
- the Administration Issues Log.

NRCs were required to keep electronic copies of all of the above documents as well as hard copies of the Student Tracking Forms and test booklets (completed and unused).

### **7.15 Data checking**

After submitting the data as described above, SurveyLang transferred the data into the database and carefully reviewed it before creating a report detailing any discrepancies between the student participation on the student tracking forms and data found in the data entry reports. For example, there were cases where it was indicated that a student was present during testing on the tracking form but no data was submitted and vice versa.

NRCs were asked to resolve any such queries as soon as possible. At times, this required the NRC to arrange additional data entry if it was found that some paper-based scripts had not been entered in the initial data submission. NRCs were formally notified when all queries regarding their data were resolved.

### **7.16 Coding of Student Questionnaires**

After the data checking process was complete, countries were sent questionnaire coding tools. These ran in Microsoft Excel and contained all of the student response data entered for the four open response questions relating to parental occupation in the Student Questionnaire. There were no open responses requiring coding in the Teacher and Principal Questionnaires.

To code parental occupation, the 1988 edition of the International Standard Classification of Occupations (ISCO 88) was used, including the Programme for International Student Assessment (PISA) modifications. The ISCO classification was developed by the International Labour Organisation (ILO 1991).

The ISCO 88 edition was used in preference to the ISCO 08 classification, which was adopted in 2007. This was primarily because PISA has used the ISCO-88 edition in its studies to date and it was necessary to link the ISCO codes used in the ESLC to the International Socioeconomic Index of Occupational Status (ISEI). This computation had not yet been done for the ISCO 08 ISEI index and the conversion tables showing the linkage between the ISCO 08 codes and ISEI index were not available at this time. This information, however, is available for the ISCO 88 codes.

Four-digit ISCO codes (ILO 1991) were assigned by coders to each of the four open responses for each student. Upon saving the file, the codes were checked

automatically and non-existing ISCO codes or missing codes were marked for the coder's attention.

**Figure 38 Open response questions in the Student Questionnaire**

---

**7 What is your mother's main job?**  
If she is not currently working, please tell us what her last main job was.  
*(Please write down the [job title], for example sales manager)*

---

**8 What does your mother do in her main job?**  
*(Please describe the kind of work she does or did in that job, for example manages a sales team)*

---

**10 What is your father's main job?**  
If he is not currently working, please tell us what his last main job was.  
*(Please write down the [job title], for example sales manager)*

---

**11 What does your father do in his main job?**  
*(Please describe the kind of work he does or did in that job, for example manages a sales team)*

---

The coding work was done by specialist coding staff appointed by the NRC. Coding guidelines, which very similar to the guidelines provided in the PISA Main Study Data Management Manual, were provided to assist in this work. These guidelines provided customised guidance for coders using the ESLC coding tools.

Additionally, NRCs were asked to refer their coders to the complete ISCO 88 Manual (ILO 1990) as well as background information, which included the complete list of codes including definitions and examples which were electronically available (in English, French and Spanish) on the ILO website (ILO 2010).

Quality control measures such as the use of a team of coders and having regular intervals where coders could discuss and analyse their use of codes were outlined in the Coding Guidelines. For this reason, five separate files for coding, each containing a random selection of 20% of all student responses, were provided.



NRCs were strongly recommended to conduct double coding for at least 10% of all student responses. For this specific purpose, an additional coding file was provided which contained a random selection of 10% of all student responses.

The ESLC coding tool enabled coders to identify cases where non-existent codes had been entered or cases where codes were missing. These had to be resolved before submitting the coding data to SurveyLang.

### **7.17 Main Study: a review**

NRCs were asked to complete a structured review of their Main Study operations. This provided NRCs with an opportunity to comment and provide feedback on various aspects of the implementation of the survey and to suggest areas that could be improved for future cycles. This was also an opportunity for NRCs to reflect on their own processes and procedures and comment on and formally record what worked well and what did not. This report was submitted to SurveyLang after the data submission process was complete.

NRCs had also been given the opportunity to comment and give feedback after the Field Trial. At this stage, a lot more detailed feedback was provided. Much of this was taken into account in SurveyLang's review of all operational processes and documentation between the Field Trial and Main Study. It was also taken into account in the feedback given by SurveyLang to NRCs in preparation for the Main Study.

### **7.18 References**

International Labour Office (1990) *International Standard Classification of Occupations, ISCO-88*, Geneva: International Labour Office.

International Labour Organization (ILO) (2010) *International Standard Classification of Occupations*, retrieved 15 January, 2011, from <http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm>

# ***Chapter 8:***

## ***Operations -***

### ***Quality monitoring***

## 8 Operations - Quality monitoring

This chapter provides an overview of the quality procedures employed for the ESLC. Note, the discussion focuses on the Main Study processes only, Field Trial processes are not discussed unless relevant.

### 8.1 *An introduction*

It is essential that users of the ESLC data have confidence that the data collection activities were undertaken to a high standard and for the purpose of creating an international dataset of a quality that will enable valid comparisons across participating educational systems. There were various methods, detailed further below, used to ensure this confidence.

The ESLC Technical Standards (which can be seen in Appendix 4) provided the set of standards upon which the data collection activities were based and were fundamental to the quality control methods employed by SurveyLang for the ESLC. There are three types of standards; each with a specific purpose. *Data Standards* ensure that all collected data can be added to the final ESLC 2011 dataset that will be released by the Commission. *Management Standards* ensure that all ESLC operational objectives are met in a timely and coordinated manner. *National Involvement Standards* ensure that the internationally developed instruments meet the highest standards of cross-national, cross-cultural and cross-linguistic validity and equivalence and that the ESLC results have the greatest possible meaning for national stakeholders.

The *Data Standards* outlined the standards for the following areas:

- target population and sampling
- translation
- test administration
- security of materials
- quality monitoring
- printing of material
- marking, coding and data entry
- data submission

All SurveyLang procedures were carefully developed and documented to ensure data of the desired quality. Quality monitoring played an important role and the implementation of the operational procedures documented in the guidelines and manuals made available to NRCs were continually monitored. In any cases where the documented operational processes were not fully implemented, these were logged and discussed with the NRC to understand the likely impact for the data. Quality monitoring was, therefore, the process of systematically observing and recording the extent to which data were collected and stored according to the procedures described in the



ESLC field operations documentation. Quality monitoring was a continuous process and was a shared responsibility between the NRC and SurveyLang.

The main elements of the quality monitoring procedures were:

- The ESLC Basecamp website: where all NRC communications, 'To Do' lists, project milestones and signed off documentation were managed
- Central Training: NRC attendance at central SurveyLang training
- Credential sheets: SurveyLang provided role definitions and criteria for the appointment of Test Administrators, Quality Monitors, Data Entry staff, Markers of Writing, and Coders for the open responses in the Student Questionnaire
- SurveyLang team: SurveyLang assisted NRCs in the planning and implementation of key processes. SurveyLang systematically monitored the key processes of sampling, translation, field operations and data entry, the marking of writing and the coding of the open responses in the Student Questionnaire
- Central Issues Log: all risks and issues relating to the implementation of operational processes were stored on a central register and regularly reviewed
- Quality Monitors: Quality Monitors observed the implementation of ESLC field operations at the educational system level, NRCs appointed a Quality Monitor to make unannounced visits to, typically, 10 schools in each educational system, interviewing both the Test Administrator and School Coordinator. The Quality Monitor also visited each National Research Centre and interviewed the NRC and five data entry staff. The lead Quality Monitor for each educational system wrote a full report on the results of their visits and interviews. Further details about their role are available below in section 8.4.
- ESLC Administration Report forms: ESLC Test Administrators completed a report after each ESLC test administration, thus providing an overview of the test administration across participating educational systems.
- ESLC Administration Issues Log: ESLC Test Administrators completed a log of all administration issues, thus providing an overview of any arising issues across participating educational systems.
- NRC report: SurveyLang developed a template report which allowed NRCs to systematically self-report on the implementation of key processes in their educational system.

## **8.2 Support for NRCs in quality monitoring**

The documentation that formed the basis for the quality monitoring procedures were:

- ESLC Technical Standards
- NRC Field Operations Manual





- Test Administrator Manuals (paper-based and computer-based including a script for test administration)
- School Coordinator Guidelines (paper-based and computer-based)
- Sampling Manual
- Translation Guidelines and Manuals
- Data Entry Guidelines (Listening, Reading and Writing and Student Questionnaires)
- Marking of Writing documentation
- Coding Guidelines
- Quality Plan for Quality Monitors and Quality Plan Report.

The quality monitoring instruments developed from these manuals and guidelines included role credential sheets, a range of sampling forms, WebTrans for the translation and verification work, a Test Administrator interview protocol for the Quality Monitor, a School Coordinator interview protocol for the Quality Monitor, a Data Entry interview protocol for the Quality Monitor, NRC interview protocol for the Quality Monitor, NRC feedback report templates, an ESLC Administration Report Form and an ESLC Administration Issues Log.

Credential sheets: as outlined above, SurveyLang provided role definitions and criteria for the appointment of Test Administrators, Quality Monitors, Data Entry staff, Markers of Writing, and Coders for the open responses in the Student Questionnaire

Sampling forms: SurveyLang developed a series of forms for collecting key data and for monitoring school and student sampling outcomes. The NRC and SurveyLang experts negotiated agreement on sampling plans and outcomes (see chapter 4).

WebTrans for translation and verification: this system managed the quality monitoring activities for translation of all documentation at the national level (see Chapter 5).

Test Administrator quality monitor interview protocol: a standard schedule was prepared by SurveyLang to systematically record the outcomes of the Quality Monitor's site visit. See Task 1 below in section 8.4 for further information.

School Coordinator quality monitor interview protocol: a standard schedule was prepared by SurveyLang to systematically record the outcomes of the Quality Monitor's site visit. See Task 1 below in section 8.4 for further information.

Data Entry quality monitor interview protocol: a standard schedule was prepared by SurveyLang to systematically record the outcomes of the Quality Monitor's site visit. See Task 2 below in section 8.4 for further information.

NRC quality monitor interview protocol: a standard schedule was prepared by SurveyLang to systematically record the outcomes of the Quality Monitor's site visit. The interview protocol recorded information on:

- the general organisation of the ESLC in that educational system

- the submission of sampling information
- the logistical arrangements
- the printing of materials
- the Helpdesk during the test administration window
- the return of assessment materials
- the data entry, marking of writing and coding activities
- the security of materials.

NRC feedback report: A standard report template was prepared by SurveyLang to systematically record the NRCs feedback on all operational processes and documentation. For example, the report template recorded information on:

- NRC structure
- staffing: the recruitment process and quality of staffing
- feedback on centralised SurveyLang training sessions
- communications with SurveyLang in different work areas
- feedback and ratings on all documentation provided by SurveyLang
- feedback on whether all Technical Standards were met
- sampling processes
- translation processes
- feedback on the questionnaire development process including agreeing the questionnaire language(s) and Localisation spreadsheet
- feedback on the Language Test development process, including the familiarisation and routing test processes
- feedback on the materials management processes
- feedback on the test administration and logistical processes including the Helpdesk
- feedback on the data entry, marking of writing and coding activities
- feedback on the schedule
- any other aspect of the project.

Administration Report Form: a form for the Test Administrator to record timing and conditions of each administration.

Administration Issues Log: a document for the Test Administrator to record any technical or administrations issues experienced.

### **8.3 Implementation of quality monitoring procedures**

The ESLC Basecamp website was fundamental to the quality monitoring process. As described in Chapter 7, all communications with NRCs took place and were stored on the ESLC Basecamp website. All final documentation was also filed there, either in



each NRC's dedicated and private area of the ESLC Basecamp website or in general areas if relevant to all NRCs.

SurveyLang also used 'To Do' lists and project milestones on the ESLC Basecamp website which were set up and then negotiated individually with each NRC as necessary to monitor the progress of each participating educational system. Main Study testing dates, country-specific requirements and SurveyLang operational imperatives provided the basis for negotiation of task deadlines and deviations from standard operating procedures. SurveyLang used the 'To Do' lists and milestones on the ESLC Basecamp website to monitor the progress of each NRC through key parts of the project and, when problems were identified, to advise on actions in order to minimise further operational problems and delays. Further information on the function of the ESLC Basecamp website can be found in section 7.3 of Chapter 7.

SurveyLang did not systematically visit all National Research Centres; however, it was necessary to visit several NRCs to spend additional time with them. This was required, for example, when the National Research Centre changed organisations and new personnel were appointed and also when a particular NRC had a substantive number of issues or queries to discuss with SurveyLang. The majority of support was provided by telephone and over the ESLC Basecamp website. Dedicated staff made regular contact with each NRC to ensure that they understood all tasks and that they were on schedule with their work. Any potential issues or risks were logged on the Central Issues Log and therefore SurveyLang knew of issues in advance and could work with the NRC to minimise the impact on upcoming requirements.

#### **8.4 ESLC quality monitors**

A detailed document for Quality Monitoring provided support for NRCs. This document described the NRC's role with respect to quality during the period directly *preceding*, *during* and *after* the Test Administration stage of the Main Study.

The procedures were designed so that improvements could be made following the feedback obtained from the Field Trial and also so that amendments could be made to procedures and processes within the Main Study test administration period. In this sense, the procedures were set not only to monitor quality, but to ensure that NRCs were supported in improving processes where possible within the test administration window.

NRCs were required to appoint a Quality Monitor for the Test Administration period. This person was appointed and paid for by the NRC. SurveyLang recommended that one person was appointed to this role with additional support in the form of further quality monitors appointed for task 1 below as necessary. Once employed, the NRC was required to send the name and contact details of their appointed lead Quality Monitor to SurveyLang.

SurveyLang provided a full credential sheet for Quality Monitors. In brief, the Quality Monitor:

- should have past experience of acting as a Quality Monitor on a similar project or in a similar role
- should have fluency in English and the questionnaire language
- should not be an employee at the same organisation as the NRC
- must not be an immediate relative of an employee at the NRC
- must not be line managed by the NRC
- must send their report directly to SurveyLang in electronic format
- must be able to and have the capacity to independently and effectively communicate with SurveyLang using email and telephone.

However any feedback that could be used by the NRC to correct the way tasks were managed *during* the Main Study were discussed with the NRC so the NRC could take quick action. Such instances were to be documented in the report.

The NRC had to meet with the Quality Monitor to:

- Train the Quality Monitor in the background of the ESLC.
- Make all operational documentation available particularly the School Coordinator Guidelines, Test Administrator Manual and the Data Entry Guidelines.
- Make background information about the project available, e.g. the Inception and Interim Reports.
- Make a plan of schools and data entry staff for the Quality Monitor to visit.
- Inform the Quality Monitor of variations agreed to standard SurveyLang procedures.
- Be available to respond to questions raised by the Quality Monitor. Where there are several Quality Monitors, one person should be appointed as the lead person and should be the key liaison with the NRC.
- Inform the Quality Monitor that they can ask questions directly of SurveyLang if they wish. Where there are several Quality Monitors, the lead person should be the contact with SurveyLang.

SurveyLang suggested that the Quality Monitor attended the NRC's in-country Test Administrator training.

The appointed Quality Monitor (or Quality Monitors in the case of Task 1 below) made unannounced visits to assess the quality of the implementation of SurveyLang processes and procedures carried out within the educational system.

Task 1: Test Administration: This task required the Quality Monitor to visit 10 schools; a mix across administration modes (computer and paper-based where both were used), regions and Test Administrators was preferred. The Quality Monitor was required to:

- Be at the school from 1 hour before the start of the administration to the end of administration when the Test Administrator and School Coordinator had packaged completed test materials up and completed the student tracking form.
- Interview the Test Administrator and School Coordinator separately for approximately 15 minutes. A template list of questions was provided for this purpose covering all aspects before, during and after the test administration focusing particularly on the logistical arrangements, materials management and clarity of documentation and processes.
- Write all responses down from the Test Administrator and School Coordinator interviews.
- Summarise the key findings from the interview in terms of what worked well and what did not work well for the Test Administrators and the School Coordinators. Were the procedures followed as specified? What problems were encountered? How were these resolved?
- Quality Monitors may also have wished to review some of the Administration Report Forms for the Test Administrations that they observed.

Task 2: Data Entry staff: This task required a quality check of the data entry work. The Quality Monitor had to review a sample of the data entry work from each data entry person employed by the NRC. The Quality Monitor was required to:

- Review a sample of five booklets per skill (Reading, Listening, Writing and the questionnaires) from the work each data entry person had performed.
- Check that the data was entered correctly for each data entry person. In the report, the number of mistakes found had to be entered. The data had to be corrected and the NRC informed so that they could take corrective measures if necessary.
- Interview each data entry person for approximately 5 minutes. A template list of questions was provided for this purpose, covering aspects such as the clarity of the guidelines, ease of using the tool and the data entry person's confidence of their accuracy.
- Write all responses down and document in the report what worked well and what didn't work well for each data entry staff member. Were the procedures followed as specified? What problems were encountered? How were these resolved?

Task 3: NRC report: The Quality Monitor was required to:

- Talk through each step in the period from mid January (including printing and the receipt of materials) to data submission for approximately 30 minutes. A template list of questions was provided for this purpose.
- Write all responses down and document in the report what worked well and what didn't work well for the NRC. Were the procedures followed as specified? What problems were encountered? How were these resolved?

## **8.5 Quality monitoring data**

The quality monitoring data collected from all of the documents and mechanisms in paragraph 0 were carefully reviewed and analysed both after the Field Trial and the Main Study.

Feedback was assessed by SurveyLang in order to improve all central and educational system processes and documentation after the Field Trial. All feedback and data was assessed again after the Main Study.

ESLC Quality monitor reports: each of the 16 educational systems submitted a Quality Monitor report on the conduct of testing sessions. The report consisted of a summary of their general observations together with a summary of the main findings from each of the different types of interviews conducted.

ESLC NRC Reports: each of the 16 NRCs submitted feedback reports on the overall processes and documentation for the Main Study.

In general, the quality monitoring reports and NRC feedback reports suggested a strong organisational base existed within educational systems for the conduct of the ESLC. The Quality Monitor reports indicated that, overall, NRC staff had a very good understanding of the operational aspects of the ESLC. The reports indicated that the ESLC administrations were conducted in a manner that was largely consistent with the documented procedures in the ESLC operations manuals.

Issues that were reported from the Main Study include:

- The required school sample size was not reached for one or both test languages (England, Greece). Note: to minimise the impact of this issue, additional students were drawn in the sample across participating schools.
- Some schools did not reach the required participation rate but did not inform the NRC and did not organise follow-up sessions (France).
- Some schools were unable to administer the Listening test in the required classroom setting because there were not enough rooms or test administrators available (German Community of Belgium).
- Data entry deadlines could not be met due to staff shortages. The team leader monitoring data input had to undertake data entry himself and therefore could not supervise others properly (Malta).
- Data entry deadlines could not be met due to staff shortages and financial restrictions which prevented the NRC from employing experienced staff (Greece).
- Data entry files became corrupted and could not be recovered requiring re-entry of data which resulted in delays (German Community of Belgium).
- There were complaints about the lack of information to parents before the administration (German Community of Belgium).



- Several issues were reported about the Student Questionnaire: a misprint and a translation issue with Question 47 (France).
- Translation inconsistencies were reported between the different operational documents (Estonia).
- The School Coordinator Guidelines were sent out to schools before the final sign-off had been given by SurveyLang (Sweden). Note: to minimise the impact of this issue, SurveyLang gave feedback and requested amendments be sent to schools.
- The required number of multiply-marked scripts for Writing was not met due to a misunderstanding between SurveyLang and the NRC (Sweden).

# ***Chapter 9: Data processing – Weighting***



## 9 Data processing - Weighting

This chapter deals with sampling weights, adjusting weights for non-response, and variance estimation.

### 9.1 *Motivation and overview*

Survey statistics are practically always **weighted**, i.e., any measurement or response for a person is given a specific weight when calculating the statistic. Weights are used for several reasons, of which the following are the most salient.

- (i) **To calibrate sample totals to the population totals.** While scientific research may be more interested in structural aspects of the data as expressed in averages, proportions, or regressions, policy makers typically have to deal with absolute numbers, i.e., population totals. In the context of SurveyLang, an example of a population total might be the number of students in a country that are studying a particular FL and have reached a certain level of proficiency. To arrive at correct estimates of the population total, the data for any individual in the sample is given a weight that is, in principle, equal to the inverse of the inclusion probability for the person (in practice, further adjustment may be needed).

**To avoid bias due to unequal sampling probabilities.** Members of the population are seldom sampled with the same probability. As long as the probability is known, we can still get unbiased results by weighting. Without weighting, statistics would assign the same weight to the each person, and that would slant the results towards persons who had a higher probability to be selected, possibly causing the sample results to deviate from what is true for the population. In this second function, weights will differ across persons, so they influence the estimation of all statistics, not just the overall magnitude of totals.

**Non-response** is an unwelcome but unavoidable complication in all surveys. It has a negative impact on both aims discussed above. When persons drop out, our estimates for the population totals will decrease. To counteract, we redistribute the weights of the non-respondents among those who did respond. Furthermore, non-responding schools and students may differ in important aspects from those who respond, so their absence may bias the results. To counteract, we try to redistribute their weights not in a general fashions, but among similar schools and students who did respond. All these operations rely on weights.

The computation of sampling weights is a rather complicated procedure that can involve many steps. Sampling itself may proceed in several stages (for instance, schools at the first stage, and students within schools at the second). In addition, there may be different adjustments for non-response, etc. A general principle is that the **final**

**weight** given to a person's data is a **product**, sometimes rather long, of various components and adjustment factors.

The sampling design of SurveyLang involves **two stages**. In the first stage, schools are sampled with a probability proportional to their size. If all students from the schools in the sample were to be tested, students from large schools would be overrepresented, a problem that can be easily fixed by using appropriate weights. However, the second stage samples the same number of students in each school, large or small. This means that, at the second stage, students are sampled with a probability inversely proportional to the school size. Following the principle that sampling weights at different stages are multiplied to produce the final weight, the inclusion probability for the individual would then be about the same in the final run. This property is described with the name **self-weighting sample**. If it really holds, we could use non-weighted statistics, and population totals could be obtained by multiplying all means and proportions with a constant factor.

Reality is not that simple because of the inevitable problems with non-response mentioned above, and as a result of **stratified sampling with disproportional allocation**. To increase precision and simplify logistics, sampling is done not from the entire population but separately for subpopulations (strata) that may have different sampling probabilities and different non-response rates. The largest "strata" are, in a way, the participating countries, which differ dramatically in population size but are represented with samples of approximately the same size. Within countries, there is stratification by school size and other school characteristics – for details on the stratified sampling design in SurveyLang, see chapter 4 on Sampling. Because of non-response and disproportional allocation, weights can vary considerably even in a design that is originally self-weighting.

So far, we have discussed the importance of weights for determining the statistic itself (the point estimate). Another important issue is how to estimate the **statistical precision** of the point estimates. The appropriate method does depend on the fact that inclusion probabilities are not necessarily equal. A standard technique involves taking many subsamples from the sample and inferring the **standard error (SE)** of the sample from the variability among the subsamples.

In this chapter, we concentrate on the following topics:

- (i) the computation of base school weights and base student weights  
the various adjustments made to the base weights to compensate for non-response at the school and student level  
the estimation of standard errors.

In general, the weighting procedures used for ESLC follow the standards of the Best Practices used for this type of complex survey. Similar procedures were used in other international studies of this nature including PISA (Programme for International Student Assessment), TIMSS (Third International Mathematics and Science Study) and PIRLS (Progress in International Reading Literacy Study).

## 9.2 *Base weights*

Since sampling for the ESLC involves two stages, the weight attached to the responses or properties of each individual student is the product of at least two components:

- A. A base weight for the school, which is inverse of the sampling probability for the school, and
- B. A base weight for the student, which is the inverse of the sampling probability for the student within his or her school.

Two additional factors are added to the product:

- C. A trimming factor for the schools is intended to compensate for imprecision's in the sampling frame
- D. The complex design of SurveyLang expects each student in the sample to complete a Student Questionnaire and to be tested in two out of three skills. As a consequence, there are four weights per student, and this fourth factor is the adjustment for sub-sampling students for two out of three cognitive tests

Prior to any adjustments for non-response, the student weight is therefore

$$A_s \times B_p \times C_s \times D_k$$

Throughout this chapter, we use the index  $s$  for schools, the index  $p$  for persons, and the index  $k$  for skills. The formulae are shown in a maximally simplified form, and the index shows the lowest (most detailed) level at which weights or adjustment factors differ. We now explain the computation of the four elements in more detail.

### 9.2.1 **A: School base weight**

The school base weight for school  $s$  is the inverse of the probability of selection of that school in the school sample. Based on a PPS (Probability Proportional to Size) sampling scheme, the school base weight can be calculated as:

$$A_s = M / (n \times \text{mos}_s)$$

where  $n$  is the sample size (of schools),  $\text{mos}_s$  is a measure of size for school  $s$ , and  $M$  is the total of the measures of size. When explicit strata are used, a sample is drawn from each stratum separately, and the formula above applies to the stratum, even if there is no explicit indexing. The formula applies to non-certainty selections, i.e., schools whose probability to be sampled was less than 1. For schools selected with certainty (i.e., schools with measure of size large enough to make the right hand side of the equation less than 1), the school base weight was set to 1. For some countries and some languages, a census was conducted and no sampling of schools was

undertaken at all. All schools from these countries (and languages) were assigned a base weight of 1.

The measure of school size ( $mos_s$ ) used for the PPS sampling is based on the estimated enrolment of eligible students in specific grades. For the purpose of completing school sampling on time, these estimates had to be generated in advance and were primarily based on similar enrolment figures from the previous year. Put simply, the number of students learning the language in the grade below the eligible grade was used as the best estimate. Obviously, such estimates cannot be completely accurate. In most countries, they were found to overestimate the population of students eligible for the survey.

### **9.2.2 Student base weight (B)**

The student base weight is the reciprocal of the probability for a student to be sampled within his or her school. In other words, it is obtained as the actual number of eligible students in the school divided by the number of sampled students in the school. Note that the student base weight is based on actual enrolment, not on the measure of size used to select the sample of schools.

When students within a school had to be further sub-stratified (in situations where there were students eligible for sampling in both languages within the same school), student base weights were calculated separately in each sub-stratum. The value of the student base weight was therefore always larger than 1 unless all students were sampled, in which case the student base weight is equal to 1.

### **9.2.3 School weight trimming factor (C)**

Trimming of school base weights was undertaken to avoid unexpectedly large values of the weights,  $A_s$ . This was found necessary for schools that turned out to be significantly larger than what was expected at the time of school sampling based on their estimated size as of that point of time. Trimming was done for schools where the revised enrolment (RENr) at the time of sampling exceeded at least 3 times the original measure of size ( $mos_s$ ), or the target cluster size (TCS) for the stratum, whichever of the two was larger. In such situations, the student base weight,  $B_s$ , might become excessively large, so it was decided to replace the original measure of size,  $mos_s$ , in the formula for the school base weight, by  $3 \times \max(TCS, mos_s)$ , whenever the conditions described above were met. Since the measure of size tended to underestimate actual enrolment in most countries, the number of instances where weight trimming had to be used was relatively small.

### **9.2.4 Student sub-sampling weight for skills (D)**

Once the student sample within the school was selected, each sampled student was assigned to two of the three skill tests (Reading Comprehension, Listening

Comprehension and Writing) at random. To account for this sub-sampling, a weighting adjustment factor,  $D_k$ , was calculated as the ratio of the total number of students sampled to the number of students assigned to a specific test,  $k$ . The value of  $D_k$  is hence about 3/2 but may vary somewhat because of the integer arithmetic involved. Since all students were expected to complete a Student Questionnaire,  $D = 1$  for this “skill”.

### 9.3 *Adjusting weights for non-response*

When schools or persons do not participate in the survey, their responses are lost. This has at least two important consequences:

- (i) Population totals will be underestimated. In a large country like France or Spain, a sampled student may represent 500 students in the population, so 20 missing students in the sample will decrease the population total by 10000. To avoid underestimation, we redistribute the weights assigned to non-responding students among those who did respond.
- (ii) Unless non-responding students are a perfectly random part of the sample, which they can hardly be expected to be, non-response can lead to biased estimates. This is a problem that affects not only totals but all kinds of statistics. To counteract, we try to redistribute the weights for the non-responding schools or students not in general, but among those schools or students that are as similar as possible to the non-responding ones.

This means that, to the four components of the individual student weight discussed above, we add four adjustment factors:

- E. A correction factor for schools that dropped out of the survey and could not be replaced;
- F. A correction factor for a small number of students who were supposed to be excluded from the sample but were nonetheless sampled;
- G. A correction factor for students who did not participate;
- H. A trimming factors for student weights.

In the final run, the formula for the individual weight becomes

$$A_s \times B_p \times C_s \times D_k \times E_s \times F_s \times G_p \times H_p$$

As explained, the index  $s$  denotes schools, the index  $p$  denotes persons, and the index  $k$  denotes skills, and we only show the index for the most detailed level at which weight components differ. We now explain the computation of the four correction factors in more detail.

### **9.3.1 Adjustment factor for school non-participation (E)**

The adjustment factor for school non-participation is based on the product (base school weight)  $\times$  (measure of school size), or  $A_s \times mos_s$ . The sum of this product for all schools in the sample is divided by the sum for the schools that did participate, yielding a result that is either 1 (when all schools responded) or larger than 1 (when some schools did not participate).

To reduce bias from non-response, the computation is not performed for the whole sample, but separately in the so-called non-response adjustment cells. These are groups of similar schools, usually based on either the explicit strata in which sampling was performed, or on the implicit strata proposed by the participating countries. Attention is focused on cells that do contain non-participating schools: these should ideally be as homogeneous as possible, which means rather small, but on the other hand they must contain a sufficient number of participating schools – otherwise, the adjusted weights may become idiosyncratic. We most often used the explicit strata, possibly merging some very small cells, but sometimes the implicit strata if they seemed a better way to provide reasonably homogeneous cells with a sufficiently large number of participating schools.

### **9.3.2 Adjustment factor for excluded students (F)**

Before adjusting student weights for non-response, we computed another adjustment to student weights to compensate for the exclusion of a limited number of ineligible students (blind, dyslectic, etc.) who had been sampled even though they were not supposed to appear in the sampling frame. As explained above, the student base weight in each school is the ratio of the number of eligible students to the number of sampled students. We adjusted for exclusion by subtracting the number of excluded students from both the numerator and the denominator.

### **9.3.3 Adjustment factor for student non-response (G)**

The adjustment factor for student non-response is calculated within each school, following the same logic as with the adjustment for non-participating schools. The sum of base student weights for all students sampled for a skill is divided by the sum of base weights for the students who were actually tested. Again, the result is 1 when all students did the test, or larger than 1 when some students did not participate.

### **9.3.4 Trimming factor for student weights (H)**

Trimming is a procedure for avoiding excessively large weights. At the second stage, student weights in a school have been trimmed by other similar surveys to four times the median weight in the explicit stratum containing the school. We have performed this trimming, which affects a very small number of students, say 50, without changing their weights dramatically.

### 9.3.5 The extent of non-response

SurveyLang has adopted rather strict quality criteria with respect to coverage and response rates at both school and student level. The Technical Standards define acceptable response rates for schools as at least 85% of sampled schools, and acceptable response rates for students as at least 80% of all sampled students.

Overall, we had a 93.6% rate of participation for schools in target language 1, and a 93.3% rate of participation for schools in target language 2. Table 28 shows more detailed data on school participation (after replacement of initial refusals) by country and target language (as far as sampling is concerned, the two target languages count as two separate surveys). Except for one country, the criterion of 85% participation at school level is comfortably met everywhere.

**Table 28 Number of participating and non-participating schools<sup>24</sup> and percentage of participating schools per country and target language**

	First target language			Second target language		
	No	Yes	%	No	Yes	%
Belgium (German Community)	0	9	100.0	0	9	100.0
Estonia	0	79	100.0	8	98	92.5
Spain	0	78	100.0	0	82	100.0
Croatia	0	75	100.0	1	76	98.7
Slovenia	2	71	97.3	3	89	96.7
Malta	2	55	96.5	2	55	96.5
Bulgaria	3	74	96.1	2	75	97.4
Portugal	3	72	96.0	0	76	100.0
Sweden	3	72	96.0	3	71	95.9
Belgium (Flemish Community)	5	70	93.3	2	72	97.3
Poland	8	81	91.0	8	71	89.9
France	7	67	90.5	4	70	94.6
Belgium (French Community) <sup>25</sup>	8	70	89.7	5	55	91.7
Netherlands	9	66	88.0	11	66	85.7
Greece	18	57	76.0	24	55	69.6

<sup>24</sup> Note that in cases where within a sampled school students completed the cognitive tests but no students completed the Questionnaires, the school is classified as non-participating.

<sup>25</sup> Note that the figures presented for the French Community of Belgium should read: for target language 2 (German) 4 schools out of 59 did not participate in the survey (and not 5 out of 60 as it is mentioned in the table).



Rates of participation are related to the magnitude of the adjustment factors for sampling weights. Studies similar to SurveyLang have used rules-of-the-thumb that the adjustment factor should not exceed a certain level. For instance, PISA sets the limit to 2, meaning that the weight of any participating school should not be more than doubled to compensate for non-participating schools; other surveys may even allow a maximum adjustment of 3. In the main survey, school adjustment factors had a median of 1.02 for target language 1, and 1.07 for target language 2. In the one difficult country, the largest adjustment factor was around 1.9 for the first target language, and 1.7 for the second target language, while the median was around 1.4 for both languages.

The average student participation rate within participating schools (excluding ineligible students) is about 90% for both target languages. Detailed data by country and skill are shown in Table 29.

Given a school participation rate of about 93% and an average student participation rate within participating schools of about 90% we have an overall student response rate of about 83.7%, well above the 80% target.

Because of the complex design for the three skills, we need a working definition of student participation within a school. We have adopted a criterion for a participating student as one who has responded to the Student Questionnaire (required of all students), and has done at least one of the two cognitive tests assigned. Based on this criterion, all schools that had not withdrawn from the survey had student participation rates above 25%.



**Table 29 Student participation rates within participating schools, excluding ineligible students, by target language, country, and skill (Q=Student Questionnaire, L=Listening, R=Reading, W=Writing)**

	First target language				Second target language			
	Q	L	R	W	Q	L	R	W
Flemish Community of Belgium	90.3	89.7	90.1	90.9	88.8	88.6	89.6	88.6
French Community of Belgium	89.9	89.7	89.0	90.9	92.4	92.6	91.9	93.9
German Community of Belgium	94.6	94.0	94.4	95.1	94.2	94.4	93.3	96.6
Bulgaria	87.2	89.5	89.9	75.2	88.8	91.5	92.1	79.1
Croatia	92.1	91.2	93.4	92.0	92.3	92.4	92.5	92.5
Estonia	92.5	92.9	93.1	93.2	92.2	92.6	92.5	92.8
France	91.1	91.2	91.5	90.9	88.6	90.3	89.7	87.6
Greece	95.0	94.7	95.7	95.5	92.9	92.8	93.1	92.6
Malta	86.2	87.4	87.5	86.5	79.9	81.6	81.4	79.5
Netherlands	87.2	87.5	87.8	88.9	90.0	90.4	89.8	90.9
Poland	85.5	85.0	86.0	85.7	87.9	87.7	87.9	87.8
Portugal	90.6	91.4	91.3	92.5	92.2	92.1	93.6	93.3
Slovenia	90.7	90.9	90.9	90.8	94.0	94.3	94.4	92.8
Spain	91.5	90.8	92.2	91.7	94.3	93.6	94.7	94.8
Sweden	87.4	90.1	90.0	89.2	85.3	85.8	86.8	84.7

Since one of the purposes of weighting and weight adjustment is to preserve the population total, it is of interest to trace the effect of these procedures on estimated population size. Summary results are shown in Table 30. At the stage when schools are sampled, the population total is estimated as the sum of the products (school weight \* measure of school size) over the whole sample. Note that this uses the measure of school size, a fallible projection of the number of students taking a target language at the specific school. When student weights have been calculated as the product of the school base weight and the student base weight, the estimate of the population total becomes the sum of student weights over the whole sample. Since student base weights are computed from actual enrolment in the target language rather than the estimated measure of school size, this brings about a change in the estimated population total. In almost all participating countries, the measure of size overestimated actual enrolment, so the estimates for the population total decrease. All other adjustments do preserve the total except for trimming, which slightly decreases the population total.

**Table 30 Projected population sizes at various stages of weight adjustment per target language**

	First target language	Second target language
A. Sum of measure of school size (MOS) for the population	2291384	1221855
B. Sum of measure of school size without exclusions	2241251	1217049
C. Sum of (school base weight * MOS) for the sample	2281699	1217624
D. Sum of (trimmed school weight * MOS) for the sample	2279061	1215502
E. Sum of (adjusted school base weight * MOS)	2281767	1217731
F. Sum of (adjusted trimmed school weight * MOS)	2279061	1215502
G. Sum of (adj. school base weight * student base weight)	2084512	1065780
H. Sum of (adj. trimmed school weight * student base weight)	2072241	1056497
I. Sum of (adj. trimmed school base weight * adj. student base weight)	2072470	1053262
J. As in J, but trimmed (see text for explanation)	2072368	1052939

All data in the table is based on weights for the Student Questionnaire. The weights for the three skills have been calibrated such that they also preserve the population totals.

For each student, we provide eight sampling weights: there are untrimmed and trimmed versions of the weights for the Student Questionnaire, and the weights for the three skills. All weights are based on the trimmed version of the base school weights, so the difference between trimmed and untrimmed student weights refers to trimming associated with adjusting for student non-response.

In addition to the cognitive tests and the Student Questionnaire, SurveyLang also includes a **Teacher Questionnaire** and a **School Principal Questionnaire**. The merging of information from the tests and the three kinds of questionnaires as part of the effort to explain student achievement results in multi-level problems that are far from trivial, especially in the context of complex sample design (Asparouhov 2006; Rabe-Hesketh and Skrondal 2006). If, on the other hand, the Teacher Questionnaire (TQ) and the Principal Questionnaire (PQ) are to be analysed separately, they can have sampling weights just like the Student Questionnaire (SQ), and adjustment for non-response follows the same logic (in fact, non-response for the TQ and the PQ tends to be more serious than for the SQ). By design, all teachers teaching the target language were supposed to fill in the TQ, so the teacher base weights to be redistributed for non-response are all equal to 1. There is only one principal per school so weights for the PQ are the same as school weights, except of course that adjustments for non-response are more noticeable as there is more non-response

among school principals than there is among schools. In practice, the adjustment of weights for the TQ and the PQ took place later than the adjustment for the SQ, and necessitated some small changes in the choice of the adjustment cells.

Non-response among teachers and school principals was much more variable, and hence sometimes higher, than among students (Table 31). Together with a previous decision on the part of the EC to disallow linking students to their teachers, this prevented a true multi-level approach to the regression analysis with indices originating from the TQ and the PQ. In the case of indices constructed from the PQ, we aggregated indices to the school level, using means for quantitative variables and the mode for a few categorical indices. The standard errors were estimated with the JK1 version of the jackknife: processing each country separately, there were as many jackknife replications as there were schools with useable teacher means, and each replication dropped one school. A similar approach was used when analyzing indices based on the PQ. In both cases, the dependent variables were plausible school means of the cognitive outcomes.

**Table 31 Teacher and Principal Questionnaire response rates**

	Teachers		Principals	
	TL1	TL2	TL1	TL2
<b>Flemish Community of Belgium</b>	83.1	75.6	82.4	78.3
<b>French Community of Belgium</b>	55.4	62.2	61.4	67.3
<b>German Community of Belgium</b>	50.9	50.0	40.0	50.0
<b>Bulgaria</b>	75.6	67.3	80.3	78.9
<b>England</b>	73.6	60.9	74.1	66.7
<b>Estonia</b>	87.0	89.0	89.2	93.2
<b>France</b>	34.9	36.4	61.2	68.6
<b>Greece</b>	66.9	71.3	80.0	83.6
<b>Croatia</b>	86.3	82.5	93.1	86.1
<b>Malta</b>	68.5	69.8	71.4	74.1
<b>Netherlands</b>	29.3	24.4	41.3	33.3
<b>Poland</b>	90.4	89.5	78.7	91.5
<b>Portugal</b>	92.2	88.4	89.7	83.8
<b>Slovenia</b>	90.3	91.1	85.3	88.2
<b>Spain</b>	69.6	80.7	82.7	84.1
<b>Sweden</b>	48.3	47.3	59.4	47.1

## 9.4 Variance estimation

Surveys like SurveyLang typically use some kind of **replication procedure** to estimate the standard error of their results (Lohr 1999; Wolter 2007). For instance, PISA uses Fay's modification of the balanced repeated replications approach (BRR). TIMSS and PIRLS rely on a variant of the jack-knife, JK2, (Westat 2007) that has sometimes been called JRR to emphasize its similarity to BRR.

Both BRR (including Fay's method) and JRR arrange primary sampling units (PSU) in variance strata containing two PSU each. To estimate standard errors, each result of the survey is computed as many times as there are variance strata, and the standard error is inferred from the variability in these replications. In each replication, the corresponding variance stratum is treated in a special way. When using **JRR, the method chosen for SurveyLang**, the data from one of the PSU in the stratum is ignored, and the data from the other one receives double weight. To compute the sampling variance of a result in the complete data, just sum the squared deviations of the replicate results from the complete data result. The estimate for the standard error is obtained as the square root of the sampling variance.

There are a number of details to consider:

- (i) The PSU in SurveyLang is usually a school. Thus, most variance strata contain a pair of schools as the variance units. The number of variance strata depends on the number of schools sampled in each country. We used a maximum number of 40 strata for target language 1, and 41 strata for target language 2. Not all strata are used in all countries. In countries where the number of sampled school is too small to fill all variance strata, the design matrix for the unused strata will only contain 1. In practice, this means that a number of extra computations will be performed, with no influence on the results.

Schools are sometimes sampled with certainty, in which case the individual student, rather than the school, becomes the primary sampling unit. Certainty schools are easily recognized by having a school base of 1. In practice, this occurred only in cases when the country sample was actually a census, as in Malta or the German Community of Belgium. In such cases, pairs of students were assigned to variance strata, using the maximum number of strata (40 for target language 1, and 41 for target language 2).

In countries where the number of participating schools is not even, one school cannot be assigned to a variance stratum. There is no generally accepted rule for such situations. For simplicity, we decided to treat the odd-one-out schools as certainty schools.

For compatibility with existing software, our data sets call the variance strata JKZONE and the variance units JKREP. The design matrix for the replicates is stored in variables with the generic name RPW#, where # stands for 1, 2 ... The actual replicate weights are obtained by multiplying each column of the design matrix with the sampling weights.

The way in which computations are organised then depends upon the software used. IEA's IDB Analyzer only requires a sampling weight (pick the appropriate one for the skill), the JKzone, and the JKrep; constructing the design matrix and multiplying its columns with the sampling weight are both performed automatically. When using the survey package in R, it is necessary to specify the sampling weights and the design matrix while multiplication is still automatic. Only the SPSS and SAS macros published by the PISA consortium seem to expect pre-multiplied replicate weights. We do not provide pre-multiplied replicate weights in the data sets because that would add at least 164 variables (41 per skill), and a potential for confusion.

In principle, there are two ways to compute the standard error of a statistic from replicates. Some authors (and computer packages) work with the squared deviations of the replicates from the mean of all replicates, while others take the squared deviations of the replicates from the complete data statistic. The former approach is closer to the idea of a standard error (SE), while the latter really estimates a mean squared error (MSE). There is no compelling theoretical reason to prefer one option over the other, and results are very similar in practice. The standard errors reported in the Final Report have been computed following the SE approach.

The estimation of the statistical margin of error for cognitive results (plausible values and plausible levels) also takes into account the measurement error. Details on this computation are the subject of Chapter 12.

## 9.5 References

- Asparouhov, T (2006). *General Multi-Level Modeling with Sampling Weights. Communications in Statistics* 35, 439-460.
- IEA (2009) *TIMSS Advanced 2008 Technical Report*
- Lohr, S (1999) *Sampling: Design and Analysis*. Duxberry: Pacific Grove.
- Lumley, T (2009) *Complex Surveys: A Guide To Analysis Using R.*, Wiley: New York.
- Monseur, C (2005). An exploratory alternative approach for student non response weight adjustment. *Studies in Educational Evaluation* 31, .129 -144
- OECD (2009). *PISA 2006: Technical Report*
- Rabe-Hesketh, S, and Skrondal, A (2006) Multilevel modelling of complex survey data, *Journal of the Royal Statistical Society. Series A, Statistics in society* 169, 805-827
- Westat (2007). *WesVar® 5.1 Computer software and manual*
- Wolter, K M (2007) *Introduction to Variance Estimation*, Springer: New York.

# ***Chapter 10: Data processing – Questionnaire indices***

## 10 Data processing - Questionnaire indices

The questionnaires yielded very rich data and numerous item response variables (326 item response variables from the Student Questionnaire, 348 item response variables from the Teacher Questionnaire, and 428 from the Principal Questionnaire). Most item response variables were not meant to be used as separate variables in the descriptive and regression analyses, but were meant to be combined in such a way that they, together, yielded a valid measurement of a single concept from the conceptual framework. This chapter describes how the item response variables from the Student Questionnaire, from the Teacher Questionnaire and the Principal Questionnaires were combined into indices for the final analyses<sup>25</sup>.

### 10.1 *Type of indices*

Because the main goal of the questionnaires was to gather empirical information on the malleable and concrete context of foreign language learning (namely the language policies within the Member States, see chapter 3), the majority of the concepts in the conceptual framework are so-called concrete concepts. These concepts refer to very concrete characteristics (e.g. class size), behaviours (e.g. the use of the testing language during the classes), situations or events. For concrete concepts simple or compound indices have been constructed.

Simple indices equal one single item response variable or a transformation of one single item response variable.

Compound indices: These indices were constructed through the arithmetical transformation of several items (such as a mean score or a sum score).

Within the conceptual framework also some abstract concepts are mentioned, such as 'perception of the foreign language lessons'. Abstract concepts are concepts which cannot be observed directly and indices (referred to as latent variables) for those abstract concepts are constructed using scaling procedures.

### 10.2 *Testing the structure of latent variables*

For these latent variables a confirmatory factor analysis was performed using LISREL (Jöreskog & Sörbom 2004) to test the theoretically expected factor structure and, if necessary, to re-specify the dimensional structure. The fit of the theoretical models

---

<sup>25</sup> The item response variables from the National Questionnaire were all meant to be used as separate variables in the description of the country profiles and some of the policy issues. Hence, for the National Questionnaire no indices were calculated.

was evaluated with an absolute fit index, the root-mean square error of approximation (RMSEA), and with the three incremental fit indices: the normed fit index (NFI), the non-normed fit index (NNFI), and the comparative fit index (CFI). For the RMSEA values lower than 0.10 are considered indicative of an acceptable fit and values lower than 0.05 of a close fit. For the incremental fit indices value above 0.90 are considered indicative of an acceptable fit and values above 0.95 of a close fit. For the results presented in this chapter, maximum likelihood estimation and covariance matrices were used for the analyses of the items (so all items were treated as continuous). The covariance matrices were obtained in the equally weighted sample. Cases with missing item responses were list wise deleted.

The reliability of the latent variables was assessed using Cronbach's alpha and an estimated Coefficient Alpha if the scale had a standard length of 10 items. In order to facilitate comparisons of scales with a different number of items, we used the Spearman-Brown prophecy formula to calculate Cronbach's Alpha for a (hypothetical) similar scale having 10 items. In the description of the indices in this chapter, for each latent variable (index for an abstract concept) the results of the confirmatory factor analysis and the reliability are presented.

### **10.3 Data preparation**

To ensure comparability of the item scores across educational systems, prior to constructing the indices the integrity and completeness of the data were checked and all item response variables were analysed in a fashion similar to the Field Trial (see chapter 3). The aim of the analysis was to detect questions that had a high item non-response internationally or locally (in a particular Educational system) and to detect potential misspecifications (in particular of the localised questions). Most, but not all, questionnaire responses were hard coded. Hard coded means that the codes used when registering the answers of the respondents matched the scoring rule mentioned in the source questionnaires. Whenever a misspecification was observed, i.e. when the registered scores did not correspond to the scoring rule of the source questionnaire, the cause was detected and the specification was corrected.

Because data obtained through open questions (like the duration of the class period) tend to display more often distributions that are problematic for analyses than data obtained with closed questions, the distribution of each open question was inspected separately. The distribution of the responses to each open question was plotted for each Educational system separately, on the basis of which the method of handling the outliers and normalising the distribution was determined. In the cases where outliers were handled prior to calculating an index, the applied method of handling outliers is described as well for each index. For the Principal Questionnaire several open questions and one closed question were excluded from further use, because they had



too high an average question non-response across countries (>5%), and/or too many outliers (yielding on average across countries less than 95% of valid answers)<sup>26</sup>.

Some questions were posed solely for enhancing the data quality and usability of the data. These questions were not combined into indices (with the exception of the questions meant to measure Economic, social and cultural status, see 0 about this index) or used for the final descriptive and regression analyses. Similarly, in cases when we measured the same concept in several questionnaires only one measurement was included in the descriptive analysis. In those instances, we used the data with the highest response rate (both at unit level and item level), to reduce the risk of a non-response bias<sup>27</sup>.

## 10.4 Student Questionnaire

### 10.4.1 Issue 1: Early language learning

#### Onset of foreign language learning (I01\_ST\_M\_S39B)

The “onset of foreign language learning” is a compound index (minimum converted score). The index equals the lowest grade selected in question SQ39 *'In which grades did you take foreign language lessons in school?'*. Prior to calculating the index the response options reflecting grades higher than the testing grade within each subsample (students sampled for the 1<sup>st</sup> and 2<sup>nd</sup> target language in each Educational system) were excluded from the calculation. The responses were converted such that they reflected comparable international grades (see Table 32) rather than national grades (1=first international grade; 2 = second international grade; 3 = third international grade up until score 11 = eleventh international grade).

---

<sup>26</sup> PQ01, PQ09, PQ10, PQ21, PQ27, PQ28, PQ29, PQ32, PQ33, PQ34, PQ35, PQ43

<sup>27</sup> Question solely meant for quality control and the enhancement of the usability of the data (e.g. comparison with other international surveys):

- SQ5, SQ9, SQ12, SQ18, SQ32, SQ53, SQ56, SQ58, SQ61 and the questions for the index of Economic, Social and Cultural Status: SQ07, SQ08, SQ10, SQ11, SQ13, SQ14, SQ19, SQ20, SQ21, SQ22
- TQ6, TQ8, TQ9, TQ10, TQ11, TQ14, TQ17, TQ20, TQ21, TQ25, TQ26, TQ27, TQ28, TQ37, TQ38, TQ44, TQ46, TQ47, TQ48, TQ52, TQ57, TQ58
- PQ05, PQ06, PQ08, PQ12, PQ13, PQ16, PQ20, PQ23, PQ24, PQ25, PQ26, PQ31, PQ37, PQ38, PQ05, PQ06, PQ08, PQ12, PQ13, PQ16, PQ20, PQ23, PQ24, PQ25, PQ26, PQ31, PQ37, PQ38

## Duration of foreign language learning (I01\_ST\_M\_S39A)

The “duration of foreign language learning” is a compound index (sum score). The index equals the total number of selected options in question SQ39 *'In which grades did you take foreign language lessons in school?'*. Prior to calculating the index the response options reflecting grades higher than the testing grade within each subsample (students sampled for the 1<sup>st</sup> and 2<sup>nd</sup> target language in each Educational system) (marked grey in Table 32) were excluded from the calculation. Given the high collinearity with the onset of foreign language learning (see 0) this measure was not used for the descriptive analyses.

**Table 32 ISCED level, testing grade and international grades in each educational system**

Country code	BE nl		BE fr	BE de		BG	UK-ENG	ES	EE	FR	EL	HR	MT	NL	PL	PT	SI	SE
	TL1	TL2		TL1	TL2													
ISCED level	2	3	3	2	3	3	3	2	2	2	2	2	2	2	2	2	2	2
Testing grade	8	10	10	8	10	10	11	10	9	9	9	8	11	9 or 10	9	9	9	9
2nd grade of ISCED3		10	10		10	10	10											
1st grade of ISCED3		9	9		9	9	9											
5th grade of ISCED2													11					
4th grade of ISCED2						8		10		9		8	10	10				
3rd grade of ISCED2						7	8	9	9	8	9	7	9	9	9	9	9	9
2nd grade of ISCED2	8	8	8	8	8	6	7	8	8	7	8	6	8	8	8	8	8	8
1st grade of ISCED2	7	7	7	7	7	5	6	7	7	6	7	5	7	7	7	7	7	7
6th grade of ISCED1	6	6	6	6	6			6	6		6		6	6	6	6	6	6
5th grade of ISCED1	5	5	5	5	5		5	5	5	5	5		5	5	5	5	5	5
4th grade of ISCED1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
3rd grade of ISCED1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2nd grade of ISCED1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1st grade of ISCED1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Before 1st grade of ISCED1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note. Flemish Community of Belgium: The 1st target language was tested in the last grade of ISCED2 and the 2nd target language in the second grade of ISCED3.

German Community of Belgium: The 1st target language was tested in the last grade of ISCED2 and the 2nd target language in the second grade of ISCED3.

Netherlands: Depending on the study program the last grade of ISCED2 is either the 9th international grade or the 10th international grade. The Netherlands the last grade of ISCED2 (the testing grade) depends on the study program. BE nl and BE de for the 1st testing language tested in last grade of ISCED2 and the 2nd in the 2nd grade of ISCED3.

### **Onset of target language learning (I01\_ST\_M\_S40B)**

The “onset of target language learning” is a compound index (minimum converted score). The index equals the lowest grade selected in question SQ40 *'In which grades did you take target language lessons in school?'*. Prior to calculating the index the response options reflecting grades higher than the testing grade within each subsample (students sampled for the 1<sup>st</sup> and 2<sup>nd</sup> target language in each Educational system) were excluded from the calculation. The responses were converted such that they reflected comparable international grades (see Table 32) rather than national grades (1=first international grade; 2 = second international grade; 3 = third international grade up until score 11 = eleventh international grade).

### **Duration of target language learning (I01\_ST\_M\_S40A)**

The “duration of target language learning” is a compound index (sum score). The index equals the number of selected options in question SQ40 *'In which grades did you take target language lessons in school?'*. Prior to calculating the index the response options reflecting grades higher than the testing grade within each subsample (students sampled for the 1<sup>st</sup> and 2<sup>nd</sup> target language in each Educational system) (marked grey in Table 32) were excluded from the calculation. Given the high collinearity with the onset of target language learning (see 0) this measure was not used for the descriptive analyses.

### **Foreign language lesson time a week (I01\_ST\_M\_S44B)**

“Foreign language lesson time a week” is a compound index (multiplied scores). The index equals the number of class periods (rounded to whole numbers) for all foreign languages together (SQ44 item 2) \* the duration of a class period/60 (SQ43). Outliers were replaced with the cut-off value 20.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. The responses to question SQ43 *'How long does a class period last at your school?'* were rounded up to the eight modes of the item response distribution (40=1 thru 40 minutes; 45=41 thru 45 minutes; 50=46 thru 50 minutes; 55=51 thru 55 minutes; 60=56 thru 60 minutes; 80=61 thru 80 minutes; 90=81 thru 90 minutes; 120=91 thru 120 minutes).

Outliers in the responses to second item SQ44 *'How many class periods do you have for the following subjects in a normal full week at school?: (2) For all foreign languages together (including Latin and ancient Greek)'* were replaced with the cut-off value 20

and the invalid answer 0 hours was removed (coded as invalid). If a response to item 2 was missing, the missing response was replaced with the response to item 1 (class periods for the subject of target language, see 0).

### **Foreign language learning time a week for homework (I01\_ST\_M\_S63B)**

“Foreign language learning time a week for homework” is a simple index (item score). The index equals the response to item 2 of question SQ63 *‘Generally, how much time do you spend each week on homework and assignments for the following subjects?: (2) For other foreign languages (including Latin and ancient Greek)’* and can have the following values: 0=Zero hours;1=Less than one hour a week;2=About one to two 3)hours a week;3=About two to three hours a week;4=More than three hours a week.

### **Target language lesson time a week (I01\_ST\_M\_S44A)**

“Target language lesson time a week” is a compound index (multiplied scores). The index equals the number of class periods (rounded to whole numbers) for the subject of target language(SQ44 item 1) \* the duration of a class period/60 (SQ43). Outliers were replaced with the cut-off value 10.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. The responses to question SQ43 *‘How long does a class period last at your school?’* were rounded up to the eight modes of the item response distribution (see 0).

Outliers in the responses to first item SQ44i01C *‘How many class periods do you have for the following subjects in a normal full week at school?: (1) For target language’* were replaced with the cut-off value 10 and the invalid answer ‘0 hours’ was removed (coded as invalid).

### **Target language learning time for tests (I01\_ST\_M\_S59A)**

“Target language learning time for tests” is a compound index (multiplied scores). The index is the response to question SQ60 *‘How much time do you usually study for a target language test?’* multiplied with the average of the responses to all items of question SQ59 *‘How often does your teacher of target language do the following?: (1) Give a target language test or assignment that is marked or scored; (2) Provide comments on a test or assignment you made’*.

### **Target language learning time a week for homework (I01\_ST\_M\_S63A)**

“Target language learning time a week for homework” is a simple index (item score). The index equals the response to item 1 of question SQ63 *‘Generally, how much time do you spend each week on homework and assignments for the following subjects?: (1) For target language’* and can have the following values: 0=Zero hours;1=Less than



one hour a week;2=About one to two hours a week;3=About two to three hours a week;4=More than three hours a week.

#### **10.4.2 Issue 2: Diversity and order of foreign language offered**

##### **Number of ancient languages learned (I02\_ST\_M\_S37A)**

The “number of ancient languages learned” is a compound index (sum score). The index equals the number of selected options referring to ancient languages in question SQ37 *'Which of the following foreign languages do you have or did you have as a subject in primary or secondary school?'*. On the basis of the localisation file (Taught Languages Table) for each country the options referring to ancient languages have been identified (see Table 33). The index can have the values 0=“No ancient languages”; 1=“One ancient language”; 2=“Two ancient languages”.

##### **Number of modern foreign languages learned (I02\_ST\_M\_S37B)**

The “number of modern foreign languages learned” is a compound index (categorised sum score). The index equals one plus the number of selected options referring to modern foreign languages other than the target language in question SQ37 *'Which of the following foreign languages do you have or did you have as a subject in primary or secondary school?'*. On the basis of the localisation file (Taught Languages Table) for each country the options referring to modern foreign languages have been identified (see Table 33). The index has the following categories: 1=“One modern foreign language” (sum score+1=1); 2=“Two modern foreign languages” (sum score +1=2); 3=“Three or more modern foreign languages” (sum score+1≥3).

##### **Number of languages studied before target language (I02\_ST\_M\_S41A)**

The “number of languages studied before target language” is a simple index (item score). The index equals the response to question SQ41 *'How many foreign languages did you study in school before you started studying target language?'* with the response scale 0=No foreign languages; 1=One foreign language; 2=Two foreign languages; 3=Three or more foreign languages.

##### **First foreign language studied in school (I02\_ST\_M\_S38A)**

The “First foreign language studied” in school is a simple index (converted item score), based on the converted responses to question SQ38 *'Which of the following foreign languages was the first foreign language that you were taught in school?'*. Based on the localisation file (taught language Table) the responses to SQ38 were converted to specific languages (see Table 33): 0=Ancient Greek; 1=Arabic; 2=Bengali; 3=Chinese; 4=Dutch; 5=English; 6=Finnish; 7=French; 8=German; 9=Hebrew; 10=Italian; 11=Japanese; 12=Latin; 13=Portuguese; 14=Russian; 15=Sami languages; 16=Spanish; 17=Swedish; 18=Turkish; 19=Urdu.

**Table 33 Ancient and modern foreign languages mentioned in question SQ37 and question SQ38 within each Educational system.**

	1st most widely taught foreign language (option 0)	2nd most widely taught foreign language (option 1)	3rd most widely taught foreign language (option 2)	4th most widely taught foreign language (option 3)	5th most widely taught foreign language (option 4)	6th most widely taught foreign language (option 5)	7th most widely taught foreign language (option 6)	8th most widely taught foreign language (option 7)	9th most widely taught foreign language (option 8)	10th most widely taught foreign language (option 9)
<b>BE nl</b>	French <sup>1</sup>	English <sup>2</sup>	German	<i>Latin</i>	<i>Ancient Greek</i>	Spanish	Italian	-	-	-
<b>BE fr</b>	Dutch	English <sup>1</sup>	German <sup>2</sup>	Spanish	Italian	<i>Latin</i>	<i>Ancient Greek</i>	Arabic	-	-
<b>BE de</b>	French <sup>1</sup>	English <sup>2</sup>	Dutch	Spanish	<i>Latin</i>	-	-	-	-	-
<b>BG</b>	English <sup>1</sup>	Russian	German <sup>2</sup>	French	Spanish	Italian	-	-	-	-
<b>UK-ENG</b>	French <sup>1</sup>	German <sup>2</sup>	Spanish	<i>Latin</i>	Urdu	Chinese	Italian	Russian	Arabic	Bengali
<b>ES</b>	English <sup>1</sup>	French <sup>2</sup>	German	<i>Latin</i>	Portuguese	Italian	<i>Ancient Greek</i>	-	-	-
<b>EE</b>	English <sup>1</sup>	Russian	German <sup>2</sup>	French	Finnish	Spanish	Swedish	Latin	Hebrew	Japanese
<b>FR</b>	English <sup>1</sup>	Spanish <sup>2</sup>	German	Italian	<i>Latin</i>	<i>Ancient Greek</i>	-	-	-	-
<b>EL</b>	<i>Ancient Greek</i>	English <sup>1</sup>	<i>Latin</i>	French <sup>2</sup>	German	Italian	Spanish	Turkish	Russian	-
<b>HR</b>	English <sup>1</sup>	German <sup>2</sup>	Italian	French	Spanish	-	-	-	-	-
<b>MT</b>	English <sup>1</sup>	Italian <sup>2</sup>	French	German	Spanish	Arabic	Russian	<i>Latin</i>	-	-
<b>NL</b>	English <sup>1</sup>	German <sup>2</sup>	French	Latin	<i>Ancient Greek</i>	Spanish	Arabic	Turkish	Russian	-
<b>PL</b>	English <sup>1</sup>	German <sup>2</sup>	French	Russian	<i>Latin</i>	Spanish	Italian	-	-	-
<b>PT</b>	English <sup>1</sup>	French <sup>2</sup>	Spanish	German	-	-	-	-	-	-
<b>SI</b>	English <sup>1</sup>	German <sup>2</sup>	French	Italian	Spanish	-	-	-	-	-
<b>SE</b>	English <sup>1</sup>	Spanish <sup>2</sup>	German	French	Finnish	Sami languages	Italian	-	-	-

Note. <sup>1</sup>=1<sup>st</sup> target language; <sup>2</sup>=2<sup>nd</sup> target language

### 10.4.3 Issue 3: Informal language learning opportunities

#### Number of first languages (I03\_ST\_A\_S04A)

The “number of first languages” is a compound index (categorised sum score). The index equals the number of selected options in question SQ4 ‘*Which language(s) did you speak at home as a small child (before the age of five)?*’. The index has the following categories: 1=“One language” (sum score=1); 2=“Two languages” (sum score=2); 3=“Three or more languages” (sum score $\geq$ 3).

#### Number of languages used at home (I03\_ST\_A\_S26A)

The “number of languages used at home” is a compound index (categorised sum score). The index equals the number of selected options in question SQ26 ‘*Which language(s) do you, yourself, speak regularly at home?*’. The index has the following categories: 1=“One language” (sum score=1); 2=“Two languages” (sum score=2); 3=“Three or more languages” (sum score $\geq$ 3).

#### Number of languages exposed to in home (I03\_ST\_A\_S25A)

The “number of languages exposed to in home” is a compound index (categorised sum score). The index equals the number of selected options in question SQ25 ‘*Which language(s) does your family speak regularly at home?*’. The index has the following categories: 1=“One language” (sum score=1); 2=“Two languages” (sum score=2); 3=“Three or more languages” (sum score $\geq$ 3). Given the high collinearity with the number of languages used in the home (see 0) this index was not described in the Final Report.

#### Parents target language knowledge (I03\_ST\_A\_S28A)

“Parents target language knowledge” is a compound index (mean score). The index equals the average of the responses to all items of question SQ28 ‘*In your opinion, how well do your parents know target language?*’

#### Target language exposure in home (I03\_ST\_A\_S25B)

“Target language exposure in home” is a simple index (item score). The index equals the selection of option 5 in question SQ25 ‘*Which language(s) does your family speak regularly at home?: (5) target language*’. When the student selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).



### **Target language as first language (I03\_ST\_A\_S04B)**

“Target language as first language” is a simple index (item score). The index equals the selection of option 5 in question SQ4 '*Which language(s) did you speak at home as a small child (before the age of five)?: (5) target language*'. When the student selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Target language use in home (I03\_ST\_A\_S26B)**

“Target language use in home” is a simple index (item score). The index equals the selection of option 5 in question SQ26 '*Which language(s) do you, yourself, speak regularly at home?: (5) target language*'. When the student selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Target language as most spoken language at home (I03\_ST\_A\_S27B)**

“Target language as most spoken language at home” is a simple index (item score). The index equals the response “target language” (response category 5) to question SQ27 '*Which language do you speak most often at home?*'. If the student answered “target language” the index has value 1, if the student did not answer “target language” the index has value 0.

### **Target language exposure through home environment (I03\_ST\_A\_S29A)**

The “target language exposure through home environment” is a compound index (sum score). The index equals the sum of all items answered with “Yes” in question SQ29 '*Do you, yourself, come into contact with target language outside school in the following ways?*'.

### **Target language use through home environment (I03\_ST\_A\_S30A)**

“Target language use through home environment” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.17) of the responses to all items of question SQ30 '*How often do you use target language outside school in the following ways?*'.

### **Target language exposure and use through visits abroad (I03\_ST\_A\_S45A)**

“Target language exposure and use through visits abroad” is a compound index (mean score). The index equals the average of the responses to items 3 and 4 of question SQ45: (3) '*How often did you go with your family to a target language speaking country?*' and (4) '*How often did you go with your family to a (non-target language speaking) country?*'.



## **Target language exposure and use through traditional and new media (I03\_ST\_A\_S31A)**

“Target language exposure and use through traditional and new media” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.11) of the responses to all items of question SQ31 *'How often do you come into contact with target language through media in the following ways?’*

### **Antecedent conditions**

#### **10.4.3.1.1 Home location (I03\_ST\_A\_S03A)**

“Home location” is a simple index (item score). The index equals the response to question SQ3 *'The place where you live is ‘* (0=A village, hamlet or rural area (fewer than three thousand people); 1=A small town (three thousand to around fifteen thousand people); 2=A town (fifteen thousand to around hundred thousand people); 3=A city (hundred thousand to around one million people); 4=A large city with over one million people’)

#### **10.4.4 Issue 4: School's foreign language specialization**

### **Participation in foreign language enrichment or remedial lessons (I04\_ST\_M\_S64B)**

“Participation in foreign language enrichment or remedial lessons” is a compound index (minimum score). The index is the minimum of the responses to items 2 and 4 of question SQ64 *'What type of extra lessons have you attended or are you attending?: (2) Enrichment lessons for other foreign languages (including for Latin and ancient Greek)’* and (4) *'Remedial lessons for other foreign languages (including for Latin and ancient Greek)’*. If the student answered both items with “No” the index has value 0 (“No”), else the index has the value 1 (“Yes”).

### **Participation in target language enrichment or remedial lessons (I04\_ST\_M\_S64A)**

“Participation in target language enrichment or remedial lessons” is a compound index (minimum score). The index equals the minimum of the responses to items 1 and 3 of question SQ64 *'What type of extra lessons have you attended or are you attending?: (1) Enrichment lessons for target language’* and (3) *'Remedial lessons for target language’*. If the student answered both items with “No” the index has value 0 (“No”), else the index has the value 1 (“Yes”).

### **10.4.5 Issue 5: Information and Communication Technology to enhance FL learning and teaching**

#### **Frequency of using ICT for foreign language learning (I05\_ST\_M\_S62A)**

The “frequency of using ICT for foreign language learning” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 1) of the responses to all items of question SQ62 *'When studying and doing homework for target language, how often do you use a computer for the following?'*.

#### **Frequency of using ICT outside school'. (I05\_ST\_A\_S24A)**

The “frequency of using ICT outside school” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.17) of the responses to all items of question SQ24 *'How often do you use a computer outside school time for the following?'*

#### **Antecedent conditions**

##### **10.4.5.1.1 ICT-facilities at home' (I05\_ST\_A\_S23A)**

The “ICT-facilities at home” is a compound index (sum score). The index equals the sum of all items answered with "Yes" in question SQ23 *'Are the following devices available for you to use at your home?'*.

### **10.4.6 Issue 6: Intercultural exchanges**

#### **Received opportunities regarding the target language for exchange visits (I06\_ST\_M\_S45A)**

The “received opportunities regarding the target language for exchange visits” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.25) of the responses to items 1, 2, 4 and 5 of question SQ45: (1) *'How often did you go on a school trip to a target language speaking country?'*, (2) *'How often did you go on a school trip to another (non-target language speaking) country'*, (4) *'How often did a school class from a target language speaking country visit your school?'* and (5) *'How often did a school class from another (non-target language) speaking country visit your school?'*

## **Received opportunities regarding the target language for school language projects (I06\_ST\_M\_S46A)**

The “received opportunities regarding the target language for school language projects” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.14) of the responses to all items of question SQ46 *'In the past three years, how often have you participated in the following activities for foreign languages at school?'.*

### **10.4.7 Issue 8: Language learning for all**

#### **Received help in mastering host language (I08\_ST\_M\_S64A)**

The “received help in mastering host language” is a simple index (item score). The index equals the response to item 5 of question SQ64 *'What type of extra lessons have you attended or are you attending?: (5) Extra lessons for questionnaire language'.* If the respondent is a native student (see 10.4.7.1.2, I08\_ST\_A\_S15A=0) the response was non-applicable (value set to 0).

#### **Received formal education in language(s) of origin (I08\_ST\_M\_S64B)**

The “received formal education in language(s) of origin” is a simple index (item score). The index equals the response to item 6 of question SQ64 *'What type of extra lessons have you attended or are you attending?: (6) Extra lessons in another language than questionnaire language that is spoken regularly at your home'.* If the respondent is a native student (see 10.4.7.1.2, I08\_ST\_A\_S15A=0) the response was non-applicable (value set to 0).

#### **Antecedent conditions**

##### **10.4.7.1.1 Gender (I08\_ST\_A\_S01A)**

“Gender” is a simple index (item score). The index equals the response to question SQ1 *'Are you female or male?'* (0=Female;1=Male).

##### **10.4.7.1.2 Age (I08\_ST\_A\_S02A)**

“Age” is a compound index (difference score). The index equals the difference between the date of the middle of the testing window in each Educational system (see Table 34) and the date of birth SQ2 *'What is your date of birth?'.*

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid years ( $\leq 1987$  and  $\geq 2000$ ), invalid months (0 and  $\geq 13$ ), and

invalid days (0 and  $\geq 31$ ) were removed (coded as invalid). Years that were written as two numbers (YY) were converted into a four numbers (YYYY).

**Table 34 Middle of the testing window in all educational systems**

BFL	Flemish Community of Belgium	BE nl	2-3-2011
BFR	French Community of Belgium	BE fr	2-3-2011
BGE	German Community of Belgium	BE de	22-3-2011
BGR	Bulgaria	BG	12-3-2011
ENG	England	UK-ENG	31-10-2011
ESP	Spain	ES	16-3-2011
EST	Estonia	EE	23-2-2011
FRA	France	FR	19-3-2011
GRC	Greece	EL	19-3-2011
HRV	Croatia	HR	15-3-2011
MLT	Malta	MT	27-1-2011
NLD	Netherlands	NL	26-2-2011
POL	Poland	PL	12-3-2011
PRT	Portugal	PT	10-3-2011
SVN	Slovenia	SI	2-3-2011
SWE	Sweden	SE	16-3-2011

#### **10.4.7.1.3 Immigration background (I08\_ST\_A\_S15A)**

The “immigration background” is a compound index (categorisation of dichotomised scores). The index is a categorisation of the dichotomised responses to three questions:

- (i) SQ15 'What country were you born in?'
- SQ16 'What country was your mother born in?'
- SQ17 'What country was your father born in?'

First, the responses of the students to those three questions were first dichotomised (0=Not born in the country; 1 = Born in the country). The dichotomised responses were combined into the following categories: (1) native students: those students who had at least one parent born in the country, (2) second generation' students: those born in the country of assessment but whose parent(s) were born in another country and (3) first-generation students: those students born outside the country of assessment and whose parents were also born in another country.

#### 10.4.8 Issue 9: Foreign language teaching approach

##### **Students' report of teacher's use of the target language during foreign language lessons (I09\_IN\_M\_S49A)**

Students' report of "Teacher's use of the target language during foreign language lessons" is a compound index (mean score). The index equals the average of the responses to all items of question SQ49 *'How often does your teacher of target language speak target language when doing the following?'*.

##### **Students' reported use of the target language during foreign language lessons (I09\_IN\_M\_S50A)**

"Students' reported use of the target language during foreign language lessons" is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.33) of the responses to all items of question SQ50 *'How often do students speak target language when doing the following in a target language lesson?'*

##### **Resource use in target language lessons (I09\_IN\_M\_S51A)**

The "resource use in target language lessons" is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.11) of the responses to all items of question SQ51: *'How often are the following resources used in your target language lessons?'*

##### **Perceived emphasis on similarities between known languages (I09\_IN\_M\_S57A)**

The "perceived emphasis on similarities between known languages" is a latent variable. The index reflects the principal component of the responses to question SQ57 *'How often does your teacher of target language point out similarities between target language and other languages when teaching the following?'* and equals the weighted sum score (rounded to a multiple of 0.583<sup>28</sup>) of the responses to all items of question SQ57 (see Table 35). Before modelling the missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model had a good fit ( $NFI = 0.99$ ;  $NNFI = 0.98$ ;  $CFI = 0.99$ ;  $RMSEA = 0.09$ ;  $RMR = 0.02$ ) and the scale had good reliability (see Table 36)

---

<sup>28</sup> For the index (and other latent variables), the sumscore was weighted with component score coefficient.

**Table 35 Component score coefficient matrix of question SQ57 '*How often does your teacher of target language point out similarities between target language and other languages when teaching the following?*' in the equally weighted sample**

		<b>Component loading</b>	<b>Component score coefficient</b>
SQt57i01	Reported frequency of emphasis between [target language] and other languages during teaching to: write in [target language]	0,84	0,16
SQt57i02	Reported frequency of emphasis between [target language] and other languages during teaching to: speak [target language]	0,88	0,17
SQt57i03	Reported frequency of emphasis between [target language] and other languages during teaching to: understand spoken [target language]	0,87	0,17
SQt57i04	Reported frequency of emphasis between [target language] and other languages during teaching: Teaching [target language] grammar	0,83	0,16
SQt57i05	Reported frequency of emphasis between [target language] and other languages during teaching to: read [target language] texts	0,88	0,17
SQt57i06	Reported frequency of emphasis between [target language] and other languages during teaching to: pronounce [target language] correctly	0,86	0,17
SQt57i07	Reported frequency of emphasis between [target language] and other languages during teaching: [target language] words	0,84	0,16

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 36 Reliability of the index “Perceived emphasis on similarities between known languages” in the equally weighted samples**

<b>Adjudicated Entity</b>	<b>Code</b>	<b>Standardized Cronbach's Alpha</b>	<b>Estimated Cronbach's Alpha<sub>10</sub></b>
Flemish Community of Belgium	BE nl	0,93	0,95
French Community of Belgium	BE fr	0,93	0,95
German Community of Belgium	BE de	0,93	0,95
Bulgaria	BG	0,96	0,97
Spain	ES	0,92	0,94
Estonia	EE	0,93	0,95
France	FR	0,92	0,94
Greece	EL	0,93	0,95
Croatia	HR	0,95	0,96
Malta	MT	0,94	0,96
Netherlands	NL	0,93	0,95
Poland	PL	0,93	0,95
Portugal	PT	0,96	0,97
Slovenia	SI	0,94	0,96
Sweden	SE	0,95	0,96

**Perceived usefulness of target language and target language learning (I09\_ST\_M\_S33B)**

The “Perception of usefulness of target language and target language learning” is a latent variable based on three components:

- (i) Component 1 is based on the responses to question SQ33
- Component 2 is based on the responses to question SQ34
- Component 3 is based on the responses to question SQ35

The index equals the weighted sum score of the three components (see the component loadings in Table 37). The composed index had an adequate reliability in all educational systems (see Table 37).

**Table 37 Component loadings and reliability of the index “Perception of usefulness of target language and target language learning” in the equally weighted sample**

Adjudicated Entity	Code	SQ33	SQ34	SQ35	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,77	0,78	0,77	0,66	0,87
French Community of Belgium	BE fr	0,81	0,78	0,79	0,71	0,89
German Community of Belgium	BE de	0,70	0,77	0,71	0,55	0,80
Bulgaria	BG	0,79	0,79	0,81	0,71	0,89
Spain	ES	0,81	0,76	0,81	0,71	0,89
Estonia	EE	0,86	0,79	0,87	0,79	0,93
France	FR	0,80	0,69	0,82	0,66	0,87
Greece	EL	0,67	0,77	0,72	0,54	0,79
Croatia	HR	0,82	0,80	0,84	0,75	0,91
Malta	MT	0,83	0,84	0,87	0,80	0,93
Netherlands	NL	0,81	0,72	0,84	0,70	0,89
Poland	PL	0,83	0,81	0,88	0,79	0,93
Portugal	PT	0,81	0,76	0,84	0,72	0,90
Slovenia	SI	0,80	0,77	0,82	0,72	0,89
Sweden	SE	0,87	0,81	0,90	0,82	0,94

The first component (a latent variable) reflects the principal component of the responses to question SQ33 *'In your opinion, how useful is target language for the following purposes?' and equals the weighted sum score of the responses to all items of question SQ33 (see Table 38). Before modelling the missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model had an adequate fit ( $NFI = 0.94$ ;  $NNFI = 0.92$ ;  $CFI = 0.94$ ;  $RMSEA = 0.16$ ) and the scale had good reliability (see Table 39). A 2<sup>nd</sup> order one factor model had better fit ( $NFI = 0.98$ ;  $NNFI = 0.97$ ;  $CFI = 0.98$ ;  $RMSEA = 0.09$ ), indicating that within the “perceived usefulness” three aspects can be distinguished: for contacts (items 1, 2, 6, 7), for the future (items 3, 4, 5) and for entertainment (items 8, 9, 10).*



**Table 38 Component score coefficient matrix of question SQ33 'In your opinion, how useful is target language for the following purposes?' in the equally weighted sample**

	Component loading	Component score coefficient
SQt33i01 Usefulness of [target language] for: travelling	0,71	0,13
SQt33i02 Usefulness of [target language] for: your personal life	0,73	0,13
SQt33i03 Usefulness of [target language] for: your further	0,80	0,14
SQt33i04 Usefulness of [target language] for: your future work	0,80	0,14
SQt33i05 Usefulness of [target language] for: getting a good job	0,79	0,14
SQt33i06 Usefulness of [target language] for: contact with foreigners	0,67	0,12
SQt33i07 Usefulness of [target language] for: your personal satisfaction	0,74	0,13
SQt33i08 Usefulness of [target language] for: the use of computers and other technical devices	0,77	0,14
SQt33i09 Usefulness of [target language] for: reading books, magazines, etc.	0,73	0,13
SQt33i10 Usefulness of [target language] for: entertainment (movies, television programmes, music, games)	0,74	0,13

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 39 Reliability of the component SQ33 'In your opinion, how useful is target language for the following purposes?' in the equally weighted samples**

Adjudicated Entity	Code	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,89	0,89
French Community of Belgium	BE fr	0,88	0,88
German Community of Belgium	BE de	0,88	0,88
Bulgaria	BG	0,93	0,93
Spain	ES	0,90	0,90
Estonia	EE	0,93	0,93
France	FR	0,88	0,88
Greece	EL	0,90	0,90
Croatia	HR	0,93	0,93
Malta	MT	0,93	0,93
Netherlands	NL	0,91	0,91
Poland	PL	0,93	0,93
Portugal	PT	0,92	0,92
Slovenia	SI	0,91	0,91
Sweden	SE	0,95	0,95

The second component (a difference score) is the difference between the response to item 6 and the responses to all other items of question SQ34 'How much do you like the following school subjects: (1) Mathematics; (2) Science subjects, e.g. physics; (3) Human and society subjects, e.g. history; (4) Culture and arts subjects, e.g. music, art

history; (5) Questionnaire language; (6) Target language; (7) Other foreign languages (including Latin and ancient Greek);(8) Vocational skills subjects;(9) Sports' . The question had the following response scale 0=Do not like at all;1=Hardly like;2=Quite like;3=Like a lot.

The third component (a difference score) is the difference between the response to item 6 and the responses to all other items of question SQ35 '*In your opinion, how useful are the following school subjects?: (1) Mathematics; (2) Science subjects, e.g. physics; (3) Human and society subjects, e.g. history; (4) Culture and arts subjects, e.g. music, art history; (5) Questionnaire language; (6) target language; (7) Other foreign languages (including Latin and ancient Greek);(8) Vocational skills subjects;(9) Sports' . The question had the following response scale 0=Not useful at all;1=Hardly useful;2=Quite useful;3=Very useful.*

### Perceived difficulty of target language learning (I09\_ST\_M\_S48A)

The “perceived difficulty of target language learning” is a latent variable. The index reflects the principal component of the responses to question SQ48 '*How difficult is it for you to learn the following?*' and equals the weighted sum score (rounded to a multiple of 0.326) of the responses to all items of question SQ48 (see Table 40). Before modelling the missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model had a good fit ( $NFI = 0.98$ ;  $NNFI = 0.97$ ;  $CFI = 0.98$ ;  $RMSEA = 0.09$ ) and the scale had good reliability (see Table 41).

**Table 40 Component score coefficient matrix of question SQ48 '*How difficult is it for you to learn the following?*' in the equally weighted sample**

	Component loading	Component score coefficient
SQt48i01 Perceived difficulty of [target language] learning to: write in [target language]	0,76	0,18
SQt48i02 Perceived difficulty of [target language] learning to: speak [target language]	0,83	0,20
SQt48i03 Perceived difficulty of [target language] learning to: understand spoken [target language]	0,77	0,19
SQt48i04 Perceived difficulty of [target language] learning: [target language] grammar	0,69	0,17
SQt48i05 Perceived difficulty of [target language] learning to: read [target language] texts	0,80	0,19
SQt48i06 Perceived difficulty of [target language] learning to: pronounce [target language] correctly	0,76	0,19
SQt48i07 Perceived difficulty of [target language] learning: [target language] words	0,75	0,18

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 41 Reliability of the index “Perceived difficulty of target language learning” in the equally weighted samples**

<b>Adjudicated Entity</b>	<b>Code</b>	<b>Standardized Cronbach's Alpha</b>	<b>Estimated Cronbach's Alpha<sub>10</sub></b>
Flemish Community of Belgium	BE nl	0,89	0,92
French Community of Belgium	BE fr	0,83	0,88
German Community of Belgium	BE de	0,86	0,90
Bulgaria	BG	0,87	0,91
Spain	ES	0,82	0,87
Estonia	EE	0,87	0,90
France	FR	0,84	0,88
Greece	EL	0,90	0,93
Croatia	HR	0,90	0,93
Malta	MT	0,93	0,95
Netherlands	NL	0,87	0,90
Poland	PL	0,88	0,91
Portugal	PT	0,91	0,93
Slovenia	SI	0,91	0,94
Sweden	SE	0,91	0,94

#### **Perception of target language lessons, teacher and textbook (I09\_ST\_M\_S52B)**

The “Perception of target language lessons, teacher and textbook” is a latent variable based on three components:

- (i) component 1 is based on the responses to question SQ52
- (ii) component 2 is based on the responses to question SQ54
- (iii) component 3 is based on the responses to question SQ55

the index equals the weighted sum score of the three components (see the factor loadings in Table 42). The index had an adequate reliability in all educational systems (see Table 42).

**Table 42 Component loadings and reliability of the index “Perception of target language lessons, teacher and textbook” in the equally weighted sample**

Adjudicated Entity	Code	SQ52	SQ54	SQ55	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,51	0,88	0,90	0,66	0,87
French Community of Belgium	BE fr	0,65	0,87	0,90	0,73	0,90
German Community of Belgium	BE de	0,53	0,88	0,90	0,68	0,87
Bulgaria	BG	0,70	0,84	0,88	0,73	0,90
Spain	ES	0,59	0,89	0,90	0,71	0,89
Estonia	EE	0,62	0,88	0,90	0,72	0,90
France	FR	0,49	0,91	0,92	0,70	0,88
Greece	EL	0,75	0,80	0,82	0,70	0,89
Croatia	HR	0,71	0,83	0,85	0,72	0,89
Malta	MT	0,63	0,88	0,90	0,73	0,90
Netherlands	NL	0,56	0,86	0,89	0,67	0,87
Poland	PL	0,67	0,90	0,91	0,77	0,92
Portugal	PT	0,62	0,88	0,90	0,73	0,90
Slovenia	SI	0,65	0,87	0,89	0,74	0,90
Sweden	SE	0,68	0,87	0,89	0,75	0,91

The first component reflects the principal component (weighted sum score) of the responses to question SQ52 '*How useful are your target language textbooks, or is your target language textbook, for the following?*' and equals the weighted sum score of the responses to all items of question SQ52 (see Table 43). Before modelling the missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model had a moderate fit ( $NFI = 0.93$ ;  $NNFI = 0.90$ ;  $CFI = 0.93$ ;  $RMSEA = ,21$ ). A 2<sup>nd</sup> order one factor model had a better fit ( $RMSEA = 0.07$ ;  $NFI = 0.99$ ;  $NNFI = 0.99$ ;  $CFI = 0.99$ ), indicating that within the “usefulness of the textbook” three aspects can be distinguished: for written communication (items 1;5), for spoken communication (items 2; 3; 6) and for grammar/vocabulary (items 4; 6). ) The scale had good reliability (see Table 44).

**Table 43 Component score coefficient matrix of question SQ52 'How useful are your target language textbooks, or is your target language textbook, for the following?' in the equally weighted sample**

	Component loading	Component score coefficient
SQt52i01 Perceived usefulness [target language] textbooks for learning to: write in [target language]	0,80	0,18
SQt52i02 Perceived usefulness [target language] textbooks for learning to: speak [target language]	0,83	0,19
SQt52i03 Perceived usefulness [target language] textbooks for learning to: understand spoken [target language]	0,80	0,18
SQt52i04 Perceived usefulness [target language] textbooks for learning : [target language] grammar	0,76	0,17
SQt52i05 Perceived usefulness [target language] textbooks for learning to: read [target language] texts	0,80	0,18
SQt52i06 Perceived usefulness [target language] textbooks for learning to: pronounce [target language] correctly	0,76	0,18
SQt52i07 Perceived usefulness [target language] textbooks for learning: [target language] words	0,77	0,18

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 44 Reliability of the component SQ52 'How useful are your target language textbooks, or is your target language textbook, for the following?' in the equally weighted samples**

Adjudicated Entity	Code	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,83	0,87
French Community of Belgium	BE fr	0,86	0,90
German Community of Belgium	BE de	0,88	0,92
Bulgaria	BG	0,94	0,96
Spain	ES	0,86	0,89
Estonia	EE	0,87	0,91
France	FR	0,90	0,93
Greece	EL	0,95	0,96
Croatia	HR	0,92	0,94
Malta	MT	0,92	0,95
Netherlands	NL	0,84	0,88
Poland	PL	0,91	0,93
Portugal	PT	0,91	0,93
Slovenia	SI	0,91	0,93
Sweden	SE	0,90	0,93

The second component reflects the principal component (weighted sum score) of the responses to question SQ54 'To what extent do you agree or disagree with the following statements about your teacher of target language?' and equals the weighted sum score of the responses to items 1, 2, 3, 4 and 5 of question SQ54 (see Table 45).

Before modelling the missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model had a good fit ( $NFI = 0.99$ ;  $NNFI = 0.98$ ;  $CFI = 0.99$ ,  $RMSEA = 0.09$ ) and the scale had good reliability (see Table 46).

**Table 45 Component score coefficient matrix of question SQ54 'To what extent do you agree or disagree with the following statements about your teacher of target language?' in the equally weighted sample**

	Component loading	Component score coefficient
SQt54i01 Perception of [target language] teacher: My teacher of [target language] is a good teacher	0,88	0,23
SQt54i02 Perception of [target language] teacher: I get along with my teacher of [target language]	0,87	0,23
SQt54i03 Perception of [target language] teacher: My teacher of [target language] makes an effort to make the lessons interesting for us	0,85	0,22
SQt54i04 Perception of [target language] teacher: My teacher of [target language] is helpful	0,89	0,23
SQt54i05 Perception of [target language] teacher: I like my teacher of [target language]	0,89	0,23

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 46 Reliability of the component SQ54 'To what extent do you agree or disagree with the following statements about your teacher of target language?' in the equally weighted samples**

Adjudicated Entity	Code	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,93	0,96
French Community of Belgium	BE fr	0,92	0,96
German Community of Belgium	BE de	0,92	0,96
Bulgaria	BG	0,93	0,96
Spain	ES	0,92	0,96
Estonia	EE	0,91	0,95
France	FR	0,92	0,96
Greece	EL	0,92	0,96
Croatia	HR	0,92	0,96
Malta	MT	0,92	0,96
Netherlands	NL	0,92	0,96
Poland	PL	0,94	0,97
Portugal	PT	0,94	0,97
Slovenia	SI	0,92	0,96
Sweden	SE	0,93	0,97

The third component reflects the principal component (weighted sum score) of the responses to question SQ55 *'To what extent do you agree or disagree with the following statements about your target language lessons?'* and equals the weighted sum score of the responses to items 1, 2, 3, 4 and 6 of question SQ55 (see Table 47). Before modelling the contra-indicative items 4 and 6 were inverted for scaling and all missing item responses have been replaced with the mean question score. A confirmatory factor analysis showed that a one factor model in which the two contra-indicative items have a correlated error had a good fit ( $NFI = 0.99$ ;  $NNFI = 0.97$ ;  $CFI = 0.99$ ;  $RMSEA = 0.09$ ) and the scale had good reliability (see Table 48).

**Table 47 Component score coefficient matrix of question SQ55 *'To what extent do you agree or disagree with the following statements about your target language lessons?'* in the equally weighted sample**

	Component loading	Component score coefficient
SQt55i01 Perception of [target language] lessons: My [target language] lessons are interesting	0,88	0,27
SQt55i02 Perception of [target language] lessons: My [target language] lessons are enjoyable	0,87	0,27
SQt55i03 Perception of [target language] lessons: My [target language] lessons are good	0,85	0,26
SQt55i04F Perception of [target language] lessons: My [target language] lessons are waste of time {RECODED}	0,89	0,20
SQt55i06F Perception of [target language] lessons: My [target language] lessons are boring {RECODED}	0,89	0,23

Note. Component score coefficients are based on pairwise deletion of missing variables.

**Table 48 Reliability of the component SQ55 'To what extent do you agree or disagree with the following statements about your target language lessons?' in the equally weighted samples**

<b>Adjudicated Entity</b>	<b>Code</b>	<b>Standardized Cronbach's Alpha</b>	<b>Estimated Cronbach's Alpha<sub>10</sub></b>
Flemish Community of Belgium	BE nl	0,86	0,93
French Community of Belgium	BE fr	0,86	0,92
German Community of Belgium	BE de	0,85	0,92
Bulgaria	BG	0,83	0,91
Spain	ES	0,88	0,93
Estonia	EE	0,87	0,93
France	FR	0,89	0,94
Greece	EL	0,82	0,90
Croatia	HR	0,87	0,93
Malta	MT	0,87	0,93
Netherlands	NL	0,84	0,91
Poland	PL	0,87	0,93
Portugal	PT	0,88	0,93
Slovenia	SI	0,86	0,93
Sweden	SE	0,85	0,92

#### 10.4.9 Organisational structure of the educational systems

##### Class size (I14\_IN\_A\_S42A)

“Class size” is a simple index (categorised item score). The index equals the categorised responses to question SQ42 *'On average, how many students are there in your classroom during the target language lessons?'*. The index “Class size” has the following categories: 5= 1 to 5 students; 10=6 to 10 students; 15=11 to 15 students; 20=16to20 students; 25=21 to 25 students; 30=26to30 students; 35=31to 35 students; 40=36 to 40 students. Prior to the categorisation of the open responses, the invalid response “zero” and outliers (scores higher than 40) were removed.

##### Program level (I14\_ST\_A\_S06A)

The “program level” is the educational level (ISCED2 or ISCED3) in which is sampled for each Educational system and target language (see Table 32).

##### Program designation (I14\_ST\_A\_S06B)

The “program designation” is a simple index (converted item score). The index equals the designation of the selected "study program" in question SQ6 *'Which one of the following programmes are you in?'*. Based on the localisation file (Study Program



Table) the selected study program in SQ6 was converted into the designation of the study program.

### **Program orientation (I14\_ST\_A\_S06C)**

The “program orientation” is a simple index (converted item score). The index equals the orientation of the selected “study program” in question SQ6 *‘Which one of the following programmes are you in?’*. Based on the localisation file (Study Program Table, the selected study program in SQ6 was converted into the orientation of the study program.

### **Compulsory target language learning (I14\_ST\_M\_S47A)**

“Compulsory target language learning” is a simple index (item score) equal to the response to question SQ47 *‘Why are you learning target language?’* and has the categories 0=Because the subject of target language is compulsory;1=Because studying a foreign language is compulsory and I chose target language;2=Because I chose target language as an optional subject.

### **10.4.10 Other indices**

All questions were used for calculating the plausible values, including the index for Economic, social and cultural status, described below, and the other questions<sup>29</sup> that were included solely for enhancing the data quality and usability (see chapter 3) and not for the description of language policies.

### **Economic, social and cultural status (ESCS) (I08\_ST\_A\_S19B)**

As in PISA 2003, PISA 2006, and PISA 2009 (OECD 2012) “Economic, social and cultural status” (ESCS) is comprised of three components:

- (i) home possessions (HOMEPOS)
- parental occupation (HISEI)
- higher parental education expressed as years of schooling (PARED)

Missing values for students with missing data for only one component were imputed with predicted values plus a random component based on a regression on the other two components. If there was missing data on more than one variable, ESCS was not computed for that case and a missing value was assigned for ESCS.

Variables with imputed values were then used for a principal component analysis. The ESCS scores were obtained as component scores for the first principal component (standardized in the entire equally weighted sample, see Table 49). A zero score on

---

<sup>29</sup> SQ5, SQ9, SQ12, SQ18, SQ32, SQ53, SQ56, SQ58, SQ61

the ESCS index refers to the score of an average respondent and one to the standard deviation. The reliability of the ESCS was good (see Table 49).

**Table 49 Component loadings and reliability of the index “Economic, social and cultural status (ESCS)” in the equally weighted sample**

Adjudicated Entity	Code	PARED	HOMEPOS	HISEI	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,79	0,73	0,81	0,67	0,87
French Community of Belgium	BE fr	0,83	0,73	0,82	0,71	0,89
German Community of Belgium	BE de	0,76	0,65	0,78	0,57	0,82
Bulgaria	BG	0,84	0,71	0,82	0,70	0,89
Spain	ES	0,85	0,73	0,85	0,74	0,91
Estonia	EE	0,81	0,70	0,81	0,67	0,87
France	FR	0,78	0,76	0,80	0,68	0,88
Greece	EL	0,82	0,73	0,83	0,71	0,89
Croatia	HR	0,83	0,70	0,84	0,70	0,89
Malta	MT	0,83	0,66	0,83	0,67	0,87
Netherlands	NL	0,79	0,73	0,78	0,64	0,86
Poland	PL	0,85	0,76	0,84	0,76	0,91
Portugal	PT	0,87	0,76	0,84	0,76	0,91
Slovenia	SI	0,85	0,47	0,84	0,57	0,82
Sweden	SE	0,76	0,68	0,77	0,58	0,82

### Home possessions (HOMEPOS)

Similar to the PISA procedure, the index “home possession” has been constructed using IRT modelling of the responses to all items of the questions related to home possession:

- (i) SQ19 *'Which of the following do you have at home?: (1) A desk to study at; (2) A room of your own; (3) A quiet place to study; (4) Books to help with your school work (for example an encyclopaedia or atlas); (5) A computer you can use for school work; (6) Educational software; (7) An internet connection; (8) A dictionary'*
- SQ20 *'Which of the following are in your home? (continued): (1) Classics from the literature of Educational system (e.g. books of Shakespeare); (2) Books of poetry; (3) Works of art (e.g. paintings); (4) A dishwasher; (5) A DVD player; (6) Country specific wealth item1; (7) Country specific wealth item2; (8) Country specific wealth item3'*
- SQ21 *'How many books are there in your home?' (0=0-10 books;1=11-25 books;2=26-100 books;3=101-200 books;4=201-500 books;5=More than 500 books)*

SQ22 *'How many of these are there in your home?: (1) Mobile phones; (2) Television sets; (3) Computers or laptops; (4) Cars; (5) Bathrooms'. (0=None;1=One;2=Two;3=Three or more)'*

For the IRT modelling the software package OPLM (Verhelst, Glas, & Verstralen 1995) was used. OPLM is an extension of a Rasch model and estimates difficulty parameters. By imputing discrimination indices as known constants, OPLM maintains the desirable characteristics of a one-parameter logistic model. Parameters are estimated by use of a conditional maximum likelihood estimation procedure. (Verhelst, Glas and Verstralen 1995). After the calibration and estimation of item parameters, person's parameters have been established by weighted maximum likelihood estimates using the item parameters produced in first stage.

Following the PISA procedure we estimated all items free across countries and tried to maintain as many items as possible in the calibration. This resulted in parameter characteristics that are comparable within countries, and only to a lesser degree between countries. During the calibration procedure it came out that variable t22i05 (How many of these are there in your home? 5) Bathrooms) in Slovenia suffered from severe misfit. Closer inspections of the Slovenian questionnaire translation revealed that the country-specific wealth items for Slovenia also included a question on the presence of bathrooms within students' households. This may have led to severe multicollinearity and therefore this item has accordingly been excluded for Slovenia from further analyses. Furthermore, in Estonia no third country specific item was asked for (t20i08; Which of the following are in your home? {Country specific wealth item3}) and has therefore been excluded for Estonia in the analyses. The R1c statistic provides a global test for model fit and is based on the differences between the observed and expected proportion of responses in homogeneous score groups. The R1c statistic value for the final model was 6530, with 1547 degrees of freedom, which is an acceptable fit given the large sample size.

The scale had a rather poor reliability in the educational systems (see Table 50). The reliability in each Educational system is very similar to the reliabilities of the "Home possessions" in previous PISA cycles, see the preliminary version of the technical report PISA 2009 (OECD 2012). The low reliability may be due to the higher degree of accessibility of household items. A very high percentage of students reported the existence of many of the household items which makes them less appropriate as indicators of wealth.

**Table 50 Reliability of “home possessions” in the equally weighted samples.**

Adjudicated Entity	Code	PARED	HOMEPOS	HISEI	Standardized Cronbach's Alpha	Estimated Cronbach's Alpha <sub>10</sub>
Flemish Community of Belgium	BE nl	0,79	0,73	0,81	0,67	0,87
French Community of Belgium	BE fr	0,83	0,73	0,82	0,71	0,89
German Community of Belgium	BE de	0,76	0,65	0,78	0,57	0,82
Bulgaria	BG	0,84	0,71	0,82	0,70	0,89
Spain	ES	0,85	0,73	0,85	0,74	0,91
Estonia	EE	0,81	0,70	0,81	0,67	0,87
France	FR	0,78	0,76	0,80	0,68	0,88
Greece	EL	0,82	0,73	0,83	0,71	0,89
Croatia	HR	0,83	0,70	0,84	0,70	0,89
Malta	MT	0,83	0,66	0,83	0,67	0,87
Netherlands	NL	0,79	0,73	0,78	0,64	0,86
Poland	PL	0,85	0,76	0,84	0,76	0,91
Portugal	PT	0,87	0,76	0,84	0,76	0,91
Slovenia	SI	0,85	0,47	0,84	0,57	0,82
Sweden	SE	0,76	0,68	0,77	0,58	0,82

### Parental occupation (HISEI)

The students' answers to the four questions about parental occupation were coded in each educational system using the International Standard Classification of Occupations (ISCO-88) developed by ILO, including the PISA modifications (see chapter 7.16):

- (i) SQ7 'What is your mother's main job?'  
SQ8 'What does your mother do in her main job?'  
SQ10 'What is your father's main job?'  
SQ11 'What does your father do in his main job?'

The codes for parental occupation (ISCO\_M “International Standard Classification of Occupation mother” and ISCO\_F “International Standard Classification of Occupation father”) were transformed into the international socio-economic index of occupational status (ISEI) (Ganzeboom & Treiman 1996). The higher ISEI scores indicated higher levels of occupational status. The component “parental occupation (HISEI)” corresponds to the higher ISEI score of either parent or the only available parent's ISEI.

### Higher parental education expressed as years of schooling (PARED)

The calculation of this component is based on a transformation of the answers to two questions:

- (i) SQ13 'What is the highest level of schooling completed by your mother?'
- SQ14 'What is the highest level of schooling completed by your father?'

The responses to these questions were converted into estimated years of schooling using the mapping of PISA 2006 (OECD 2007) with a few small changes (see Table 51), because not all educational systems participating in the ESLC were represented in the PISA table. The component "higher parental education expressed as years of schooling" (PARED) corresponds to the higher PARED score of either parent or the only available parent's PARED.

**Table 51 Mapping of ISCED to accumulated years of education**

Adjudicated Entity	ISCED 1 not completed or never went to school (score 7)	ISCED1 (score 6)	ISCED2 (score 5)	ISCED 3B/3C (score 4)	ISCED 3A (score 3) or ISCED 4 (score 2)	ISCED 5B (score 1)	ISCED 5A or 6 (score 0)
BE nl	0	6	9	12	12	14,5	17
BE fr	0	6	9	12	12	14,5	17
BE de	0	6	9	12	12	14,5	17
BG	0	4	8	12	12	15	17,5
EN	0	6	9	12	13	15	16
ES	0	5	8	10	12	13	16,5
EE	0	4	9	12	12	15	16
FR	0	5	9	12	12	14	15
EL	0	6	9	11,5	12	15	17
HR	0	4	8	11	12	15	17
MT*	0	5	10	12	12	15	16
NL	0	6	10		12	15*	16
PL	0	6	8	11	12	15	16
PT	0	6	9	12	12	15	17
SI	0	4	8	11	12	15	16
SE	0	6	9	11,5	12	14	15,5

Note: \*MT was not represented in the PISA table. The information of Malta is based on the educational structure as reported by Eurydice (The structure of the European education systems 2010/11: schematic diagrams 2010).

## 10.5 Teacher Questionnaire

### 10.5.1 Issue 4: School's foreign language specialisation

#### Target language class size (I04\_IN\_A\_T39\_\_)

"Target language class size" is a simple index (categorised item score). The index equals the categorised response to question TQ39 'In general, how many students are there in your classroom during target language lessons?'. The index has the following categories: 5=1 to 5 students, 10=6 to 10 students, 15=11 to 15 students, 20=16 to 20

students, 25=21 to 25 students, 30=26 to 30 students, 35=31 to 35 students and 40=36 to 40 students.

Prior to the categorisation of the open responses, invalid answers (higher than 40) were removed (coded as invalid).

### **10.5.2 Issue 5: Information and Communication technology to enhance FL learning and teaching**

#### **Number of different ICT-facilities in school (I05\_ED\_A\_T43\_\_)**

The “number of different ICT-facilities in school” is a compound index (sum of dichotomised scores). The index equals the sum of the dichotomised responses to all items of question TQ43 *‘How often do you use the following devices at school for teaching target language?’*. Prior to calculating the index the item responses of question TQ43 were dichotomised: 0=Not available (score 0) and 1=Available (score  $\geq 1$ ).

#### **Frequency of using ICT outside lessons for teaching (I05\_IN\_M\_T05\_\_)**

The “frequency of using ICT outside lessons for teaching” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.2) to all items of question TQ5 *‘How often do you use a computer outside your lessons (at home or elsewhere) for the following?’*.

#### **Frequency of using ICT devices when teaching (I05\_IN\_M\_T43B\_)**

The “frequency of using ICT devices when teaching” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.14) to all items of question TQ43 *‘How often do you use the following devices at school for teaching target language?’*.

#### **Frequency using web content for teaching (I05\_IN\_M\_T45\_\_)**

The “frequency using web content for teaching” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.11) to all items of question TQ45 *‘In general, how often do you or your students use the following ICT facilities for a target language class you teach?’*.

#### **Antecedent conditions**

##### **10.5.2.1.1 Number of different ICT-devices at the teacher’s home (I05\_IN\_A\_T04\_\_)**

The “number of different ICT-devices at home” is a compound index (sum score). The index equals the sum of all items answered with "Yes" in question TQ4 ‘*Do you have the following devices at home?*’.

### **10.5.3 Issue 6: Intercultural exchanges**

#### **Created opportunities for exchange visits (I06\_ED\_M\_T41\_\_)**

“Created opportunities for exchange visits” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.25) to all items of question TQ41 ‘*During the past three years, how often were you involved in the organisation of the following?*’.

#### **Created opportunities for school language projects (I06\_ED\_M\_T42\_\_)**

“Created opportunities for school language projects” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.14) to all items of question TQ42 ‘*In the past three years, how often were you involved in the organisation of the following activities at school?*’.

### **10.5.4 Issue 7: Staff from other language communities**

#### **Number of teacher’s first languages (I07\_IN\_A\_T07\_\_)**

The “number of teacher’s first languages” is a compound index (categorised sum score). The index equals the number of selected options in question TQ7 ‘*Which language(s) did you speak at home as a small child (before the age of five)?*’. The index has the following categories: 1=“One language” (sum score=1); 2=“Two languages” (sum score=2); 3=“Three or more languages” (sum score≥3).

#### **Target language as teacher’s first language (I07\_IN\_A\_T0705)**

“Target language as teacher’s first language” is a simple index (item score). The index is the selection of option 5 in question TQ7 ‘*Which language(s) did you speak at home as a small child (before the age of five)?: (5) target language*’. When the teacher selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).



## **Received training to teach target language as a foreign language (I07\_IN\_M\_T1505)**

“Training to teach target language as a foreign language” is a compound index (minimum score). The index equals the minimum of the responses to item 5 of question TQ15 and item 5 of question 32 (0=“No” and 1=“Yes”):

- (i) TQ15 *‘Did you receive instruction in the following language related subjects during your initial training as a teacher?: (5) Teaching target language as a foreign language’*
- (ii) TQ32 *‘In the past five years, have you, as a teacher, participated in in-service training covering any of the following language related themes?: (5) Teaching target language as a foreign language’*

### **Antecedent conditions**

#### **10.5.4.1.1 Born in Educational system (I07\_IN\_A\_T03\_\_)**

The index “born in Educational system” is a simple index (dichotomised item score) which equals the dichotomised responses to TQ3 ‘What country were you born in?’. The response to question TQ3 was dichotomised into two categories 0=“Born abroad” and 1=“Born in Educational system”.

### **10.5.5 Issue 9: Foreign language teaching approach**

#### **Emphasis on similarities between known languages (I09\_IN\_M\_T54\_\_)**

“Emphasis on similarities between known languages” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.14) of the responses to all items of question TQ54 *‘In general, how often do you point out similarities between target language and other languages (Including questionnaire language) when teaching the following to one of your classes?’*.

#### **Emphasis on the four language skills and other aspects of language learning**

#### **Emphasis on Writing target language (I09\_IN\_M\_T5301)**

“Emphasis on Writing target language” is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about “Writing” (item 1) in four questions:

- (i) TQ53 *‘In general, how often do you teach the following to a target language class?’*
- TQ55 *‘In your opinion, how important is it that your students learn the following?’*





TQ56 *'In general, how often do you give a [target language] class homework or assignments aimed at the following?'*

TQ59 *'How important are the following when you determine a mark for the final grade of students for the subject of target language?'*

As we are interested in the relative emphasis a teacher places on Writing compared to other aspects of language learning<sup>30</sup>, the item responses to the four questions were rescaled prior to calculating the index. The item responses of each question were rescaled such that the mean question score was zero and the question standard deviation one in each subsample (teachers of respectively the 1<sup>st</sup> target language and 2<sup>nd</sup> target language in each Educational system).

A negative value on the index means that the target language teacher relatively places less emphasis on writing and a positive value that the target language teacher relatively places more emphasis on writing.

### **Emphasis on speaking target language (I09\_IN\_M\_T5302)**

"Emphasis on speaking target language" is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about "speaking" (item 2) in the same four questions as used for "Emphasis on writing target language" (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Emphasis on understanding spoken target language (I09\_IN\_M\_T5303)**

"Emphasis on understanding spoken target language" is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about "understanding spoken target language" (item 3) in the same four questions as used for "Emphasis on writing target language" (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Emphasis on Reading target language texts (I09\_IN\_M\_T5305)**

"Emphasis on Reading target language texts" is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about "Reading target language texts" (item 5) in the same four questions as used for "Emphasis on writing target language" (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

---

<sup>30</sup> Each of the four questions (TQ53, TQ55, TQ56 and TQ59) addressed eight aspects of language learning: (1) Writing; (2) speaking; (3) Listening; (4) Reading; (5) grammar; (6) pronunciation; (7) vocabulary; (8) culture and literature.

### **Emphasis on target language grammar (I09\_IN\_M\_T5304)**

“Emphasis on target language grammar” is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about “speaking” (item 4) in the same four questions as used for “Emphasis on writing target language” (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Emphasis on pronouncing target language correctly (I09\_IN\_M\_T5306)**

“Emphasis on pronouncing target language correctly” is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about “pronouncing target language correctly” (item 6) in the same four questions as used for “Emphasis on writing target language” (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Emphasis on target language vocabulary (I09\_IN\_M\_T5307)**

“Emphasis on target language vocabulary” is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about “target language vocabulary” (item 7) in the same four questions as used for “Emphasis on writing target language” (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Emphasis on target language culture or literature (I09\_IN\_M\_T5308)**

“Emphasis on target language culture or literature” is a compound index (mean rescaled score). The index equals the average of the rescaled responses to the items about “target language culture or literature” (item 8) in the same four questions as used for “Emphasis on writing target language” (TQ53, TQ55, TQ56, TQ59). The rescaling was identical as well (see 0).

### **Use of the target language during foreign language lessons by students (I09\_IN\_M\_T50\_\_)**

Teachers’ reported “Use of the target language during foreign language lessons by student” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.33) of the responses to all items of question TQ50 ‘*In general, how often do your students speak target language when they do the following in a target language lesson?*’.

## **Use of the target language during foreign language lessons by teacher (I09\_IN\_M\_T49\_\_)**

Teachers' reported "Use of the target language during foreign language lessons by teacher" is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.5) of the responses to all items of question TQ49 '*In general, how often do you speak target language when you do the following in a target language lesson?*'.

### **10.5.6 Issue 10: Teachers' access to high quality initial and continuous training**

#### **Educational level of teacher (I10\_IN\_M\_T13\_\_)**

The "educational level of teacher" is a simple index (categorised inverted item score). The index equals the categorised and inverted response to questions TQ13 '*What is the highest level of education that you have completed?*'. The index has the following categories 0='ISCED3 OR 4' (item score  $\geq 3$ ); 1='ISCED5B' (item score 2); 2='ISCED 5A' (item score 1); 3='ISCED 6' (item score 0).

#### **Certification for target language (I10\_IN\_M\_T19\_\_)**

"Certification for target language" is a simple index (item score). The index equals the response to question TQ19 '*What kind of certification for teaching target language do you currently hold?*' with the following response categories: 0=No certificate;1=Temporary or emergency certification;2=Provisional certificate, e.g. Newly Qualified teacher;3=Full certificate;4=Other certificate.

#### **Qualified to teach target language (I10\_IN\_M\_T24\_\_)**

"Qualified to teach target language" is a simple index (item score). The index equals the selection of the option referring to the target language in question TQ24 '*Which language(s) are you qualified to teach?*'. When the teacher selected the option the index has the value one ("selected"), else the index has the value zero ("unselected"). On the basis of the localisation file (Taught Languages Table) for each country and questionnaire version the option referring to the target language has been identified (see Table 52).

**Table 52 The options in TQ24 referring to the target languages**

	Questionnaire version	
	For the 1st target language	For the 2nd target language
BE nl	TQt24i06	TQt24i07
BE fr	TQt24i07	TQt24i08
BE de	TQt24i06	TQt24i07
BG	TQt24i06	TQt24i08
UK-ENG	TQt24i06	TQt24i07
ES	TQt24i06	TQt24i07
EE	TQt24i06	TQt24i08
FR	TQt24i06	TQt24i07
EL	TQt24i07	TQt24i09
HR	TQt24i06	TQt24i07
MT	TQt24i06	TQt24i07
NL	TQt24i06	TQt24i07
PL	TQt24i06	TQt24i02
PT	TQt24i06	TQt24i07
SI	TQt24i06	TQt24i07
SE	TQt24i06	TQt24i07

### Language specialisation (I10\_IN\_M\_T22\_\_)

“Language specialization” is a compound index (combination of question scores). The index is a combination of three aspects:

- (i) The number of languages a teacher is qualified to teach, which equals the response to question TQ23 ‘*How many languages are you qualified to teach?*’

The number of other subjects than languages the teacher is qualified to teach (TQ22<sub>0</sub>), which equals the number of items referring to other subjects than languages (1, 2, 3, 4, 8, 9) answered with “yes” in question TQ22 ‘*Which school subjects are you qualified to teach?: (1) Mathematics, (2) One or more science subjects, e.g. physics, (3) One or more Human and society subjects, e.g. history, (4) One or more Culture and arts subjects, e.g. music, art history, (8) One or more vocational skills subjects, (9) Sports.*

Whether the teacher is qualified to teach target language (Qualified to teach target language (I10\_IN\_M\_T24\_\_), see 0).

Those aspects were combined into the following categories:

- 0=No qualification for any subject (neither for languages, nor for other subjects than language)
- 1=Not qualified for languages, but only qualified for other subjects than languages

- 2=Generalist: qualified for language(s) and for more than two other subjects
- 3=Semi-specialized in languages: qualified for language(s) (but not only for target language) and for two other subjects
- 4=Semi-specialized in target language: qualified for target language (but not for other languages) and for two other subjects
- 5=Specialized in languages: qualified for language(s) (but not only for target language) and one other subject
- 6=Specialized in target language: qualified for target language (but not for other languages) and one other subject
- 7=Completely specialized in languages (no other subjects): qualified for language(s) (but not only for target language) and for no other subject
- 8=Completely specialized in target language (no other subjects): qualified for target language only (not for other languages or other subjects)

### **Participation in in-service training is a right for teachers (I10\_ED\_M\_T3001)**

“Participation in in-service training is a right for teachers” is a simple index (item score). The index equals the selection of the option 1 in question TQ30 *‘Is participation in in-service training an obligation, a right or an option for you?: (1) Participation in in-service training is a right for teachers’*. When the teacher selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Participation in in-service training is required for promotion (I10\_ED\_M\_T3002)**

“Participation in in-service training is required for promotion” is a simple index (item score) which equals the selection of the option 2 in question TQ30 *‘Is participation in in-service training an obligation, a right or an option for you?: (2) Participation in in-service training is required for promotion’*. When the teacher selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Participation in in-service training is optional (I10\_ED\_M\_T3003)**

“Participation in in-service training is optional” is a simple index (item score) which equals the selection of the option 3 in question TQ30 *‘Is participation in in-service training an obligation, a right or an option for you?: (3) Participation in in-service training is optional’*. When the teacher selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Number of different financial incentives for in-service training (I10\_ED\_M\_T34\_\_)**

The “number of different financial incentives for in-service training” is a compound index (sum score). The index equals the number of items answered with “Yes” in question TQ34 *‘Which of the following financial compensations can you get for participation in in-service training?’*

### **Organisation of in-service training (I10\_ED\_M\_T35\_\_)**

“Organisation of in-service training” is a simple index (item score). The index equals the response to question TQ35 ‘*When are you normally allowed to participate in in-service training?*’ which has the following the response categories 0=During your working hours with a substitute teacher for your classes; 1=During your working hours but not during teaching hours (a substitute teacher for your classes is not organised); 2=Only outside your working hours.

### **Participation in in-service training is an obligation for teachers (I10\_ED\_M\_T3000)**

“Participation in in-service training is an obligation for teachers” is a simple index (item score) which equals the selection of the option 0 in question TQ30 ‘*Is participation in in-service training an obligation, a right or an option for you?: (0) Participation in in-service training is an obligation for teachers*’. When the teacher selected the option the index has the value one (“selected”), else the index has the value zero (“unselected”).

### **Number of times the teacher participated in in-service training through different modes (I10\_IN\_M\_T310A)**

The “number of times the teacher participated in in-service training through different modes” is a compound index (sum of dichotomised scores). The index equals the sum of the dichotomised responses to all items of question TQ31 ‘*In the past five years, how often have you participated in an in-service training at the following places?*’. Outliers in the sum scores (values > 6) were removed.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 1000) were removed (coded as invalid) and the responses were dichotomised: 0=“No” (score 0) and 1=“Yes” (scores ≥1).

### **Participated in an in-service training at least once (I10\_IN\_M\_T310B)**

“Participated in an in-service training at least once” is a compound index (sum of dichotomised scores). The index equals the sum of the dichotomised responses to question TQ31 ‘*In the past five years, how often have you participated in an in-service training at the following places?*’ categorised into the following categories: 0=“No” (sum score 0) and 1=“Yes” (sum score ≥ 1 and < 6). The index equals the sum of the dichotomised responses to all items of question TQ31 ‘*In the past five years, how often have you participated in an in-service training at the following places?*’.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 1000) were removed (coded as invalid) and the responses were dichotomised: 0=“No” (score 0) and 1=“Yes” (scores ≥1).

### **Focus of in-service training on languages or teaching related subjects (I10\_IN\_M\_T32\_\_)**

“Focus of in-service training on languages or teaching related subjects” is a compound index (difference between mean scores). The index equals the average of the responses to all items of question TQ32 minus the average of the responses to all items of question TQ33.

- (i) TQ32 *‘In the past five years, have you, as a teacher, participated in in-service training covering any of the following language related themes?’*
- (ii) TQ33 *‘In the past five years, have you, as a teacher, participated in in-service training treating any of the following themes related to the theory and practice of teaching in general?’*

A value zero on the index means that the teacher has followed the same amount of in-service training with language related themes as training with teaching related themes. A negative value indicates that the teacher followed relatively more training with teaching related subjects and a positive value means that the teacher followed relatively more training in language related subjects.

### **Mode of in-service training (I10\_IN\_M\_T3101, I10\_IN\_M\_T3102, I10\_IN\_M\_T3103, I10\_IN\_M\_T3104, I10\_IN\_M\_T3105)**

The “mode of in-service training” is assessed with five simple indices (dichotomised item score), each equalling one dichotomised item response to question TQ31 *‘In the past five years, how often have you participated in an in-service training at the following places?’*:

- (i) Participated in an in-service training at the school where you teach (I10\_IN\_M\_T3101)
- Participated in an in-service training at another institute in Educational system (I10\_IN\_M\_T3102)
- Participated in an in-service training at an institute in a target language speaking educational system (I10\_IN\_M\_T3103)
- Participated in an in-service training at an institute in a non-target language speaking educational system other than Educational system (I10\_IN\_M\_T3104)
- Participated in an in-service training online (I10\_IN\_M\_T3105)

Prior to calculating the indices the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 1000) were removed (coded as invalid) and the responses were dichotomised: 0=“No” (score 0) and 1=“Yes” (scores ≥1).

### **Antecedent conditions**

#### **10.5.6.1.1 Teachers age group (I10\_IN\_A\_T02\_\_)**



“Teachers age group” is a simple index (item score) which equals the response to TQ2 ‘*How old are you?*’ with the response scale: 0=Under 25; 1=25-34; 2=35-44; 3=45-54; 4=55 or older.

#### **10.5.6.1.2 Teachers gender (I10\_IN\_A\_T01\_\_)**

“Teachers gender” is a simple index (item score) equal to the response to question TQ1 ‘*Are you female or male?*’ with the response scale 0=Female; 1=Male.

### **10.5.7 Issue 11: A period of work or study in another country for teachers**

#### **Stays in target culture for different reasons (I11\_IN\_M\_T12\_\_)**

“Stays in target culture for different reasons” is a compound index (sum of dichotomised scores). The index is the sum of the dichotomised responses to all items of question TQ12 ‘*How often have you stayed more than one month in a target language speaking country for the following reasons?*’.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 100) to the items were removed (coded as invalid) and the responses to the items were dichotomised (score 0 and scores  $\geq 1$ ).

#### **Number of long stays in target culture (I11\_IN\_M\_T120B)**

The “number of long stays in the target culture” is a compound index (sum score). The index equals the sum of the responses to all items of question TQ12 ‘*How often have you stayed more than one month in a target language speaking country for the following reasons?*’.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 100) to the items were removed (coded as invalid) and outliers were replaced with the cut-off value 10. Due to the high collinearity with stays in the target culture for different reasons (see 0) this index has not been used in the Final Report.

#### **A stay in target culture for longer than one month (I11\_IN\_M\_T120A)**

“A stay in target culture for longer than one month” is a compound index (minimum dichotomised score). The index equals the minimum of the dichotomised responses to all items of question TQ12 ‘*How often have you stayed more than one month in a target language speaking country for the following reasons?*’. The index has the following categories: 0=“No” and 1=“Yes”.

Prior to calculating the index the open responses were prepared for the arithmetical transformation. Invalid answers (higher than 100) to the items were removed (coded



as invalid) and the responses to the items were dichotomised (score 0 and scores  $\geq 1$ ). Due to the high collinearity with stays in the target culture for different reasons (see 0) this index has not been used in the Final Report.

### 10.5.8 Issue 12: Use of existing European language assessment tools

#### Use of CEFR (I12\_IN\_M\_T40\_\_)

“Use of the CEFR” is a compound index (mean score). The index equals the average of the responses to all items of question TQ40 ‘*How often have you used the Common European Framework of Reference for the following?*’ with the response scale 0=Never;1=Sometimes;2=Quite often;3=Very often.

#### Received training about CEFR (I12\_IN\_M\_T1509)

“Received training about CEFR” is a compound index (minimum score). The index equals the minimum number of two items (of two) answered with “Yes”.

- (i) TQ15 ‘*Did you receive instruction in the following language related subjects during your initial training as a teacher?: (9) The Common European Framework of Reference*’ (0=No;1=Yes)
- (ii) TQ32 ‘*In the past five years, have you, as a teacher, participated in in-service training covering any of the following language related themes?: (9) The Common European Framework of Reference*’ (0=No;1=Yes)

#### Use of Language Portfolio (I12\_IN\_M\_T4507)

“Use of Language Portfolio” is a simple index (dichotomised item score). The index equals the dichotomised response (score 0 and scores  $\geq 1$ ) to item 7 of TQ45 ‘In general, how often do you or your students use the following ICT facilities for a target language class you teach?: (7) Online portfolio’. The index has the values 0=“No” and 1=“Yes”.

#### Received training in use of Portfolio (I12\_IN\_M\_T1510)

“Received training in use of Portfolio” is a compound index (minimum score). The index equals the minimum number of two items (of two) answered with “Yes”. The index has the values 0=“No” and 1=“Yes”.

- (i) TQ15 ‘*Did you receive instruction in the following language related subjects during your initial training as a teacher?: (10) The use of a Portfolio, e.g. the European Language Portfolio*’ (0=No;1=Yes)
- TQ32 ‘*In the past five years, have you, as a teacher, participated in in-service training covering any of the following language related themes?: (10) The use of a Portfolio, e.g. the European Language Portfolio*’ (0=No;1=Yes)

## 10.5.9 Issue 13: Practical experience

### Duration of in-school teaching placement (I13\_IN\_M\_T1801)

“Duration of in-school teaching placement” is a simple index (categorised item score). The index equals the categorised response to the first item of question TQ18 ‘*How long were the following phases during your initial training as a teacher?: (1) In-school teaching placements*’. The index has the following categories: 0=0 months; 1=1 month; 2=2 thru 3 months; 3=4 thru 6 months; 4=7 thru 12 months; 5=13 thru 29 months.

Prior to the categorization of the open responses, invalid answers (higher than 30) were removed (coded as invalid).

### Experience in teaching target language (I13\_IN\_M\_T2901)

“Experience in teaching target language” is a simple index (item score). The index equals the response to the first item of question TQ29 ‘*By the end of this school year, how many years will you have been teaching the following?: (1) Target language*’. Invalid open responses (more than 70 years) were removed (coded as invalid).

### Experience in teaching languages other than target language (I13\_IN\_M\_T2902)

“Experience in teaching languages other than target language” is a simple index (item score). The index equals the response to the second item of question TQ29 ‘*By the end of this school year, how many years will you have been teaching the following?: (2) Other languages than target language, including ancient languages*’. Invalid open responses (more than 70 years) were removed (coded as invalid).

### Experience in teaching other subjects than languages (I13\_IN\_M\_T2903)

“Experience in teaching other subjects than languages” is a compound index (difference score). The index equals the response to the third item of question TQ29 ‘*By the end of this school year, how many years will you have been teaching the following?: (3) All subjects, including languages, together (total)*’ minus the sum of item 1 and 2 (see 0 and 0). Invalid open responses (more than 70 years) were removed (coded as invalid).

### Number of languages taught in the past five years (I13\_IN\_M\_T360A)

The “number of languages taught in the past five years” is a compound index (sum score). The index equals the number of the selected options in question TQ36 ‘*Which of the following languages have you taught during the past five years?*’.

## 10.5.10 Organisational structure of the educational systems

### Within class ability grouping (setting)

“Within class ability grouping” is a simple index (item score). The index equals the response to the second item of question TQ51 ‘*In general, how often do you do the following during a target language lesson?: (2) Let the students work in same-ability groups*’ with the response scale (0=Never; 1=Hardly ever; 2=Every now and then; 3=Usually; 4=Always).

## 10.6 Principal Questionnaire

### 10.6.1 Issue 2: Diversity and order of foreign language offered

#### Number of foreign and ancient languages on offer in school (I02\_ED\_M\_P220)

The “number of foreign and ancient languages on offer in school” is a compound index (sum score). The index equals the number of selected options referring to the most widely taught languages (options 5 to 10) in question PQ22 ‘*Which of the following languages can students study in your school?*’. On the basis of the localisation file (Taught Languages Table) for each country the options referring to the most widely taught foreign or ancient languages have been identified (see Table 53)

**Table 53 The options in PQ22 referring to the most widely taught foreign and ancient languages.**

BE nl	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
BE fr	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
BE de	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
BG	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
UK-ENG	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
ES	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
EE	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
FR	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
EL	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
HR	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
MT	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
NL	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
PL	PQt22i05	PQt22i01	PQt22i07	PQt22i08	PQt22i09	PQt22i10
PT	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10
SI	PQt22i05	PQt22i06	PQt22i07	PQt22i02	PQt22i09	PQt22i10
SE	PQt22i05	PQt22i06	PQt22i07	PQt22i08	PQt22i09	PQt22i10

## 10.6.2 Issue 4: School's foreign language specialization

### Content and Language Integrated Learning (I04\_ED\_M\_P3601)

“Content and Language Integrated Learning” is a simple index (item score) equal to the response to item 1 of question PQ36 *'Does your school offer the following to encourage language learning?: (1) Content and Language Integrated Learning (CLIL)'*. The index has the values 0=“No” and 1=“Yes”.

### Specialist language profile (I04\_ED\_M\_P360)

“Specialist language profile” is a compound index (sum score). The index equals the number of items answered with “Yes” in question PQ36 *'Does your school offer the following to encourage language learning?'*.

### Provision of target language enrichment or remedial lessons (I04\_ED\_M\_P4001)

“Provision of target language enrichment or remedial lessons” is a compound index (maximum score). The index is the maximum of the responses to the items 1 and 3 of question PQ40 *'What type of extra lessons does your school offer to students?: (1) Enrichment lessons for target language ;(3) Remedial lessons for target language'* (0=No;1=Yes). If at least one of the items is answered with “Yes” the index has the value 1 (“Yes”), else the index has the value zero (“No”).

### Provision of foreign language enrichment or remedial lesson (I04\_ED\_M\_P4002)

“Provision of foreign language enrichment or remedial lessons” is a compound index (maximum score). The index is the maximum of the responses to the items 2 and 4 of question PQ40 *'What type of extra lessons does your school offer to students?: (2) Enrichment lessons for other foreign languages (including for Latin and ancient Greek) ; (4) Remedial lessons for other foreign languages (including for Latin and ancient Greek)'*. If at least one of the items is answered with “Yes” the index has the value 1 (“Yes”), else the index has the value zero (“No”).

## 10.6.3 Issue 5: Information and communication technology to enhance FL learning and teaching

### Availability of ICT in classrooms (I05\_ED\_A\_P440)

“Availability of ICT in classrooms” is a compound index (rounded mean score). The index equals the mean (rounded to a multiple of 0.25) of the responses to all items of question PQ44 *'Are the following devices available in the classrooms?'*.

### **Availability of a multimedia (language) lab (I05\_ED\_A\_P4501)**

“Availability of a multimedia (language) lab” is a compound index (categorised scores). The index is based on the categorisation of the responses to items 1 and 2 of question PQ45 *'Does your school have the following ICT facilities?: (1) Multimedia language lab (teacher PC and student PCs with specific language learning software) ; (2) Multimedia lab (teacher PC and student PCs without specific language learning software)'* The index has the following categories:

- 0='No' when both items were answered with “No” (item<sub>1</sub>=0 & item<sub>2</sub>=0)
- 1='Not language specific' when only the second item was answered with “Yes” (item<sub>1</sub>=0 & item<sub>2</sub>=1) and
- 2='Yes, language-specific' when the first item was answered with “Yes” (item<sub>1</sub>=1)

### **Presence of a virtual learning environment (I05\_ED\_A\_P4503)**

“Presence of a virtual learning environment” is a simple index (item score) which equals the response to item 3 of question PQ45 *'Does your school have the following ICT facilities?: (3) A virtual learning environment to support teaching and learning, e.g. Moodle, WebCT, Blackboard, Fronter, Sakai'* (0=“No”; 1=“Yes”).

### **Level of availability of software for language assessment or language teaching (I05\_ED\_A\_P4506)**

The “level of availability of software for language assessment or language teaching” is a compound index (categorised mean score). The index equals the categorised average of the responses to items 6, 7 and 8 of question PQ45 *'Does your school have the following ICT facilities?: (6) Software or tools developed in house for learning and teaching languages; (7) Digital student portfolio; (8) Software for language assessment'*. The index has the following categories 0=“Low” (0 ≤ mean score < 0.25); 1=“Medium” (0.25 ≤ mean score < 0.75); 2=“High” (0.75 ≤ mean score ≤ 1).

### **Level of access to websites useful for (target) language learning (I05\_ED\_A\_P4504)**

The “level of access to websites useful for (target) language learning” is a compound index (categorised mean score). The index equals the categorised average of the responses to items 4, 5, 9, 10 and eleven of question PQ45 *'Does your school have the following ICT facilities?: (4) Software or access to websites specifically designed for learning languages; (5) Software for communication tools; (9) Access to online dictionaries and other reference works; (10) Access to online news media (TV, radio, newspapers) in target language ; (11) Access to other websites on life and culture in target language speaking country/countries'*. The index has the following categories

0="Low" ( $0 \leq \text{mean score} < 0.25$ ); 1="Medium" ( $0.25 \leq \text{mean score} < 0.75$ ); 2="High" ( $0.75 \leq \text{mean score} \leq 1$ ).

#### **10.6.4 Issue 6: Intercultural exchanges**

##### **Funding of student exchanges (I06\_ED\_M\_P390)**

"Funding of student exchanges" is a compound index (rounded mean score). The index is the average (rounded) of the responses to all items of question PQ39 '*To what extent are intercultural exchanges for students (such as exchange visits) funded in the following ways?*'.

#### **10.6.5 Issue 7: Staff from other language communities**

##### **Guest target language teachers participating in exchange visits (I07\_ED\_M\_P1801)**

"Guest target language teachers participating in exchange visits" is a simple index (dichotomised item score). The index equals the dichotomised response to item 1 of question PQ18 '*In the previous school year, how many teachers from abroad came to work in your school for longer than one month?: (1) Guest teachers of target language*' and has the categories 0 "No" (item score 0) and 1 "Yes" (item scores  $\geq 1$ ).

##### **Guest teachers participating in exchange visits (I07\_ED\_M\_P1804)**

"Guest teachers participating in exchange visits" is a simple index (dichotomised item score). The index equals the dichotomised response to item 4 of question PQ18 '*In the previous school year, how many teachers from abroad came to work in your school for longer than one month?: (4) TOTAL number of guest teachers*' and has the categories 0 "No" (item score 0) and 1 "Yes" (item scores  $\geq 1$ ).

#### **10.6.6 Issue 8: Language learning for all**

##### **Provision of formal education in language(s) of origin (I08\_ED\_M\_P4006)**

"Provision of formal education in language(s) of origin" is a simple index (item score). The index equals the response to item 6 of question PQ40 '*What type of extra lessons does your school offer to students?: (6) Extra lessons in students' home language if this is a different language to questionnaire language*' (0="No"; 1="Yes").



### **Provisions for help in mastering host language (I08\_ED\_M\_P4005)**

“Provisions for help in mastering host language” is a simple index (item score). The index equals the response to item 5 of question PQ40 *'What type of extra lessons does your school offer to students?: (5) Extra questionnaire language lessons for students with a different home language to questionnaire language'* (0=“No”; 1=“Yes”).

### **10.6.7 Issue 10: Teachers’ access to high quality initial and continuous training**

#### **Teacher shortage (I10\_ED\_O\_P110)**

“Teacher shortage” is a compound index (categorised scores). The index is based on the responses to all items of question PQ11 *'During the past five years, have you had difficulty in filling teaching vacancies or covering for absent teachers for the following subjects?'*. The index has the following categories:

- 0='No' when all items were answered with “No” (all item scores 0)
- 1='Only other subjects than languages' when all items referring to languages (item 1 to 4, maximum item score 0) was answered with “No” and the item referring to other subjects (fifth item) was answered with “Yes” (item score 1).
- 2='Only languages' when at least one of the item referring to languages (item 1 to 4, maximum item score 1) was answered with “Yes” and the item referring to other subjects (fifth item) was answered with “No” (item score 0).
- 3='Languages & other subjects' when at least one of the item referring to languages (item 1 to 4, maximum item score 1) was answered with “Yes” and the item referring to other subjects (fifth item) was answered with “Yes” (item score 1).

#### **Target language teacher shortage (I10\_ED\_O\_P1103)**

“Target language teacher shortage” is a simple index (item score). The index equals the response to item 3 of question PQ11 *'During the past five years, have you had difficulty in filling teaching vacancies or covering for absent teachers for the following subjects?: (3) For target language'* (0=“No”;1=“Yes”).

#### **Number of different financial incentives for in-service training from other sources than the school (I10\_ED\_M\_P150)**

The “number of different financial incentives for in-service training from other sources than the school” is a compound index (sum score). The index equals the number of items answered with “Yes” in question PQ15 *'Which of the following financial compensations can teachers get for participation in in-service training from sources other than your school, for example, funds or the national or local government?'*.

## **Number of different financial incentives for in-service training from school (I10\_ED\_M\_P140)**

The “number of different financial incentives for in-service training from school” is a compound index (sum score). The index equals the number of items answered with “Yes” in question PQ14 *'Which of the following financial compensations can teachers get from your school for participation in in-service training?'*

## **10.6.8 Issue 11: A period of work or study in another country for teachers**

### **Target language teachers participating in exchange visits (I11\_ED\_M\_P1701)**

“Target language teachers participating in exchange visits” is a simple index (dichotomised item score). The index equals the dichotomised response to item 1 of question PQ17 *'In the previous school year, how many teachers participated in teacher exchange visits to work or study in another country for longer than one month?: (1) Teachers of target language'*. The index has the categories 0 “No” (score 0) and 1 “Yes” (scores  $\geq 1$ ).

### **Teachers participating in exchange visits (I11\_ED\_M\_P1704)**

“Teachers participating in exchange visits” is a simple index (dichotomised item score). The index equals the dichotomised response to item 4 of question PQ17 *'In the previous school year, how many teachers participated in teacher exchange visits to work or study in another country for longer than one month?: (4) TOTAL number of teachers'* and has the categories 0 “No” (score 0) and 1 “Yes” (scores  $\geq 1$ ).

### **Funding for exchange visits (I11\_ED\_M\_P190)**

“Funding for exchange visits” is a compound index (minimum score). The index equals the minimum number of all items answered with “Yes” in question PQ19 *'In the previous school year, did any of the teachers or guest teachers receive funding for exchange visits in the following ways?'* The index has the categories 0=“No” and 1=“Yes”.

## **10.6.9 Issue 12: Use of existing European language assessment tools**

### **Frequency of student assessment (I12\_ED\_O\_P300)**

The “frequency of student assessment” is a compound index (rounded mean score). The index equals the average (rounded to a multiple of 0.2) of the responses to all items of PQ30 *'Generally, in your school, how often are students in all grades of ISCED2 and the first two years of ISCED3 assessed using the following methods? (as far as these grades are present in your school)'*.



### 10.6.10 Issue 13: Practical experience (no indices at school level)

### 10.6.11 Organisational structure of the educational systems

#### Ability grouping (I14\_ED\_A\_P040)

Ability grouping is a compound index (categorised maximum score). The index equals the categorised maximum of the responses to all items of question PQ4 'What is your school's policy on organising instruction for students with different abilities?'. The index has the following categories:

- 0='No form of ability grouping' (if all three forms of grouping are not applied for any subject (maximum = 0))
- 1='A form of ability grouping for some subjects' ( $0 < \text{maximum} < 2$ )
- 2='A form of ability grouping for all subjects' (maximum = 2).

#### Resources: Sufficient qualified teachers (I14\_ED\_O\_P4101)

"Resources: Sufficient qualified teachers" is a compound index (categorised mean score). The index equals on the categorised average of the (inverted) responses to items 1 through 5 of question PQ41 'Is your school's capacity to provide instruction hindered by any of the following issues?: (1) A lack of qualified language teachers; (2) A lack of qualified questionnaire language teachers; (3) A lack of qualified target language teachers; (4) A lack of qualified teachers of foreign languages (including ancient languages) other than target language ; (5) A lack of qualified teachers of subjects other than languages'.

As all items were contra-indicative, they were inverted before calculating the index. The index has the following categories:

- 0='No' when there is a lack of qualified teachers for all five subjects (mean inverted item score  $< 0.25$ )
- 1='Some' when there is a lack of qualified teachers for some subjects, but not all ( $0.25 \leq \text{mean inverted item score} < 0.75$ )
- 2='Yes' when there is no lack of qualified teachers (mean inverted item score  $\geq 0.75$ ).

#### Resources: Sufficient support personnel (I14\_ED\_O\_P4106)

"Resources: Sufficient support personnel" is a compound index (categorised mean score). The index equals on the categorised average of the (inverted) responses to items 6 and 7 of question PQ41 'Is your school's capacity to provide instruction hindered by any of the following issues?: (6) A lack of library staff; (7) A lack of other support personnel'.

As all items were contra-indicative, they were inverted before calculating the index. The index has the following categories:

- 0='No' when there is a lack of both types of support personnel (mean inverted item score  $< 0.25$ )
- 1='Some' when there is a lack of one type of support personnel, but not both ( $0.25 \leq$  mean inverted item score  $< 0.75$ )
- 2='Yes' when there is no lack of both types of support personnel (mean inverted item score  $\geq 0.75$ ).

#### **Resources: Sufficient ICT (I14\_ED\_O\_P4109)**

“Resources: Sufficient ICT” is a compound index (categorised mean score). The index equals on the categorised average of the (inverted) responses to items 9, 10 and 11 of question PQ41 *'Is your school's capacity to provide instruction hindered by any of the following issues?: (9) Shortage or inadequacy of computers for instruction; (10) Lack or inadequacy of Internet connectivity; (11) Shortage or inadequacy of computer software for instruction'.*

As all items were contra-indicative, they were inverted before calculating the index. The index has the following categories:

- 0='No' when all three ICT resources are lacking (mean inverted item score  $< 0.25$ )
- 1='Some' when some of the three ICT resources are lacking, but not all ( $0.25 \leq$  mean inverted item score  $< 0.75$ )
- 2='Yes' when all three ICT resources are sufficient (mean inverted item score  $\geq 0.75$ ).

#### **Resources: Sufficient instructional, library or audio-visual materials (I14\_ED\_O\_P4108)**

“Resources: Sufficient instructional, library or audio-visual materials” is a compound index (categorised mean score). The index equals the categorised average of the inverted responses to items 8, 12 and 13 of question PQ41 *'Is your school's capacity to provide instruction hindered by any of the following issues?: (8) Shortage or inadequacy of instructional materials (e.g. textbooks) ; (12) Shortage or inadequacy of library materials; (13) Shortage or inadequacy of audio-visual resources'.*

As all items were contra-indicative, they were inverted before calculating the index. The index has the following categories:

- 0='No' when all types of material are lacking (mean inverted item score  $< 0.25$ )
- 1='Some' when some of the materials are lacking, but not all ( $0.25 \leq$  mean inverted item score  $< 0.75$ )

- 2='Yes' when all materials are sufficiently present (mean inverted item score  $\geq 0.75$ ).

**Resources: sufficient audio-visual material in target language (I14\_ED\_O\_P4201)**

“Resources: sufficient audio-visual material in target language” is a compound index (categorised mean score). The index equals the categorised average of the responses to items 1, 2, 3 and 4 of question PQ42 *'In your opinion, are the following resources in your school sufficient to support the instruction in foreign languages?: (1) Audio cassettes, CDs or other audio-material spoken in target language; (2) Audio cassettes, CDs or other audio-material spoken in foreign languages other than target language; (3) Video cassettes, DVDs, video clips from YouTube or other audio-visual material spoken in target language; (4) Video cassettes, DVDs, video clips from YouTube or other audio-visual material spoken in foreign languages other than target language'.*

The index has the following categories:

- 0='No' when all four audio-visual resources are insufficient (mean item score  $< 0.25$ )
- 1='Some' when some of audio-visual resources are sufficient, but not all ( $0.25 \leq$  mean item score  $< 0.75$ )
- 2='Yes' when all four audio-visual resources are sufficient (mean item score  $\geq 0.75$ ).

**Resources: sufficient lesson material for target language (I14\_ED\_O\_P4207)**

“Resources: sufficient lesson material for target language” is a compound index (categorised mean score). The index equals the categorised average of the responses to items 7, 8, 11 and 12 of question PQ42 *'In your opinion, are the following resources in your school sufficient to support the instruction in foreign languages?: (7) Textbook(s) for target language ; (8) Textbook(s) for foreign languages other than target language ; (11) Lesson materials prepared by the teacher(s) of target language (e.g. hand-outs, Reading texts) ; (12) Lesson materials prepared by the teacher(s) of other foreign languages than target language (e.g. hand-outs, Reading texts)'.*

The index has the following categories:

- 0='No' when all four types of lesson material for target language are insufficient (mean item score  $< 0.25$ )
- 1='Some' when some of the four types of lesson material for target language are sufficient, but not all ( $0.25 \leq$  mean item score  $< 0.75$ )
- 2='Yes' when all four types of lesson material for target language are sufficient (mean item score  $\geq 0.75$ ).

### **Resources: sufficient library material in target language (I14\_ED\_O\_P4205)**

“Resources: sufficient library material in target language” is a compound index (categorised mean score). The index equals the categorised average of the responses to items 5, 6, 9 and 10 of question PQ42 *'In your opinion, are the following resources in your school sufficient to support the instruction in foreign languages?: (5) Newspapers, magazines, comics or song texts written in target language ; (6) Newspapers, magazines, comics or song texts written in foreign languages other than target language; (9) Books written in target language for extensive Reading, e.g. fiction; (10) Books written in foreign languages other than target language for extensive Reading e.g. fiction'.*

The index has the following categories:

- 0='No' when all four types of library material for target language are insufficient (mean item score  $< 0.25$ )
- 1='Some' when some of the four types of library material for target language are sufficient, but not all ( $0.25 \leq \text{mean item score} < 0.75$ )
- 2='Yes' when all four types of library material for target language are sufficient (mean item score  $\geq 0.75$ ).

### **School responsibility for curriculum (I14\_ED\_O\_P4610)**

“School responsibility for curriculum” is a compound index (proportion). The index is based on the responses to item 10 through 13 of question PQ46 *'For your school, who has considerable responsibility for the following tasks?: (10) Choosing which textbooks are used; (11) Determining course content; (12) Determining which foreign languages are offered as a subject; (13) Determining in which order foreign languages are offered'.*

For each entity (principals, teachers, the school governing board, the regional or local education authority, the national education authority) the number of tasks (item 10 through 13) for which they hold responsibility were counted. The index equals the sum of tasks for which the school principal has responsibility and of tasks for which the teachers have responsibility divided by the total counted tasks for principals, for teachers, for the school governing board, for the regional or local education authority, and for the National education authority together. Higher values on the scale indicate relatively higher levels of school responsibility for the curriculum. A value zero means that the school (principals and teachers) has no responsibility for the curriculum and a value one means that only the school (principals and teachers) has responsibility for the curriculum.

### **School responsibility for resource allocation (I14\_ED\_O\_P4601)**

“School responsibility for resource allocation” is a compound index (proportion). The index is based on the responses to item 1 through 6 of question PQ46 *'For your*

*school, who has considerable responsibility for the following tasks?: (1) Selecting teachers for hire; (2) Firing teachers; (3) Establishing teachers' starting salaries; (4) Determining teachers' salary increases; (5) Formulating the school budget; (6) Deciding on budget allocations within the school.*

For each entity (principals, teachers, the school governing board, the regional or local education authority, the national education authority) the number of tasks (item 1 through 6) for which they hold responsibility were counted. The index equals the sum of tasks for which the school principal has responsibility and of tasks for which the teachers have responsibility divided by the total counted tasks for principals, for teachers, for the school governing board, for the regional or local education authority, and for the National education authority together. Higher values on the scale indicate relatively higher levels of school responsibility for resource allocation. A value zero means that the school (principals and teachers) has no responsibility for resource allocation and a value one means that only the school (principals and teachers) has responsibility for resource allocation.

### **School selectivity (I14\_ED\_A\_P030)**

“School selectivity” is a compound index (categorised maximum score). The index equals the categorised maximum of the responses to items 3 and 4 of question PQ3 '*How often are the following factors considered when students are admitted to your school?: (3) Student's record of academic performance (including placement tests) in all subjects; (4) Recommendation of feeder schools*' (0=Never; 1=Hardly ever; 2=Every now and then; 3=Usually; 4=Always).

The index has the following categories:

- 0='Neither of the two aspects is considered for student admittance' (maximum item score =0)
- 1='At least one of the two aspects is considered for student admittance' (0 < maximum item score <4).
- 2='At least one of the two aspects is always considered for student admittance' (maximum item score = 4)

### **School type (I14\_ED\_O\_P070)**

“School type” is a simple index (item score). The index equals the response to question PQ7 '*Is your school a public or a private school?*', which has the response options 0=A public school; 1=A private school.

### **Target language as a compulsory subject (I14\_ED\_O\_P2601)**

“Target language as a compulsory subject” is a compound index (categorisation of scores). The index is based on the categorisation of the responses to item 1 and 1 of question PQ26 '*What is the status of target language in the school's curriculum?: (1)*

*Target language is a compulsory subject in the curriculum of all students; (2) Target language is a compulsory subject in the curriculum of some students'*. The index has the following categories:

- 0='Not for any student' when both items were answered with "No" ( $item_1=0$  &  $item_2=0$ )
- 1='For some students' when only the second item was answered with "Yes" ( $item_1=0$  &  $item_2=1$ )
- 2="For all students" when the first item was answered with "Yes" ( $item_1=1$ ).

## 10.7 References

- Eurydice (2008) *Key Data on Teaching Languages at School in Europe - 2008 Edition*.
- Eurydice (2010) *The structure of the European education systems 2010/11: schematic diagrams*. European Commission.
- Eurydice. (sd) *Eurypedia - The European Encyclopedia on National Education Systems*. Opgehaald van Eurydice:  
[http://eacea.ec.europa.eu/education/eurydice/eurybase\\_en.php](http://eacea.ec.europa.eu/education/eurydice/eurybase_en.php)
- Ganzeboom, H B, and Treiman, D (1996) *Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations*, *Social Science Research* 25, 201-239.
- Jöreskog, K, and Sörbom, D (2004) LISREL 8.7 for Windows [Computer Software]. Lincolnwood, IL: Scientific Software International Inc.
- OECD. (2007). *PISA 2006 Technical report*. Paris: OECD.
- OECD. (2012). *PISA 2009 Technical Report (preliminary version)*. Opgehaald van Programme for International Student Assessment (PISA):  
[http://www.oecd.org/document/19/0,3746,en\\_2649\\_35845621\\_48577747\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/19/0,3746,en_2649_35845621_48577747_1_1_1_1,00.html)
- Verhelst, N., Glas, C and Verstralen, H (1995). *One parameter logistic model (OPLM)*. Arnhem: CITO.

# ***Chapter 11: Setting standards in relation to the CEFR***



## 11 Setting standards in relation to the CEFR

Standard setting can be regarded as both a procedural and a conceptual issue. Kane (1994) distinguishes ‘the *passing score*, defined as a point on a score scale, and the *performance standard*, defined as the minimally adequate level of performance for some purpose,’ where the former captures the operational level, and the latter the conceptual one. Conceptually, therefore, standard setting represents an *inference* from test performance to the proficiency a student could demonstrate in some purposeful domain of activity. This recalls current conceptions of *validity* in assessment, as the degree to which test results justify inferences (Cizek and Bunch 2007:17, Messick 1989). In the context of the ESLC the domain to which test performance is referred is the framework of proficiency levels described by the CEFR.

The standard setting literature is careful to reject the idea that there is a reality behind standards. ‘Standard setting does not seek to find some pre-existing or ‘true’ cutting score that separates real, unique categories on a continuous underlying trait (such as ‘competence’)’ (Cizek and Bunch 2007:18). How should we relate this statement to the context of the ESLC? Clearly the purpose of the CEFR and of those who work with it is to enable different contexts of learning and different languages to benefit from a shared understanding of levels. The CEFR identifies and describes discrete levels as an aid to developing such a shared understanding, and an important aim of the ESLC is to provide tests which operationalise these levels. We cannot deny that current understanding of the CEFR may still reflect local standards and custom, and that comparability across contexts of learning has its inevitable limits (see Chapter 3.2 of the Final Report); at the same time the purpose of the ESLC requires us to make every effort to ensure that the standards set reflect a common interpretation of the CEFR levels, and are as comparable across languages as possible. Without insisting on any absolute truth we must aim at consistency.

The problem of linking tests in five languages to the CEFR becomes much easier if these tests are comparable with each other and relate validly to the CEFR. As described more fully in Chapter 2 the language tests developed for the survey set out to satisfy these requirements.

The design and implementation of the language tests not only allow some confidence that the tests constructed for each language and skill relate validly to the CEFR, but also that the difficulty of the tasks should be broadly comparable across languages. As described further below, this provides important evidence for setting standards.

Many methods are available for setting standards. Different methods differ both with respect to the information provided to, and with respect to the type of judgment required from, panel members. Regardless of which method is chosen, each method will show some degree of disagreement among panel members, and hence requires reconciliation to produce a standard that has support from as broad a range of panel members as possible.

## 11.1 Outline of the standard setting conference (Sept 26-30 2011)

This event took place in Cambridge from September 26<sup>th</sup> to September 30<sup>th</sup> 2011. Standard setting was done by panels of judges. There was a panel per language. Table 54 shows the number of panel members for each of the five languages. These are maximum numbers, but the actual number per round is never less than in Table 54 minus 1, or 2 in the case of English.

**Table 54 Number of panel members per language**

Language	# panel members
English	21
French	9
German	14
Italian	8
Spanish	12

Standards were set by language, with the five panels working separately.

Standards were set for Listening, Reading and Writing. The standards were set for each skill three times, in three different rounds. These rounds can be described as follows:

- (i) Round 1: individual standard setting after having taken the test (as a student)
- (ii) Round 2: individual standard setting after discussion of the results of round 1. The results are said to provide normative information: individual members can see their position in the set of individual standards set in the first round. The discussion (in small groups) is aimed at clarifying differences and to find out if these differences are due to either a misunderstanding of the CEFR or a misinterpretation of the demands of the tasks. It is expected that the inter-rater differences will decrease in the second round.
- (iii) Round 3. This round is intended as a validation procedure. Panel members set their individual standards again but with a variation in the task they have to fulfil. This variation is described in detail below.

The week's activities were organized approximately as follows:

- Monday: general introduction to SurveyLang, familiarisation to the CEFR, and introduction to the methods of standard setting to be followed.
- Tuesday and Wednesday morning: round 1 for all skills
- Wednesday afternoon and Thursday morning: round 2 for all skills
- Thursday afternoon: round 3

- Friday morning: roundup, presentation of results and filling out the feedback questionnaire.

## 11.2 Standard setting methodology

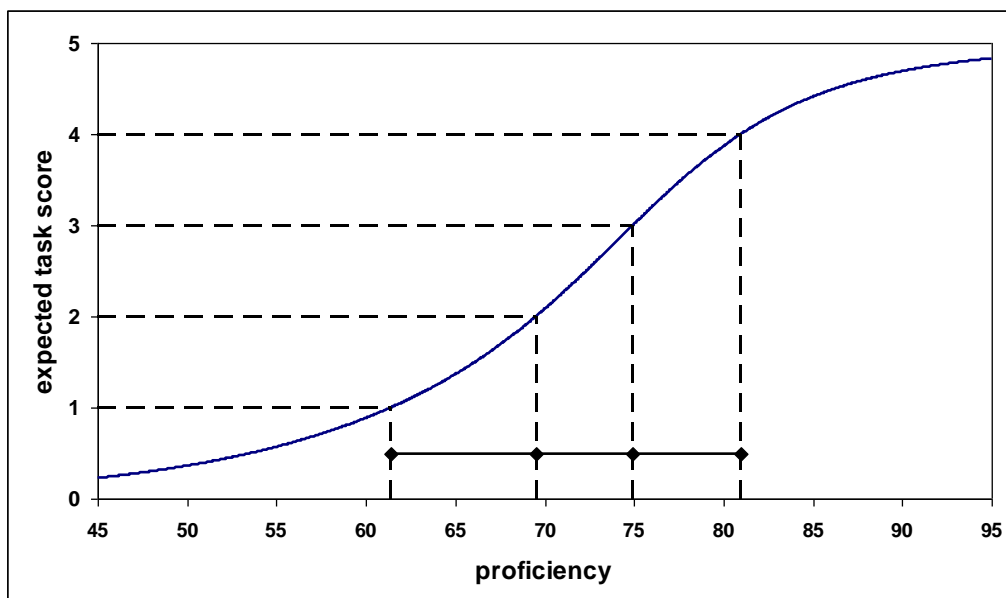
### 11.2.1 Reading and Listening

For the objectively-marked skills of Reading and Listening, the approach is a task-centred one. A partial credit IRT model was used for ESLC tasks, to avoid effects caused by item dependencies within tasks. In consequence there are no item parameters. To deal with this it was decided to use a variation on the Van der Schoot method (Van der Schoot 2009). This is similar to the Cito variation on the bookmark method described in the *Manual*, (Council of Europe 2009), which requires item parameters.

The model used considers the task as the basic unit and defines the score on a task as the number of correctly answered items (within a task). It fits the collected data reasonably well.

Once the parameters are estimated, one can construct the task response function. This function relates the latent variable (the proficiency) to the expected score on the task. (In fact, it is the regression function of the task score on the latent variable). An example (task EL231: a Listening task for English at level B1) is given in Figure 39.

**Figure 39 A task response curve**



Here are some comments to Figure 39:

- The curve is S-shaped. This is necessarily the case because the proficiency scale is unbounded and the score is bounded from below and from above. This means also that the expected score is zero or its maximal value (5) only if the proficiency is  $-\infty$  or  $+\infty$ , respectively.
- The proficiency scale itself is an interval scale where origin and unit are arbitrary. Origin and unit have been chosen in such a way that no negative or fractional numbers appear in the figure.
- The dashed lines show the relationship between proficiency and score on the task: for example, if the proficiency value is 75 (approximately), then the expected score on the task is 3.
- The horizontal line just above the horizontal axis summarizes the whole curve: the diamonds represent the position on the proficiency axis which lead to an expected score of 1, 2, 3 and 4 respectively. All proficiency values left of the first diamond lead to an expected score of less than 1 point and all proficiency values to the right of the rightmost diamond lead to an expected score less than one unit away of the maximum score. This horizontal line is an effective way of displaying information on several items together, Figure 40 illustrates for three tasks.

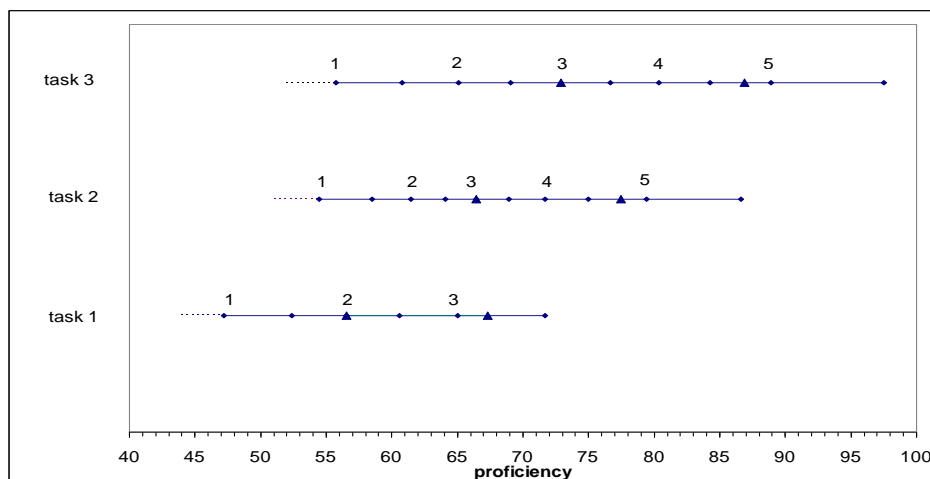
The form of presentation used in standard setting is similar to Figure 39. It has these features:

- On the horizontal lines, expected scores are displayed at every half unit, starting from 1 and continuing until the maximum score minus a half. Integer values are indicated by the corresponding number.
- In the instructions for the panel members three levels of mastery of a task were defined, analogous to the Van der Schoot method:
- Full mastery is defined as an expected score at least 80% of the maximum.
- Moderate mastery is defined as an expected score between 50% and 80% of the maximum score.
- An expected score of less than 50% counts as no mastery.
- The points which represent 50% and 80% of the maximum score are indicated by a triangle on the line representing the task. Example: task 3 in Figure 40 has a maximum score of 6: the triangles are located at the proficiency values giving an expected score of  $0.5 \times 6 = 3$  and  $0.8 \times 6 = 4.8$ .

When displaying the lines of several items in a single display, one has to decide on the order in which they are displayed. In all displays used in the standard setting the displays are ordered according to the 50% point of mastery. This means that going

from the bottom task to the top task in Figure 40, the leftmost triangles are ordered in terms of proficiency (they go from left to right).

**Figure 40 Summary of task response curves for three tasks**



Some of the tasks consist of two-choice MC items (true/false). It was explained to the panel members that blind guessing in such a case would lead to an expected score of 50% of the maximum, but that in this case a 50% score could not be considered as moderate mastery<sup>31</sup>.

Central to this procedure is the concept of a borderline person. In the example the panel members were asked to imagine a person (or several persons) who are borderline B1, i.e., their proficiency is certainly top-A2 but not yet convincingly B1.

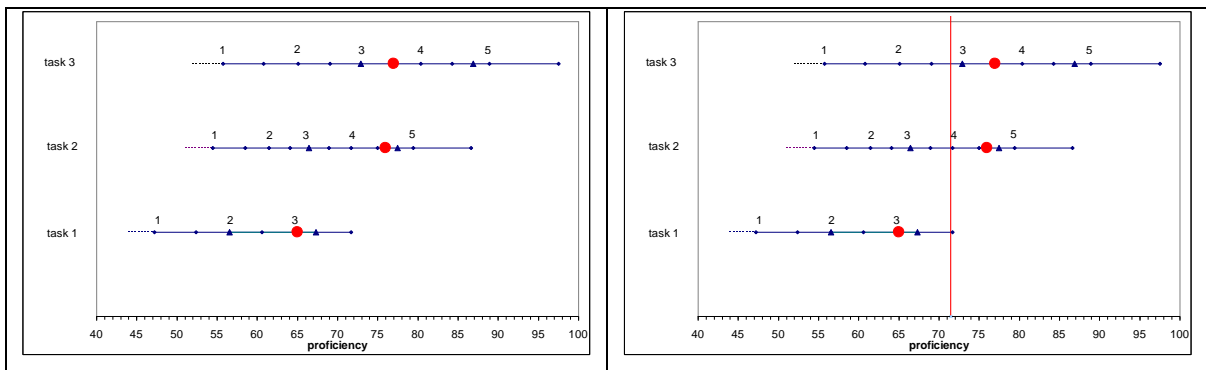
The task for the panel members then consists of two steps:

In step 1 they consider each task (a task booklet is available) and decide what the expected score is for such a borderline person and to indicate this on the graph for this task (by drawing a cross). This is illustrated in the left-hand panel of Figure 41.

Each marked position (one per task) is in fact a standard setting, but if the marks are not set on a vertical line, some compromise has to be found, either by reconsidering the requirement for single tasks or by just choosing a kind of average position. This compromise is indicated by drawing a vertical line through the display as shown in the right-hand panel of Figure 41. Panel members had set-squares to help them draw a vertical line.

<sup>31</sup> It would have been a good idea to indicate the expected score under a strategy of blind guessing on the displays. This may be a useful suggestion for future replications of standard setting using this method.

**Figure 41 The two steps in the standard setting procedure**

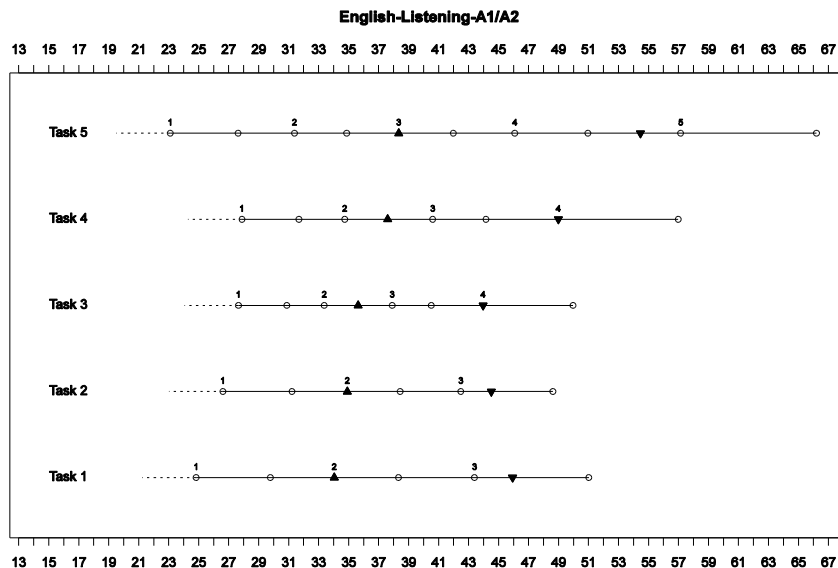


In round 1 and round 2, four standards were set. This required three different subsessions, each employing a different answer sheet.

- For A1 and A2, the sheet displayed the tasks with intended levels A1 and A2. An example is given in Figure 42 (English Listening).
- For B1 the sheet displayed all the A2 and B1 tasks.
- For B2 the sheet displayed all the B1 and B2 tasks.
- As a consequence the A2 and B1 tasks were displayed twice.

For each of the sub-rounds a separate text-and-item booklet was constructed. The panel members were told that the A2 and B1 tasks would appear twice, although the intended level for each task was not indicated. In the example of Figure 42, there are five tasks: two A1 tasks (task 1 and task 3) and 3 A2 tasks (tasks 2, 4 and 5).

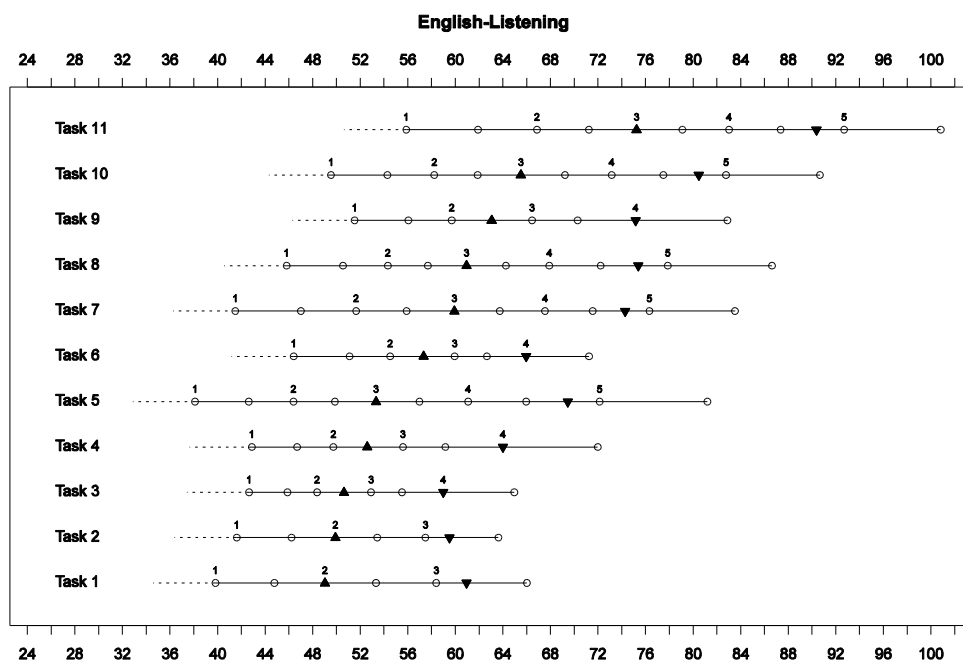
**Figure 42 An example of an answer sheet (English Listening, standards A1 and A2)**



The individual standard (see right panel of Figure 41) was read from the response sheet to the nearest half unit and these are the basic data for the analysis.

In round 3 (the validation round), the same kind of graphs were used, but with *all* tasks displayed on a single sheet. The task for the panel members was to set standards for the four levels on this sheet. An example (for English Listening) is given as Figure 43. For the third round a table was available for the panel members where the intended level of each task was indicated.

**Figure 43 Example of the answer sheet for Round 3**



### 11.2.1 Writing

For Writing a student centred method was used. The complete writing performances of a number of students per language (i.e. 2 or 3 tasks, depending on the assigned test level) were sampled and transcribed. For a fuller description of the selection procedure see section 11.4.1 below.

All 8 tasks used in the Main Study were sampled. (For Italian: only 3 tasks were used, one at each of the levels A2, B1 and B2; for this language no students were assigned to the low level.)

12 performances were selected for each of the 8 tasks (or 3 for Italian). The number of different students selected was 52 for English, 50 for French and German, 20 for Italian and 53 for Spanish.

For rounds 1 and 2, a variation of the Body of Work (BoW) method was used. For each task the 12 performances were to be sorted into passes or fails, bearing in mind the intended level of the task, i.e. those judged to have dealt adequately with the demands of the task and those who did not.

A task-level approach was chosen for two reasons:



- CEFR levels are usefully understood in relation to the tasks a student can perform at a level. The challenge presented by a task influences students' performance as well as raters' perceptions of it. Thus it is important for raters to be aware of, and judge in relation to, the demands of the task.
- The task-level approach was coherent with that used in the Writing Alignment study (section 11.4) and was thus hoped to elicit comparable behaviour.

Round 1 responses were captured and presented in the form of a table of tasks by raters, with a zero or one in each cell (passed or failed) and marginal percentage totals. This was provided as normative information in Round 2, where raters were asked to consider and discuss their ratings, and if they wished, change them.

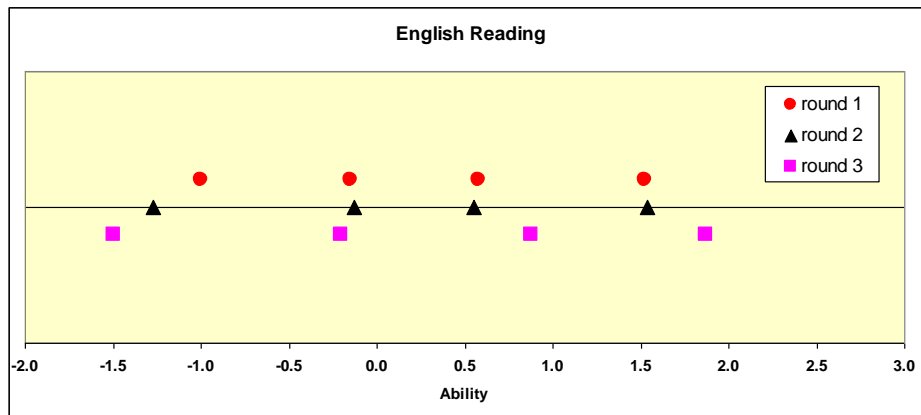
In round 3, the BoW method was used at student level: for each of 30 students the panel member had to assign a CEFR level on the basis of a student's complete set of performances (2 or 3 tasks). Raters were offered 8 categories (a higher and lower category for each of A1, A2, B1), and these were collapsed to 5 (Pre-A1 to B2) for analysis. The analysis of these data will also be discussed in the results section.

### **11.3 The standard setting conference - results**

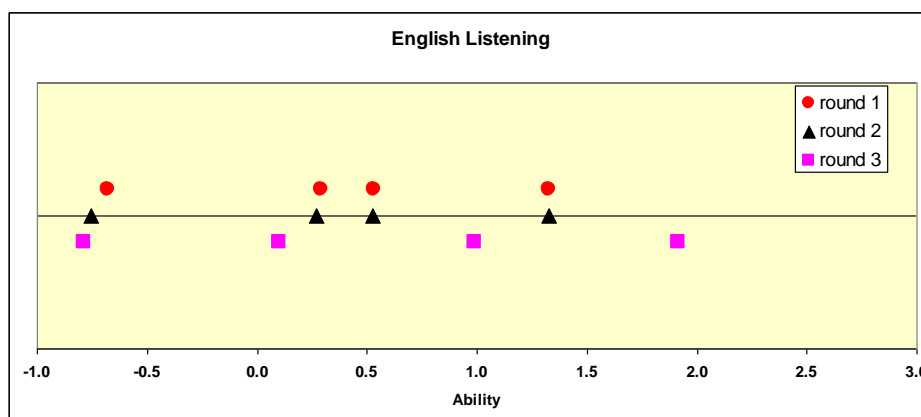
Not unexpectedly, the standard setting produced mixed results in terms of coherence and interpretability. For all skills, significant differences were found for the method used in rounds 1 and 2, and that used in round 3. Also within methods, significant differences were found among panel members and/or tasks. It is well understood that different standard setting procedures tend to produce different results. It is also better to have identified the differences than to have used one method only and remained quite unaware of the issues.

Reconciling the outcomes requires a mix of informed judgment and statistical evidence. The aim of the reconciliation described below is to identify a standard that reflects as well as possible the individual judgments of as many of the panel members as possible.

**Figure 44 CEFR level cut-offs from 3 rounds of standard setting (English Reading)**



**Figure 45 CEFR level cut-offs from 3 rounds of standard setting (English Listening)**

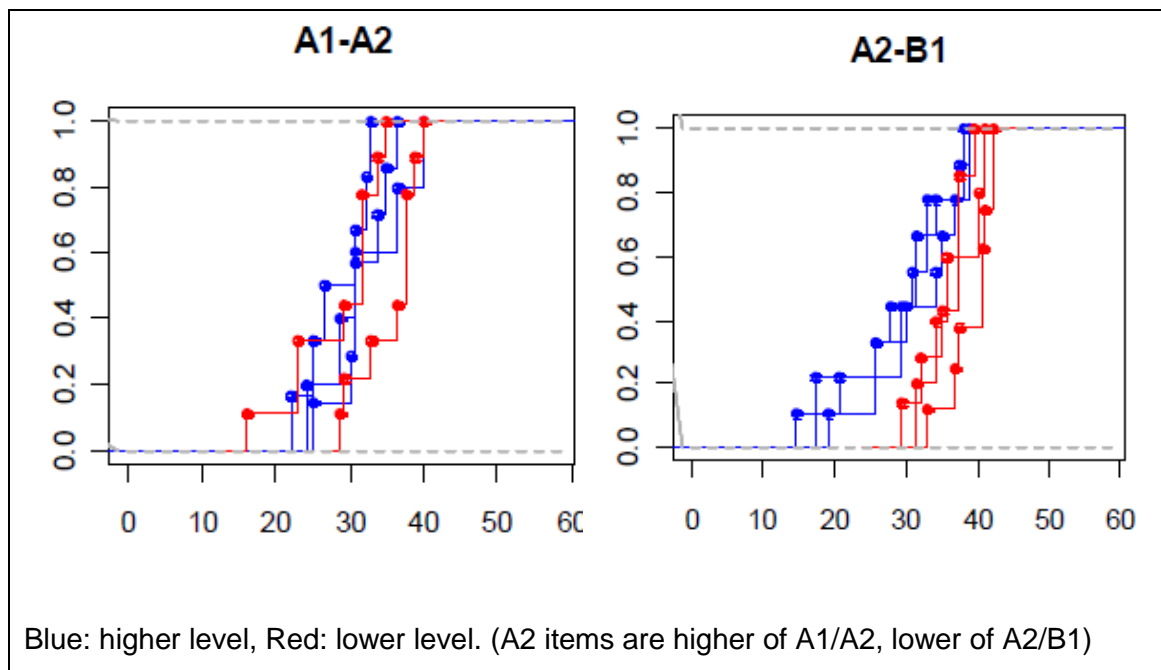


### 11.3.1 Reading and Listening: rater behaviour

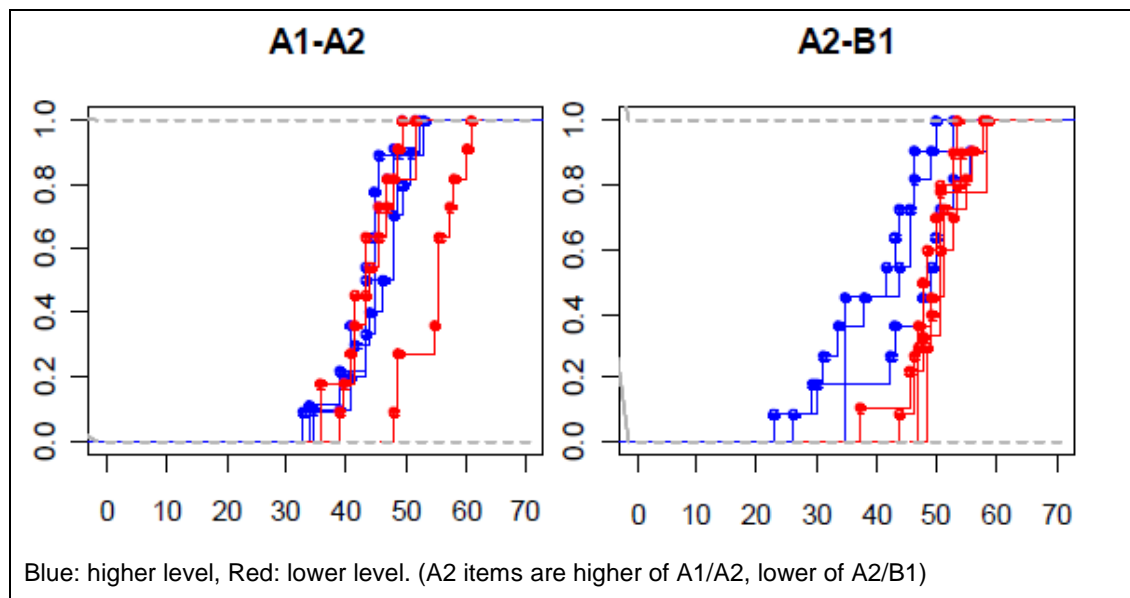
At first glance, the standards derived from the different rounds produced some unexpected outcomes. This was most evident in the relative placement of the A1/A2 and the A2/B1 cut-offs. These in several cases fell quite close together, and in two were actually reversed, i.e. a lower standard of performance was expected at B1 than at A2.

A close study of individual judgments reveals a characteristic pattern of behaviour associated with many of these cases. Figure 46 and Figure 47 illustrate for the cases of French Listening and Spanish Reading.

**Figure 46 French Listening**



**Figure 47 Spanish Reading**



The two testing levels A1/A2 and A2/B1 are linked by the common A2 tasks. A2 is the higher level in the first, (shown red), the lower level in the second (shown blue).

Every line in these Figures relates to a particular task, and gives the distribution of the scores judges assigned to the borderline candidate for the particular task. All judgements were translated to the expected score metric on the full set of tasks. This allows for the largest degree of cross language comparability.

. Figure 46 shows that for French listening the A2 tasks, which provide the anchor across levels, discriminate the levels reasonably well: the average level demanded is about 30 for the A1/A2 cut-off, and about 35 for the A2/B1 cut-off. However, for the B1 tasks a significantly lower ability is demanded, so that the A2/B1 cut-off overall is set too low. The same pattern is observed in Figure 47 for Spanish Reading.

These figures clearly identify differences between tasks, across the judges, in terms of coherence, rather than systematic differences between judges, across the tasks.

This study motivates a decision to ignore those tasks that are systematically different from the standard implied by the majority tasks.

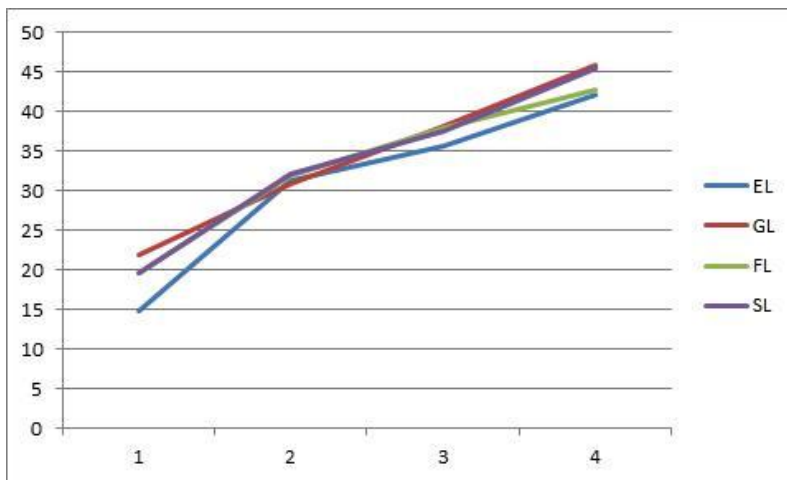
While this pattern characterises a number of cases, it is by no means general. Also, other effects may be at work. For example, the fact that the pre-A1 and A1 cut-offs were set on the same set of tasks, (the A1/A2 level test), was likely to lead judges to make a clearer separation of these cut-offs, and perhaps shift the A1/A2 cut-off higher than they might otherwise have done. This would be in line with round 3 outcomes, where all levels and tasks were shown on a single sheet of paper, and where a greater discrimination of levels was also observed.

Reconciling the judgments provided by panel members for the different tasks is not an "algorithmic" affair. There is no one rule that will always produce the right, or even the best, standard. Rather an informed judgment is needed to set a standard that does justice to as many individual judgments as possible. In the next section we consider how the cross-language perspective adds to the credibility of the (provisional) standards obtained for each of the languages.

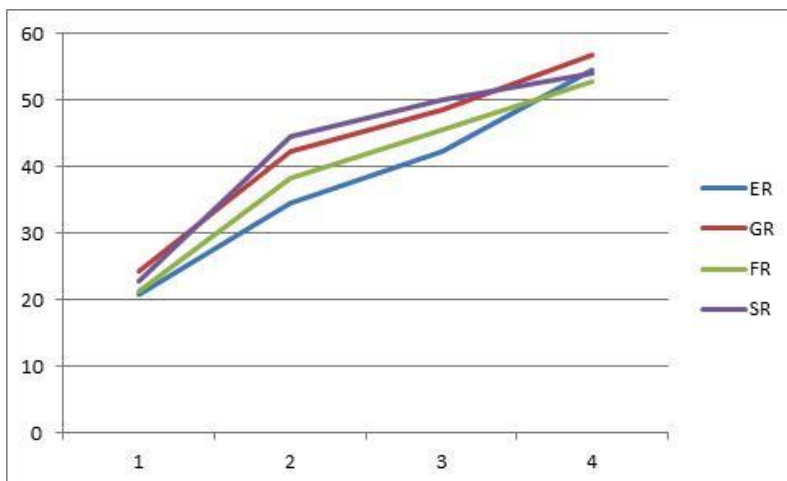
### **11.3.2 Reading and Listening: cross-language comparison**

From the cleaned set of rater judgments a composite summary view of the cut-offs for four languages was produced (Figure 48 and Figure 49). Italian was omitted at this stage as the tests comprised a subset of the tasks used by the other languages.

**Figure 48 Listening: summary of cut-offs for English, German, French and Spanish**



**Figure 49 Reading: summary of cut-offs for English, German, French and Spanish**



In the figures above the x axes represent the four cut-offs A1, A2, B1, B2.

The vertical axes represent scores on the total test, i.e. the expected scores of a student who took every task included in the ESLC. This transformation of the latent trait ability values is useful because it enables languages to be directly compared. If the test tasks at each level are of comparable difficulty and the same standard is set for each language, these lines should lie close together.

As noted in Chapter 2, considerable effort went into constructing comparable tests, so the assumption of similar overall difficulty deserves serious consideration. Figure 48 (Listening) shows quite good agreement, with only English A1 diverging significantly. Figure 49 (Reading) shows more divergence at A2 and B1. A decision is needed

whether to accept these standards as representing substantive differences between the tests, accurately diagnosed by the raters; or to weight more heavily the argument for comparability, and treat differences as random error on the part of five teams, working separately on essentially the same problem.

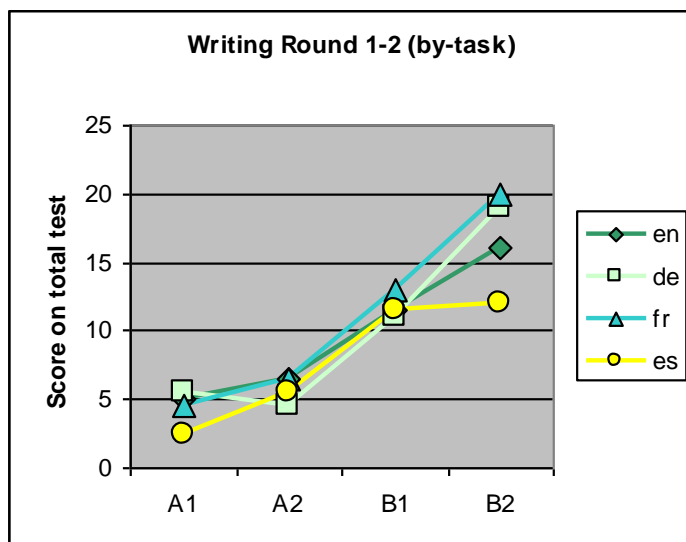
The second approach was chosen, and the final standards were set by averaging for each skill across the four languages (in the case of Listening, after raising English A1 to 20). Italian was subsequently fitted to the other languages through reference to the common task set.

### 11.3.3 Writing – results

For Writing the judgments were modelled using logistic regression with language proficiency as predictor. As noted in section 11.3 above, the different methods used in round 1-2 and in round 3 produced different results: the by-student judgments stretched the distance between A1 and B2 more than did the by-task approach. It is not clear whether this reflects differences in the logistic models used, in the degree of agreement elicited by the two methods, or substantive effects in how raters view a student's complete body of work or a single task. Some approach to reconciling these two sources of data is needed.

As with Reading and Listening, standards were finalized by comparing across languages and using a score metric – the modelled total score were a student to complete all tasks. This enables a direct comparison across languages on the assumption that scores should be broadly comparable.

**Figure 50 Writing cut-offs from round 1-2 (by-task)**



**Figure 51 Writing cut-offs from round 3 (by-student)**

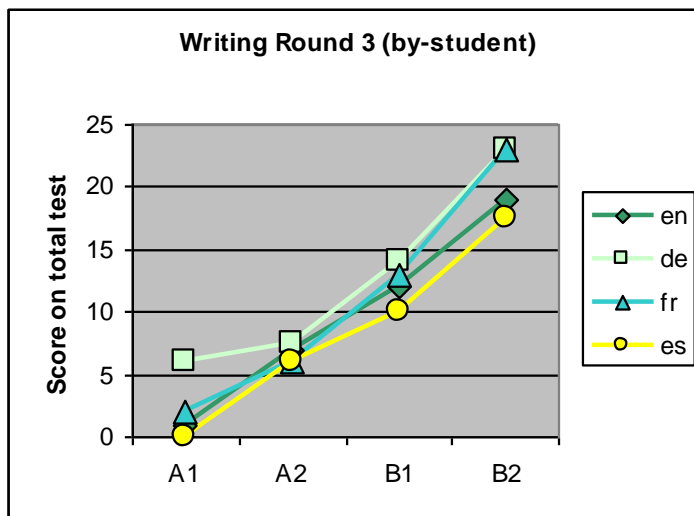


Figure 50 and Figure 51 show for four languages the cut-offs derived from logistic regression (by-task) or polytomous regression (by-student). The A2 and B1 cut-offs are similar for both methods. The B2 cut-off varies substantially for the by-task method, whereas it is more consistent for the by-student method. The A1 cut-off is poorly discriminated from the A2 cut-off in the by-task method, and is consistently lower in the by-student method.

The A2 and B1 cut-offs were set by averaging over the two methods and the four languages. The A1 and B2 cut-offs were set to reconcile differences between languages and weighting the outcomes of the by-student method.

The A2 and B1 cut-offs were set by averaging over the two methods and the four languages. The A1 and B2 cut-offs were set to reconcile differences between languages and weighting the outcomes of the by-student method. Italian, which used a reduced task set, was subsequently fitted to the other four languages.

For Writing the cross-language comparability of these standards could be validated against the outcomes of the Writing Alignment study, which had been completed prior to the standard-setting conference (section 11.4 below).

## **11.4 The writing alignment study (August 2011)**

### **11.4.1 The study design**

For the performance skill of Writing samples of performance in different languages can be directly compared, and in this way inform an alignment of standards. The eight ESLC Writing tasks used in the Main Study are essentially cloned across the five languages, making comparison by task straightforward.

The purpose of the study was to inform alignment of standards set. It was not intended to impact at all on the underlying scale construction for each language and skill from Main Study response data.

The approach was based on ranking: raters with a competence in two languages were asked to rank order sets of samples, half in each language. The use of ranking in a multilingual study was validated at Sevres in 2008 at a conference to set common standards for Speaking in five languages. This used both ranking and rating, and indicated that ranking produced very similar results, in terms of rank order, to the more familiar rating approach (Jones 2009). As ranking involves relative rather than absolute judgment it is well suited to the purpose of aligning performances across languages onto a common scale, thus enabling a check on standards set separately per language at the standard setting conference.

The study design was as follows:

- Students were selected from the multiply- and centrally-marked Main Study scripts, randomly from each level of a distribution stratified by booklet and average mark. This ensured linking across levels, and a gradation of levels of performance. Selection was done for each language by the appropriate language partner, who verified the suitability of the performances of each selected student and could select an alternative if necessary. All the performances of selected students (2 or 3 tasks) were transcribed and entered into a database. The samples used for standard setting were selected from these. The alignment study samples are a subset of the standard setting samples.
- From the larger list of transcribed tasks, samples were selected evenly across all the test booklets in which each task was included, to ensure linking across levels.
- For each task in each language a total of between 9 and 13 samples were selected, the difference reflecting the number of booklets in which the task appeared. These were grouped into 2 or 3 non-overlapping allocations per task.
- The allocation design ensured equal use of all allocations in different permutations, ensuring linking across allocations and languages.
- Each rater was assigned a set of 8-10 samples, half in one language, half in another, for each of the eight tasks. To lessen the workload each rater was instructed to rate each task on only one of the two criteria (odd tasks on language, even tasks on communication, or vice versa).

The study was conducted by email. Raters were recruited through requests to NRCs, invitation to experts due to attend the standard setting conference, through SurveyLang partners, and other ways. While most raters accepted one assignment (i.e. two languages), a few accepted more. Assignments were sent to more than 100 raters. The completed responses of 80 raters were finally available for analysis.



Raters were asked to print out their allocated samples, rank them for each task from highest to lowest, and record their rankings on paper. They could then upload their responses to an online survey portal using a personalised link identifying each rater by their email address.

### **11.5 Writing alignment study outcomes**

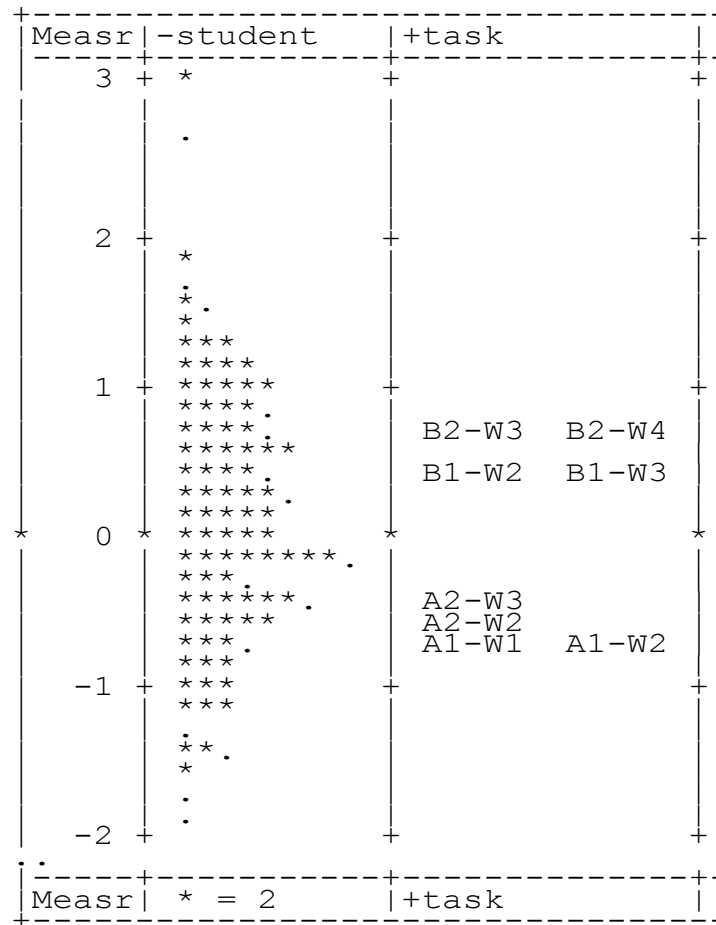
The data were analysed using multi-faceted Rasch analysis (Linacre 2011). Features of the analysis design include:

- The 8 tasks were treated as the same across languages. The absence of significant task by language interaction effects confirmed that this was reasonable.
- Linking across tasks is provided by the students, who each respond to 2 or 3 tasks. It is the student abilities which are modelled rather than each individual performance. These abilities can be compared with those from the Main Study analysis. The productions of 189 students were included in the analysis, each receiving on average 20 rankings.

Figure 52 illustrates outcomes graphically. The student abilities are reasonably well distributed, but over quite a short scale. The tasks rank in their intended order, with B2 the most difficult, A1 the easiest. There is good separation of A2 and B1 tasks, but less separation of A1 and A2 tasks. The variance explained by Rasch measures in this analysis is 51.9%, indicating a degree of noise in the ranking data.

The reliability of the student estimates is 0.95. The mean standard error of student estimates is 0.19

**Figure 52: Summary output from FACETS analysis of alignment data**



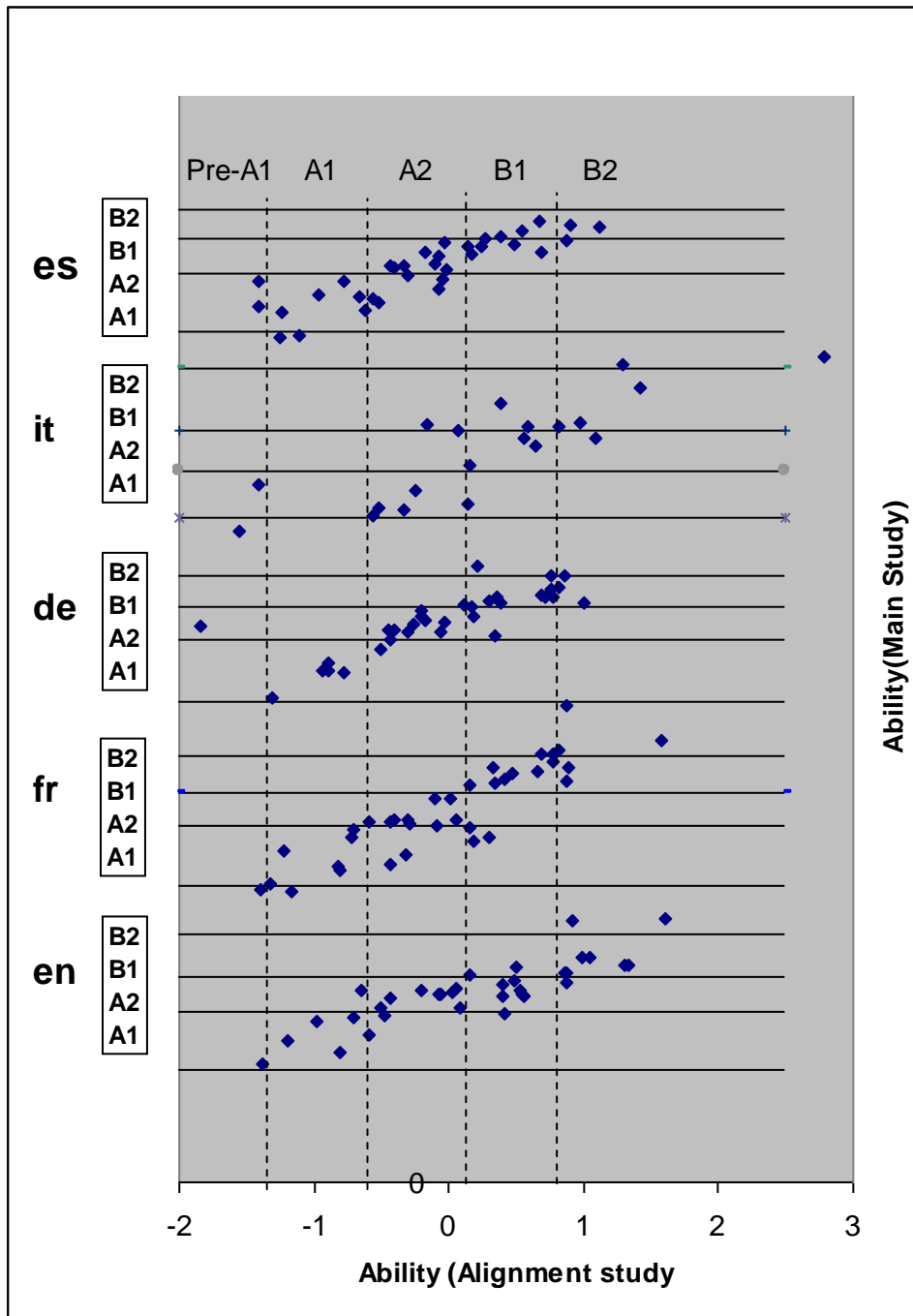
The correlation between ability estimates from the Alignment Study (AS) and Main Study (MS) is shown in Table 55.

**Table 55 R<sup>2</sup> of AS and MS ability estimates**

Language	R <sup>2</sup>	N
en	0.811	35
fr	0.811	34
de	0.851	29
it	0.708	19
es	0.804	32

Figure 53 compares student abilities as estimated from MS and AS. The x-axis shows the student abilities on the common metric provided by the AS. The y-axis shows, for each language, estimated abilities in relation to the cut-offs set for that language.

**Figure 53 Main Study and Alignment study abilities**



The number of data points potentially informative about each cut-off and each language is quite small, and there is variation between the two sets of estimates. Visual inspection is the most appropriate means of evaluation.

The vertical dotted lines show subjectively placed CEFR cut-offs on the AS scale. It is possible to place these lines so that the transitions between level indicated by the MS cut-offs on the vertical axis correspond quite well to the transitions indicated on the AS scale. This is confirmatory evidence that the final standards set per language are comparable. Certainly there is no evidence of gross variation.

Figure 53 shows the final standards. These reflect the one change made on the evidence of the alignment study: English B2 was slightly lowered.

The alignment study is an innovative aspect of the Survey. It addresses the empirical validation of standard setting procedures, the importance of which is discussed in the *manual for relating examinations to the CEFR* (Council of Europe 2009, Chapter 7) under the heading of external validation.

## **11.6 The status of the standards**

This chapter describes a complex process which has involved considering and weighing different forms of evidence. The process ends with a set of standards which are defensible and should be practically useful. However, there must remain some uncertainty about their status. The uncertainty starts with the CEFR itself, the nature of its levels and the different interpretation placed on them by different users, depending on local standards and custom. The aim of the ESLC standard setting is to set a common standard which may promote convergence of use in future, but the standard setting event and the data collected inevitably reflect this uncertainty.

The defining feature of standard setting for the ESLC is that it involves five languages. All standard setting is essentially arbitrary, that is, based on judgment. In a single language context it may be immaterial how the standard might relate to any other context. But the ESLC standard setting, although done separately by language; could not be completed without comparison across languages. If the goal is to implement a common interpretation of CEFR levels it is necessary to validate the judgments of individual language teams against each other.

In finalizing standards significant weight has been given to the test materials themselves, and the argument that the test construction process was such as to have produced tests in each language broadly comparable in terms of the construct tested and the absolute level of the tasks. Thus the very final step in reconciling standards across languages has been, for each skill, to compare standards across languages on a whole-test score metric, and to address discrepancies by averaging or subjectively imposing a common standard at each level.

There are two arguments for this approach:

- It should apportion uncertainty about the “true” standard more equally across languages.
- It ensures that the proportional size of each level is similar across languages. This satisfies an important requirement of a language indicator. It is highly



desirable that the proportion of students achieving each CEFR level in any iteration of the ESLC should depend solely on levels of achievement, and not on variations in the proportional placement of cut-offs. Imposing consistency at this stage should simplify interpretation in future.

The manual for relating examinations to the CEFR (Council of Europe 2008), which was referred to in designing the standard setting approach, distinguishes standard setting (chapter 6) from empirical validation (chapter 7). The latter should provide evidence, possibly over a longer timeframe, for the validity of the standard set. Within the timeframe of the ESLC there is limited scope for external validation; however, two aspects of the ESLC can be seen to fall under this heading:

- the Alignment Study for Writing offers independent empirical verification of the comparability of standards across languages. As described in section 11.5 above, it provides confirmatory evidence that these standards are indeed comparable
- the can-do statements included in the Student Questionnaire.

A discussion of the can-do statements is given in chapter 2 of the Final Report. Analysis of these interestingly demonstrates how far understanding of CEFR levels varies across countries, and so contributes to placing the ESLC findings in context. For this reason however they were deemed problematic to integrate into an already complex standard setting process, and thus were not directly used in finalising standards.

There is scope for experimental research following up the ESLC to provide further empirical validation of the standards set and possibly inform the design of a future round.

## 11.7 References

Council of Europe (2008) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) A Manual*. Strasbourg, France: Council of Europe. Retrieved from:

[http://www.coe.int/t/dg4/linguistic/manuel1\\_en.asp](http://www.coe.int/t/dg4/linguistic/manuel1_en.asp)

Communication from the Commission to the Council. (2007) Framework for the European survey on language competences. COM(2007) 184 final. Brussels

Jones, N (2009) A comparative approach to constructing a multilingual proficiency framework: constraining the role of standard setting, in Figueras N and Noijons J (Eds) *Linking to the CEFR levels: Research Perspectives*, Cito, Council of Europe, EALTA.

Retrieved from [http://www.coe.int/t/dg4/linguistic/Proceedings\\_CITO\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/Proceedings_CITO_EN.pdf)

Linacre, M (2011) FACETS. <http://www.winsteps.com/facets.htm>



Van der Schoot (2009) Cito variation on the bookmark method, in Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) A Manual, Strasbourg, France: Council of Europe. Retrieved from: [http://www.coe.int/t/dg4/linguistic/manuel1\\_en.asp](http://www.coe.int/t/dg4/linguistic/manuel1_en.asp)

# ***Chapter 12:***

# ***Analyses***

## 12 Analyses

### 12.1 Introduction

In this chapter we discuss the analyses that underlie the results presented in final report on the ESLC. The focus of this chapter is on the methodological underpinning, that is on the ‘what’ and ‘why’, rather than on the ‘how’. There is one exception. The algorithm used for generating plausible values was specifically developed for the ESLC and a short overview is included as an Appendix to this chapter, as it is one of the outcomes of the ESLC in its own right.

The data collected for the ESLC extends well beyond the direct policy questions the ESLC seeks to answer. These data are available for secondary analyses. The interested reader may consult, for instance, the PISA Data Analysis Manual (OECD, 2009) that gives a good introduction to analyzing survey data (involving plausible values and sampling weights) using SPSS and/or SAS.

Formally, the purpose of a survey is to characterize the conditional distribution of item responses  $\mathbf{x}$  conditionally on key policy indicator  $\mathbf{y}$ ,  $f(\mathbf{x}|\mathbf{y})$ .

Like most large scale international surveys, the ESLC employs an incomplete design where not every student answers all questions related to his/her language proficiency. Having an incomplete design serves two distinct purposes. First, in order to cover the intended construct (in this case the CEFR), a large number of items and tasks is needed. Second, for practical reasons every student can only be administered a limited number of items and tasks. Balancing these two contrasting goals is elegantly achieved with an incomplete design in which every student only answers to a subset, from a sufficiently large collection, of the items and tasks.

The ESLC is administered as a multistage test, consisting of two stages. The first stage (routing test) only serves to assign students to three broad ranges of language proficiency. Such a two stage test has a number of advantages. First, because students are given questions that roughly correspond to their language proficiency they should find the test not too difficult, nor too easy. This reduces both boredom and frustration, and hence improves the quality of the data. Second, because the item (roughly) match the proficiency of the students, the information obtained about his/her ability increases, relative to a test consisting of more widely spread questions. The efficiency of the multistage administration is illustrated with Figure 55, which will be considered in more detail later on.

The cost of balancing the need for a broad range of tasks and items and a limited testing time per student is that responses, and scores, are not directly comparable



across students that were administered different subsets of items and tasks. It is for this reason that some form of score equating is called for. The purpose of score equating is to render scores comparable across different subsets of tasks and items.

Many methods exist for rendering scores on different sets of items comparable. By and large, the use of an Item Response Theory (IRT) model has been the most popular method for the past couple of decades.

An IRT model postulates a relationship between the probability to give one of the possible responses to a question and the latent ability of the student answering the question. If we use  $X_{pi}$  to denote a random variable indicating the response of person  $p$  to item  $i$ , and  $\theta_p$  to denote the latent ability of person  $p$  we obtain the following expression:

$$P(X_{pi} = j | \theta_p) = f_{ij}(\theta_p)$$

If the functions  $f_{ij}$  are known for all items and responses, we can infer the value of  $\theta_p$  from the responses to any subset of the items. It is in this way that we can render responses to different sets of items (that all relate to the same underlying ability) comparable.

The use of an IRT model allows us to consider the following characterisation of the distribution of item responses conditionally on policy indicators:

$$f(\mathbf{x} | \mathbf{y}) = \int_{\mathcal{R}} p(\mathbf{x} | \theta) f(\theta | \mathbf{y}) d\theta$$

where the relation between  $\mathbf{x}$  and  $\mathbf{y}$  runs *via* the unobserved latent trait  $\theta$ . As in most surveys, an IRT model is used for characterizing the distribution of item responses conditionally on ability, and a regression model for characterizing the distribution of ability conditionally on the policy indicators. In the following sections we explain how the IRT and regression models used for the ESLC were chosen.

For estimating ability distributions, a Bayesian approach has become the method of choice. If we use  $f$  to denote the true distribution of ability, and  $g$  to denote a prior ability distribution, we obtain that

$$\square \quad f(\theta) = \int_{\mathcal{R}} \int_{\mathcal{R}} g(\theta | \mathbf{x}) p(\mathbf{x} | \theta_{true}) f(\theta_{true}) d\mathbf{x} d\theta_{true}$$

with  $g(\theta | \mathbf{x})$  denoting the posterior distribution of ability

$$g(\theta | \mathbf{x}) = \frac{p(\mathbf{x} | \theta) g(\theta)}{\int_{\mathcal{R}} p(\mathbf{x} | \theta) g(\theta) d\theta}$$

One can show that for an increasing number of items, the distribution  $\hat{f}$  converges to the true distribution  $f$ . This reflects the well known fact from Bayesian theory that for an increasing sample size (here the number of administered items), the likelihood (here the distribution of the item responses conditionally on ability) will dominate the prior distribution (here  $g$ ).

The draws from the posterior distribution of ability  $g(\theta | \mathbf{x})$  are commonly referred to as plausible values. Even though the prior distribution need not match the true ability distribution, the distribution of the plausible values will more closely resemble the true ability distribution. It is this remarkable property which makes plausible values such a powerful tool for survey research.

## 12.2 Item Response Theory

Selecting an appropriate measurement model for equating scores from a complex survey such as the ESLC requires striking a detailed balance between model complexity and model fit.

### 12.2.1 Reading and Listening

The test forms for Reading and Listening each consist of a number of tasks. For instance, with a single text, a student has to answer a number of questions. With such an approach it is not clear that the responses for the same student to different questions relating to the same text would be independent. That is, we run the risk of violating the conditional independence assumption that is crucial to (almost) all IRT models.

For that reason, the data from the ESLC are analysed at the task level. That is, the basic observation is the number of correctly answered items for a task, rather than the item responses themselves. Such an approach deals elegantly with possible violations of the assumption of conditional independence.

Based on the analyses of the Field Trial data, a Partial Credit Model (PCM) was selected for modelling the task scores. The PCM models the probability that a person with ability  $\theta$  answers  $j$  out of  $J$  items correctly as follows:

$$P(X_i = j | \theta) = \frac{\exp(\sum_{l=1}^j [\theta - \beta_{il}])}{1 + \sum_{k=1}^J \exp(\sum_{l=1}^k [\theta - \beta_{il}])}$$

From the expression above it is immediately clear that the  $\theta$  and  $\beta_{ij}$  parameters are not identifiable. Specifically, the same constant can be added to both the  $\theta$  and the

$\beta_{ij}$  parameters without changing any of the probabilities. For that reason, one of the  $\beta_{ij}$  parameters is fixed to an arbitrary constant, to make the parameters identifiable.

**Table 56 an illustrative summary of the data collected in the Field Trial for one skill (Reading), and one language (English).**

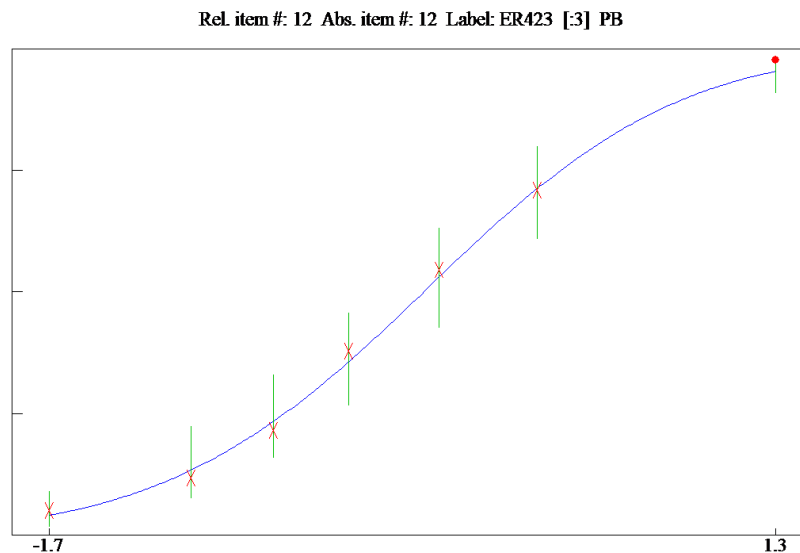
item	label	a	N	P	0	1	2	3	4	5	6
1	ER111	1	513	0.901	8	15	24	79	387		
2	ER112	1	509	0.729	10	36	107	190	166		
3	ER211	1	511	0.666	18	56	132	179	126		
4	ER212	1	524	0.719	11	67	104	135	207		
5	ER221	1	836	0.632	38	84	126	210	215	163	
6	ER223	1	775	0.507	92	157	135	147	124	120	
7	ER311	1	543	0.916	4	7	4	19	129	380	
8	ER312	1	518	0.843	4	11	23	52	170	258	
9	ER321	1	756	0.483	102	127	131	104	83	97	112
10	ER323	1	848	0.543	63	115	127	136	159	136	112
11	ER422	1	770	0.685	14	62	110	174	217	193	
12	ER423	1	777	0.494	67	152	189	165	126	78	
13	ER522	1	747	0.774	16	59	127	180	365		
14	ER523	1	793	0.726	20	72	153	268	280		
15	ER532	1	1102	0.631	41	150	178	215	263	255	
16	ER533	1	1088	0.717	43	107	133	140	217	448	
17	ER631	1	1069	0.692	30	95	113	137	159	141	394
18	ER633	1	1127	0.622	30	80	180	216	212	186	223
19	ER642A	1	552	0.604	26	46	68	103	118	110	81
20	ER642B	1	552	0.582	22	51	99	97	104	104	75
21	ER643A	1	366	0.471	35	64	76	61	52	40	38
22	ER643B	1	366	0.389	42	82	88	71	39	36	8
23	ER731	1	1046	0.636	73	189	223	218	343		
24	ER733	1	1096	0.568	116	228	238	269	245		
25	ER741	1	551	0.650	17	40	56	85	124	127	102
26	ER742	1	372	0.551	13	50	70	68	69	53	49
27	ER841	1	546	0.500	52	100	88	94	77	38	97
28	ER843	1	371	0.609	24	46	35	40	72	93	61

Table 56 gives for every task in the Field trial, the number of observations (N) and the distribution across the different scores. Observe that even though the Field trial had a substantial sample size, the number of observations per task is not huge, taken into

account that for every score (other than zero) on any task a threshold parameter  $\beta_{ij}$  has to be estimated.

From a purely statistical point of view, one has to conclude that the model does not fit the data. The R1c statistic equals 3671.479 for English Reading, which with 2315 degrees-of-freedom reflects a statistically significant violation of the model (Verhelst, Glas, & Verstralen, 1993). However, evaluating the size of the model misfit from a visual inspection of the discrepancy between observed and expected category response functions did not reveal substantial misfit. An illustration of such a visual inspection is given in Figure 54. See the OPLM manual (Verhelst, Glas, & Verstralen, 1993) for more information.

**Figure 54 Illustration of size of misfit between observed and expected empirical category response functions.**



The curves relate to the probability to score in category 3 or higher, for task ER423. The solid line gives expected and the red dots observed proportions.

An important advantage of the PCM compared to alternative IRT models for polytomous responses is that the sum score (in this case the total number of correctly answered questions) is a sufficient statistic for ability. As a consequence, the task parameters  $\beta$  can be estimated using a conditional maximum likelihood (CML) approach. The main advantage of using CML, is that it is possible to estimate the task parameters and evaluate the fit of the PCM, without the need for simultaneously estimating a structural model for ability. As a consequence, incorrect assumptions about the population model can not lead to bias in the estimated threshold parameters.

An important threat to the validity of the substantive outcomes of any survey is item bias. For the ESLC, the impact of item bias is reduced by analysing data at the task level. Furthermore, task parameters were estimated for every language and skill in every educational system separately. Even though the sample size per educational system, implies that detailed comparisons of task parameters across countries have only limited statistical power, gross forms of task bias would have been detected.

**Table 57 The sufficient statistics for English reading for the threshold parameters in the PCM model, across all the countries.**

TASK	N	1	2	3	4	5	6	7	8	9	10	11	12
ER112	3248	300	746	1129	1073								
ER211	3187	365	833	1119	870								
ER223	4927	1019	1020	931	992	965							
ER312	3295	64	121	369	1055	1686							
ER321	4860	1070	949	776	645	626	794						
ER423	5063	1047	1224	1092	940	760							
ER523	5173	453	919	1915	1886								
ER532	5837	684	1020	1174	1436	1523							
ER631	5901	480	700	914	998	929	1880						
ER642	3185	42	94	173	222	267	314	382	341	366	374	323	287
ER731	5480	1104	1411	1393	1572								
ER741	3123	178	295	465	720	866	599						
ER841	2988	550	599	597	479	271	492						

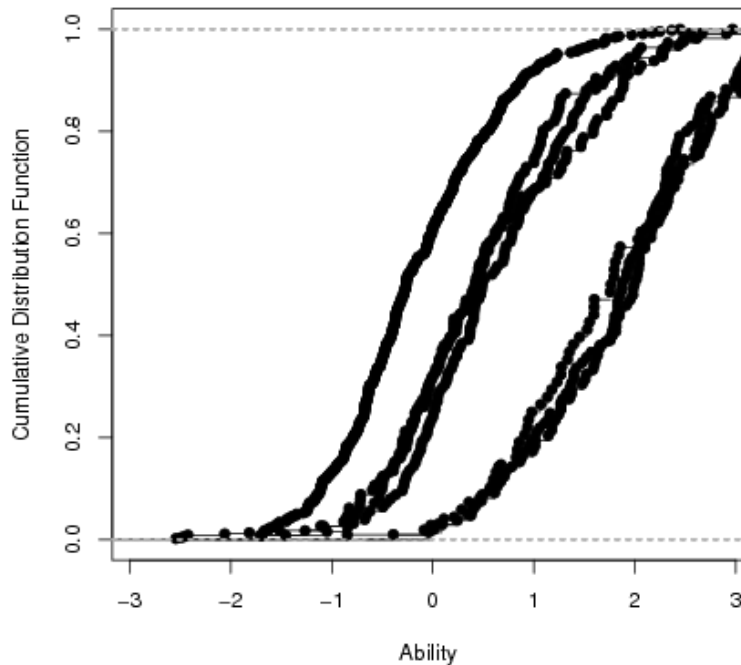
For comparison, Table 58 below gives the same information for one particular educational system. Taking into account that from every entry (in the columns labelled 1, 2, etc.) a parameter needs to be estimated, it should be clear that many of these can not be estimated with any degree of accuracy. As a consequence, comparisons across different countries have only limited value.

**Table 58 The sufficient statistics for English reading for the threshold parameters in the PCM model, across one educational system**

TASK	N	1	2	3	4	5	6	7	8	9	10	11	12
ER112	363	35	85	131	112								
ER211	386	33	91	154	108								
ER223	398	127	95	68	68	40							
ER312	359	5	13	42	134	165							
ER321	413	115	96	85	56	20	41						
ER423	443	140	137	91	49	26							
ER523	439	53	109	155	122								
ER532	250	55	64	66	42	23							
ER631	278	47	43	61	61	27	39						
ER642	32	2	4	2	3	4	4	3	1	5	3	1	
ER731	228	99	66	39	24								
ER741	25	3	4	4	4	7	3						
ER841	32	11	6	5	6	1	3						

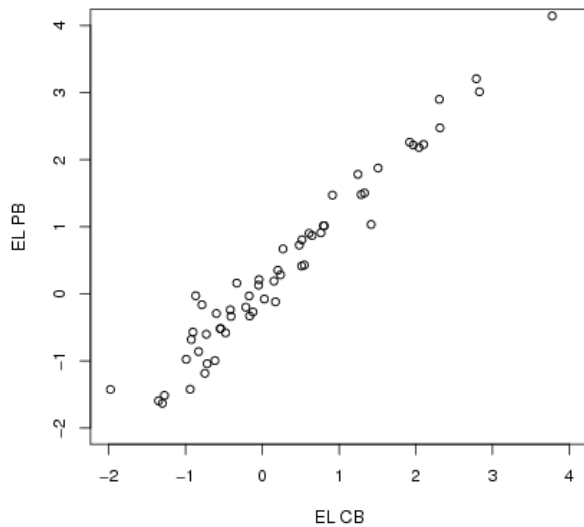
The ESLC consists of test forms at three levels. Because there are multiple test forms, consisting of different combinations of tasks, and distributed randomly to students assigned to the particular level in every educational system, we can meaningfully compare the equated score distributions across test forms in every educational system. Using the plausible values, to be described in more detail later on, one looks at the cumulative proficiency distribution across the different test forms. In the absence of task bias, the distributions for different test forms relating to the same level should overlap. These checks revealed no serious task bias, as is illustrated in Figure 55 which shows one particular such comparison. The distribution corresponding to the three levels are easily distinguished (illustrating the efficiency of the multi-stage approach) across the levels, but not within the levels. Based on the data sets for secondary analyses, it is possible to reproduce figures like the one in Table 58 for all language-skill-educational system combinations. As these amount to 15 educational systems, 4 skills (two aspects of writing), and 2 languages per educational system, not all figures are included in this report.

**Figure 55 Distribution of ability for different test forms, for one language and skill in one educational system.**

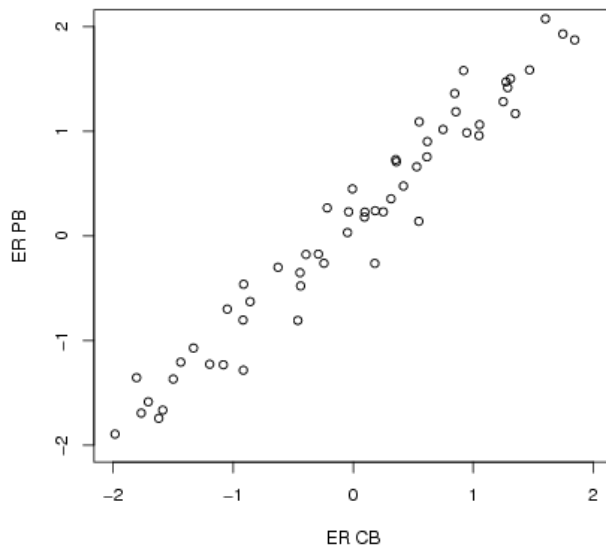


For reading and listening, the survey is administered in two modes of administration, paper-based and computer-based. A comparison between the two modes is relevant, but complicated by the fact that the mode of administration is confounded with educational system. A further complication arises from the unequal sample sizes (overall) for the different languages. For English, a comparison between students from one or the other administration mode is possible. Figures 3 and 4 show the relation between the parameters for English reading and listening estimated from the paper-based and those estimated from the computer-based administration.

**Figure 56 Estimated threshold parameters for English listening in the paper-based and the computer-based test administration.**



**Figure 57 Estimated threshold parameters for English Reading in the paper-based and the computer-based test administration.**



Although the parameters are of course not exactly the same, it is clear from Figure 56 and Figure 57 that the mode of administration did not have a substantial impact on the values of the parameters.

Because task bias, with respect to educational system or mode of administration, did not seem to pose a serious threat to the substantive outcomes of the survey, an international calibration was carried out, where all task parameters were estimated per language and skill across countries, to be used for the remainder of the analyses.



Finally, observe that Figure 55 clearly demonstrates the efficiency of the multi-stage administration of the survey based on a short routing test. The distribution of ability for students assigned to each of the three levels are clearly separated and ordered, according to expectation. Observe, furthermore, that even though the distributions are ordered, they also overlap one another. This implies that even a student assigned to the lowest level can demonstrate ability more typical for students assigned at one of the higher levels. This to, is entirely according to expectation.

### 12.2.2 Writing

Compared to Reading and Listening, the analysis for Writing is more complicated, due to the impact of human judges. The analysis is further complicated by the fact that writing samples from educational system A are only judged by judges from educational system A, and (for a sub-sample of students) by a team of central judges. A final complicating factor is that two aspects of writing were being judged, using the same writing samples. As is the case for multiple judges judging the quality of the same sample, we can not simply assume that judgements with respect to different aspects based on same sample are conditionally independent. For that reason it was decided to analyse the two aspects separately.

For writing the following variant of the PCM is used (for each aspect):

$$P(X_{pir} = j | \theta_{pr}) = \frac{\exp(\sum_{l=1}^j [\theta_{pr} - \beta_{il}])}{1 + \sum_{k=1}^J \exp(\sum_{l=1}^k [\theta_{pr} - \beta_{il}])}$$

That is, for equating purposes, different judgements for the same writing sample are treated as if they are from different persons.

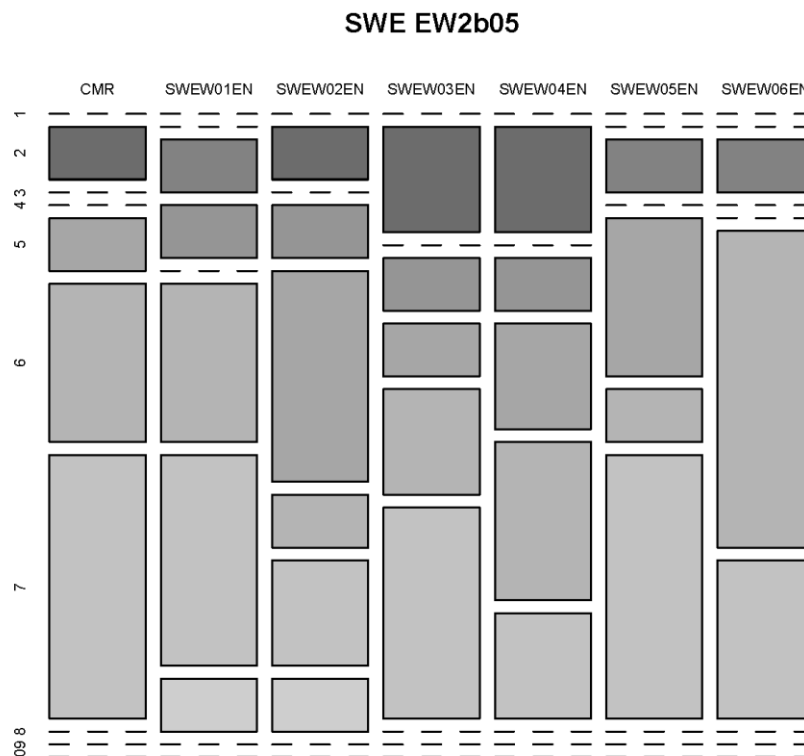
Of course this model does not account correctly for the potential violation of conditional independence resulting from the fact that multiple judges judge the same writing sample.

For the purpose of estimating individual ability, this PCM is not ideally suited. It implies, for instance, that with increasing the number of judges, we can increase the reliability of the estimated ability without bound, which does not make sense. However, for the purpose of estimating the  $\beta$  parameters, and for the purpose of population inferences they should not lead to substantial bias.

Analyses of the Field Trial data revealed that judgements were effected by substantial educational system specific context effects. For that reason it was decided to adjust the equated scores (being plausible values) for every educational system in the direction of the central judges. This adjustment only took place when the latent regression models were estimated, and the plausible values generated, so more information on the adjustment will be given later on.

To get insight into the level of consistency of marking within countries and the level of consistency between markers and the central markers connecting the countries to each other, figures such as the one in Figure 58 are insightful.

**Figure 58 Mosaic plot giving the score distributions per test form, language, educational system for each of the markers.**



The first column, labelled CMR, refers to the central marker. The title refers to the educational system, SWE=Sweden, language and skill, EW=English Writing, level and booklet number.

Because there are substantial differences between the languages and countries, these are contained in an Appendix to this chapter for all test forms, educational systems and languages.

### **12.3 Linking language proficiency to key policy indicators: Latent regression**

The ESLC produces two main outcomes. First is the distribution of students across the CEFR levels in every educational system for every skill and each of the tested languages, in the educational system. Second is the distribution of ability for each

language and skill in relation to a number of key policy indicators. Here we consider how the relation of language proficiency and policy indicators is determined.

First, there are policy indicators at the level of students, teachers, schools, and educational systems, coming from the Student, Teacher, Principal, and National Research Coordinator Questionnaires. In theory one big multi-level latent regression analyses allows for modelling all effects at once. However, even though students are nested under teachers that are nested under schools, etc., the link between students and teachers is not observed. That is, we do not know which teachers teach which students. Moreover, at the level of teachers and principals there were substantial amounts of missing responses that can not simply be assumed to be missing at random. Both of these problems complicate a joint analysis of all indicators at once. For that reason it was decided to aggregate information from the Teacher Questionnaire to the school level.

The latent regression part of the analyses consists of three separate parts that we describe below. For generating plausible values, and for estimating conditional effects all students that were administered a test form for the language and skill in question were included. Students with more than 3 missing values on the Questionnaire indices (that were included in the regression model) were excluded from the estimation of conditional effects.

### **12.3.1 Generating plausible values**

As observed before, plausible values are a powerful tool for studying the relationship between ability and contextual information. Formally, plausible values are draws from the posterior distribution of ability for every student, where the prior distribution is adjusted (in a hierarchical Bayesian fashion) for every individual based on information from the Questionnaires, using a latent regression model.

The model used as a prior distribution for ability, contains the Student Questionnaire indices together with a school mean and a school variance. By including a school mean and a school variance, any effect of Teacher, School, and higher level indices is correctly accounted for. In particular, by including school means, also the intraclass correlation is fully taken into account.

For generating the plausible values, the newly developed algorithm described in the Interim Report is used, which is included as an Appendix to this chapter.

For the analyses of Writing, plausible values were adjusted towards the central judgements. A representative sample of students was judged by multiple judges, and by a central judge. The central judges serve as the only connection between judgements in different countries. In every educational system, and for every judge, a linear transformation was applied to (all) the plausible values, such that the transformed plausible values for every judge match the mean and variance of the plausible values for the central judge on the sample of students that was multiply judged.

### 12.3.2 Conditional effects

One not only wants to know what the relationship between a particular index and language proficiency looks like. But also how the relationship looks like taking into account all other indices. That is, we want to answer the question, all other things being equal, what is the effect of this particular index.

For this purpose a latent regression model is estimated, similar to the model estimated for generating the plausible values. However, where for the generation of plausible values the parameters of the regression model were of no interest, they are now the essential outcome of the analyses. For that reason, more care is needed in setting up the regression model.

Because the indices are based around policy issues and (partly) derive from the same (or closely related) sources of information, we need to deal with the potential co-linearity problems. For that reason a substantively motivated selection of the indices was made to capture the full width of the Questionnaire framework, while at the same time resolving the co-linearity.

Observe that the school means and school variances can be considered as plausible values, pertaining to schools rather than students, in their own right. These can be used for evaluating the effect of school level policy indices.

### 12.3.3 Marginal effects

The main outcomes of the survey are the relationships between each of the policy indices and language proficiency in isolation from all other effects. These we will refer to as marginal effects.

A trivial example relates to the distribution of language proficiency for boys and girls, respectively, in a particular educational system. Plausible values, in combination with sampling weights, allow us to correctly estimate these distributions (and their respective standard errors).

For indices that are nominal, or ordinal with a limited number of categories, the distribution for each of the values of the index can be directly estimated. For real valued indices, a simple linear regression model with the index in question as the single predictor is estimated.

## 12.4 Appendix: Generating plausible values

Conceptually, the most unrestricted model is one in which every person, indicated with a subscript  $p$ , is a draw from a person specific distribution, denoted as  $f_p$ , hence:

$$f_p(\theta | \mathbf{x}_p) \propto p_{\delta, \mathbf{a}}(\mathbf{x} | \theta) f_p(\theta)$$

where  $\theta$  denotes the unobserved ability of person  $p$ ,  $\mathbf{x}_p$  contains the response of the  $p$ -th person to the language test, and  $\delta$  and  $\mathbf{a}$  are the item difficulty and discrimination parameters, respectively. Most implementations will place restrictions on the person specific distributions  $f_p$ . These restrictions serve to reduce the complexity of the statistical problem. Without such restriction no statistical analysis is possible. Important and popular restrictions are that:

- (a)  $f_p$  is assumed to belong to a parametric family of distributions (e.g., normal distribution with unknown mean and variance), and/or that
- (b) the distribution of ability is assumed to be the same for different persons (e.g., the same distribution for all boys, and another one common to all girls).

Another popular restriction that is often added to both of these is that the parameters characterizing the distribution of ability within different groups are related to each other. Suppose we assume that ability is normally distributed in different age groups with a common variance  $\sigma^2$  and a mean that depends on age as follows:

$$\mu(\text{age}) = \alpha + \beta \text{age}$$

where  $\alpha$  and  $\beta$  are parameters that are to be estimated together with  $\sigma^2$ . Here we make the assumptions that:

- (a) ability is normally distributed,
- (b) persons of the same age come from the same distribution, and
- (c) the variance is common to all age groups, and the mean depends in a linear fashion on age.

One of the main outcomes of a survey is to describe how characteristics of the ability distribution (e.g., the mean) vary over different groups (e.g., age groups). It is in answering such questions that plausible values play an important role.

The statistical model used for analysing data from a complex educational survey is typically complex and tools for statistical inference for such complex models are not widely available, certainly not to educational researchers wishing to do secondary analyses. The main problem is that the relation between item responses and student, school, and educational system characteristics runs via the unobserved ability of a student. If ability were known, standard methods implemented in statistical software packages such as GNU-R, SPSS, and SAS could be used for data analysis. The approach taken in many large scale educational surveys is to supplement the observed variables with imputed values for the missing ability parameters, in such a way that standard software can be used for statistical analyses from these completed data files. In order to correctly account for the uncertainty associated with each of the imputed ability values, a number of draws from their posterior distribution is generated.

That is, rather than a single completed data file, a number of completed data files are produced, such that from a same analysis with each of the data files the correct point estimates and, more importantly, their correct standard errors can be computed. This approach goes back to Mislevy (1991), and has been successfully implemented in many large scale educational surveys since that time (e.g. PISA).

In order for the analyses reported on in the main report of the ESLC to be reproducible, these will be based on the same plausible values that are released for secondary analyses.

#### 12.4.1 Multi-level regression with latent dependent variables

Typically in a large scale international educational survey, information is gathered about the characteristics of students, schools, and educational systems. In relating these characteristics to, in this case, language proficiency, we need to correctly account for the fact that students are nested within schools, and schools are nested within educational systems. Hence, for all students in the same school, the same school characteristics apply. Typically, characteristics of schools and educational systems are not estimated separately for every school; but rather the distribution of these characteristics is estimated across schools. That is, they constitute a random effect. Correctly accounting for random effects that are, for instance, constant across students within the same school can be achieved with a multi-level statistical model.

A simple example of a multi-level regression model is the following.

$$\theta_{ps} = m_s + \varepsilon_{ps}$$

where  $\theta_{ps}$  is the proficiency of student  $p$  in school  $s$ ,  $m_s$  is the average proficiency for school  $s$  and  $\varepsilon_{ps}$  is assumed to be a standard normal distributed deviance for student  $p$  in school  $s$ . The school means  $\mathbf{m}$  are not estimated for every school, but are assumed to be random variables themselves that are independent and normally distributed with a mean  $\mu$  and variance  $\sigma^2$ :

$$m_s = \mu + \sigma v_s$$

where  $v_s$  is the standard normally distributed deviance for school  $s$ .

The multi-level regression model can be written more compactly in the following way:

$$\theta_{ps} = \mu + \sigma v_s + \varepsilon_{ps}$$

in which we recognize an ordinary (single-level) regression model with two residuals, one relating to the school a student attends and one particular for the student himself.

In general, every multi-level regression model can be written in the form of an ordinary regression model with more than one residual for every observation:

$$\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\lambda} + \mathbf{B}\boldsymbol{\Sigma}\boldsymbol{\Sigma} + \boldsymbol{\eta}\mathbf{R}$$

where both  $\mathbf{E}$  and  $\mathbf{R}$  are independent and identically normal variates with mean equal to zero and variance equal to one. The design matrix  $\mathbf{A}$  relates to fixed effect parameters  $\boldsymbol{\lambda}$ ; whereas the design matrix  $\mathbf{B}$  relates to the random effects  $\mathbf{E}$ , which are characterized by the diagonal matrix with standard deviations  $\boldsymbol{\Sigma}$ . Hence, if the operator  $\text{diag}$  takes a vector as argument and produces a diagonal matrix with this vector as diagonal, we may write  $\boldsymbol{\Sigma}$  as  $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})$ . In this way we can easily impose equality constraints on the standard deviations of the random effects, needed in most multi-level models:  $\boldsymbol{\Sigma} = \text{diag}(\mathbf{C}\boldsymbol{\tau})$ . Moreover, we may rewrite our model equivalently in the following way:

$$\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\lambda} + \mathbf{B}\text{diag}(\mathbf{E})\boldsymbol{\sigma} + \boldsymbol{\eta}\mathbf{R}$$

which proves convenient later on. From this formulation of the model we readily obtain that:

$$\mathbf{E}(\boldsymbol{\theta} | \mathbf{E}) = \mathbf{A}\boldsymbol{\lambda} + \mathbf{B}\boldsymbol{\Sigma}\boldsymbol{\Sigma} \quad ,$$

$$\perp\!\!\!\perp_p \boldsymbol{\theta}_p | \mathbf{E} \quad ,$$

and

$$\mathbf{V}(\boldsymbol{\theta} | \mathbf{E}) = \boldsymbol{\eta}$$

In principle it is possible to obtain point estimates of all the parameters in the multi-level regression model. However, in contrast to the item parameters  $\mathbf{a}$  and  $\boldsymbol{\delta}$  for which point estimates are used, the parameters in the multi-level regression model are estimated differently. Specifically, a Bayesian approach is used for these parameters. In this way we can account for the uncertainty with which the parameters of the multi-level regression model are estimated in the generation of plausible values. The reason for this different approach is the following. Before the language tests and questionnaires are administered, we know how often an item will be administered. For the parameters of the multi-level regression model, however, this is not the case. For instance, if we want to estimate the interaction between gender and month of birth in reading proficiency we need both boys and girls for every month of the year. Beforehand, however, it is not known how large these samples will be. In a more complicated multi-level regression model there will be, possibly many, sub-populations from which only a small number of students participated in the survey. A direct consequence of this is that some of the parameters of the multi-level regression model are only weakly determined from the observations. Discarding the uncertainty with which parameters are estimated in such a situation may lead to bias in the substantive outcomes of the survey.

In the following a Markov chain Monte Carlo method will be described for drawing a sample from the joint posterior distribution of the latent dependent variable  $\boldsymbol{\theta}$ , and the parameters  $\boldsymbol{\lambda}$ ,  $\boldsymbol{\Sigma}$ , and  $\boldsymbol{\eta}$ .

### 12.4.2 A Gibbs sampler

The posterior distribution of the latent dependent variable  $\theta$ , the parameters  $\lambda$ ,  $\Sigma$ , and  $\eta$ , and the latent residuals  $e$  can be written as follows:

$$f(\theta, \lambda, \Sigma, e, \eta | \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \delta)$$

$$\propto \left[ \prod_p p_{\mathbf{a}, \delta}(\mathbf{x}_p | \theta_p) \right]$$

$$\frac{1}{\sqrt{2\pi\eta}} \exp \left( - \frac{(\theta_p - \sum_j A_{pj} \lambda_j - \sum_k D_{pk} \sigma_k e_k)^2}{2\eta^2} \right)$$

$$[f_0(\eta)] \left[ \prod_j f_1(\lambda_j) \right] \left[ \prod_k f_2(\sigma_k) f_3(e_k) \right] \quad (1)$$

If a sample from this posterior distribution can be obtained, we can derive from it a sample from the posterior distribution of the latent dependent variable  $\theta$ , the parameters  $\lambda$ ,  $\Sigma$ , and  $\eta$  by discarding the sampled values for the latent residual  $e$ . If from such a sample, we only retain the values for the latent dependent variable  $\theta$ , we have the plausible values needed for further data analysis. If on the other hand the plausible values  $\theta$  are discarded, we have a sample from the posterior distribution of the parameters of the multi-level regression model.

The Gibbs sampler is an abstract divide-and-conquer approach for generating a dependent sample from a complex multivariate distribution. Formally, the Gibbs sampler generates a Markov chain for which the (posterior) distribution from which a sample is desired is the invariant distribution. In general, the Gibbs sampler reduces a complex multivariate sampling problem to a series of univariate (or lower dimensional multivariate) sampling problems. A Markov chain for sampling from a joint distribution  $f(x, y)$  is constructed in the following way:

$$f(x, y) = \int_R \int_R f(x | y) f(y | x^*) f(x^*, y^*) dx^* dy^*$$

It is a routine exercise to show that the transition kernel  $f(x | y) f(y | x^*)$  indeed has the joint distribution  $f(x, y)$  as its invariant distribution if  $f(x | y)$  and  $f(y | x)$  are the conditional distribution associated with the joint distribution  $f(x, y)$ . The conditional distributions are commonly referred to as *full conditional distributions* to stress the fact that we consider the distribution of one (set of) variable(s) conditionally on *all* the other ones. The interested reader is referred to Casella and George (1992), Patz and Junker (1999), Béguin and Glas (2001), Fox and Glas (2001), and Maris and Maris (2002) for more information on Markov chain Monte Carlo methods for statistical inference in the context of



educational measurement, and to Metropolis, Rosenbluth, Rosenbluth, Teller and Teller (1953), Casella and George (1992), Chib and Greenberg (1995), Tierney (1994), Gelman, Carlin, Stern and Rubin (1995) for general information on Markov chain Monte Carlo methods for statistical inference.

From the full posterior distribution in Equation (2) we see that it naturally leads to the following full conditional distributions:

- The posterior distribution of ability, given responses to the language tests and the multi-level regression model; and
- the posterior distribution of the parameters of the multi-level regression model, given ability.

In this way we obtain that the full conditional distribution for the parameters of the multi-level model are of the same form as the full posterior we would have obtained had the latent ability parameters been observed:

$$f(\lambda, \Sigma, \mathbf{e}, \eta | \boldsymbol{\theta}, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \boldsymbol{\delta})$$

$$\propto \left[ \prod_p \frac{1}{\sqrt{2\pi}\eta} \exp \left( -\frac{(\theta_p - \sum_j A_{pj}\lambda_j - \sum_k D_{pk}\sigma_k e_k)^2}{2\eta^2} \right) \right]$$

$$[f_0(\eta)] \left[ \prod_j f_1(\lambda_j) \right] \left[ \prod_k f_2(\sigma_k) f_3(e_k) \right] \quad (2)$$

This is a standard model that is dealt with in many texts and articles on Bayesian statistical inference (e.g., chapter 13, Gelman et. al., 1995). For that reason we only briefly go into the details for sampling from this full conditional distribution. The full conditional distribution for the ability parameters constitutes the complicated part of multi-level regression with a latent dependent variable. For that reason more attention is devoted to explaining how one can sample from this full conditional distribution.

### 12.4.3 Multi-level regression model

For sampling from the full conditional distribution in Equation (2) a Gibbs sampler is used as well. It is readily seen from the full conditional distribution in Equation (2) that the full conditional distributions are of the same form for each of the  $\lambda_j$ 's,  $\sigma_k$ 's, and  $e_k$ 's. Hence, two classes of full conditional distributions need to be specified:

$$f(\eta | \boldsymbol{\theta}, \lambda, \Sigma, \mathbf{e}, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \boldsymbol{\delta})$$

$$f(\lambda | \boldsymbol{\theta}, \Sigma, \mathbf{e}, \eta, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \boldsymbol{\delta}), \text{ with the same form of full conditional obtained for } \Sigma$$

and  $\mathbf{E}$

We will find in the following that these two full conditional distributions pose few difficulties and are of the exact same form they would have had the dependent variable  $\theta$  actually been observed.

#### 12.4.4 Sampling the variance of person specific residuals

The full conditional distribution for the variance of person specific residuals  $\eta^2$  is the following:

$$f(\eta^2 | \theta, \lambda, \Sigma, \mathbf{e}, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \delta) \propto \left[ \prod_p \frac{1}{\sqrt{2\pi\eta}} \exp \left( - \frac{(\theta_p - \sum_j A_{pj} \lambda_j - \sum_k D_{pk} \sigma_k e_k)^2}{2\eta^2} \right) \right] [f_0(\eta^2)] \quad (3)$$

If the prior distribution of  $\eta^2$  is assumed to be an inverse chi-square distribution with  $\nu$  degrees of freedom, the full conditional distribution is readily seen to be an inverse chi square distribution as well.

#### Sampling $\lambda$ , $\Sigma$ , and $\mathbf{E}$

The full conditional distribution of, say,  $\lambda$  can be written as follows:

$$f(\lambda | \theta, \Sigma, \mathbf{e}, \eta, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \delta) \propto \exp \left( - \sum_p \frac{(\theta_p - \sum_j A_{pj} \lambda_j - \sum_k D_{pk} \sigma_k e_k)^2}{2\eta^2} \right) \left[ \prod_j f_1(\lambda_j) \right] \quad (4)$$

If an improper prior distribution is assumed for each of the  $\lambda_j$ 's (i.e.,  $f_1(\lambda_j) \propto 1$ ), this full conditional may be written as follows:

$$f(\lambda | \theta, \Sigma, \mathbf{e}, \eta, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \delta) \propto \exp \left( - \sum_p \frac{(\theta_p - \sum_j A_{pj} \lambda_j - \sum_k D_{pk} \sigma_k e_k)^2}{2\eta^2} \right)$$

where we see that the argument of the exp -function is quadratic in each of the  $\lambda_j$ 's. That is, this full conditional distribution reduces to a multivariate normal distribution,

with a known mean and variance/covariance matrix. The whole computational complexity for this full conditional distribution is in correctly computing the mean and variance/covariance matrix. It is an easy exercise to see that the mean is equal to  $(A^T A)^{-1} A^T (\boldsymbol{\theta} - \mathbf{B} \boldsymbol{\Sigma} \boldsymbol{\Sigma})$  and (co-)variance matrix equal to  $(A^T A)^{-1} \boldsymbol{\eta}$ .

#### 12.4.5 A rejection sampler for plausible values

Here an algorithm is developed for drawing a sample from

$f(\boldsymbol{\theta} | \boldsymbol{\lambda}, \boldsymbol{\Sigma}, \mathbf{e}, \boldsymbol{\eta}, \mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{a}, \boldsymbol{\delta})$ . The algorithm is built up in a number of steps.

Suppose we want to draw from the distribution of ability conditionally on the responses of a person, which is assumed to be drawn randomly from some well defined population characterized by some completely specified distribution  $f$ :

$$f(\boldsymbol{\theta} | \mathbf{x}, \boldsymbol{\delta}, \mathbf{a}, \boldsymbol{\lambda}) \propto p(\mathbf{x} | \boldsymbol{\theta}; \boldsymbol{\delta}, \mathbf{a}) f(\boldsymbol{\theta})$$

By far the easiest method is the following rejection algorithm (written in pseudo-code):

```
repeat
   $\boldsymbol{\theta} \sim f(\boldsymbol{\theta})$ 
   $\mathbf{y} \sim p(\mathbf{x} | \boldsymbol{\theta}; \boldsymbol{\delta}, \mathbf{a})$ 
until  $\mathbf{y} = \mathbf{x}$ 
return  $\boldsymbol{\theta}$ 
```

It is obvious that for a test of any non-trivial length this procedure is completely useless. The number of repetitions of the repeat-until loop in the algorithm before a value is generated is astronomical. The procedure is similar to throwing a coin which falls head with probability  $p(\mathbf{x})$  (a very small number) until it actually falls head for the first time, a number roughly inverse proportional to  $p(\mathbf{x})$ <sup>32</sup>. Since there are  $2^n$  possible response patterns  $\mathbf{x}$  in a test of  $n$  binary items, the reader may appreciate that  $p(\mathbf{x})$ , for almost any particular  $\mathbf{x}$  will be very close to zero.

However in our context great improvements in efficiency are possible. First, the values of all the discrimination parameters are known beforehand, and moreover are integer valued. Hence, there are much fewer different scores than there are response patterns<sup>33</sup>, and since the score is sufficient<sup>34</sup> it is all that we ever need. For that

---

<sup>32</sup> to be exact its expectation is  $\frac{1 - p(\mathbf{x})}{p(\mathbf{x})}$ .

<sup>33</sup> With 20 binary items there are  $2^{20} = 1048576$  different response patterns to consider, but if the Rasch model is assumed only 21 different scores.

<sup>34</sup> This means that the posterior  $f(\boldsymbol{\theta} | \mathbf{x}, \boldsymbol{\lambda}) = f(\boldsymbol{\theta} | \mathbf{y}, \boldsymbol{\lambda})$  whenever  $\mathbf{x}$  and  $\mathbf{y}$  yield the same score.

reason the following variant of our basic algorithm achieves the same goal in a more efficient manner.

```

repeat
   $\theta \sim f(\theta|\lambda)$ 
   $\mathbf{y} \sim p(\mathbf{x}|\theta; \delta, \mathbf{a})$ 
until  $\mathbf{y}_+ = \mathbf{x}_+$ 
return  $\theta$ 

```

Obviously, the waiting time here is also (approximately) inverse proportional to  $p(\mathbf{x}_+)$ , a number which usually compares favorably to the probabilities in our first naive algorithm. Currently, this algorithm is implemented in GNU-R (R Development Core Team, 2008), see the Appendix for some illustrative code, and can deal with the size and complexity of the ESLC analyses. In the following subsections, we shortly consider how further improvements in efficiency can be realized.

#### 12.4.6 Improvements

Further progress can be made. We can develop a strategy where the intermediate results (those values of  $\theta$  that did not generate the appropriate score) can be recycled. In any iteration we need to generate a predetermined number of independent draws from  $f(\theta|\mathbf{x}_+, \lambda)$  for every value of  $\mathbf{x}_+$ . The joint posterior distribution of  $\theta$  conditionally on the observations  $\mathbf{y}$  and assuming both the item parameters and the population parameters known is the following:

$$f(\theta|\mathbf{x}, \delta, \mathbf{a}, \lambda) = \prod_p \frac{p(\mathbf{x}_p | \theta_p, \delta, \mathbf{a}) f(\theta_p | \lambda)}{\int_R p(\mathbf{x}_p | \theta, \delta, \mathbf{a}) f(\theta | \lambda) d\theta} = \prod_p f(\theta_p | \mathbf{x}_p, \delta, \mathbf{a}, \lambda)$$

and conditionally on the observations  $\mathbf{y}$ , assuming both the item parameters and the population parameters known, we find that the ability parameters of different persons are independent. So, once we know for every possible score, how many students actually got that score (denoted by  $n_{\mathbf{x}_+}$ ), we may proceed in the following manner:

- (a)  $\forall \mathbf{x}_+$ : generate an iid sample of size  $n_{\mathbf{x}_+}$  from  $f(\theta|\mathbf{x}_+, \delta, \mathbf{a}, \lambda)$
- (b)  $\forall \mathbf{x}_+$ : distribute the values generated in the previous step randomly to the  $n_{\mathbf{x}_+}$  students that got that particular score.

In every step of our earlier algorithm, a  $\theta$  value as well as a score are generated. Either we still need realizations from the posterior distribution for that particular score, in which case we assign the generated  $\theta$  value to a particular student with that score, or we no longer need realizations from the posterior distribution for that particular

score, in which case we generate the next  $\theta$  value and score. In pseudo code, this leads to the following algorithm, where we store the  $\theta$  values in a matrix  $\beta$ :

```

repeat
   $\theta \sim f(\theta|\lambda)$ 
   $\mathbf{y} \sim p(\mathbf{x}|\theta; \delta, \mathbf{a})$ 
   $n_{\mathbf{y}_+} = n_{\mathbf{y}_+} + 1$ 
   $\beta_{\mathbf{y}_+, n_{\mathbf{y}_+}} = \theta$ 
until  $\forall j : n_j \geq N_j$ 
return  $\beta$ 

```

The only thing left to do now is to assign the generated values  $\beta$  to the real test takers. This algorithm is both very efficient and computationally easy to implement. Moreover, the efficiency increases as the number of subjects increases.

Observe that the computational complexity of this algorithm is quite independent of the exact choice for the population distribution  $f(\theta|\lambda)$ , as long as we can generate independent and identically distributed samples from it. However, it is dependent on how many subpopulations we wish to distinguish. Because for every subpopulation we need to set up such an algorithm, and hence if the number of distinct subpopulations becomes very large (as is the case with continuous indicators), the efficiency of the algorithm is greatly reduced. This problem is considered in the next subsection.

In order to be able to deal with a large number of distinct subpopulations a further variation on our basic rejection sampling algorithm is considered. It is important to stress that for this situation it is necessary to construct an explicit model of how the mean and variance (in case a normal distribution is assumed for every subpopulation) change with the continuous indicator of the subpopulation.

The basic idea we propose is best illustrated with the first naive rejection sampling algorithm we proposed. This algorithm uses

```

repeat
   $\theta \sim f(\theta|\lambda)$ 
   $\mathbf{y} \sim p(\mathbf{x}|\theta; \delta, \mathbf{a})$ 
until  $\mathbf{y} = \mathbf{x}$ 
return  $\theta$ 

```

to sample from

$$f(\theta | \mathbf{x}, \delta, \mathbf{a}, \lambda) \propto p(\mathbf{x} | \theta; \delta, \mathbf{a}) f(\theta | \lambda)$$

Suppose that instead of drawing from the assumed distribution of ability in a particular population  $f(\theta|\lambda)$  we draw from a different distribution  $g(\theta)$ . It is clear that this algorithm will *not* generate plausible values from the correct posterior distribution. However, we may use rejection sampling again to correct for the specification error, resulting in the following algorithm:

```
repeat
   $\theta \sim g(\theta)$ 
   $\mathbf{y} \sim p(\mathbf{x}|\theta; \delta, \mathbf{a})$ 
until  $\mathbf{y} = \mathbf{x}$  and  $U \leq c \frac{f(\theta|\lambda)}{g(\theta)}$ 
return  $\theta$ 
```

where  $c$  is a known constant such that  $c \frac{f(\theta|\lambda)}{g(\theta)}$ , for every value of  $\theta$ , is less than or equal to one<sup>35</sup> and  $U$  is an independent standard uniform variable. It should be clear that with this adapted algorithm, proposal values are rejected more often, and hence the efficiency will be lower compared to the original algorithm. On the other hand, the algorithm allows for greater flexibility. It should also be clear that the refinements of the original naive rejection sampling algorithm can be carried through in the same way for this new algorithm.

#### 12.4.7 Code for generating plausible values

```
Score=function(t)
{
  sum(a*(rlogis(nI,0,1)<=log(b*t^a)))
}
# generate one set of plausible values
temp=A%%elambda+B%%Sigma%%eE
for (p in 1:nP){
  repeat{
    rt[p]=exp(rnorm(1,temp[p],1))
    ascore=Score(rt[p])
    if (ascore==score[p]) break
  }
}
```

### 12.5 References

Albert, J. H. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of Educational Statistics*, 17(3), 251-269.

Béguin, A. A., & Glas, C. A. W. (2001). MCMC estimation and some model-fit

---

<sup>35</sup>Technically, the distribution  $g$  needs to dominate the target distribution  $f(\theta|\lambda)$ , meaning that their ratio is bounded from above.



- analysis of multidimensional IRT models. *Psychometrika*, 66(4), 541-562.
- Casella, G., & George, E. I. (1992). Explaining the gibbs sampler. *The American Statistician*, 46(3), 167-174.
- Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327-335.
- Fox, J. P., & Glas, C. A. W. (2001). Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika*, 66, 269-286.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman & Hall.
- Maris, G., & Maris, E. (2002). A MCMC-method for models with continuous latent responses. *Psychometrika*, 67, 335-350.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087-1091.
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196.
- Patz, R. J., & Junker, B. W. (1999). A straightforward approach to Markov chain Monte Carlo methods for item response models. *Journal of Educational and Behavioral Statistics*, 24(2), 146-178.
- R Development Core Team. (2008). *R: A language and environment for statistical computing* [Computer software manual]. Vienna, Austria. Available from <http://www.R-project.org> (ISBN 3-900051-07-0)
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22(4), 1701-1762.
- Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1993). *OPLM: One parameter logistic model*. Computer program and manual. Arnhem: Cito.

# ***Chapter 13: Data processing – Data sets***



## 13 Data processing - Data sets

This chapter details the contents of the ESCL data sets.

The ESLC international data sets consist of seven data files: four student-level files, one teacher-level file and two school-level files.

### **13.1 *The Student Questionnaire and performance data file***

Filename: INT\_stu.txt

For each student who participated in the assessment the following information is available:

- Identification variables for the educational system, school, target language and student
- The student responses on the questionnaire
- The students' indices derived from the original questions in the questionnaire
- Plausible values for the students' performance scores in Listening, Reading and Writing (only for the two skills out of three, for which each student was sampled)
- The students' weights and replicates for the computation of the sampling variance estimates

### **13.2 *Language assessment items data files***

#### **13.2.1 Scored responses**

Filename: INT\_cogn\_sco.txt

For each student who participated in the cognitive assessment the following information is available:

- Identification variables for the educational system, school, target language, student and marker
- The students' scored responses to Listening and Reading items
- The students' marked responses for Writing items. In case a student's Writing booklet was marked by a central marker this file contains the marked responses from the central marker. In case a student's booklet was marked by more than one marker, but not a central marker, the file contains the marks by a randomly selected marker

### **13.2.2 Raw responses**

Filename: INT\_cogn\_raw.txt

For each student who participated in a Listening or Reading test, the following information is available:

- Identification variables for the educational system, school, target language and student
- The students raw responses to Listening and Reading items

### **13.2.3 Multiple marking**

Filename: INT\_cogn\_mm.txt

For each Writing booklet which was marked more than once the following information is available:

- Identification variables for the educational system, school, target language, student and marker
- Marked responses

## **13.3 Teacher Questionnaire data file**

Filename: INT\_tea.txt

For each teacher who filled out the questionnaire the following information is available:

- Identification variables for the educational system, school, target language and teacher
- The teachers' responses to the questionnaire
- The teachers' indices derived from the original responses in the questionnaire
- The teachers' weights and replicates for the computation of the sampling variance estimates

## **13.4 School Questionnaire data files**

File names: INT\_sch\_TL1.txt, INT\_sch\_TL2.txt

For each school that participated in the survey the following information is available:

- Identification variables for the educational system, implicit and explicit strata, school, target language and principal
- The principals' responses to the questionnaire, only when applicable to the target language

- School plausible values for Listening, Reading and Writing and standard errors for the school plausible values
- School weights

The school dataset is divided separate in files for the first target language and the second target language. If a school participated for two target languages, the school is present in both files. Since only one principal responded per school the principal responses and indices are replicated in both files as far as they are applicable to both target languages.

### **13.5 Records in the data sets**

#### **Student level**

- All students who attended at least one questionnaire or test booklet session

#### **Teacher level**

- All teachers who responded to the questionnaire

#### **School level**

- All schools for which at least one student attended a questionnaire or test booklet session

### **13.6 Records excluded from the datasets**

The following data is excluded from the datasets

- Students that did not participate in any session, either because they were ineligible, excluded or absent
- Teachers that did not respond to the questionnaire
- Schools for which no students attended a questionnaire or test booklet session.

### **13.7 Weights in the datasets**

All schools for which any student participated in the survey are in the datasets. However, only students and schools that meet the formal criteria for participation have a weight in the datasets.

A participating student is defined as one who has responded to the Student Questionnaire (required of all students), and has done at least one of the two cognitive tests assigned.

A participating school is defined as a school where at least 25% of the sampled students have completed the questionnaire and at least one test booklet. Based on this criterion four schools (two in the first target language sample and two in the second target language sample) did not get a weight because all questionnaires for these schools were lost.

In Spain and the Flemish Community of Belgium, a number of schools took part that were not part of the sample. These schools can be identified through the code 'EXTRA' in the variable 'main\_study\_sample'. These schools and student respondents from these schools do not have weights.

### **13.8 Representing missing data**

Missing responses were coded to distinguish between four types of missing data<sup>36</sup>:

- Not applicable: 77 for closed questions and 7777 in open questions. This code is used for items or options in the questionnaires that were not administered to respondents, mainly due to the localisation (see Chapter 3).
- Not applicable: 78 for closed questions and 7778 in open questions. This code is used for items or options in the Principal Questionnaire that were not applicable for the target language because the principal responded to the other target language version of the questionnaire.
- Invalid: 88 for closed questions and 8888 in open questions. This code is used when a respondent gave an invalid answer, for example selected several answers when only one answer was expected.
- Missing: 99 for closed questions and 9999 for open questions. This code is used when the respondent did not provide an answer to the questions.

### **13.9 Identification of respondents, schools and markers**

The following identifiers were used:

- Educational system identification variable named educational\_system\_id. The educational system codes used in ESLC are the educational system codes of the European Commission
- The school identification variable named school\_id. This consists of the letters 'SC' followed by a randomly assigned 8 digit code
- The respondent identification variable named respondent\_id. Unique randomly assigned number for identification of students, teachers and principals
- The marker identification variable called marker\_id. This is a string consisting of a three letter educational system identification variable (ISO 3166, with BGE, BFL, BFR for the German, Flemish and French Communities of Belgium

---

<sup>36</sup> Note that as far as the indices are concerned, each missing value is a true missing value



and 'CMR' for central marker), concatenated with a five letter string to identify the marker.

- Full details of all identifiers and codes used can be found in the codebook made available with the data sets.

Note: since some schools participated for two target languages, merging the student files with the teacher or school files, through what is known as an 'inner join', should always be done on two variables: `school_id` and `targetLanguage_id`.

# *Appendices*

## 14 Appendices

### 14.1 *Example language test task types*

This appendix illustrates each of the language test task types used in the ESLC. Task types were selected and used consistently to test different aspects of the constructs of Listening, Reading and Writing. See Table 6 to Table 8 in section 2.2.5 for a description of the task types in terms of their formal features, and of the testing focus and text type specified for each task type.

Below one example is provided to illustrate each task type, selected from the range of tested languages. These tasks were developed for but not used in the Main Study.

Chapter 2 of the Final Report provides further illustrations of task types and how they reflect progression across CEFR levels. Four task types are illustrated for each skill – one at each CEFR target level. Examples of each task type are provided for all languages. The relative position of these tasks on a proficiency scale is reported, illustrating the level of performance required on a task to achieve a CEFR level. We do not provide difficulties for the examples in this appendix, but Chapter 2 illustrates the relative difficulty of other examples of the same task type.

The tasks are shown here in a condensed layout in order to conserve space. This is not the form in which they were presented to students in the computer-based or paper-based modes of administration. Tapescripts are provided for the Listening passages.

### 14.1.1 Listening

Four listening task types were used in the final design for the ESLC.

#### 14.1.1.1 Listening task type 1 (target level A1)

##### SL111 Amigos-vacaciones

Vas a escuchar a dos jóvenes que hablan de las vacaciones de su familia y amigos.  
¿Qué van a hacer cada uno de ellos en vacaciones?

Para las siguientes 5 preguntas, selecciona la respuesta (A–G). Selecciona sólo una letra cada vez.

- 
- |   |         |
|---|---------|
| 1 | Roberto |
| 2 | Lucía   |
| 3 | Claudia |
| 4 | Marcos  |
| 5 | Nuria   |

#### Options





## Tapescript

<b>Javier::</b>	<i>¡Qué bien, tenemos vacaciones! ¿Qué vas a hacer, Cristina?</i>
<b>Cristina:</b>	<i>Pues, quedarme en casa.</i>
<b>Javier::</b>	<i>¿No vas a salir de vacaciones con tu novio?</i>
<b>Cristina:</b>	<i>No, porque Roberto va a ir a la playa con sus amigos. Quieren hacer surf y nadar. A mí no me gusta la playa.</i>
<b>Javier:</b>	<i>¿Y con tu hermana Lucía?</i>
<b>Cristina:</b>	<i>Con Lucía ¡imposible! Va a Madrid para hacer un curso de pintura.</i>
<b>Javier:</b>	<i>Oye, ¿y tu amiga Claudia?</i>
<b>Cristina:</b>	<i>Claudia se va al campo a hacer fotografías para una revista.</i>
<b>Javier:</b>	<i>¡Vaya! ¡Lo siento!</i>
<b>Cristina:</b>	<i>¿Y tú? ¿vas a ir a algún lugar?</i>
<b>Javier:</b>	<i>Sí, cinco días a las montañas con mi hermano Marcos.</i>
<b>Cristina:</b>	<i>A Marcos le encanta la aventura ¡Qué bien!</i>
<b>Javier:</b>	<i>Sólo son cinco días, porque tengo que ayudar a mi prima Nuria.</i>
<b>Cristina:</b>	<i>¿Por qué?</i>
<b>Javier:</b>	<i>Tiene que estudiar, ha tenido malas notas en los exámenes.</i>
<b>Cristina:</b>	<i>Entonces, voy a llamarla para salir un día juntas.</i>

## Key

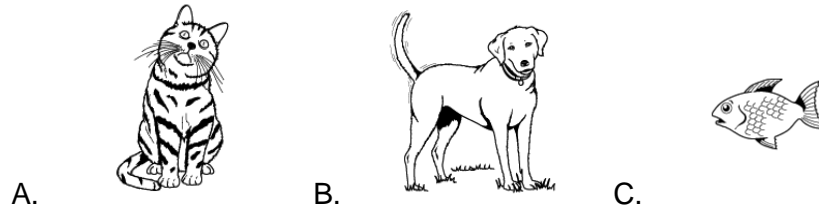
<b>1</b>	<b>F</b>
<b>2</b>	<b>G</b>
<b>3</b>	<b>E</b>
<b>4</b>	<b>C</b>
<b>5</b>	<b>A</b>

### 14.1.1.2 Listening task type 2 (target level A1)

#### GL211 Banane, TV

Du hörst Menschen in verschiedenen Situationen. Wähle bei jeder Aufgabe die richtige Lösung A, B oder C.

#### 1 Welchen Film möchte Sandra heute Abend sehen?



#### Tapescript

- T: Vati, darf ich heute im Fernsehen den Film über Hunde sehen?  
V: Sandra, ich glaube, der ist morgen Abend. Heute gibt es zwei andere Filme über Tiere - über Katzen und über Fische.  
T: Dann möchte ich den Film über Fische sehen - Katzen mag ich nicht.

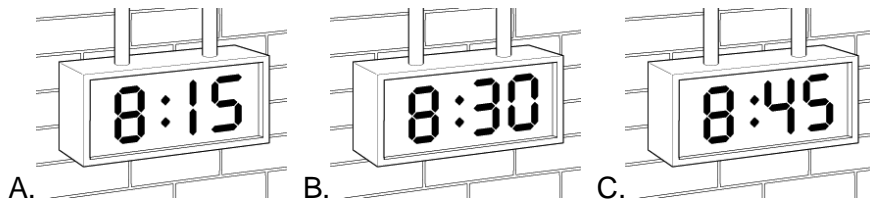
#### 2 Was braucht Peter?



#### Tapescript

Hallo Mama, Peter hier. Ich brauche für meine Hausaufgaben für morgen Sportbilder. Ich habe noch Fotos von unserem Skiurlaub auf einer CD. Kannst du mir bitte noch eine Fußballzeitung kaufen?

#### 3 Wann fährt der Zug nach Hamburg ab?



### Tapescript

Achtung! Der IC 375 um 8.15 Uhr nach Hamburg hat eine Verspätung von dreißig Minuten. Voraussichtliche Abfahrt ist um 8.45 Uhr. Wir bitten die Fahrgäste um Entschuldigung.

### 4 Wo ist Lisas Tasche?



### Tapescript

T: Vati, wo ist meine Tasche? Sie ist nicht auf dem Tisch!  
 V: Tja, Lisa, such sie doch mal in deinem Schlafzimmer, vielleicht ist sie unter dem Bett.  
 T: Ach, da ist sie ja, auf dem Stuhl im Wohnzimmer.

### Key

1	C
2	B
3	C
4	A



### 14.1.1.3 Listening task type 3 (target level A2)

#### IL321 Le braccia di Diego

Ascolterai due amici, Diego e Paola, che parlano di un problema che Diego ha alle braccia.

Per le 6 domande seguenti rispondi Sì o No.

---

- 1 *Diego si è bruciato le braccia durante una partita di calcio.*  
A. Sì      B. No
- 2 *Questa notte Diego ha avuto dolore alle braccia.*  
A. Sì      B. No
- 3 *Diego oggi ha un appuntamento dal dottore.*  
A. Sì      B. No
- 4 *Diego ha avuto problemi per vestirsi stamattina.*  
A. Sì      B. No
- 5 *Paola ha deciso di comprare una crema in farmacia.*  
A. Sì      B. No
- 6 *Paola dice che Diego sta bene e che oggi può andare a lezione.*  
A. Sì      B. No

#### Tapescript

<b>Paola:</b>	<b><i>Ciao Diego, ma che è successo alle tue braccia?</i></b>
<b>Diego:</b>	<b><i>Ciao Paola, beh, ieri c'era molto sole e ho giocato a omain in spiaggia tutta la mattina, e quando sono omain a casa per pranzo, ho visto che avevo tutte le braccia rosse.</i></b>
<b>Paola:</b>	<b><i>Accidenti! Ti fanno molto male?</i></b>
<b>Diego:</b>	<b><i>Beh sì, molto, infatti stanotte non ho potuto dormire per il dolore e ora sono molto stanco. Credo che non andrò a lezione!</i></b>
<b>Paola:</b>	<b><i>Perché non vai dal dottore?</i></b>
<b>Diego:</b>	<b><i>Oggi non è omain e. Mia madre ha chiamato per prendere un appuntamento, ma il dottore era già occupato per tutto il giorno. Devo aspettare fino a omain.</i></b>
<b>Paola:</b>	<b><i>Però se stai male, potresti andare all'ospedale.</i></b>
<b>Diego:</b>	<b><i>Beh! Male male non sto. Posso aspettare fino a omain. È vero, stamattina ho avuto qualche problema a mettermi la maglietta, ma poi l'ho fatto.</i></b>
<b>Paola:</b>	<b><i>Perché non vai in farmacia? Lì ti possono dare una crema specifica.</i></b>
<b>Diego:</b>	<b><i>Mi sono già messo una crema stamattina, ma per il momento non va meglio.</i></b>
<b>Paola:</b>	<b><i>Beh, comunque hai ragione, non devi andare a lezione oggi, ma non ti preoccupare, poi ti spiego tutto io.</i></b>

### Key

<b>1</b>	<b>A</b>
<b>2</b>	<b>A</b>
<b>3</b>	<b>B</b>
<b>4</b>	<b>A</b>
<b>5</b>	<b>B</b>
<b>6</b>	<b>B</b>

#### 14.1.1.4 Listening task type 4 (target level B2)

##### FL441 Cascadeur

Voici une conversation entre une journaliste, Marie, et un cascadeur, Jérôme, qui présente son métier.

Pour les 6 questions suivantes, répons A, B ou C.

---

- 1 *Qui sont les cascadeurs selon la journaliste ? Des personnes qui*
  - A. font des actions extraordinaires.
  - B. jouent avec leur vie.
  - C. aiment vivre au jour le jour.
  
- 2 *Comment Jérôme est-il devenu cascadeur ?*
  - A. En travaillant dans un club sportif.
  - B. En postulant à un poste.
  - C. En aidant un groupe d'artistes.
  
- 3 *Jérôme pense que sa formation scolaire*
  - A. peut encore beaucoup l'aider.
  - B. est très éloignée de son métier.
  - C. est juste un parcours obligatoire.
  
- 4 *Selon Jérôme, pour être un bon cascadeur, le principal c'est d'être*
  - A. courageux.
  - B. déterminé.
  - C. sportif.
  
- 5 *D'après Jérôme, qu'est-ce qu'un régleur ? La personne qui*
  - A. surveille les cascades.
  - B. met en scène les cascades.
  - C. finance les cascades.
  
- 6 *Que conseille Jérôme à tous ceux qui veulent devenir cascadeurs ?*
  - A. De pratiquer un sport intensément et de bien savoir conduire.
  - B. De commencer à travailler et de se spécialiser plus tard.
  - C. De se former auprès des professionnels et d'avoir un diplôme.

## Tapescript

<b>Marie:</b>	<i>Bonjour, aujourd'hui nous recevons un homme qui depuis 10 ans saute de toits en toits, conduit à toute vitesse dans les rues des villes, se jette dans du feu et fait plein d'autres choses insensées. Bonjour Jérôme.</i>
<b>Jérôme:</b>	<i>Bonjour, Marie:, je vous remercie de votre invitation.</i>
<b>Marie:</b>	<i>Comment êtes-vous devenu cascadeur ?</i>
<b>Jérôme:</b>	<i>Par hasard ! Comme beaucoup. En 1998, je cherchais du travail. Un ami cascadeur qui s'entraînait dans la même salle de gym que moi, m'a proposé de rejoindre son équipe pour travailler à Disney sur le spectacle Pocahontas. Le contrat s'étendait sur 8 mois, ça ne m'engageait à rien, j'ai donc décidé de partir à l'aventure... sachant que ma formation scolaire était plutôt tournée vers l'informatique. Ca n'avait rien à voir ...</i>
<b>Marie:</b>	<i>Ah, non, effectivement. Quelles sont les qualités pour devenir cascadeur ?</i>
<b>Jérôme:</b>	<i>Avant tout, il faut être en très bonne condition physique, donc la pratique de la gymnastique, des art martiaux, natation... est un minimum, après toutes les autres activités peuvent être des avantages, comme l'équitation, l'escrime, le plongeon de haut vol...</i>
<b>Marie:</b>	<i>Il existe plusieurs spécialisations dans le métier de cascadeur, je crois. Vous pouvez nous en parler ?</i>
<b>Jérôme:</b>	<i>Alors, il y a les cascadeurs auto-motos, dits mécaniques, ceux qui conduisent des voitures, des motos, ... d'autres sont spécialisés avec des chevaux, ce sont les cascadeurs équestres. Il y a encore ceux spécialisés dans le cinéma. D'ailleurs, le cinéma est notre principal employeur.</i>
<b>Marie:</b>	<i>En effet, très peu de cascadeurs réussissent à vivre de leur métier. On en compte une cinquantaine en France.</i>
<b>Jérôme:</b>	<i>C'est vrai. Pour trouver du travail, il vaut mieux appartenir à une équipe dirigée par un régleur-chorégraphe, la personne qui est contactée par le cinéma ou la télévision et qui va coordonner les cascades.</i>
<b>Marie:</b>	<i>Votre métier est dangereux et à haut risque. Tout bon sportif ne peut pas s'improviser cascadeur ?</i>
<b>Jérôme:</b>	<i>Oui, bien sûr, car ce métier demande des connaissances et des performances dans différents sports, surtout la gymnastique et les arts martiaux et nécessite une extrême rigueur.</i>
<b>Marie:</b>	<i>Quels conseils donneriez-vous à un jeune qui souhaite devenir cascadeur ?</i>
<b>Jérôme:</b>	<i>Je lui conseille d'essayer de s'entraîner avec des cascadeurs professionnels, ça lui permettra de se former et aussi de rester au courant des tournages en préparation. Et surtout d'avoir un diplôme pour donner des cours de sport, par exemple ... Bref, il faut absolument conserver une activité à côté de la cascade.</i>
<b>Key</b>	
<b>1</b>	<b>A</b>

<b>2</b>	<b><i>C</i></b>
<b>3</b>	<b><i>B</i></b>
<b>4</b>	<b><i>C</i></b>
<b>5</b>	<b><i>B</i></b>
<b>6</b>	<b><i>C</i></b>



### 14.1.1.5 Listening task type 5 (target level A1)

#### EL532 Music Festival

You will hear a teacher called Mrs Smith telling a group of students about a Music Festival their college is planning.  
For the next 6 questions, answer A, B or C.

---

- 1      *What will be included for the first time at this year's festival?*
  - A      Classial music
  - B      Jazz music
  - C      Rock music
  
- 2      *When will the festival take place?*
  - A      in May
  - B      in June
  - C      in July
  
- 3      *Mrs Smith is very pleased that this year's prizes will be*
  - A      sums of money
  - B      new instruments
  - C      cups
  
- 4      *How does Mrs Smith feel about organising the food?*
  - A      annoyed that the pizza restaurant can't run a stall
  - B      worried that she won't have enough parents to help
  - C      anxious about whether there will be a lot of garbage
  
- 5      *Mrs Smith says the organiser's office will be*
  - A      in the English department
  - B      in a first-floor classroom
  - C      in the library
  
- 6      *Mrs Smith says that the first thing students should start doing is*
  - A      ordering tickets
  - B      writing letters



## C designing posters

### Tapescript

Good morning. As you know, all students are invited to give their ideas about the Music Festival and are also expected to help with the organisation. Now, to make a change from the usual classical pieces performed by school orchestras, we're going to also have rock bands. I know that a few of you wanted to include jazz but the staff feel that there wouldn't be enough interest from other schools in that.

There have been some problems with possible dates. Traditionally, the festival takes place in the first week in May, but, this year, exams will still be on, so it was decided to move it. We looked at two options: June and July and it was agreed the earlier date would be better for everyone, although some teachers preferred the later date. The principal has now confirmed the dates, which are the 21st to the 23rd, so please put that in your diary.

Now, the school usually gives the winning performance a silver cup but, last year, it was suggested that we ask local businesses if they would like to give something instead – for example, new instruments. But I'm delighted to announce that the local travel agency has agreed to give a cash prize to each winner. The singer, Paul Dobson, will give out the prizes on the day.

Now, we should think about food because we need to provide for possibly 300 people. We've depended on parents to help in the past, but I'm concerned that not many will be willing to help again as it's a lot to ask. We'll have a pizza stall but we'll have to run it ourselves because the local restaurant is too busy - but I'm sure we'll manage. There will, of course, be a lot of garbage to clear up - so we'll have a team of students dealing with this.

Now, the event organiser's office. Last year we used a first-floor classroom but that's not an option this year, so Dr Twining has kindly agreed that we can use the English department. The library will be closed. The performance areas, will, of course, be in the school hall and theatre.

Could you all sign up for what you'd be interested in doing? I'll need people to make a start writing letters – these need to go out next month. But the most urgent thing is designing posters – these will be placed in shop windows all over town. The Principal has ordered the tickets, so we don't have to worry about that.

Now, is there ... (fade)

## Key

<b>1</b>	<b>C</b>
<b>2</b>	<b>B</b>
<b>3</b>	<b>A</b>
<b>4</b>	<b>B</b>
<b>5</b>	<b>A</b>
<b>6</b>	<b>C</b>

## 14.1.2 Reading

Eight Reading task types were used in the final design for the ESLC.

### 14.1.2.1 Reading task type 1 (target level A1)

#### IR112 Cartolina di Lucia

Leggerai una cartolina. Lucia parla delle sue vacanze.  
Per le 4 domande seguenti scegli la risposta A, B o C.

Cara Emma,

Qui il tempo è bello e vado tutti i giorni in spiaggia. La sera, con i miei genitori, andiamo a mangiare in un ristorante tipico molto carino. Porto spesso quel vestito nero che abbiamo comprato insieme.

Ho un nuovo amico. Si chiama Karl, è simpatico e molto bello. È un ragazzo tedesco.

E tu? Sei andata in montagna, sulle Alpi, con tua cugina?

Quando torno a casa ti telefono. Ho tante cose da raccontarti!

Baci,

Lucia

#### 1 *Dov'è Lucia?*



#### 2 *Che cosa fa Lucia la sera?*



**3** *Che cosa porta spesso Lucia?*



A.



B.



C.

**4** *Al suo ritorno a casa, Lucia...*



A.



B.



C.

### Key

1	A
2	A
3	B
4	C

#### 14.1.2.2 Reading task type 2 (target level A1)

##### SR212 Hacer amigos

Vas a leer un anuncio sobre diferentes actividades escolares.  
Para las siguientes 4 preguntas, selecciona A, B o C.

---

### ¿Eres nuevo?

### ¿Quieres hacer nuevos amigos?

- El grupo de fútbol (chicos y chicas de 14 años) juega los lunes a las 17 h. en el gimnasio de la escuela.
- La reunión del grupo de cine es los miércoles. Hablar con Verónica. Teléfono: 686 345632.
- Necesitamos personas para trabajar en nuestro periódico de la escuela. Contactar con joaquin@moreno.es.
- ¿Tocas la guitarra o algún instrumento? La orquesta de la escuela practica los viernes a las 18:30 h. Contactar con la Señora García.
- Puedes ir a la discoteca una vez al mes. Encontrarás una invitación en tu clase.

- 1      *¿Cuándo puedes hacer deporte?*
  - A      una vez a la semana
  - B      dos veces al mes
  - C      tres veces al mes
  
- 2      *Si te gusta el cine, puedes hablar con*
  - A      Joaquín
  - B      Verónica
  - C      La señora García
  
- 3      *El grupo de música toca los*
  - A      lunes
  - B      miércoles
  - C      viernes
  
- 4      *Si te gusta escribir, puedes participar en*

- A la orquesta
- B el grupo de cine
- C el periódico de la escuela

**Key**

<b>1</b>	<b>A</b>
<b>2</b>	<b>B</b>
<b>3</b>	<b>C</b>
<b>4</b>	<b>C</b>



### **14.1.2.3 Reading task type 3 (target level A2)**

#### **ER323 Adverts on the notice-board**

You will read some advertisements from a library noticeboard.  
For the next 6 questions, choose the answer (A–H). Use each letter once only.

---

- 1**        You can find out which sport is best for you.
- 2**        Telephone this number if you want a new job.
- 3**        Go here if you want to learn a language.
- 4**        We can help you improve the way you learn.
- 5**        You can go online here.
- 6**        Join this to practise a sport.



**A STUDY ROOM**

.

Mornings: 09:00 to 13:00

Afternoons: 14:00 to  
19:00

Computers available

Sorry – we are closed in  
July

**B SUMMER CLASSES**

.

French and Spanish

Beginners

Small groups

Excellent prices

Call Westside College on  
0115 4432 556

- C** We need girls/boys to  
make  
a basketball team

**Seagrove Sports Club**

Wednesdays, 19.00-  
20.00

**D MATHS TEACHER  
NEEDED**

.

We can offer:

- Attractive salary
- Very good working  
conditions

tel. 0115 995 6781

**E STUDY SKILLS  
CENTRE**

.

Courses start in  
October

Study better, not  
harder!

tel. 0115 894 2583

**F TALK :**

.

"How to choose a  
Sport"

Speaker: Sara Smith

22 January, 19:30

All welcome

**G I CAN TEACH YOU TO  
DANCE!**

.

You can move to the music!

Mike: tel. 0115 338 3765

**H "UNDERSTANDING THE  
LANGUAGE OF DREAMS"**

.



Organised by visiting  
professor, P. Lewis

(tel. 0115 781 0401)

## Key

<b>1</b>	<b><i>F</i></b>	
<b>2</b>	<b><i>D</i></b>	
<b>3</b>	<b><i>B</i></b>	
<b>4</b>	<b><i>E</i></b>	
<b>5</b>	<b><i>A</i></b>	
<b>6</b>	<b><i>C</i></b>	

#### 14.1.2.4 Reading task type 4 (target level A2)

##### GR423 Schreiben macht den Kopf frei

Du liest den Bericht einer jungen Autorin.

Wähle bei den folgenden 5 Aufgaben die richtige Lösung A, B oder C.

---

##### Beim Schreiben bekomme ich einen klaren Kopf

Ich heiße Lisa Rotfuß und bin 16 Jahre alt. 2009 kam mein erster Roman „Die vergessene Insel“ heraus. Ich verkaufte viele tausend Stück. Ein wunderbares Glück, ich kann es immer noch nicht glauben!

Als ich klein war, hat mir Schreiben in der Schule großen Spaß gemacht. Ich habe immer gern gelesen und schon früh meine eigenen Geschichten erzählt. Ich erzähle immer von meinem eigenen Leben. Das Leben von anderen Personen finde ich nicht so spannend.

Jeden Tag schreibe ich ein bis zwei Stunden. Oft höre ich dabei Musik. Danach fühle ich mich sehr gut. Das Schreiben ist für mich wie Joggen oder Schwimmen.

Bis ein Buch fertig ist, dauert es sehr lange. Ich mache es immer so:

Zuerst erzähle ich meiner Freundin Sophie meine Idee. Dann schreibe ich die Geschichte in ein Heft. Meine Mutter schreibt die Texte dann mit dem Computer ab. Mit ihrer Hilfe kann ich meine Texte auch verbessern. Sie sagt mir, was ihr gefällt und was sie nicht versteht.

Mein zweiter Roman ist schon fast fertig. Er ist ganz anders als der erste. Hoffentlich gefällt er meinen Lesern genauso gut wie der erste.

**1** *Beim Schreiben benutzt Lisa Erfahrungen aus ...*

- A anderen Büchern
- B dem Leben ihrer Freunde
- C ihrer Kindheit

**2** *Schreiben ist für Lisa wie ...*

- A Musik machen
- B Sport machen
- C eine Geschichte lesen

**3** *Wer hört Lisas Geschichten zuerst?*

- A ihre Freundin
- B ihre Mutter
- C andere Personen

- 4 *Lisa denkt, dass ihr neuer Roman ...*
- A besser ist als der erste
  - B anders ist als der erste
  - C länger ist als der erste
- 5 *Was erfahren wir in diesem Text über Lisa?*
- A Wo sie ihre Romane schreibt
  - B Wann sie schreibt
  - C Wie sie ihre Bücher schreibt

Key

1	C
2	B
3	A
4	B
5	C



#### 14.1.2.5 Reading task type 5 (target level A1)

##### FR522 Mes amies

Tu vas lire un mél que Peter a reçu de son amie Cathie.  
Pour les 4 questions suivantes, réponds A, B ou C.

Salut Peter,

Tu m'as demandé de te parler de mes amies. Je le fais avec plaisir. Barbara est depuis longtemps ma meilleure amie. Dès le premier cours, elle s'est assise à côté de moi et depuis on fait tout ensemble. Les devoirs, les anniversaires, les excursions.

Dans ma classe il y a aussi une fille que j'aime bien. Elle s'appelle Léa. On est toutes les trois souvent ensemble. On fait beaucoup de choses ensemble, par exemple on fait les boutiques et on essaye des vêtements.

Léa est très sportive et fait du jogging tous les matins. Barbara aime bien ça mais moi pas du tout. Je préfère faire du vélo ou bien danser.

Le week-end dernier, nous sommes parties ensemble au bord de la mer mais l'eau était encore trop froide, nous avons donc bronzé au soleil. La prochaine fois, nous prendrons un ballon de volley pour pouvoir jouer. Est-ce que toi aussi tu aimes être en plein air ?

Merci beaucoup pour ta photo : votre maison a vraiment l'air très grande. Maintenant tu dois aussi m'envoyer une photo de ton chien et je connaîtrais tout le monde !

Maintenant c'est à toi de me parler de tes amis. Je suis déjà curieuse !

Cathie

1 Où est-ce que Cathie a fait connaissance de ses amies ?



A.



B.



C.

2 Qu'est-ce que les filles aiment faire ensemble ?



A.



B.

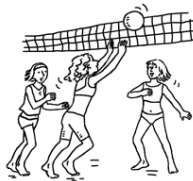


C.

3 Qu'est-ce que les filles ont fait au bord de la mer ?



A.



B.



C.

4 Qu'est-ce qu'il y a sur la photo de Peter ?



A.



B.



C.

### Key

1	A
2	C
3	C
4	B

### 14.1.2.6 Reading task type 6 (target level B1)

#### GR631 Europäischer Schüleraustausch

Du liest drei Berichte über einen europäischen Schüleraustausch.  
Wähle bei den folgenden 6 Aufgaben die richtige Lösung A, B oder C.

- 
- |   |   |
|---|---|
| 1 | Welche Teilnehmerin hatte am Anfang mehrere Probleme mit dem Projekt?                     |
| 2 | Welche Teilnehmerin wird nächstes Jahr die Schule verlassen?                              |
| 3 | Welche Teilnehmerin wird an ihrem Projekt weiterarbeiten?                                 |
| 4 | Welche Teilnehmerin will auch privat mit den Mitgliedern ihrer Gruppe in Kontakt bleiben? |
| 5 | Welche Teilnehmerin hatte keine Schwierigkeiten mit der Fremdsprache?                     |
| 6 | Welche Teilnehmerin darf in ihrem Heimatland nicht mehr so viele Schulstunden verpassen?  |

#### Europäischer Schüleraustausch – ein voller Erfolg!

**Options** A. Simone, 16, Frankreich

Es war super! Ich habe viele nette Leute kennengelernt, denen ich schreibe und die ich vielleicht sogar in den Ferien wieder treffen werde. Da ich gerne koche, habe ich mich für das Projekt „Essen international“ entschieden. Wir haben gemeinsam das Essen für das Abschlussfest vorbereitet. Probleme gab es überhaupt keine – wenn man praktisch zusammenarbeitet, ist die Sprache nicht so wichtig. Leider werde ich in Zukunft nicht mehr dabei sein können, weil meine Eltern es mir nicht mehr erlauben, eine Woche im Unterricht zu fehlen.

B. Lisa, 16, Finnland

Während der Projektwoche habe ich bei der Fotogruppe mitgemacht. Das war eine wunderschöne Erfahrung. Ich hätte nie gedacht, dass es so wenige Sprachprobleme geben würde – wir kamen immerhin aus fünf verschiedenen Ländern. Meistens haben wir Englisch gesprochen, manchmal auch Deutsch. Wir haben unser Projekt ganz selbstständig entwickelt. Dabei ist ein wirklich witziger Foto-Roman entstanden. Ich finde es wunderbar, dass wir uns nächstes Jahr wieder treffen werden, da schreiben wir dann die Fortsetzung.

C. Jasmin, 17, Deutschland

Ich hatte mir die Zeitungsgruppe ausgesucht. Das war gar nicht so einfach, weil es ja am ersten Tag noch nichts gab, worüber wir schreiben konnten. Zuerst gab es auch viele Missverständnisse in der Gruppe. Da habe ich mich wirklich gefragt, ob wir es schaffen, eine ganze Zeitung zu produzieren. Aber dann hat doch alles super geklappt, und alle konnten als



Erinnerung eine Zeitung mit nach Hause nehmen. Es hat wirklich unheimlich Spaß gemacht, aber leider war es für mich das letzte Mal, weil ich nächstes Jahr Abitur mache.

### Key

<b>1</b>	<b>C</b>
<b>2</b>	<b>C</b>
<b>3</b>	<b>B</b>
<b>4</b>	<b>A</b>
<b>5</b>	<b>B</b>
<b>6</b>	<b>A</b>

#### 14.1.2.7 Reading task type 7 (target level B1)

##### FR732 Littérature de jeunesse

Tu vas lire un texte sur les livres pour enfants.  
Pour les 4 questions suivantes, réponds A, B ou C.

Beaucoup d'enfants écrivent, mais personne ne voit jamais leurs textes... Un jour d'avril 2008, notre maison d'édition a reçu un texte manuscrit intitulé « les Fous », une version moderne d'un conte de fée qui nous a fait beaucoup rire. Quelle invention ! Quel style ! Ce jour-là, nous avons décidé de consacrer une collection entière écrite par des enfants ! Pour que tous ces enfants qui écrivent deviennent des auteurs comme les autres, parfois meilleurs !

Il fallait illustrer « les Fous » et nous avons tout de suite pensé à Armel qui nous disait après la lecture : « J'ai déjà dans les yeux et dans la tête quelques idées... Je sens que je vais créer quelque chose d'assez nouveau. J'ai envie d'y aller à fond tout en respectant ce texte qui est excellent ! »

Et le livre a pris forme jusqu'à se retrouver dans les bacs des librairies. D'un simple texte écrit à la main par une petite fille sur du papier à carreaux, nous avons fait un album qui restera dans l'histoire de la littérature jeunesse.

**1** *Que dit le texte sur le livre « Les Fous ? »*

- A Il est basé sur la vie de l'auteur
- B Il raconte une histoire imaginaire
- C C'est un recueil d'histoire drôles

**2** *Quelle a été la réaction d'Armel ?*

- A Il a hésité avant d'accepter de coopérer avec l'auteur
- B Il a immédiatement su comment il allait illustrer
- C Il a discuté avec l'auteur pour connaître ses idées

**3** *La maison d'édition*

- A accepte les textes sous n'importe quelle forme
- B demande des textes tapés à l'ordinateur
- C préfère les textes présentés de manière originale

4 Pourquoi le livre « les Fous » marquera-t-il la littérature jeunesse ?

- A Parce que son histoire s'adresse aux enfants et aux adultes
- B Parce que toutes ses illustrations sont de très grande qualité
- C Parce qu'il est le premier d'une collection écrite par des enfants

**Key**

<b>1</b>	<b>B</b>
<b>2</b>	<b>B</b>
<b>3</b>	<b>A</b>
<b>4</b>	<b>C</b>

### 14.1.2.8 Reading task type 8 (target level B2)

#### ER843 Teens and technology

You will read a magazine article about a teenage girl who tried to live without technology for a day.

Six sentences have been removed from the text below. Read the text and choose the sentence (A-G) which fits each gap. There is one extra sentence which you do not need to use.

---

#### A day without technology

Computing Today Magazine *asked 16-year-old Maya Ford to take a 24-hour break from digital technology and to report back to us on her experiences.*

My day didn't start that well. I normally have the radio on in the morning so I can catch up on the news and weather forecast. [ ...1... ] I managed it though and by 7.30am I was on the school bus. One of my friends was finishing his homework on his laptop and another was talking on her mobile phone, no doubt to a friend whom she would be seeing shortly anyway. [ ...2... ] I spent the journey looking out of the window and wondering how my technology-free day would turn out.

My first failure of the day came during my technology skills class. [ ...3.. ] I have to admit I wasn't disappointed – it would have been a shame to miss a very interesting lesson on graphics and digital photography.

During the lunch break, many students made use of the computer areas. Normally, I would have been one of them, checking my email or looking something up on the Internet. [ ...4... ] Many of them were rushing around frantically trying to finish assignments due next class. Luckily, I had remembered to print everything out at home, so I could relax and enjoy my lunch, not something that often happens!

At 3.30 pm school was finished. Normally, I would slip my earphones on and listen to some music on my way home. [ ...5... ] I often listen to music between lessons as well, and I was finding it a real struggle to get through a whole day without music.

At around 8.00 pm I sat down to do my homework. Without music playing in the background I wasn't sure I'd be able to concentrate, but in fact I was fine. [ ...6... ] This probably has quite a lot to do with the fact that I would normally get distracted by friends who are online at the same time as me. Maybe some of our chat is a bit unnecessary.

I decided to go to bed quite a bit earlier than normal. I set my mobile phone in its holder to be recharged overnight and dreamt of the next day, when I would once more be connected to the digital world. A technology-free life is not for me!

- A. This helps me to relax after working hard all day.
- B. It was a real challenge getting ready in total silence.
- C. I actually finished in far less time than it normally takes me.
- D. Since I couldn't participate, I felt like a bit of an outsider.
- E. The younger ones all seemed to be playing on hand-held gaming devices.
- F. This is compulsory for all students, and I was not allowed to skip it – experiment or no experiment.
- G. Instead I took the chance to sit back and observe my friends' behaviour.

**Key**

<b>1</b>	<b>B</b>
<b>2</b>	<b>E</b>
<b>3</b>	<b>F</b>
<b>4</b>	<b>G</b>
<b>5</b>	<b>A</b>
<b>6</b>	<b>C</b>

### 14.1.3 Writing

Four writing task types were used in the final design for the ESLC.

#### 14.1.3.1 Writing task type 1 (target level A1)

##### FW113 photo de famille



Tu as passé tes vacances dans une famille française. Tu écris un email à un ami et tu envoies cette photo.

Tu dis 3 choses sur cette photo :

- tu parles de la famille,
- de la profession des parents,
- de la maison.

Tu écris 20 à 30 mots.



### 14.1.3.2 Writing task type 2 (target level A2)

#### SW221 Resultados del examen

Has hecho un examen de español. Los resultados salen hoy en el tablón de anuncios de tu escuela pero no puedes ir a verlos.

Escribe un *e-mail* a un amigo para:

- explicarle por qué no puedes ir a ver los resultados;
- pedirle que te informe sobre tus resultados;
- darle las gracias y proponerle un plan.

Escribe 25-35 palabras.



### 14.1.3.3 Writing task type 3 (target level B1)

#### IW334 Rispondere ad una email

Questa è una parte della email che hai ricevuto da un amico.  
Scrivigli una email e rispondi alle sue domande.

*Io faccio parte della squadra di nuoto della mia scuola e abbiamo vinto il campionato.*

*Tu che sport fai nella tua scuola?*

*Qual è il tuo sport preferito e perché ti piace tanto?*

Scrivi 80 - 100 parole.

### 14.1.3.4 Writing task type 4 (target level B2)

#### EW442 Teens and going out

You recently had a class discussion about whether or not teenagers should be allowed to go out at night.

Now your teacher has asked you to write an essay answering the following question:

*Is it important for teenagers to be allowed to go out at night with their friends?*

Write an essay giving your opinion.

Write 120–180 words.



# *Appendix 2: Questionnaires*

## 14.2 Questionnaires

Student Name/ID:

ESLC ID:

Date:



# Student Questionnaire for the European Survey on Language Competences 2011 Main Study

## About this full note version

Accompanying each question are:

- **Notes for WebTrans** indicating the recurring question elements that have to be linked in WebTrans, so they have to be translated only once.
- **Notes for the NRC** clarifying terms and options, noting where localisations should be made, and providing a rationale for the question's inclusion.
- **Notes for the translator** clarifying terms and options, noting where response categories and/or terms should not be translated, because they have to be localised.
- **Notes for the Test Administrators** giving some guidance in how to answer questions that students might pose during the administration. Note that where an Educational system has adapted a question, it may be necessary for the NRC to add notes that will assist the Test Administrator to answer student questions.

## Conventions in this document

Terms in curly brackets { } should not be translated, but localised. The NRC is asked to instruct the reconciler which localised (a term which is appropriate for the country) terms to insert.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.

The term [target language] refers to the language for which the students will be tested prior to filling out the questionnaire, in other words the 1<sup>st</sup> most widely taught language among English, French, German, Spanish and Italian. Please instruct the translator which language to fill in when the term [target language] appears.

The curly brackets { } and square brackets [ ] should not appear in the translated text. Please instruct the translators and reconciler to remove these brackets from the translated text.

Please note that all questions should be translated even when it is expected that all respondents in your country will give the same answer. Question order is known to have an effect upon the answers. Removal of questions will compromise the comparability across countries and the comparability with future cycles.

Overall we have tried to prevent questions that require an open-ended text response as much as possible, as the coding of such questions (an NRC task) is very time consuming and costly.

## **Translator note**

Throughout the questionnaire the informal address is used.

“You” is singular unless otherwise indicated.

Terms in curly brackets { } should not be translated, but localised. The reconciler and/or NRC should insert a phrase or word which is appropriate for the country.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.

Please remove all curly brackets { } and square brackets [ ] in the translated text.



In this questionnaire you will find questions about:

- You
  - Your family and your home
  - Computers in your home
  - Languages in your home environment
  - Your opinion about foreign languages
- 
- Your school subjects
  - Learning foreign languages in school
  - Your [target language] lessons
  - Tests and assignments for the subject of [target language]
  - Studying and doing homework for foreign languages out of school time
  - Your skills in [target language]

In this questionnaire, there are no right or wrong answers. Your answers should be the ones that are right for you. Please read the questions carefully. You may ask for help if you do not understand something or are not sure how to answer a question.

It is important that you answer all questions.

All your answers will be kept confidential and secret.

**Translator note**

Please check that the list of subjects is consistent with the section headings.

home - refers to that place where the student usually resides, not including boarding school. The term used should connote a family or domestic setting.

[Target language] - Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

all - Please underline the appropriate text in the translation.

## About you

- 1 Are you female or male? Female Male
- ☐ 0 ☐ 1

### NRC note

This question assesses the gender of students.

- 2 What is your date of birth?
- (Please write down the day, month and year in which you were born)* Day Month Year

### NRC note

This question assesses the age of students.

- 3 The place where you live is *(Please select only one answer)*
- A village, hamlet or rural area (fewer than three thousand people) ----- ☐ 0
- A small town (three thousand to around fifteen thousand people) ----- ☐ 1
- A town (fifteen thousand to around hundred thousand people) ----- ☐ 2
- {A city (hundred thousand to around one million people)} ----- ☐ 3
- {A large city with over one million people} ----- ☐ 4

### WebTrans Note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

### NRC note

This question assesses the home location of students. The home location of students is likely to impact the informal language learning opportunities of students (e.g. in a large city the potential of coming into contact with other languages is bigger). The number of inhabitants mentioned in each response option follow the convention of PISA.

Some of the items might need to be localised. Please instruct the reconciler which items should be used.



{A city (hundred thousand to around one million people)} - If a town of this size does not exist in the Educational system, this response option can be omitted.

{A large city with over one million people} - If a town of this size does not exist in the Educational system, this response option can be omitted.

Please provide the test administrator several examples of towns and cities for the response categories and instruct the test administrator to look up the size of the town where the school is located.

“A small town (three thousand to around fifteen thousand people)” - e.g. {Wells in Somerset}

“A town (fifteen thousand to around hundred thousand people)”, e.g. {Bath}

“A city (hundred thousand to around one million people)”, e.g. {Nottingham}

“A large city with over one million people”, e.g. {London}

### **Administrator note**

If a student does not know whether he/she lives in a small town, town, city or large city, you can help the student with asking questions such as “Do you live in the town where the school is located?” and with giving the student some examples of towns and cities of different sizes.

4	Which language(s) did you speak at home as a small child (before the age of five)?			
<i>(Please tick as many boxes as applicable)</i>				
	{most widely spoken indigenous language1} -----	<input type="checkbox"/> 0		
	{most widely spoken indigenous language2} -----	<input type="checkbox"/> 1		
	{most widely spoken indigenous language3} -----	<input type="checkbox"/> 2		
	{most widely spoken indigenous language4} -----	<input type="checkbox"/> 3		
	{most widely spoken indigenous language5} -----	<input type="checkbox"/> 4		
	[target language] -----	<input type="checkbox"/> 5		
	{most widely spoken non-indigenous language 1} ---	<input type="checkbox"/> 6		
	{most widely spoken non-indigenous language 2} ---	<input type="checkbox"/> 7		
	{most widely spoken non-indigenous language 3} ---	<input type="checkbox"/> 8		
	{most widely spoken non-indigenous language 4} ---	<input type="checkbox"/> 9		
	{most widely spoken non-indigenous language 5} ---	<input type="checkbox"/> 10		
	Other European language(s) -----	<input type="checkbox"/> 11		
	Other non-European language(s) -----	<input type="checkbox"/> 12		

### WebTrans note

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ4, SQ25, SQ26 and SQ27 are identical.

Please note that the number of response categories can differ between countries (but the maximum is 13 - the number of response categories displayed).

### NRC note

SQ4, SQ25, SQ26, SQ27 and SQ28 provide information about the languages in the home environment. This question inquires after the 1<sup>st</sup> language(s) of the student. Please note that the term “mother tongue” is not used, because “for a considerable number of people in Europe, the notion of “mother tongue” has lost its meaning” (High Level Group on Multilingualism, 2007, p. 6) for example when respondents grew up in mixed language families or multilingual environments.

The response categories of this question need to be localised. The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.



A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

#### **Translator note**

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

5 Which [grade] are you in?

*[grade]*

---

#### **NRC note**

Grade - refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

In countries where students may be in different grades/year levels for different courses, an adequate adaptation of this question that can be understood by students is advisable.

#### **Translator note**

[Grade] - refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.



## 6 Which one of the following [programmes] are you in?

*(Please select only one answer)*

- {Programme 1}----- ☐ 0
- {Programme 2}----- ☐ 1
- {Programme 3}----- ☐ 2
- {Programme 4}----- ☐ 3
- {Programme 5}----- ☐ 4
- {Programme 6}----- ☐ 5
- {Programme 7}----- ☐ 6
- {Programme 8}----- ☐ 7
- {Programme 9}----- ☐ 8
- {Programme 10}----- ☐ 9
- {Programme 11}----- ☐ 10
- {Programme 12}----- ☐ 11
- {Programme 13}----- ☐ 12
- {Programme 14}----- ☐ 13
- {Programme 15}----- ☐ 14

### WebTrans Note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

Please note that the number of response categories can differ between countries (but the maximum is 15).

### NRC note

This questions assesses the study program the student is following.

The response categories of this question need to be localised. The study programmes presented as response categories should correspond to the study programmes at ISCED2 and ISCED3 level in the Localisation file (Study Program Table). It is necessary to phrase study programme labels in such a way that students will easily understand.

In some countries there may not be an administrative or structural boundary between some successive ISCED levels (between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels (see the Instruction for the Main Study Localisation file).



In countries where students are grouped in school according to their ability, an adequate adaptation of this question that can be understood by students is advisable.

Please instruct the reconciler which study programs should be used as response categories.

**Translator note**

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

[Programmes] - the study programmes the student can follow in secondary education.

## About your family and your home

*In this section you will be asked some questions about your family and your home.*

*Some of the following questions are about your mother and father or those persons who are like a mother or father to you — for example, guardians, step-parents, foster parents, etc.*

*If you share your time with more than one set of parents or guardians, please answer the following questions for those parents/guardians you spend the most time with.*

### **NRC note**

The questions in this section are necessary for a measurement of socio-economic status that is comparable to PISA and for a measurement of immigrant status.

The same questions, items and response scales are used for calculating the measure of socio-economic status as in PISA. In some instances the exact wording has been improved whenever the English partners of the Consortium or the cognitive labs indicated that the original (PISA) wording was awkward.

### **Translator note**

home - refers to that place where the student usually resides, not including boarding school. The term used should connote a family or domestic setting.

### **Administrator note**

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer 'Home' questions for the chosen household.

Some of the following questions are about the mother and father or the student or those persons who are like mother or father to the student — for example, guardians, stepparents, foster parents, etc.

If the student shares his/her time with more than one set of parents or guardians, the questions should be answered for those parents/guardians the student spends the most time with.

If one of the student's parents (or equivalent guardian) is deceased, then the information (last job and highest level of schooling) can still be provided - but if this is unknown the question should be left blank.

If a student has only one parent (or equivalent guardian), then only for this parent can information be provided and the questions about the other parent should be left blank.

---

## 7 What is your mother's main job?

If she is not currently working, please tell us what her last main job was.

*(Please write down the [job title], for example sales manager)*

---

### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

### Translator note

[job title] - is the common name of the job. Please use an appropriate term.

main job - is the job in which most time is spent, not necessarily the highest earning job.

### Administrator note

If the mother has more than one job, her 'main job' is the job in which most time is spent, not necessarily the highest earning job.

If the student's mother (or equivalent female guardian) is deceased, then the last job can still be provided, but if this is unknown the question should be left blank.

If the student has only a father (or equivalent male guardian), then the questions should be left blank.

Encourage students to answer this question. A general description such as 'works in an office' is better than nothing written at all.

---

## 8 What does your mother do in her main job?

*(Please describe the kind of work she does or did in that job, for example manages a sales team)*

---

### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

### Translator note

main job - is the job in which most time is spent, not necessarily the highest earning job.

### Administrator note

If the mother has more than one job, her 'main job' is the job in which most time is spent, not necessarily the highest earning job.



If the student's mother (or equivalent female guardian) is deceased, then the last job can still be provided, but if this is unknown the question should be left blank.

If the student has only a father (or equivalent male guardian), then the questions should be left blank.

Encourage students to answer this question. A general description such as 'works in an office' is better than nothing written at all.

9	How is your mother currently employed?				
	<i>(Please select only one answer)</i>				
	Working in full-time paid employment -----	<input type="radio"/> 0			
	Working in part-time paid employment -----	<input type="radio"/> 1			
	Not working, but looking for a job -----	<input type="radio"/> 2			
	Other, e.g. home duties, retired -----	<input type="radio"/> 3			

#### WebTrans Note

The same response instruction "(Please select only one answer)" occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

#### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

#### Administrator note

Students should answer this question with regard to their mother's main occupation. If e.g. the mother is working part-time but also doing home duties, students should choose "Working in part-time paid employment".

If the student's mother (or equivalent female guardian) is deceased, then the questions should be left blank.

If the student has only a father (or equivalent male guardian), then the questions should be left blank.

Please note that students should tick only one box.

---

## 10 What is your father's main job?

If he is not currently working, please tell us what his last main job was.

*(Please write down the [job title], for example sales manager)*

---

### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

### Translator note

[job title] - is the common name of the job. Please use an appropriate term.

main job - is the job in which most time is spent, not necessarily the highest earning job.

### Administrator note

If the father has more than one job, his 'main job' is the job in which most time is spent, not necessarily the highest earning job.

If the student's father (or equivalent male guardian) is deceased, then the last job can still be provided, but if this is unknown the question should be left blank.

If the student has only a mother (or equivalent female guardian), then the questions should be left blank.

Encourage students to answer this question. A general description such as 'works in an office' is better than nothing written at all.

---

## 11 What does your father do in his main job?

*(Please describe the kind of work he does or did in that job, for example manages a sales team)*

---

### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

### Translator note

main job - is the job in which most time is spent, not necessarily the highest earning job.

### Administrator note

If the father has more than one job, his 'main job' is the job in which most time is spent, not necessarily the highest earning job.

If the student's father (or equivalent male guardian) is deceased, then the last job can still be provided, but if this is unknown the question should be left blank.

If the student has only a mother (or equivalent female guardian), then the questions should be left blank.

Encourage students to answer this question. A general description such as 'works in an office' is better than nothing written at all.

12	How is your father currently employed?			
	<i>(Please select only one answer)</i>			
	Working in full-time paid employment -----	<input type="radio"/> 0		
	Working in part-time paid employment -----	<input type="radio"/> 1		
	Not working, but looking for a job -----	<input type="radio"/> 2		
	Other, e.g. home duties, retired -----	<input type="radio"/> 3		

#### WebTrans Note

The same response instruction "(Please select only one answer)" occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

#### NRC note

SQ7, SQ8, SQ9, SQ10, SQ11, SQ12 are similar to the PISA questions about parental occupational and employment status.

#### Administrator note

Students should answer this question with regard to their father's main occupation. If e.g. the father is working part-time but also doing home duties, students should choose "Working in part-time paid employment".

If the student's father (or equivalent male guardian) is deceased, then the questions should be left blank.

If the student has only a mother (or equivalent female guardian), then the questions should be left blank.

Please note that students should tick only one box.

### 13 What is the highest level of schooling completed by your mother?

If you are not sure which answer to choose, please ask the test administrator for help.

*(Please select only one answer)*

- {ISCED 5A,6}----- ☐ 0
- {ISCED5B}----- ☐ 1
- {ISCED4}----- ☐ 2
- {ISCED3A}----- ☐ 3
- {ISCED 3B,3C}----- ☐ 4
- {ISCED2}----- ☐ 5
- {ISCED1}----- ☐ 6
- She did not complete {ISCED1} or she never went to school ----- ☐ 7

#### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The clarification “If you are not sure which answer to choose, please ask the test administrator for help.” is identical in SQ13 and SQ14.

The response categories of SQ13 and SQ14 are identical, except for the last response category (“She did not complete {ISCED1} or she never went to school”)

Please note that the number of response categories can differ between countries (but the maximum is 8 - the number of response categories displayed).

#### NRC note

SQ13 and SQ14 are similar to the four PISA questions about parental educational status.

The response categories of this question need to be localised. Most EU countries have officially classified their educational system using the ISCED classification of educational levels (levels (see Classifying Educational Programmes — Manual for ISCED-97 Implementation in OECD Countries, 1999 Edition, OECD). These country-specific classifications can be also be found on the website of Eurydice (<http://eacea.ec.europa.eu/portal/page/portal/Eurydice/>).

Please make sure that each ISCED-level is represented (ISCED1, ISCED2, ISCED3, ISCED4 and ISCED5) as a response category. It is necessary to phrase the response categories in such a way that students will easily understand. When a ISCED level with a particular orientation, e.g. ISCED3B, does not exist in the country the corresponding response category should be omitted.

In some countries there may not be an administrative or structural boundary between some successive ISCED levels (for example between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels (see the Instruction for the Main Study Localisation file).

Changes in the educational system - several countries have changed their educational system throughout the years. Please make sure that also previous study programs are included.





Qualifications obtained abroad - increasingly, students have parents whose qualifications were obtained abroad and these may not match the nationally specific categories listed in the questionnaire. The student has been instructed to ask the test administrator if they have any doubt about which option to choose.

Please provide the test administrator some guidelines in their training on the equivalence of local qualifications to those obtained abroad. The match does not have to be exact. It is more important to try to distinguish between the three general levels: ISCED 5A and above, ISCED 3, and below ISCED 3. The test administrator should ask the student appropriate questions to identify which of the three levels most closely corresponds.

Please instruct the reconciler which terms should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets and do not translate the term {ISCED1} in the last response category. The reconciler and/or NRC is asked to add the appropriate response categories and terms.

### Administrator note

If the student's mother (or equivalent female guardian) is deceased, then the highest level of schooling can still be provided - but if this is unknown the question should be left blank.

If the student has only a father (or equivalent male guardian), then the questions should be left blank.

If the student's mother obtained her qualifications abroad, please help the student to choose the response from the list that is closest. Asking questions like "How long did she go to school for" and "Did she go to university?" should help clarify.

---

## 14 What is the highest level of schooling completed by your father?

If you are not sure which answer to choose, please ask the test administrator for help.

*(Please select only one answer)*

- |   |                         |
|---|-------------------------|
| {ISCED 5A,6}-----   | <input type="radio"/> 0 |
| {ISCED5B}-----  | <input type="radio"/> 1 |
| {ISCED4}-----   | <input type="radio"/> 2 |
| {ISCED3A}-----  | <input type="radio"/> 3 |
| {ISCED 3B,3C}-----  | <input type="radio"/> 4 |
| {ISCED2}-----   | <input type="radio"/> 5 |
| {ISCED1}-----   | <input type="radio"/> 6 |
| He did not complete {ISCED1} or he never went to school ----- | <input type="radio"/> 7 |

---

### WebTrans note

The same response instruction "(Please select only one answer)" occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.



The clarification “If you are not sure which answer to choose, please ask the test administrator for help.” is identical in SQ13 and SQ14.

The response categories of SQ13 and SQ14 are identical, except for the last response category (“She did not complete {ISCED1} or she never went to school”)

Please note that the number of response categories can differ between countries (but the maximum is 8 - the number of response categories displayed).

### **NRC note (identical to note with SQ13)**

SQ13 and SQ14 are similar to the four PISA questions about parental educational status.

The response categories of this question need to be localised. Most EU countries have officially classified their educational system using the ISCED classification of educational levels (levels (see *Classifying Educational Programmes — Manual for ISCED-97 Implementation in OECD Countries*, 1999 Edition, OECD). These country-specific classifications can be also be found on the website of Eurydice (<http://eacea.ec.europa.eu/portal/page/portal/Eurydice/>).

Please make sure that each ISCED-level is represented (ISCED1, ISCED2, ISCED3, ISCED4 and ISCED5) as a response category. It is necessary to phrase the response categories in such a way that students will easily understand. When a ISCED level with a particular orientation, e.g. ISCED3B, does not exist in the country the corresponding response category should be omitted.

In some countries there may not be an administrative or structural boundary between some successive ISCED levels (for example between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels (see the Instruction for the Main Study Localisation file).

Changes in the educational system - several countries have changed their educational system throughout the years. Please make sure that also previous study programs are included.

Qualifications obtained abroad - increasingly, students have parents whose qualifications were obtained abroad and these may not match the nationally specific categories listed in the questionnaire. The student has been instructed to ask the test administrator if they have any doubt about which option to choose. Please provide the test administrator some guidelines in their training on the equivalence of local qualifications to those obtained abroad. The match does not have to be exact. It is more important to try to distinguish between the three general levels: ISCED 5A and above, ISCED 3, and below ISCED 3. The test administrator should ask the student appropriate questions to identify which of the three levels most closely corresponds.

Please instruct the reconciler which terms should be used as response categories.

### **Translator note**

Please do not translate the response categories in curly brackets and do not translate the term {ISCED1} in the last response category. The reconciler and/or NRC is asked to add the appropriate response categories and terms.

### **Administrator note**

If the student's mother (or equivalent female guardian) is deceased, then the highest level of schooling can still be provided - but if this is unknown the question should be left blank.

If the student has only a father (or equivalent male guardian), then the questions should be left blank.

If the student's mother obtained her qualifications abroad, please help the student to choose the response from the list that is closest. Asking questions like “How long did she go to school for” and “Did she go to university?” should help clarify.

15	What country were you born in?			
	<i>(Please select only one answer)</i>			
	[Educational system] -----	<input type="radio"/> 0		
	{Country A} -----	<input type="radio"/> 1		
	{Country B} -----	<input type="radio"/> 2		
	{Country C} -----	<input type="radio"/> 3		
	{Country D} -----	<input type="radio"/> 4		
	{Country E} -----	<input type="radio"/> 5		
	{Country F} -----	<input type="radio"/> 6		
	{Country G} -----	<input type="radio"/> 7		
	Other European country -----	<input type="radio"/> 8		
	Other non-European country -----	<input type="radio"/> 9		

### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The response categories of SQ15, SQ16 and SQ17 are identical

Please note that the number of response categories can differ between countries (but the maximum is 10 - the number of response categories displayed).

### NRC note

SQ15, SQ16, SQ17 and SQ18 provide information on immigrant background.

The response categories of this question need to be localised. The countries presented as response categories should correspond to the countries mentioned in the Localisation file (Country Table).

The countries listed should include the countries of origin of the largest immigrant groups in your Educational system. A maximum number of seven countries (other than your Educational system) can be included. Less than seven countries can be included when less than seven immigrant groups of substantial size reside in your Educational system (see the Instruction for the Main Study Localisation file).

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by students is advisable.

Please instruct the reconciler which countries should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

Educational system - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

16	What country was your mother born in?			
	<i>(Please select only one answer)</i>			
	[Educational system] -----	<input type="radio"/> 0		
	{Country A} -----	<input type="radio"/> 1		
	{Country B} -----	<input type="radio"/> 2		
	{Country C} -----	<input type="radio"/> 3		
	{Country D} -----	<input type="radio"/> 4		
	{Country E} -----	<input type="radio"/> 5		
	{Country F} -----	<input type="radio"/> 6		
	{Country G} -----	<input type="radio"/> 7		
	Other European country -----	<input type="radio"/> 8		
	Other non-European country -----	<input type="radio"/> 9		

### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The response categories of SQ15, SQ16 and SQ17 are identical

Please note that the number of response categories can differ between countries (but the maximum is 10 - the number of response categories displayed).

### NRC note (identical to note with SQ15)

SQ15, SQ16, SQ17 and SQ18 provide information on immigrant background.

The response categories of this question need to be localised. The countries presented as response categories should correspond to the countries mentioned in the Localisation file (Country Table).

The countries listed should include the countries of origin of the largest immigrant groups in your Educational system. A maximum number of seven countries (other than your Educational system) can be included. Less than seven countries can be included when less than seven immigrant groups of substantial size reside in your Educational system (see the Instruction for the Main Study Localisation file).

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by students is advisable.

Please instruct the reconciler which countries should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.



[Educational system] - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

17	What country was your father born in?			
	<i>(Please select only one answer)</i>			
	[Educational system] -----	<input type="radio"/> 0		
	{Country A} -----	<input type="radio"/> 1		
	{Country B} -----	<input type="radio"/> 2		
	{Country C} -----	<input type="radio"/> 3		
	{Country D} -----	<input type="radio"/> 4		
	{Country E} -----	<input type="radio"/> 5		
	{Country F} -----	<input type="radio"/> 6		
	{Country G} -----	<input type="radio"/> 7		
	Other European country -----	<input type="radio"/> 8		
	Other non-European country -----	<input type="radio"/> 9		

#### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The response categories of SQ15, SQ16 and SQ17 are identical

Please note that the number of response categories can differ between countries (but the maximum is 10 - the number of response categories displayed).

#### NRC note (identical to note with SQ15)

SQ15, SQ16, SQ17 and SQ18 provide information on immigrant background.

The response categories of this question need to be localised. The countries presented as response categories should correspond to the countries mentioned in the Localisation file (Country Table).

The countries listed should include the countries of origin of the largest immigrant groups in your Educational system. A maximum number of seven countries (other than your Educational system) can be included. Less than seven countries can be included when less than seven immigrant groups of substantial size reside in your Educational system (see the Instruction for the Main Study Localisation file).

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by students is advisable.

Please instruct the reconciler which countries should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

[Educational system] - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

18	How many years have you been living in [Educational system]?			
	(Please select only one answer)			
	Less then one year -----	<input type="radio"/> 0		
	One to two years -----	<input type="radio"/> 1		
	Three to four years-----	<input type="radio"/> 2		
	Five to six years -----	<input type="radio"/> 3		
	Seven to eight years-----	<input type="radio"/> 4		
	Nine to ten years -----	<input type="radio"/> 5		
	Eleven years or more/All your life -----	<input type="radio"/> 6		

### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

### NRC note

SQ15, SQ16, SQ17 and SQ18 provide information on immigrant background.

In educational systems that are not an entire country an adequate adaptation of this question that can be understood by students is advisable. Please instruct the reconciler how to adapt the question.

### Translator note

[Educational system] - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

## 19 Which of the following do you have at home?

(Please select in each row No or Yes)

	No	Yes
1) A desk to study at .....	<input type="radio"/> 0	<input type="radio"/> 1
2) A room of your own .....	<input type="radio"/> 0	<input type="radio"/> 1
3) A quiet place to study .....	<input type="radio"/> 0	<input type="radio"/> 1
4) Books to help with your school work (for example an encyclopaedia or atlas) .....	<input type="radio"/> 0	<input type="radio"/> 1
5) A computer you can use for school work .....	<input type="radio"/> 0	<input type="radio"/> 1
6) Educational software .....	<input type="radio"/> 0	<input type="radio"/> 1
7) An internet connection .....	<input type="radio"/> 0	<input type="radio"/> 1
8) A dictionary .....	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The same response instruction “(Please select in each row No or Yes)” occurs in SQ19, SQ20, SQ23, SQ29, SQ64.

The response scales (No-Yes) of SQ19, SQ20, SQ23, SQ29, and SQ64 are identical.

### NRC note

SQ19, SQ20, SQ21 and SQ22 provide information about home possessions and are similar to the PISA questions about home possessions. SQ19 and SQ20 were presented as one question in PISA, but this question was too long to fit on one computer screen. SQ21 is identical to the question used in PISA about the number of books and similar to the question about the number of books in PIRLS and TIMMS. SQ22 is similar to the PISA question about the number of home possessions.

### Translator note

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

### Administrator note

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer ‘Home’ questions for the chosen household.

## 20 Which of the following are in your home? (continued)

(Please select in each row No or Yes)

	No	Yes
1) Classics from the literature of [Educational system] (e.g. books of {Shakespeare}) .....	<input type="radio"/> 0	<input type="radio"/> 1
2) Books of poetry .....	<input type="radio"/> 0	<input type="radio"/> 1
3) Works of art (e.g. paintings) .....	<input type="radio"/> 0	<input type="radio"/> 1
4) A dishwasher .....	<input type="radio"/> 0	<input type="radio"/> 1
5) A DVD player .....	<input type="radio"/> 0	<input type="radio"/> 1
6) {Country specific wealth item1} .....	<input type="radio"/> 0	<input type="radio"/> 1
7) {Country specific wealth item2} .....	<input type="radio"/> 0	<input type="radio"/> 1
8) {Country specific wealth item3} .....	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The same response instruction “(Please select in each row No or Yes)” occurs in SQ19, SQ20, SQ23, SQ29, SQ64.

The response scales (No-Yes) of SQ19, SQ20, SQ23, SQ29, and SQ64 are identical.

### NRC note

SQ19, SQ20, SQ21 and SQ22 provide information about home possessions and are similar to the PISA questions about home possessions. SQ19 and SQ20 were presented as one question in PISA, but this question was too long to fit on one computer screen. SQ21 is identical to the question used in PISA about the number of books and similar to the question about the number of books in PIRLS and TIMMS. SQ22 is similar to the PISA question about the number of home possessions.

Some of the items of this question need to be localised.

{Country specific wealth item} - NRCs are requested to add three other indicators of wealth that suit the national context. Between 20 and 80 percent of students in the Educational system should be expected to have these items at home. To allow comparison with PISA we strongly recommend to use the same country specific wealth items as are used in PISA.

{Shakespeare} - needs to be substituted by a relevant classical author in the country's language.

Please instruct the reconciler which wealth items and classical author should be used.

### Translator note

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - (in “your home”) should be plural.

[Educational system] - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.





{Shakespeare} - the reconciler has to substitute “Shakespeare” by a relevant classical author in the country’s language.

DVD player - Digital Video Disc, please use terms common in your country

Please do not translate the items and terms in curly brackets. The reconciler and/or NRC is asked to add the appropriate items.

### Administrator note

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer ‘Home’ questions for the chosen household.

---

## 21 How many books are there in your home?

Generally, there are about 40 books on a bookshelf of one meter. Do not count newspapers, magazines and schoolbooks.

*(Please select only one answer)*

0-10 books ----- ☐ 0

11-25 books ----- ☐ 1

26-100 books ----- ☐ 2

101-200 books ----- ☐ 3

201-500 books ----- ☐ 4

More than 500 books ----- ☐ 5

---

### WebTrans Note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

### NRC note

SQ19, SQ20, SQ21 and SQ22 provide information about home possessions and are similar to the PISA questions about home possessions. SQ19 and SQ20 were presented as one question in PISA, but this question was too long to fit on one computer screen. SQ21 is identical to the question used in PISA about the number of books and similar to the question about the number of books in PIRLS and TIMMS. SQ22 is similar to the PISA question about the number of home possessions.

### Translator note

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - (in “your home”) should be plural.

### Administrator note

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer 'Home' questions for the chosen household.

## 22 How many of these are there in your home?

(Please select one answer from each row)

	None	One	Two	Three or more
1) Mobile phones .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) Television sets .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) Computers or laptops .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) Cars .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) Bathrooms .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction "(Please select one answer from each row)" occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

### NRC note

SQ19, SQ20, SQ21 and SQ22 provide information about home possessions and are similar to the PISA questions about home possessions. SQ19 and SQ20 were presented as one question in PISA, but this question was too long to fit on one computer screen. SQ21 is identical to the question used in PISA about the number of books and similar to the question about the number of books in PIRLS and TIMMS. SQ22 is similar to the PISA question about the number of home possessions.

### Translator note

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - (in "your home") should be plural.

Bathroom - A room with a bath or shower, not a toilet.

How many - Please underline the appropriate text in the translation.

### Administrator note

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer 'Home' questions for the chosen household.

## About computers in your home

### Translator note

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - (in “your home”) should be plural.

### 23 Are the following devices available for you to use at your home?

(Please select in each row No or Yes)

	No	Yes
1) <u>Your own</u> computer, laptop, or notebook -----	<input type="radio"/> 0	<input type="radio"/> 1
2) Access to the internet -----	<input type="radio"/> 0	<input type="radio"/> 1
3) A printer -----	<input type="radio"/> 0	<input type="radio"/> 1
4) A CD or DVD writer -----	<input type="radio"/> 0	<input type="radio"/> 1
5) A scanner -----	<input type="radio"/> 0	<input type="radio"/> 1
6) A USB (memory) stick -----	<input type="radio"/> 0	<input type="radio"/> 1
7) A video games console, such as {Play Station, Nintendo, Wii} --	<input type="radio"/> 0	<input type="radio"/> 1
8) <u>Your own</u> iPod, Mp3 player or similar -----	<input type="radio"/> 0	<input type="radio"/> 1
9) <u>Your own</u> mobile phone-----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The same response instruction “(Please select in each row No or Yes)” occurs in SQ19, SQ20, SQ23, SQ29, SQ64.

The response scales (No-Yes) of SQ19, SQ20, SQ23, SQ29, and SQ64 are identical.

### NRC note

This question provides information on the ICT facilities in the students’ home.

Some of the terms in this question might need to be localised. Please instruct the reconciler which terms should be used.

video game console - Refers to an interactive entertainment computer or electronic device that manipulates the video display signal of a display device (a television, monitor, etc.) to display a game. National Project Managers should insert a phrase that is appropriate for their country and that is understood by the students.

{Play Station, Nintendo, Wii} - Please use examples for game consoles that will be understood by students in your country.

### Translator note

Please do not translate terms in curly brackets. The reconciler and/or NRC is asked to add the appropriate terms.

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - (in “your home”) should be plural.

Your own - Please underline the appropriate text in the translation.

“Access to the internet” - This item should inquire whether the student him/herself has access to internet, not to the mere presence of an internet connection.

### Administrator note

‘Available’ means that the student can use the device. It does not mean the mere presence of the device in the home.

“Access to the internet” - This item inquires whether the student him/herself has access to internet, not to the mere presence of an internet connection.

If a student belongs to two households then the questions refer to the household in which he or she spends most time. If a student says he or she spends equal time in two households then he or she may choose either household, but consistently answer ‘Home’ questions for the chosen household.

24	How often do you use a computer outside school time for the following?	(Please select one answer from each row)				
		Never or hardly ever	A few times a year	A few times a month	A few times a week	(Almost) every day
1)	For homework or school assignments ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	For homework or assignments for the subject of [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	For finding information -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	For games -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	For entertainment (e.g. music, movies, video clips) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	For contact with others (e.g. email, chatting, blogging, {MySpace, Skype})----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never or hardly ever-A few times a year-A few times a month-A few times a week-(Almost) every day) of SQ24 and SQ62 are identical.



### **NRC note**

SQ24, SQ51 and SQ62 provide information on the use of ICT for foreign language learning. This question provides information on the use of ICT at home.

Some of the terms in this question might need to be localised. Please instruct the reconciler which terms should be used.

Skype - Skype is a software application that allows users to make voice calls over the Internet.

MySpace - is an example of a social networking website with an interactive, user-submitted network of friends, personal profiles, blogs, groups, photos, music, and videos for teenagers and adults. Please use an example of a social networking website that is most widely known in your country (e.g. Friendster, Twitter).

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

### **Translator note**

Please do not translate terms in curly brackets. The reconciler and/or NRC is asked to add the appropriate terms.

home - refers to that place where the student and his/her family usually resides, not including boarding school. The term used should connote a family or domestic setting.

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About languages in your home environment

### Translator note

home - refers to that place where the student usually resides, not including boarding school. The term used should connote a family or domestic setting.

Your - the word your (in “your home environment”) should be singular.

25	Which language(s) does your family speak <u>regularly</u> at home?			
	(Please tick as many boxes as applicable)			
	{most widely spoken indigenous language1} -----	<input type="checkbox"/> 0		
	{most widely spoken indigenous language2} -----	<input type="checkbox"/> 1		
	{most widely spoken indigenous language3} -----	<input type="checkbox"/> 2		
	{most widely spoken indigenous language4} -----	<input type="checkbox"/> 3		
	{most widely spoken indigenous language5} -----	<input type="checkbox"/> 4		
	[target language] -----	<input type="checkbox"/> 5		
	{most widely spoken non-indigenous language 1} ---	<input type="checkbox"/> 6		
	{most widely spoken non-indigenous language 2} ---	<input type="checkbox"/> 7		
	{most widely spoken non-indigenous language 3} ---	<input type="checkbox"/> 8		
	{most widely spoken non-indigenous language 4} ---	<input type="checkbox"/> 9		
	{most widely spoken non-indigenous language 5} ---	<input type="checkbox"/> 10		
	Other European language(s) -----	<input type="checkbox"/> 11		
	Other non-European language(s) -----	<input type="checkbox"/> 12		

### WebTrans note

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ4, SQ25, SQ26 and SQ27 are identical.

Please note that the number of response categories can differ between countries (but the maximum is 13 - the number of response categories displayed).

## NRC note

SQ4, SQ25, SQ26, SQ27 and SQ28 provide information about the languages in the home environment. This question provides information on the languages the student is exposed to in his or her home environment. In mixed language families or immigrant families the languages spoken by the family (SQ25) and the languages spoken by the student him or herself (SQ26) can differ.

The response categories of this question need to be localised (see SQ4). The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.

A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

Languages that the family only speaks once in a while or only in special circumstances should not be ticked (see Administrator note). If necessary the sentence “Languages that your family only speaks once in a while or only in special circumstances should not be ticked” can be added to the instruction (similar to TQ9).

## Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

home - refers to that place where the student usually resides, not including boarding school. The term used should connote a family or domestic setting.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

regularly - Please underline the appropriate text in the translation.

## Administrator note

Students should only tick the languages that the family speaks regularly in their private life. Languages that the family only speaks once in a while or only in special circumstances should not be ticked.

26	Which language(s) do you, yourself, speak <u>regularly</u> at home?			
	(Please tick as many boxes as applicable)			
	{most widely spoken indigenous language1} -----	<input type="checkbox"/> 0		
	{most widely spoken indigenous language2} -----	<input type="checkbox"/> 1		
	{most widely spoken indigenous language3} -----	<input type="checkbox"/> 2		
	{most widely spoken indigenous language4} -----	<input type="checkbox"/> 3		
	{most widely spoken indigenous language5} -----	<input type="checkbox"/> 4		
	[target language] -----	<input type="checkbox"/> 5		
	{most widely spoken non-indigenous language 1} ---	<input type="checkbox"/> 6		
	{most widely spoken non-indigenous language 2} ---	<input type="checkbox"/> 7		
	{most widely spoken non-indigenous language 3} ---	<input type="checkbox"/> 8		
	{most widely spoken non-indigenous language 4} ---	<input type="checkbox"/> 9		
	{most widely spoken non-indigenous language 5} ---	<input type="checkbox"/> 10		
	Other European language(s) -----	<input type="checkbox"/> 11		
	Other non-European language(s) -----	<input type="checkbox"/> 12		

### WebTrans note

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ4, SQ25, SQ26 and SQ27 are identical.

Please note that the number of response categories can differ between countries (but the maximum is 13 - the number of response categories displayed).

### NRC note

SQ4, SQ25, SQ26, SQ27 and SQ28 provide information about the languages in the home environment. This question provides information on the languages the student uses in his or her home environment. In mixed language families or immigrant families the languages spoken by the family (SQ25) and the languages spoken by the student him or herself (SQ26) can differ.

The response categories of this question need to be localised (see SQ4). The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.





A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

Languages that the family only speaks once in a while or only in special circumstances should not be ticked (see Admin Note). If necessary the sentence "Languages that your family only speaks once in a while or only in special circumstances should not be ticked" can be added to the instruction (similar to TQ9).

### Administrator note

Students should only tick the languages that the student speaks regularly in his/her private life. Languages that the student only speaks once in a while or only in special circumstances should not be ticked.

27	Which language do you speak <u>most often</u> at home?			
	(Please select only one answer)			
	{most widely spoken indigenous language1} -----	<input type="radio"/> 0		
	{most widely spoken indigenous language2} -----	<input type="radio"/> 1		
	{most widely spoken indigenous language3} -----	<input type="radio"/> 2		
	{most widely spoken indigenous language4} -----	<input type="radio"/> 3		
	{most widely spoken indigenous language5} -----	<input type="radio"/> 4		
	[target language] -----	<input type="radio"/> 5		
	{most widely spoken non-indigenous language 1} ---	<input type="radio"/> 6		
	{most widely spoken non-indigenous language 2} ---	<input type="radio"/> 7		
	{most widely spoken non-indigenous language 3} ---	<input type="radio"/> 8		
	{most widely spoken non-indigenous language 4} ---	<input type="radio"/> 9		
	{most widely spoken non-indigenous language 5} ---	<input type="radio"/> 10		
	Other European language-----	<input type="radio"/> 11		
	Other non-European language-----	<input type="radio"/> 12		

### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The response categories of SQ4, SQ25, SQ26 and SQ27 are identical.



Please note that the number of response categories can differ between countries (but the maximum is 13 - the number of response categories displayed).

### NRC note

SQ4, SQ25, SQ26, SQ27 and SQ28 provide information about the languages in the home environment. This question (SQ27) is similar to the question in PISA about language use. This question allows us to compare our findings with the results reported in PISA and in the Eurydice Key data report on teaching languages at school (2008).

The response categories of this question need to be localised (see SQ4). The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.

A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets . The reconciler and/or NRC is asked to add the appropriate response categories.

most often - Please underline the appropriate text in the translation.

[Target language] –Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

Please note that students should tick only one box. Students should only tick the one language that the student speaks most in his/her private life.

## 28 In your opinion, how well do your parents know [target language]?

*(Please select one answer from each row)*

	<i>Not at all</i>	<i>A little</i>	<i>Quite well</i>	<i>Very well</i>
1) How well does your father know [target language]? –	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) How well does your mother know [target language]? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

## NRC note

SQ4, SQ25, SQ26, SQ27 and SQ28 provide information about the languages in the home environment. This question (SQ28) provides information about the students' (subjective) perception of his/her parents [target language] skills. The perceived parental skills might impact the students' perception of his/her own [target language] skills, his/her expectations regarding his/her about [target language] learning and his/her [target language] skills.

## Translator note

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

29 Do you, yourself, come into contact with [target language] outside school in the following ways?

(Please select in each row No or Yes)

	No	Yes
1) Through a friend who writes to you in [target language], for example through email, MSN, letters? -----	<input type="radio"/> 0	<input type="radio"/> 1
2) Through relatives living in a [target language] speaking country? -----	<input type="radio"/> 0	<input type="radio"/> 1
3) Through friends living in a [target language] speaking country? -----	<input type="radio"/> 0	<input type="radio"/> 1
4) Through [target language] speaking tourists who visit the place where you live? -----	<input type="radio"/> 0	<input type="radio"/> 1
5) Through [target language] speaking people who live in your place of residence? -----	<input type="radio"/> 0	<input type="radio"/> 1
6) Through people on the internet who talk to you in [target language], for example when playing online games? -----	<input type="radio"/> 0	<input type="radio"/> 1
7) Through [target language] speaking people you meet during holidays? -----	<input type="radio"/> 0	<input type="radio"/> 1

## WebTrans note

The same response instruction "(Please select in each row No or Yes)" occurs in SQ19, SQ20, SQ23, SQ29, SQ64.

The response scales (No-Yes) of SQ19, SQ20, SQ23, SQ29, and SQ64 are identical.

## NRC note

SQ29 and SQ30 assess the informal language learning opportunities in students' living environment. SQ29 assesses the exposure to [target language] and Q30 the use of [target language].

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.



In Educational systems that are not an entire country an adequate adaptation of items 2) and 3) that can be understood by students might be required. Please instruct the reconciler how to adapt the question.

### Translator note

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

yourself - Please underline the appropriate text in the translation.

## 30 How often do you use [target language] outside school in the following ways?

(Please select one answer from each row)

	Never	A few times a year	About once every month	A few times a month	A few times a week
1) How often do you write [target language] with friends (for example MSN, email, letters)? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) How often do you speak [target language] with relatives? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) How often do you speak [target language] with friends? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) How often do you speak [target language] with people living in your place of residence? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) How often do you speak [target language] with tourists? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) How often do you use [target language] with people on the internet, for example when playing online games? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-A few times a year-About once every month-A few times a month-A few times a week) of SQ30 and SQ31 are identical.



### **NRC note**

SQ29 and SQ30 assess the informal language learning opportunities in students' living environment. SQ29 assesses the exposure to [target language] and SQ30 the use of [target language].

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

### **Translator note**

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

use - Please underline the appropriate text in the translation.

31 How often do you come into contact with [target language] through media in the following ways?

(Please select one answer from each row)

	Never	A few times a year	About once every month	A few times a month	A few times a week
1) How often do you listen to songs in [target language]?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) How often do you watch movies spoken in [target language] <u>without</u> subtitles?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) How often do you watch movies spoken in [target language] <u>with</u> subtitles?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) How often do you watch television programmes (not movies) spoken in [target language] <u>without</u> subtitles?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) How often do you watch television programmes (not movies) spoken in [target language] <u>with</u> subtitles?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) How often do you play computer games spoken in [target language]?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7) How often do you read books written in [target language]?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8) How often do you read a magazine or a comic written in [target language]?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9) How often do you visit websites written in [target language]?-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

**WebTrans note**

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-A few times a year-About once every month-A few times a month-A few times a week) of SQ30 and SQ31 are identical.

**NRC note**

This question assesses the informal language learning opportunities through media.

**Translator note**

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



## Your opinion about foreign languages

32 In your opinion, how good are the following people at learning foreign languages?

*(Please select one answer from each row)*

	Very bad	Quite bad	Quite good	Very good
1) In general, how good are people from [Educational system] at learning foreign languages? .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) How good is your father at learning foreign languages? .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) How good is your mother at learning foreign languages? .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) How good are you, yourself, at learning foreign languages? .....	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

### NRC note

SQ32, SQ33, SQ56, SQ34 and SQ35 provide information about the students’ (subjective) perception of foreign language(s) and foreign language learning. This question provides information on the (subjective) view in the country on people’s capability of learning foreign languages.

### Translator note.

[Educational system] - is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.



### 33 In your opinion, how useful is [target language] for the following purposes?

(Please select one answer from each row)

	Not useful at all	Hardly useful	Quite useful	Very useful
1) For travelling -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) For your personal life -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) For your further education -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) For your future work -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) For getting a good job -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) For contact with foreigners -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) For your personal satisfaction -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8) For the use of computers and other technical devices -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9) For reading books, magazines, etc.	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
10) For entertainment (movies, television programmes, music, games)	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Not useful at all-Hardly useful-Quite useful-Very useful) of SQ33, SQ35, and SQ52 are identical.

#### NRC note

SQ32, SQ33, SQ56, SQ34 and SQ35 provide information about the students’ (subjective) perception of foreign language(s) and foreign language learning. This question provides information on the usefulness of [target language].

#### Translator note.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About your school subjects

### 34 How much do you like the following school subjects?

*(Please select one answer from each row)*

	<i>Do not like at all</i>	<i>Hardly like</i>	<i>Quite like</i>	<i>Like a lot</i>
1) {Mathematics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) {Science subjects, e.g. physics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) {Human and society subjects, e.g. history}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) {Culture and arts subjects, e.g. music, art history}----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) [Questionnaire language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Other foreign languages (including Latin and ancient Greek)-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8) {Vocational skills subjects}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9) {Sports}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The items of Q34, Q35 and Q36 are identical.

#### NRC note

SQ32, SQ33, SQ56, SQ34 and SQ35 provide information about the students’ (subjective) perception of foreign language(s) and foreign language learning. This question allows us to see how students’ perception of language learning compares to students’ perception of other school subjects.

The items of this question might need to be localised. It is necessary to phrase the groups of school subjects and examples in a way that students will easily understand. Please instruct the reconciler which items should be used.

Please make sure that the items include all or most of the subjects which are part of the official curriculum (at ISCED2 and ISCED3 level).

#### Translator note

Please do not translate the items in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate items.

[Questionnaire language] - Please substitute this term with the name of the language into which the questionnaire is being translated.



[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

Encourage students to express their opinion. If a student does not have an opinion because the student has never had a particular subject, no answer has to be given for that subject.

## 35 In your opinion, how useful are the following school subjects?

(Please select one answer from each row)

	Not useful at all	Hardly useful	Quite useful	Very useful
1) {Mathematics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) {Science subjects, e.g. physics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) {Human and society subjects, e.g. history}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) {Culture and arts subjects, e.g. music, art history}----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) [Questionnaire language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Other foreign languages (including Latin and ancient Greek)-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8) {Vocational skills subjects}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9) {Sports}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The items of Q34, Q35 and Q36 are identical.

The response scales (Not useful at all-Hardly useful-Quite useful-Very useful) of SQ33, SQ35, and SQ52 are identical.

### NRC note (identical to note with Q34)

SQ32, SQ33, SQ56, SQ34 and SQ35 provide information about the students’ (subjective) perception of foreign language(s) and foreign language learning. This question allows us to see how students’ perception of language learning compares to students’ perception of other school subjects.



The items of this question might need to be localised. It is necessary to phrase the groups of school subjects and examples in a way that students will easily understand. Please instruct the reconciler which items should be used.

Please make sure that the items include all or most of the subjects which are part of the official curriculum (at ISCED2 and ISCED3 level).

**Translator note**

Please do not translate the items in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate items.

[Questionnaire language] - Please substitute this term with the name of the language into which the questionnaire is being translated.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

**Administrator note**

Encourage students to express their opinion. If a student does not have an opinion because the student has never had a particular subject, no answer has to be given for that subject.

### 36 How much time do you spend on homework and study for the following school subjects in a normal school week?

(Please select one answer from each row)

	No time at all	Relatively little time	Quite some time	A lot of time
1) {Mathematics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) {Science subjects, e.g. physics}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) {Human and society subjects, e.g. history}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) {Culture and arts subjects, e.g. music, art history}---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) [Questionnaire language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Other foreign languages (including Latin and ancient Greek)-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8) {Vocational skills subjects}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9) {Sports}-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The items of Q34, Q35 and Q36 are identical.

#### NRC note

This question provides information on the effort students make for foreign language learning and other school subjects. This question allows us to see how, in the perception of the student, the amount of time spent on homework for foreign languages (also assessed in Q63) compares to the amount of time spent on homework for other subjects.

The items of this question might need to be localised (identical to SQ34 and SQ35). It is necessary to phrase the groups of school subjects and examples in a way that students will easily understand. Please instruct the reconciler which items should be used.

Please make sure that the items include all or most of the subjects which are part of the official curriculum (at ISCED2 and ISCED3 level).

#### Translator note

Please do not translate the items in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate items.

[Questionnaire language] - Please substitute this term with the name of the language into which the questionnaire is being translated.



[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

**Administrator note**

If a student does not have a particular subject mentioned in the question, for that subject no answer has to be given.

## About learning foreign languages in school

37	Which of the following foreign languages do you have or did you have as a subject in {primary or secondary school}?			
	<i>(Please tick as many boxes as applicable)</i>			
	{1st most widely taught foreign language in the country}-----	<input type="checkbox"/> 0		
	{2nd most widely taught foreign language in the country}-----	<input type="checkbox"/> 1		
	{3rd most widely taught foreign language in the country}-----	<input type="checkbox"/> 2		
	{4th most widely taught foreign language in the country}-----	<input type="checkbox"/> 3		
	{5th most widely taught foreign language in the country}-----	<input type="checkbox"/> 4		
	{6th most widely taught foreign language in the country}-----	<input type="checkbox"/> 5		
	{7th most widely taught foreign language in the country}-----	<input type="checkbox"/> 6		
	{8th most widely taught foreign language in the country}-----	<input type="checkbox"/> 7		
	{9th most widely taught foreign language in the country}-----	<input type="checkbox"/> 8		
	{10th most widely taught foreign language in the country}-----	<input type="checkbox"/> 9		
	Other foreign language(s)-----	<input type="checkbox"/> 10		

### WebTrans note

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ37 and SQ38 are identical.

Please note that the number of response categories can differ between countries, but the maximum is 11 - including “Other foreign language(s)”.

### NRC note

This question provides information about the foreign languages the student has learned.



The response categories of this question need to be localised. The languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table). The languages that are most widely taught in secondary education level should be included (including ancient languages). A maximum number of 10 foreign languages can be included. If less than 10 foreign languages can be studied in secondary education fewer response categories can be used (see the Instruction for the Main Study Localisation file).

Please instruct the reconciler which languages should be used as response categories.

**Translator note**

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.



38	Which of the following foreign languages was the <u>first</u> foreign language that you were taught in school?				
	<i>(Please select only one answer)</i>				
	{1st most widely taught foreign language in the country}-----	<input type="radio"/> 0			
	{2nd most widely taught foreign language in the country}-----	<input type="radio"/> 1			
	{3rd most widely taught foreign language in the country}-----	<input type="radio"/> 2			
	{4th most widely taught foreign language in the country}-----	<input type="radio"/> 3			
	{5th most widely taught foreign language in the country}-----	<input type="radio"/> 4			
	{6th most widely taught foreign language in the country}-----	<input type="radio"/> 5			
	{7th most widely taught foreign language in the country}-----	<input type="radio"/> 6			
	{8th most widely taught foreign language in the country}-----	<input type="radio"/> 7			
	{9th most widely taught foreign language in the country}-----	<input type="radio"/> 8			
	{10th most widely taught foreign language in the country}-----	<input type="radio"/> 9			
	Other foreign language -----	<input type="radio"/> 10			

### WebTrans note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

The response categories of SQ37 and SQ38 are identical.

Please note that the number of response categories can differ between countries, but the maximum is 11 - including “Other foreign language(s)”.

### NRC note

This question provides information about the foreign languages the student has learned

The response categories of this question need to be localised (identical to Q37). The languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table). The languages that are most widely taught in secondary education level should be included (including ancient languages). A maximum number of 10 foreign languages can be included. If less than 10 foreign languages can be studied in secondary education fewer response categories can be used (see the Instruction for the Main Study Localisation file).



Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

first - Please underline the appropriate text in the translation.

---

## 39 In which [grades] did you take foreign language lessons in school?

*(Please tick as many boxes as applicable)*

- {2nd grade of ISCED3} ..... ☐ 0
- {1st grade of ISCED3} ..... ☐ 1
- {6th grade of ISCED2} ..... ☐ 2
- {5th grade of ISCED2} ..... ☐ 3
- {4th grade of ISCED2} ..... ☐ 4
- {3rd grade of ISCED2} ..... ☐ 5
- {2nd grade of ISCED2} ..... ☐ 6
- {1st grade of ISCED2} ..... ☐ 7
- {6th grade of ISCED1} ..... ☐ 8
- {5th grade of ISCED1} ..... ☐ 9
- {4th grade of ISCED1} ..... ☐ 10
- {3rd grade of ISCED1} ..... ☐ 11
- {2nd grade of ISCED1} ..... ☐ 12
- {starting grade of ISCED1} ..... ☐ 13
- Before {the starting grade of ISCED1} ..... ☐ 14

---

### WebTrans note

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ39 and SQ40 are identical

Please note that the number of response categories can differ between countries (but the maximum is 15).



### **NRC note**

This question assesses the onset and duration of foreign language learning.

The response categories of this question need to be localised. All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3 should be presented as response categories. It is necessary to phrase the response categories in such a way that students will easily understand.

The number of response categories equals the total duration of primary education (ISCED1), lower secondary education (ISCED2) and the first two years of upper secondary education (ISCED3) plus one for the pre-ISCED1 period. For example, if in the Educational system both ISCED1 and ISCED2 last four years, then the four grades of ISCED1, the four grades of ISCED2 and two grades of ISCED3 should be presented as response categories (in total 11 response categories) .

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

[Grade] - refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

40 In which [grades] did you take [target language] lessons in school?

(Please tick as many boxes as applicable)

- {2nd grade of ISCED3} ..... ☐ 0
- {1st grade of ISCED3} ..... ☐ 1
- {6th grade of ISCED2} ..... ☐ 2
- {5th grade of ISCED2} ..... ☐ 3
- {4th grade of ISCED2} ..... ☐ 4
- {3rd grade of ISCED2} ..... ☐ 5
- {2nd grade of ISCED2} ..... ☐ 6
- {1st grade of ISCED2} ..... ☐ 7
- {6th grade of ISCED1} ..... ☐ 8
- {5th grade of ISCED1} ..... ☐ 9
- {4th grade of ISCED1} ..... ☐ 10
- {3rd grade of ISCED1} ..... ☐ 11
- {2nd grade of ISCED1} ..... ☐ 12
- {starting grade of ISCED1} ..... ☐ 13
- Before {the starting grade of ISCED1} ..... ☐ 14

**WebTrans note**

The same response instruction “(Please tick as many boxes as applicable)” occurs in SQ4, SQ25, SQ26, SQ37, SQ39, SQ40.

The response categories of SQ39 and SQ40 are identical

Please note that the number of response categories can differ between countries (but the maximum is 15).

**NRC note**

This question assesses the onset and duration of [target language] learning.

The response categories of this question need to be localised (identical to SQ39). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3 should be presented as response categories. It is necessary to phrase the response categories in such a way that students will easily understand.



The number of response categories equals the total duration of primary education (ISCED1), lower secondary education (ISCED2) and the first two years of upper secondary education (ISCED3) plus one for the pre-ISCED1 period. For example, if in the Educational system both ISCED1 and ISCED2 last four years, then the four grades of ISCED1, the four grades of ISCED2 and two grades of ISCED3 should be presented as response categories (in total 11 response categories).

Please instruct the reconciler which response categories to use.

### Translator note

Please do not translate the response categories in curly brackets. The reconciler and/or National Research Coordinator is asked to add the appropriate response categories.

[grades] - refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[target language] lessons - Please underline the appropriate text in the translation.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

## 41 How many foreign languages did you study in school before you started studying [target language]?

*(Please select only one answer)*

No foreign languages ----- ☐ 0

One foreign language ----- ☐ 1

Two foreign languages ----- ☐ 2

Three or more foreign languages ----- ☐ 3

---

### WebTrans Note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

### NRC note

This question (in combination with SQ38) provides information about the learning order of foreign languages.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

before - Please underline the appropriate text in the translation.

---

42 On average, how many students are there in your classroom during the [target language] lessons?

*Students*

---

**NRC note**

This question provides information about class size. This information is crucial for calculating the teaching time a week.

**Translator note**

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

43 How long does a class period last at your school?

*Minutes*

---

**NRC note**

SQ43 and SQ44 allow us to calculate the foreign language learning time a week (lessons).

---

44 How many class periods do you have for the following subjects in a normal full week at school?

*Number of  
class periods*

1) For [target language] -----

2) For all foreign languages together (including Latin and ancient Greek)-----

3) For all subjects together (in total)-----

---

**NRC note**

SQ43 and SQ44 allow us to calculate the foreign language learning time a week (lessons).

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

all; all subjects - Please underline the appropriate text in the translation.

## 45 How often have you travelled abroad or had visitors from abroad in the past three years?

(Please select one answer from each row)

	Never	Once	Twice	Three times or more
1) How often did you go on a school trip to a [target language] speaking country? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) How often did you go on a school trip to another (non-[target language] speaking) country? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) How often did you go <u>with your family</u> to a [target language] speaking country? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) How often did you go <u>with your family</u> to a (non-[target language] speaking) country? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) How often did a school class from a [target language] speaking country visit your school? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) How often did a school class from another (non-[target language]) speaking country visit your school? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-Once-Twice-Three times or more) of SQ45 and SQ46 are identical.

### NRC note

SQ45 and SQ46 provide information about the received opportunities for exchange visits and school language projects. SQ45-3 provides information on informal language learning opportunities through visits abroad.

In Educational systems that are not an entire country an adequate adaptation of items 2) and 3) that can be understood by students might be required. Please instruct the reconciler how to adapt the question.



### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

with your family - Please underline the appropriate text in the translation.

## 46 In the past three years, how often have you participated in the following activities for foreign languages at school?

(Please select one answer from each row)

	Never	Once	Twice	Three times or more
1) Collaboration project with schools abroad-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) Language clubs-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) Language competition -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) European Day of Languages -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) Extracurricular language projects -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) Writing through mail, email or MSN with students from abroad-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Excursions and field trips related to foreign language education-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-Once-Twice-Three times or more) of SQ45 and SQ46 are identical.

### NRC note

SQ45 and SQ46 provide information about the received opportunities for exchange visits and school language projects.

Please provide the test administrator with a description and several examples of language clubs and language competitions.

Language clubs - A group of students that meets regularly and organized for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition - an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.





European Day of Languages - The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

### **Translator note**

Language clubs - A group of students that meets regularly and organized for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition - an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

European Day of Languages - The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

MSN - is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

### **Administrator note**

If a student does not know an activity mentioned in the question, please describe the activity and give examples.

Language clubs - A group of students that meets regularly and organized for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition - an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

European Day of Languages - The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

## About your [target language] lessons

### Translator note

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

### 47 Why are you learning [target language]?

*(Please select only one answer)*

Because the subject of [target language] is compulsory ----- ☐ 0

Because studying a foreign language is compulsory and I chose [target language] -- ☐ 1

Because I chose [target language] as an optional subject ----- ☐ 2

---

### WebTrans Note

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

### NRC note

This question assesses whether the target language is compulsory or not for the individual student.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## 48 How difficult is it for you to learn the following?

(Please select one answer from each row)

	Very easy	Quite easy	Quite difficult	Very difficult
1) Learning to write in [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) Learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) Learning to understand spoken [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) Learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) Learning to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) Learning to pronounce [target language] correctly ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Learning [target language] words -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The items of Q48 and Q59 are identical

The response scales (Very easy-Quite easy-Quite difficult-Very difficult) of SQ48 and SQ56 are identical.

### NRC note

SQ48, SQ52, SQ58 and SQ61 provide information on the perceived emphasis on the four communicative skills and language content within the teaching activities and resources used.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

49 How often does your teacher of [target language] speak [target language] when doing the following?

(Please select one answer from each row)

	Never	Hardly ever	Every now and then	Usually	Always
1) When he or she speaks to the whole class -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) When he or she talks to one or two students -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-Hardly ever-Every now and then-Usually-Always) of SQ49 and SQ50 are identical.

### NRC note

SQ49 and SQ50 provide information on the use of the target language during foreign language lessons.

### Translator note

Teacher of [target language] - The teacher who teaches the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching other subjects.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer ‘teacher’ questions for the chosen teacher.

50 How often do students speak [target language] when doing the following in a [target language] lesson?

(Please select one answer from each row)

	Never	Hardly ever	Every now and then	Usually	Always
1) When students speak to the teacher of [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) When students work in groups and speak together -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) When students speak in front of the whole class -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

**WebTrans note**

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never-Hardly ever-Every now and then-Usually-Always) of SQ49 and SQ50 are identical.

**NRC note**

SQ49 and SQ50 provide information on the use of the target language during foreign language lessons.

**Translator note**

Teacher of [target language] - The teacher who teaches about the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching about other subjects.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

**Administrator note**

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer ‘teacher’ questions for the chosen teacher.

## 51 How often are the following resources used in your [target language] lessons?

*(Please select one answer from each row)*

	<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1) Audio-cassettes, CDs or other audio-material in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) Video cassettes, DVDs, video clips from YouTube or other audio-visual material---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) Newspapers, magazines, comics or song texts written in [target language] ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) Internet-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) Computer programmes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) Language laboratory (student PCs with specific language software)-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7) Textbook for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8) Books written in [target language] for extensive reading e.g. fiction -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9) Lesson materials prepared by your teacher of [target language] (e.g. hand-outs, reading texts) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) of SQ51, SQ58, and SQ59 are identical.

### NRC note

SQ24, SQ51 and SQ62 provide information on the use of ICT for foreign language learning. This question provides information on the use of ICT and other resources during the [target language] lessons.

### Translator note

Teacher of [target language] - The teacher who teaches about the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching about other subjects.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer ‘teacher’ questions for the chosen teacher.

## 52 How useful are your [target language] textbooks, or is your [target language] textbook, for the following?

(Please select one answer from each row)

	Not useful at all	Hardly useful	Quite useful	Very useful
1) For learning to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) For learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) For learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) For learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) For learning to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) For learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) For learning [target language] words -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Not useful at all-Hardly useful-Quite useful-Very useful) of SQ33, SQ35, and SQ52 are identical.

### NRC note

SQ48, SQ52, SQ58 and SQ61 provide information on the perceived emphasis on the four communicative skills and language content within the teaching activities and resources used.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

If students do not have a textbook for [target language], but use a reader or handouts made by their teacher of [target language], then they can answer how useful this reader or handouts are for learning [target language].

## 53 How often does the following happen during your [target language] lessons?

(Please select one answer from each row)

	Never	Hardly ever	Every now and then	Usually	Most of the time
1) The students work in groups.-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) The students work individually. -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) A group of students speaks in front of the whole class.-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) An individual student speaks in front of the whole class.-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) The teacher speaks to the whole class. --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) The teacher speaks with one or two students -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

### NRC note

This question provides information about the within class ability grouping (setting).

### Translator note

Your - (in “your [target language] lessons”) should be plural.

Teacher of [target language] - The teacher who teaches about the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching about other subjects.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



### Administrator note

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer 'teacher' questions for the chosen teacher.

## 54 To what extent do you agree or disagree with the following statements about your teacher of [target language]?

(Please select one answer from each row)

	Disagree	Slightly disagree	Slightly agree	Agree
1) My teacher of [target language] is a good teacher ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) I get along with my teacher of [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) My teacher of [target language] makes an effort to make the lessons interesting for us -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) My teacher of [target language] is helpful -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) I like my teacher of [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) My teacher of [target language] is strict -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction "(Please select one answer from each row)" occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Disagree-Slightly disagree-Slightly agree-Agree) of SQ54, and SQ55 are identical.

### NRC note

Q54 and Q55 provide information about the students' (subjective) perception of [target language] lessons and the [target language] teacher(s).

### Translator note

your teacher of [target language]- Please underline the appropriate text in the translation.

Teacher of [target language] - The teacher who teaches about the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching about other subjects.

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer 'teacher' questions for the chosen teacher.

55 To what extent do you agree or disagree with the following statements about your [target language] lessons?

(Please select one answer from each row)

	Disagree	Slightly disagree	Slightly agree	Agree
1) My [target language] lessons are interesting -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) My [target language] lessons are enjoyable -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) My [target language] lessons are good -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) My [target language] lessons are waste of time -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) My [target language] lessons are easy-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) My [target language] lessons are boring-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Disagree-Slightly disagree-Slightly agree-Agree) of SQ54, and SQ55 are identical.

#### NRC note

Q54 and Q55 provide information about the students’ (subjective) perception of [target language] lessons and the [target language] teacher(s).

#### Translator note

[target language] lessons - Please underline the appropriate text in the translation.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

56 In your opinion, how difficult is it for the following people to learn [target language]?

*(Please select one answer from each row)*

	Very easy	Quite easy	Quite difficult	Very difficult
1) For people who usually speak [questionnaire language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) For the students in your [target language] class -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Very easy-Quite easy-Quite difficult-Very difficult) of SQ48 and SQ56 are identical.

### NRC note

SQ32, SQ33, SQ56, SQ34 and SQ35 provide information about the students’ (subjective) perception of foreign language(s) and foreign language learning. This question provides information on the existing view in the country on people’s capability of learning foreign languages.

This question allows us also to see whether students have the opinion that learning [target language] is more difficult for themselves (see SQ48) than for others.

### Translator note

[Questionnaire language] - Please substitute this term with the name of the language into which the questionnaire is being translated.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

57 How often does your teacher of [target language] point out similarities between [target language] and other languages when teaching the following?

(Please select one answer from each row)

	Never	Sometimes	Quite often	Very often
1) Teaching to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) Teaching to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) Teaching to understand spoken [target language] --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) Teaching [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) Teaching to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) Teaching to pronounce [target language] correctly -	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Teaching [target language] words -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

### NRC note

This question provides information on the perceived emphasis on similarities between languages.

### Translator note

Teacher of [target language] - The teacher who teaches about the subject of [target language]. It should not refer to teachers from a [target language] speaking country or to teachers who speak [target language] when teaching about other subjects.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### Administrator note

If a student has more than one teacher for the subject of [target language] then the questions refer to the teacher that teaches the student most often. If a student says the teachers teach equally often then he or she may choose either teacher, but consistently answer ‘teacher’ questions for the chosen teacher.

## 58 How often do you do the following during [target language] lessons?

*(Please select one answer from each row)*

	<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1) Learning to write in [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) Learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) Learning to understand spoken [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) Learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) Learning to read [target language] texts --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) Learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7) Learning [target language] words -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) of SQ51, SQ58, and SQ59 are identical.

The items of Q48 and Q59 are identical

### NRC note

SQ48, SQ52, SQ58 and SQ61 provide information on the perceived emphasis on the four communicative skills and language content within the teaching activities and resources used.

### Translator note

You - should be plural.

[target language] lesson - fixed period of time at school when students are taught the subject of [target language].

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About tests and assignments for the subject of [target language]

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### 59 How often does your teacher of [target language] do the following?

*(Please select one answer from each row)*

	<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1) Give a [target language] test or assignment that is marked or scored -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) Provide comments on a test or assignment you made-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) of SQ51, SQ58, and SQ59 are identical.

### NRC note.

Q59, Q60 and Q63 provide information on the time spent on homework and study for the subject of [target language].

### Translator note

You - should be plural.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

60 How much time do you usually study for a [target language] test?

*(Please select only one answer)*

No time at all ----- ☐ 0

Less than one hour ----- ☐ 1

About one to two hours ----- ☐ 2

About two to three hours ----- ☐ 3

More than three hours ----- ☐ 4

---

**WebTrans Note**

The same response instruction “(Please select only one answer)” occurs in SQ3, SQ6, SQ9, SQ12, SQ13, SQ14, SQ15, SQ16, SQ17, SQ18, SQ21, SQ27, SQ38, SQ41, SQ47, SQ60.

**NRC note.**

Q59, Q60 and Q63 provide information on the time spent on homework and study for the subject of [target language].

**Translator note**

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

61 How important are the following in order to get a good final grade for the subject of [target language]?

(Please select one answer from each row)

	<i>Not important at all</i>	<i>Hardly important</i>	<i>Quite important</i>	<i>Very important</i>
1) Writing [target language] well-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2) Speaking [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3) Understanding spoken [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4) Knowing [target language] grammar well-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5) Reading [target language] well-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6) Pronouncing [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7) Knowing many words in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans Note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

#### NRC note

SQ48, SQ52, SQ58 and SQ61 provide information on the perceived emphasis on the four communicative skills and language content within the teaching activities and resources used.

#### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



## About studying and doing homework for foreign languages out of school time

62 When studying and doing homework for [target language], how often do you use a computer for the following?

*(Please select one answer from each row)*

	<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>A few times a week</i>	<i>(Almost) every day</i>
1) For finding information for [target language] homework or assignments -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) For [target language] homework or assignments -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) For learning to write in [target language] -	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) For learning to speak [target language]---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) For learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6) For learning [target language] grammar --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7) For learning to read [target language]texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8) For learning to pronounce [target language] correctly-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9) For learning [target language] words -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

The response scales (Never or hardly ever-A few times a year-A few times a month-A few times a week-(Almost) every day) of SQ24 and SQ62 are identical.

### NRC note

SQ24, SQ51 and SQ62 provide information on the use of ICT for foreign language learning. This question provides information the perceived emphasis on the four communicative skills and language content when using ICT.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

63 Generally, how much time do you spend each week on homework and assignments for the following subjects?

(Please select one answer from each row)

	Zero hours	Less than one hour a week	About one to two hours a week	About two to three hours a week	More than three hours a week
1) For [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) For other foreign languages (including Latin and ancient Greek) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction “(Please select one answer from each row)” occurs in SQ22, SQ24, SQ28, SQ30, SQ31, SQ32, SQ33, SQ34, SQ35, SQ36, SQ45, SQ46, SQ48, SQ49, SQ50, SQ51, SQ52, SQ53, SQ54, SQ55, SQ56, SQ57, SQ58, SQ59, SQ61, SQ62, SQ63.

### NRC note.

Q59, Q60 and Q63 provide information on the time spent on homework and study for the subject of [target language].

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

### Translator note

each week - Please underline the appropriate text in the translation.

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## 64 What type of extra lessons have you attended or are you attending?

(Please select in each row No or Yes)

	No	Yes
1) {Enrichment lessons} for [target language]-----	<input type="radio"/> 0	<input type="radio"/> 1
2) {Enrichment lessons} for other foreign languages (including for Latin and ancient Greek) -----	<input type="radio"/> 0	<input type="radio"/> 1
3) {Remedial lessons} for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
4) {Remedial lessons} for other foreign languages (including for Latin and ancient Greek) -----	<input type="radio"/> 0	<input type="radio"/> 1
5) Extra lessons for [questionnaire language] -----	<input type="radio"/> 0	<input type="radio"/> 1
6) Extra lessons in another language than [questionnaire language] that is spoken regularly at your home -----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The same response instruction “(Please select in each row No or Yes)” occurs in SQ19, SQ20, SQ23, SQ29, SQ64.

The response scales (No-Yes) of SQ19, SQ20, SQ23, SQ29, and SQ64 are identical.

### NRC note

This question provides information on students’ participation in out-of-school-time lessons and is similar to the PISA question about out-of-school-time lessons. Schools with a foreign language specialisation usually offer more opportunities for out-of-school-time lessons, than schools without such a specialisation.

Item 5) “Extra lessons for [questionnaire language]” allows us to assess whether immigrant students received help in mastering the host language.

Item 6) “Extra lessons in another language than [questionnaire language] that is spoken regularly at your home” allows us to assess whether immigrant students received formal education in language(s) of origin.

The items of this question might need to be localised. It is necessary to use terms that students will easily understand. Please instruct the reconciler which terms should be used.

extra lessons — Any lessons in school subjects that the students is learning at school, that he or she spends extra time learning outside of normal school hours. The lessons might be held at school, at home, or elsewhere.

{Remedial lessons} — Any lessons in addition to regular lessons designed to help students with learning difficulties.

{Enrichment lessons} — Any lessons in addition to regular lessons designed to extend abilities of more able students.

### Translator note

Please do not translate the terms in curly brackets. The reconciler and/or NRC is asked to add the appropriate terms.

home - refers to that place where the student usually resides, not including boarding school. The term used should connote a family or domestic setting.



[Questionnaire language] - Please substitute this term with the name of the language into which the questionnaire is being translated.

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

Extra - Please underline the appropriate text in the translation.

## Your skills in [target language]

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

*In this section please say how you judge your own skills in [target language].*

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### A1 Can you do the following when reading [target language]?

*(Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described.)*

	No, not yet	Yes
1) I can understand familiar words and very simple sentences, for example on posters. ....	<input type="radio"/> 0	<input type="radio"/> 1
2) I can find information in for example advertisements, timetables. ....	<input type="radio"/> 0	<input type="radio"/> 1
3) I can find the main points in simple newspaper articles on familiar subjects. ....	<input type="radio"/> 0	<input type="radio"/> 1
4) I can read quickly through long and complex texts, locating the details I need to find. ....	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The response instruction (Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described ) is identical for the Questions A1, A2, A3 and A4.

The response scales (No, not yet-Yes) of Questions A1, A2, A3 and A4 are identical.

Notice that the 'A' preceding the response number is required.

### Translator note

[Target language] - Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## A2 Can you do the following when listening to [target language]?

(Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described.)

	No, not yet	Yes
1) I can understand questions and instructions if people speak carefully and slowly. ....	<input type="radio"/> 0	<input type="radio"/> 1
2) I can understand what is said to me, if the speaker can take the trouble to speak clearly, slowly and directly. ....	<input type="radio"/> 0	<input type="radio"/> 1
3) I can understand the main points in radio or TV programmes about familiar subjects if the speakers do not speak too fast. ....	<input type="radio"/> 0	<input type="radio"/> 1
4) I can understand long complicated speeches and lectures if the topic is reasonably familiar. ....	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans note

The response instruction (Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described ) is identical for the Questions A1, A2, A3 and A4.

The response scales (No, not yet-Yes) of Questions A1, A2, A3 and A4 are identical.

Notice that the 'A' preceding the response number is required.

### Translator note

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

### A3 Can you do the following when writing [target language]?

(Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described.)

	No, not yet	Yes
1) I can write a few words and phrases that relate to myself, my family, where I live, my school. ....	<input type="radio"/> 0	<input type="radio"/> 1
2) I can write very short, basic descriptions of events and personal experiences. ....	<input type="radio"/> 0	<input type="radio"/> 1
3) I can write detailed letters about my experiences, feelings and about events. ....	<input type="radio"/> 0	<input type="radio"/> 1
4) I can write a detailed review of a film, book or play. ....	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans note

The response instruction (Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described ) is identical for the Questions A1, A2, A3 and A4.

The response scales (No, not yet-Yes) of Questions A1, A2, A3 and A4 are identical.

Notice that the 'A' preceding the response number is required.

#### Translator note

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

#### A4 Can you do the following when speaking [target language]?

(Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described.)

	No, not yet	Yes
1) I can ask and answer simple questions on very familiar topics. -----	<input type="radio"/> 0	<input type="radio"/> 1
2) I can tell a story or describe something in short simple sentences. -----	<input type="radio"/> 0	<input type="radio"/> 1
3) I can maintain a conversation about familiar topics and express personal opinions. -----	<input type="radio"/> 0	<input type="radio"/> 1
4) I can explain my viewpoint on a current issue, for example by giving advantages and disadvantages of various options. -----	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans note

The response instruction (Select one response for each statement: YES if the statement describes something you feel you can already do. NO if you CAN'T yet do what is described ) is identical for the Questions A1, A2, A3 and A4.

The response scales (No, not yet-Yes) of Questions A1, A2, A3 and A4 are identical.

Notice that the 'A' preceding the response number is required.

#### Translator note

For item A4, 4): 'current' means: that which currently has the attention in the media and is widely discussed.

[Target language] - Please replace the term "target language" according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## Thank you!

#### WebTrans note

This should appear in the WebTrans version to be translated.





Teacher Questionnaire  
for the European Survey on Language  
Competences 2011  
Main Study





## About this full note version

Accompanying each question are:

- **Notes for WebTrans** indicating the recurring question elements that have to be linked in WebTrans, so they have to be translated only once.
- **Notes for the NRC** clarifying terms and options, noting where localisations should be made, and providing a rationale for the question's inclusion.
- **Notes for the translator** clarifying terms and options, noting where response categories and/or terms should not be translated, because they have to be localised.

## Conventions in this document

Terms in curly brackets { } should not be translated, but localised. The NRC is asked to instruct the reconciler which localised (a term which is appropriate for the country) terms to insert.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.

The curly brackets { } and square brackets [ ] should not appear in the translated text. Please instruct the translators and reconciler to remove these brackets from the translated text.

The term [target language] refers to the language for which the students will be tested prior to filling out the questionnaire, in other words the 1<sup>st</sup> most widely taught language among English, French, German, Spanish and Italian. Please instruct the translator which language to fill in when the term [target language] appears.

The term "SQ" is used as a reference to a question in the Student Questionnaire, "TQ" as a reference to the Teacher Questionnaire and "PQ" as a reference to the Principal Questionnaire.

Please note that all questions should be translated even when it is expected that all respondents in your country will give the same answer. Question order is known to have an effect upon the answers. Removal of questions will compromise the comparability across countries and the comparability with future cycles.

Overall we have tried to prevent questions that require an open-ended text response as much as possible, as the coding of such questions (an NRC task) is very time consuming and costly.

The current lay-out of the questions is not actual lay-out in the Web survey, because the computer screen has a different size than a Word-document. The current lay-out is added to ease checking of the translation.

## **Translator note**

Throughout the questionnaire the formal address is used.

"You" is singular unless otherwise indicated.

Terms in curly brackets { } should not be translated, but localised. The reconciler and/or NRC should insert a phrase or word which is appropriate for the country.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.

Your school has agreed to participate in the European survey on Language Competences, a large European study of student learning of foreign languages, launched by the European Commission.

This survey aims to provide Member States, policy makers, teachers and practitioners with information on the foreign language competence of students enrolled in



{secondary education} and to provide insight into the progress towards the objective of improving foreign language learning.

In this questionnaire you will find questions about:

- Yourself
- The languages you speak
- Stays abroad
- Your initial training as a teacher
- Your qualifications as a teacher
- Your current employment
- In-service training
- Teaching foreign languages
- The available resources for your [target language] classes
- Your [target language] classes
- Homework and assessment

As one of the teachers of the sampled students, your responses to these questions are very important in helping to describe the context of foreign language learning in Europe and to establish how this context is related to student achievement.

Please read the questions carefully and answer each question. If you do not know a precise answer, your best estimate will be adequate for the purposes of the study.

Completing the questionnaire will take between 30 and 45 minutes. If you want to stop and continue filling out the questionnaire at a later time, please press the button “SAVE”. The next time you will log on, you can continue where you left off.

If you have completed the entire questionnaire please press “SEND”. Please do NOT press the button “SEND” before you have completed the questionnaire, as you will not be able to complete the questionnaire after sending it.

All your answers will be kept confidential and secret. It will be impossible to identify individuals from the combined responses.

#### **NRC note**

The term {secondary education} in this question might need to be localised. Please instruct the reconciler which terms should be used.

**Translator note**

Please do not translate terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

each - Please underline the appropriate text in the translation

“SAVE” - Please underline the appropriate text in the translation

## About yourself

1 Are you female or male?

Female

Male

☐ 0

☐ 1

---

**NRC note (identical to SQ1)**

This question assesses the gender of the teacher.

---

2 How old are you?

*(Please select only one answer)*

Under 25 ☐ 0

25-34 ☐ 1

35-44 ☐ 2

45-54 ☐ 3

55 or older ☐ 4

---

**WebTrans Note**

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

**NRC note**

This question assesses the age of the teachers. The national recommendations regarding the initial training and qualifications might have changed in the course of years, causing older teachers to have different training, qualifications and experiences than young teachers.

### 3 What country were you born in?

*(Please select only one answer)*

- [Educational system] ----- ☐ 0
- {Country A} ----- ☐ 1
- {Country B} ----- ☐ 2
- {Country C} ----- ☐ 3
- {Country D} ----- ☐ 4
- {Country E} ----- ☐ 5
- {Country F} ----- ☐ 6
- {Country G} ----- ☐ 7
- Other European country ----- ☐ 8
- Other non-European country ----- ☐ 9

#### WebTrans Note

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

Please note that the number of response categories can differ between countries.

#### NRC note (similar to SQ15)

This question provides information about the staff from other language communities, such as guest teachers. This question provides information on the teachers' country of birth.

The response categories of this question need to be localised (identical to the localisation of SQ15). The countries presented as response categories should correspond to the countries mentioned in the Localisation file (Country Table).

The countries listed should include the countries of origin of the largest immigrant groups in your Educational system. A maximum number of seven countries (other than your Educational system) can be included. Less than seven countries can be included when less than seven immigrant groups of substantial size reside in your Educational system (see the Instruction for the Main Study Localisation file).

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by teachers is advisable.

Please instruct the reconciler which countries should be used as response categories.

#### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response.

[Educational system] – is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

#### 4 Do you have the following devices at home?

(Please select in each row No or Yes)

	No	Yes
1) A computer or laptop .....	<input type="radio"/> 0	<input type="radio"/> 1
2) Access to internet .....	<input type="radio"/> 0	<input type="radio"/> 1
3) A printer .....	<input type="radio"/> 0	<input type="radio"/> 1
4) A CD or DVD writer .....	<input type="radio"/> 0	<input type="radio"/> 1
5) A scanner .....	<input type="radio"/> 0	<input type="radio"/> 1
6) A USB (memory) stick .....	<input type="radio"/> 0	<input type="radio"/> 1
7) A video games console, such as {Play Station, Nintendo, Wii} ..	<input type="radio"/> 0	<input type="radio"/> 1
8) Your own iPod, Mp3 player or similar .....	<input type="radio"/> 0	<input type="radio"/> 1
9) Your own mobile phone.....	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

#### NRC note (similar to SQ23)

This question provides information about the ICT facilities in the teacher's home. The use of ICT for language teaching might vary due to different ICT-facilities in school and in the home environment of teachers and pupils. This question also allows comparing the ICT facilities in the teacher's home with the ICT facilities in the students' homes.

The items are identical to the items with SQ23, however none of the words are underlined like in the Student Questionnaire and in the first item (a) "Your own" has been dropped.

Some of the terms in this question might need to be localised (identical to SQ23). Please instruct the reconciler which terms should be used.

video game console - Refers to an interactive entertainment computer or electronic device that manipulates the video display signal of a display device (a television, monitor, etc.) to display a game. National Project Managers should insert a phrase that is appropriate for their country and that is understood by the students.

{Play Station, Nintendo, Wii} - Please use examples for game consoles that will be understood by students in your country.



### Translator note

Please do not translate terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

home - refers to that place where the teacher and his/her family usually resides.

## 5 How often do you use a computer outside your lessons (at home or elsewhere) for the following?

(Please select only one answer from each row)

	Never or hardly ever	A few times a year	A few times a month	A few times a week	(Almost) every day
1) For checking students' homework or assignments from your [target language] classes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2) For preparing [target language] lessons --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3) For administrative tasks related to your [target language] classes (e.g. registering absence or marks) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4) For creating instructional material for your [target language] classes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5) For finding authentic material to use in your [target language] classes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-A few times a month-A few times a week-(Almost) every day) occurs in TQ5, TQ6, and TQ47.

### NRC note

TQ5, TQ6, TQ43, TQ44, TQ45 and TQ47 provide information on the use of ICT in foreign language teaching. This question provides information on the use of ICT outside the lessons for foreign language teaching.

### Translator note

home – refers to that place where the teacher and his/her family usually resides.

homework or assignments – All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

6	How often do you use a computer at home or elsewhere for the following personal purposes?	<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>A few times a week</i>	<i>(Almost) every day</i>
1)	For finding information -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	For games -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	For entertainment (e.g. music, movies, video clips) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	For contact with others (e.g. email, chatting, blogging, {MySpace}, {Skype}) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	For online shopping -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	For personal administration and finances -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-A few times a month-A few times a week-(Almost) every day) occurs in TQ5, TQ6, and TQ47.

### NRC note

TQ5, TQ6, TQ43, TQ44, TQ45 and TQ47 provide information on the use of ICT in foreign language teaching. This question provides information on the use of ICT outside the lessons for personal use.

Some of the terms in this question might need to be localised (see also SQ24). Please instruct the reconciler which terms should be used.

Skype - Skype is a software application that allows users to make voice calls over the Internet.

MySpace – is an example of a social networking website with an interactive, user-submitted network of friends, personal profiles, blogs, groups, photos, music, and videos for teenagers and adults. Please use an example of a social networking website that is most widely known in your country (e.g. Friendster, Twitter).

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

**Translator note**

Please do not translate terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

home - refers to that place where the teacher and his/her family usually resides.

## About the languages you speak

- 7 Which language(s) did you speak at home as a small child (before the age of five)?

*(Please tick as many boxes as applicable)*

- {most widely spoken indigenous language1} ----- ☐ 0
- {most widely spoken indigenous language2} ----- ☐ 1
- {most widely spoken indigenous language3} ----- ☐ 2
- {most widely spoken indigenous language4} ----- ☐ 3
- {most widely spoken indigenous language5} ----- ☐ 4
- [target language] ----- ☐ 5
- {most widely spoken non-indigenous language 1} --- ☐ 6
- {most widely spoken non-indigenous language 2} --- ☐ 7
- {most widely spoken non-indigenous language 3} --- ☐ 8
- {most widely spoken non-indigenous language 4} --- ☐ 9
- {most widely spoken non-indigenous language 5} --- ☐ 10
- Other European language(s) ----- ☐ 11
- Other non-European language(s) ----- ☐ 12

### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

The same response categories (from {most widely spoken indigenous language1} through {most widely spoken non-indigenous language 5}” occur in TQ7, TQ8, and TQ9.

Please note that the number of response categories can differ between countries.

### NRC note (similar to SQ4)

TQ7, TQ8 and TQ9 provide information about the language(s) the teachers from other language communities speak. This question inquires after the 1<sup>st</sup> language(s) of the teacher.

Please note that the term “mother tongue” is not used, because “for a considerable number of people in Europe, the notion of “mother tongue” has lost its meaning” (High Level Group on Multilingualism, 2007, p. 6) for example when respondents grew up in mixed language families or multilingual environments.



The response categories of this question need to be localised identical to SQ4. The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.

A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

### **Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

home – refers to that place where the teacher and his/her family usually resides.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## 8 Which language(s) do you speak regularly at home now?

*(Please tick as many boxes as applicable. Languages that you only speak once in a while or only in special circumstances should not be ticked)*

- {most widely spoken indigenous language1} ----- ☐ 0
- {most widely spoken indigenous language2} ----- ☐ 1
- {most widely spoken indigenous language3} ----- ☐ 2
- {most widely spoken indigenous language4} ----- ☐ 3
- {most widely spoken indigenous language5} ----- ☐ 4
- [target language] ----- ☐ 5
- {most widely spoken non-indigenous language 1} --- ☐ 6
- {most widely spoken non-indigenous language 2} --- ☐ 7
- {most widely spoken non-indigenous language 3} --- ☐ 8
- {most widely spoken non-indigenous language 4} --- ☐ 9
- {most widely spoken non-indigenous language 5} --- ☐ 10
- Other European language(s) ----- ☐ 11
- Other non-European language(s) ----- ☐ 12

### WebTrans Note

The same response categories (from {most widely spoken indigenous language1} through {most widely spoken non-indigenous language 5}) occur in TQ7, TQ8, and TQ9.

Please note that the number of response categories can differ between countries.

### NRC note (similar to SQ26)

TQ7, TQ8 and TQ9 provide information about the language(s) the teachers from other language communities speak. This question provides information on the languages the teacher uses currently in his or her home environment.

The response categories of this question need to be localised identical to SQ4 (and TQ7). The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.

A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.



Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

home - refers to that place where the teacher usually resides.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

regularly - Please underline the appropriate text in the translation

---

## 9 Which language do you speak most often at home?

(Please select only one answer)

{most widely spoken indigenous language1} ----- ☐ 0

{most widely spoken indigenous language2} ----- ☐ 1

{most widely spoken indigenous language3} ----- ☐ 2

{most widely spoken indigenous language4} ----- ☐ 3

{most widely spoken indigenous language5} ----- ☐ 4

[target language] ----- ☐ 5

{most widely spoken non-indigenous language 1} ---- ☐ 6

{most widely spoken non-indigenous language 2} ---- ☐ 7

{most widely spoken non-indigenous language 3} ---- ☐ 8

{most widely spoken non-indigenous language 4} ---- ☐ 9

{most widely spoken non-indigenous language 5} ---- ☐ 10

Other European language----- ☐ 11

Other non-European language----- ☐ 12

---

### WebTrans Note

The same response categories (from {most widely spoken indigenous language1} through {most widely spoken non-indigenous language 5}” occur in TQ7, TQ8, and TQ9.

Please note that the number of response categories can differ between countries.

### NRC note (similar to SQ27)

TQ7, TQ8 and TQ9 provide information about the language(s) the teachers from other language communities speak. This question is similar to the question in PISA about language use of students. This question allows us to compare our findings with the results of the students and the results reported in PISA and in the Eurydice Key data report on teaching languages at school (2008).

The response categories of this question need to be localised identical to SQ4 (and TQ7). The languages presented as response categories should correspond to the languages in the Localisation file (Language Table). The most widely spoken “indigenous languages” (national and regional) should be included and the most widely spoken “non-indigenous languages” (see the Instruction for the Main Study Localisation file). Please make sure that the languages of the largest immigrant groups in your Educational system (see SQ15, SQ16, SQ17) are among the “non-indigenous” languages. The included languages do not have to be official languages.

A maximum number of five “indigenous” languages can be included and a maximum number of five “non-indigenous” languages. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

home – refers to that place where the teacher and his/her family usually resides.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

most often – Please underline the appropriate text in the translation.

only one – Please underline the appropriate text in the translation.

---

10 How many languages – including ancient languages - have you studied as a foreign language either at school or otherwise?

*(Please write down the number of all the foreign languages you have studied)*      Languages

---

### NRC note

TQ10, TQ11 and TQ14 provide information on the (foreign) languages the teacher himself or herself has studied. TQ10 provides information on the number of languages the teacher has studied in school or in his/her leisure time. TQ11 and TQ14 provide information on which languages the teacher has studied in school (TQ11) or as part of their initial training (TQ14).

### Translator note

How many - Please underline the appropriate text in the translation.



11 Which of the following languages have you studied as a foreign language either at school or otherwise?

(Please tick as many boxes as applicable)

- {1<sup>st</sup> most widely taught foreign language in the country} ..... ☐ 0
- {2<sup>nd</sup> most widely taught foreign language in the country} ..... ☐ 1
- {3<sup>rd</sup> most widely taught foreign language in the country} ..... ☐ 2
- {4<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 3
- {5<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 4
- {6<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 5
- {7<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 6
- {8<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 7
- {9<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 8
- {10<sup>th</sup> most widely taught foreign language in the country} ..... ☐ 9
- Other foreign language(s) ..... ☐ 10

### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

Please note that the number of response categories can differ between countries.

### NRC note (similar to SQ37)

TQ10, TQ11 and TQ14 provide information on the (foreign) languages the teacher himself or herself has studied. TQ10 provides information on the number of languages the teacher has studied in school or in his/her leisure time. TQ11 and TQ14 provide information on which languages the teacher has studied in school (TQ11) or as part of their initial training (TQ14).

The response categories of this question need to be localised identical to SQ37. The languages presented as response categories should correspond to the taught languages listed in the Taught language Table of the Localisation file.



The languages that are most widely taught in secondary education level should be included (including ancient languages). A maximum number of 10 foreign languages can be included. If less than 10 foreign languages can be studied in secondary education fewer response categories can be used (see the Instruction for the Main Study Localisation file).

Please note that the list of languages in TQ11 and TQ14 cannot be exhaustive, given the number of languages in the world.

Please instruct the reconciler which languages should be used as response categories.

**Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

Which – Please underline the appropriate text in the translation.

## About stays abroad

12	How often have you stayed more than one month in a [target language] speaking country for the following reasons?	
	<i>(If you have never stayed <u>more</u> than one month in a [target language] speaking country for a reason listed, please write down 0).</i>	
		<i>Times</i>
1)	For holiday -----	<input type="text"/>
2)	For study or to take a course -----	<input type="text"/>
3)	For teaching -----	<input type="text"/>
4)	For other jobs than teaching -----	<input type="text"/>
5)	For living with your family -----	<input type="text"/>

### NRC note

This question provides information about the stays abroad and the reason for the stays abroad (work, study, or other). Stays in a country where the language taught is spoken are promoted in the EU. This question provides information about the number of times and the reason that teachers stayed in a [target language] speaking country for an extended period.

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by teachers is advisable. Please instruct the reconciler how to adapt the question.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

more - Please underline the appropriate text in the translation.

## About your initial training as a teacher

*In this section you will be asked some questions about your initial training as a teacher.*

*With initial training is meant both the study of the subject(s) to be taught, for example [target language] and the training in the theoretical and practical skills needed to be a teacher.*

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

### 13 What is the highest level of education that you have completed?

*(Please select only one answer.*

*If you completed your education abroad, please select the level that is the closest match)*

- {ISCED 6} ----- ☐ 0
- {ISCED 5A} ----- ☐ 1
- {ISCED5B} ----- ☐ 2
- {ISCED4} ----- ☐ 3
- {ISCED3A} ----- ☐ 4
- {ISCED 3B} ----- ☐ 5
- {ISCED 3C} ----- ☐ 6

---

### WebTrans Note

Please note that the number of response categories can differ between countries.

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers' initial training. This question provides information on teachers' highest level of education.

The response categories of this question need to be localised. Most EU countries have officially classified their educational system using the ISCED classification of educational levels (see Classifying Educational Programmes — Manual for ISCED-97 Implementation in OECD Countries, 1999 Edition, OECD). These country-specific classifications can be also be found on the website of Eurydice (<http://eacea.ec.europa.eu/portal/page/portal/Eurydice/>).



Please make sure that each ISCED-level is represented (ISCED3, ISCED4, ISCED5 and ISCED6) as a response category. It is necessary to phrase the response categories in such a way that teachers will easily understand. The response categories should be consistent with the response categories of Q25.

When an ISCED level with a particular orientation, e.g. ISCED3B, does not exist in the country the corresponding response category should be omitted.

Changes in the educational system – Several countries have changed their educational system throughout the years. Please make sure that also previous study programs are included.

Please instruct the reconciler which terms should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories and term.

---

## 14 Which language(s) did you study as a foreign language during your initial training as a teacher?

Initial training encompasses both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher.

*(Please tick as many boxes as applicable)*

- {none}----- ☐ 0
- {most widely spoken indigenous language1}----- ☐ 1
- {most widely spoken indigenous language2}----- ☐ 2
- {most widely spoken indigenous language3}----- ☐ 3
- {most widely spoken indigenous language4}----- ☐ 4
- {most widely spoken indigenous language5}----- ☐ 5
- {1<sup>st</sup> most widely taught foreign language in the country}- ☐ 6
- {2<sup>nd</sup> most widely taught foreign language in the country}----- ☐ 7
- {3<sup>rd</sup> most widely taught foreign language in the country} ☐ 8
- {4<sup>th</sup> most widely taught foreign language in the country}- ☐ 9
- {5<sup>th</sup> most widely taught foreign language in the country}- ☐ 10
- Other language(s) ----- ☐ 11

---

### WebTrans Note

The same clarification ("With initial training is meant both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher") occurs in TQ14, TQ15, TQ16, and TQ17.



The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

The response categories of TQ14 and TQ24 are identical. The second until the last response categories of TQ14 and TQ24 are identical to the response categories of TQ36.

Please note that the number of response categories can differ between countries.

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers' initial training. TQ10, TQ11 and TQ14 provide information on the (foreign) languages the teacher himself or herself has studied. TQ10 provides information on the number of languages the teacher has studied in school or in his/her leisure time. TQ11 and TQ14 provide information on which languages the teacher has studied in school (TQ11) or as part of their initial training (TQ14).

This question provides information on the language(s) teachers studied as part of their teacher training. The teachers in the EU can be specialised to teach one foreign language, several foreign languages or two subjects, one of which is a foreign language.

The response categories of this question need to be localised. The response categories 2 through 6 (the most widely spoken "indigenous languages") should be identical to the first five response categories of TQ8 (and SQ4). The most widely spoken "indigenous" languages presented as response categories should correspond to the languages in the Localisation file (Language Table).

The most widely spoken "indigenous" languages (national and regional) should be included. A maximum number of five "indigenous" languages can be included. If a country has less than five "indigenous" languages less response categories can be used. If a country has more than five "indigenous" languages please use a more generic description, such as "One of the Sami languages".

The response categories 7 and up (the most widely taught languages, including ancient languages) should be identical to the first response categories of TQ11 (and SQ37). The most widely taught languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table).

Please note that the list of languages in TQ11 and TQ14 cannot be exhaustive, given the number of languages in the world.

In total a maximum number of 10 languages can be included.

Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

[Target language] – Please replace the term "target language" according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

15	Did you receive instruction in the following language related subjects during your initial training as a teacher?
	Initial training encompasses both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher.

		<i>(Please select in each row No or Yes)</i>	
		No	Yes
1)	[Target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
2)	Other languages than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
3)	[Target language] literature -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	[Target language] culture -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	Teaching [target language] as a foreign language -----	<input type="radio"/> 0	<input type="radio"/> 1
6)	Didactics and methodology of [target language] teaching -----	<input type="radio"/> 0	<input type="radio"/> 1
7)	ICT for language teaching -----	<input type="radio"/> 0	<input type="radio"/> 1
8)	{Content and Language Integrated Learning (CLIL)} -----	<input type="radio"/> 0	<input type="radio"/> 1
9)	The Common European Framework of Reference -----	<input type="radio"/> 0	<input type="radio"/> 1
10)	The use of a Portfolio, e.g. the European Language Portfolio -----	<input type="radio"/> 0	<input type="radio"/> 1
11)	Development of reflective practice and action research skills -----	<input type="radio"/> 0	<input type="radio"/> 1
12)	Classroom activities for language learning -----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same clarification (“With initial training is meant both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher”) occurs in TQ14, TQ15, TQ16, and TQ17.

The items 3) through 12) are identical in TQ15 and TQ32.

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers’ initial training. This question provides information on the language related subjects teachers studied as part of their teacher training. The items of this question are similar to the items in TQ32.

- 15\_5 (and 32\_5) provide information about the training of teachers from other language communities to teach [target language] as a foreign language.
- 15\_10 (and 32\_10) provide information on the received training in using a Portfolio, like the European Language Portfolio.
- 15\_9 (and 32\_9) provide information on the use and purpose of use of CEFR.
- 15\_8 (and 32\_8) provide information on the training in Content and language integrated learning (CLIL).

Some of the items of this question need to be localised.



Content and Language Integrated Learning (CLIL) - involves teaching a curricular subject through the medium of a language other than that normally used. The subject can be entirely unrelated to language learning, such as history lessons being taught in English in a school in Spain”

Please instruct the reconciler which item should be used.

### **Translator note**

Please do not translate the item in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate item.

The Common European Framework of Reference (CEFR) – A guideline developed by the Council of Europe which provides “a practical tool for setting clear standards to be attained at successive stages of learning and for evaluating outcomes in an internationally comparable manner. [...] The Common European Framework of Reference (CEFR) provides a basis for the mutual recognition of language qualifications, thus facilitating educational and occupational mobility. It is increasingly used in the reform of national curricula and by international consortia for the comparison of language certificates” ([http://www.coe.int/t/dg4/linguistic/CADRE\\_EN.asp](http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp)).

Portfolio – “a language portfolio is a document (...) in which individual learners (...) can assemble over a period of time, and display in a systematic way, a record of their qualifications, achievements and experiences in language learning, together with samples of work they have themselves produced” (Trim, 1997, p.3).

Please note that in item 1) 3) and 4) the adapted and translated word for [target language] starts with an uppercase letter.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



16	During your initial training did you receive education or training in the following subjects related to the theory and practice of teaching in general?		
	Initial training encompasses both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher.		
	(Please select in each row No or Yes)		
		No	Yes
1)	Intercultural education-----	<input type="radio"/> 0	<input type="radio"/> 1
2)	Special needs education-----	<input type="radio"/> 0	<input type="radio"/> 1
3)	Dealing with mixed ability in the class -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	Behaviour management and school discipline -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	Development of course materials-----	<input type="radio"/> 0	<input type="radio"/> 1
6)	Tests or student assessment -----	<input type="radio"/> 0	<input type="radio"/> 1
7)	School management-----	<input type="radio"/> 0	<input type="radio"/> 1
8)	Communication skills or public relations -----	<input type="radio"/> 0	<input type="radio"/> 1
9)	Mentoring and coaching -----	<input type="radio"/> 0	<input type="radio"/> 1
10)	Accessing support networks and professional associations -----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same clarification (“With initial training is meant both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher”) occurs in TQ14, TQ15, TQ16, and TQ17.

The items of TQ16 and TQ33 are identical.

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers’ initial training. This question provides general information on education or training in subjects related to the theory and practice of teaching in general.

The items in this question are identical to the items in TQ33.

### Translator note

Special needs education – Education and the provision of additional educational resources for students with special educational needs.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

17	How much of your initial training was devoted to the following?	
	Initial training encompasses both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher.	
	<i>(Please write down a rough estimate of the percentage)</i>	
		%
1)	<u>Approximately</u> , how much of your initial training was devoted to courses related to the subject(s) you were going to teach when qualified? -----	<input type="text"/>
2)	<u>Approximately</u> , how much of your initial training was devoted to the training in the theoretical and practical skills needed to be a teacher? -----	<input type="text"/>
	Total	100%

### WebTrans Note

The same clarification (“With initial training is meant both the study of the subject(s) to be taught, for example [target language], and the training in the theoretical and practical skills needed to be a teacher”) occurs in TQ14, TQ15, TQ16, and TQ17.

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers’ initial training. This question provides a global indication of the proportion of teachers’ training devoted to training in the theoretical and practical skills needed to be a teacher. This proportion varies considerably across Europe. This question allows us to compare our findings with the results reported in the Eurydice report about the initial training of teachers in Europe (Key topics in education in Europe, vol. 3, report I, 2003).

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

Approximately – Please underline the appropriate text in the translation.

18	How long were the following phases during your initial training as a teacher?	
	<i>(Please write down the number of months. If the following phases were not part of your initial training please write down 0)</i>	
		<i>Months</i>
1)	In-school teaching placements -----	<input type="text"/>
2)	A stay in a [target language] speaking country -----	<input type="text"/>
3)	Your initial training in total -----	<input type="text"/>

### NRC note

The questions in this section provide information about the level, duration, and specialisation of teachers' initial training. This question provides information the duration of their initial training and of two recommended phases in the initial training (an in-school teaching placement and a stay in a country where the language to be taught is spoken).

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About your qualifications as a teacher

19 What kind of [certification] for teaching [target language] do you currently hold?

(Please select only one answer)

No certificate ----- ☐ 0

{Temporary or emergency certification} ----- ☐ 1

{Provisional certificate, e.g. Newly Qualified teacher}--- ☐ 2

{Full certificate}----- ☐ 3

Other certificate ----- ☐ 4

### WebTrans Note

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

### NRC note

The questions in this section provide information on the qualifications of the teachers. This question provides information on the kind of certification the teachers hold.

The response categories of this question need to be localised. Please use response categories that are appropriate in your [Educational system] and that teachers will easily understand.

{Full certificate} – credentials from an authoritative source, such as the government or a higher education institution that allow teachers to teach in schools which require authorization in general, as well as allowing educators to teach in particular content areas and across the curriculum. In some countries different kinds of certification are required for teaching ISCED2 and teaching in ISCED3. In these cases both kinds of certificates kind be presented as a separate response category.

{Provisional certificate, e.g. {Newly Qualified teacher}: In some countries, teacher training for general lower secondary education ends with a final “on-the-job” qualifying phase (also known as induction period). During this “on-the-job” qualifying phase teachers hold a provisional certificate, e.g. Newly Qualified teacher in England.

{Temporary or emergency certification}: Many countries have developed several emergency measures to deal with teacher shortage. One of such measures is providing teachers a temporary certificate that allows teachers to teach a subject for which they are not yet fully qualified (Eurydice, 2003, report Key topics in education in Europe, vol. 3, report II, 2003).

If a particular form of certificate, such as a provisional certificate, does not exist in your [Educational system] and teachers from abroad with such qualifications are not employed in your [Educational system] the response category can be left out.

Please instruct the reconciler which response categories should be used.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.



[Certification] – some form of credentials from an authoritative source, such as the government or a higher education institution that allow teachers to teach in schools which require authorization in general, as well as allowing educators to teach in particular content areas and across the curriculum.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[target language] – Please underline the appropriate text in the translation

20	Which of the following educational levels are you qualified to teach?		
	<i>(Please select in each row No or Yes)</i>		
		No	Yes
1)	{ISCED 5A,6} -----	<input type="radio"/> 0	<input type="radio"/> 1
2)	{ISCED5B} -----	<input type="radio"/> 0	<input type="radio"/> 1
3)	{ISCED4} -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	{ISCED3A} -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	{ISCED 3B,3C} -----	<input type="radio"/> 0	<input type="radio"/> 1
6)	{ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1
7)	{ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

Please note that the number of items can differ between countries.

#### NRC note

The questions in this section provide information on the qualifications of the teachers. This question provides information on the educational level the teacher is qualified to teach.

The response categories of this question need to be localised similar to SQ13. Most EU countries have officially classified their educational system using the ISCED classification of educational levels (levels (see Classifying Educational Programmes — Manual for ISCED-97 Implementation in OECD Countries, 1999 Edition, OECD). These country-specific classifications can be also be found on the website of Eurydice (<http://eacea.ec.europa.eu/portal/page/portal/Eurydice/>).

Please make sure that each ISCED-level is represented (ISCED1, ISCED2, ISCED3, ISCED4 and ISCED5) as a response category. It is necessary to phrase the response categories in such a way that teachers will easily understand.



When an ISCED level with a particular orientation, e.g. ISCED3B, does not exist in the country the corresponding response category should be omitted. The response categories should be consistent with the response categories of TQ13.

In some countries there may not be an administrative or structural boundary between some successive ISCED levels (for example between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels.

Please instruct the reconciler which terms should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { } and do not translate the term {ISCED1} in the last response category. The reconciler and/or NRC is asked to add the appropriate response categories and term.

---

## 21 How many school subjects (including languages) are you qualified to teach?

*(Please select only one answer)*

None ----- ☐ 0

One school subject ----- ☐ 1

Two school subjects ----- ☐ 2

Three school subjects ----- ☐ 3

Four school subjects or more ----- ☐ 4

---

### NRC note

The questions in this section provide information on the qualifications of the teachers. TQ21, TQ22, TQ23 and TQ24 provide information about the specialisation of teachers and allow comparing our findings with the results reported in the Eurydice report on Key Data on Teaching Languages at School in Europe, 2008 Edition. Some teachers are qualified to teach two different subjects, one of which is a foreign language, or qualified solely to teach foreign languages. Other teachers are qualified to teach a group of at least three different subjects, one or more of which is foreign languages and some teachers are qualified to teach all or almost all subjects, including foreign language(s). TQ21 provides information on the number of subjects a teacher is qualified to teach and TQ23 the number of languages. TQ22 provides information on which subjects a teacher is qualified to teach and TQ24 which languages.

Please note that the number of subjects cannot be deduced from the next question, because in the next question several school subjects are grouped together, to prevent a too long list of possible subjects.

22	Which school subjects are you qualified to teach?		
	<i>(Please select in each row No or Yes)</i>		
		No	Yes
1)	{Mathematics} -----	<input type="radio"/> 0	<input type="radio"/> 1
2)	{One or more science subjects, e.g. physics}-----	<input type="radio"/> 0	<input type="radio"/> 1
3)	{One or more Human and society subjects, e.g. history} -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	{One or more Culture and arts subjects, e.g. music, art history} -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	[Questionnaire language] -----	<input type="radio"/> 0	<input type="radio"/> 1
6)	[Target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
7)	One or more other foreign languages (including ancient languages) -----	<input type="radio"/> 0	<input type="radio"/> 1
8)	{One or more vocational skills subjects} -----	<input type="radio"/> 0	<input type="radio"/> 1
9)	{Sports} -----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

### NRC note

The questions in this section provide information on the qualifications of the teachers. This question provides information on which subjects a teacher is qualified to teach.

TQ21, TQ22, TQ23 and TQ24 provide information about the specialisation of teachers and allow comparing our findings with the results reported in the Eurydice report on Key Data on Teaching Languages at School in Europe, 2008 Edition. Some teachers are qualified to teach two different subjects, one of which is a foreign language, or qualified solely to teach foreign languages. Other teachers are qualified to teach a group of at least three different subjects, one or more of which is foreign languages and some teachers are qualified to teach all or almost all subjects, including foreign language(s). TQ21 provides information on the number of subjects a teacher is qualified to teach and TQ23 the number of languages. TQ22 provides information on which subjects a teacher is qualified to teach and TQ24 which languages.

The items of this question might need to be localised similar to SQ34. It is necessary to phrase the groups of school subjects and examples in a way that teachers will easily understand. Please instruct the reconciler which items should be used. Please note that in the items several school subjects are grouped together, to prevent a too long list of possible subjects. The number of items should remain 10.

Please make sure that the items include all or most of the subjects which are part of the official curriculum (at ISCED2 and ISCED3 level).

### Translator note

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.



[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

## 23 How many languages are you qualified to teach?

*(Please select only one answer)*

- None ----- ☐ 0
- One language ----- ☐ 1
- Two languages ----- ☐ 2
- Three languages ----- ☐ 3
- Four languages or more ----- ☐ 4

---

### NRC note

The questions in this section provide information on the qualifications of the teachers. This question provides information on which subjects a teacher is qualified to teach.

TQ21, TQ22, TQ23 and TQ24 provide information about the specialisation of teachers and allow comparing our findings with the results reported in the Eurydice report on Key Data on Teaching Languages at School in Europe, 2008 Edition. Some teachers are qualified to teach two different subjects, one of which is a foreign language, or qualified solely to teach foreign languages. Other teachers are qualified to teach a group of at least three different subjects, one or more of which is foreign languages and some teachers are qualified to teach all or almost all subjects, including foreign language(s). TQ21 provides information on the number of subjects a teacher is qualified to teach and TQ23 the number of languages. TQ22 provides information on which subjects a teacher is qualified to teach and TQ24 which languages.

Please note that the number of languages cannot be deduced from the next question, because list of languages in the next question cannot be exhaustive, given the number of languages that teachers in the [Educational system], including those from abroad, might be qualified to teach.



## 24 Which language(s) are you qualified to teach?

(Please tick as many boxes as applicable)

{none}----- ☐ 0

{most widely spoken indigenous language1}----- ☐ 1

{most widely spoken indigenous language2}----- ☐ 2

{most widely spoken indigenous language3}----- ☐ 3

{most widely spoken indigenous language4}----- ☐ 4

{most widely spoken indigenous language5}----- ☐ 5

{1<sup>st</sup> most widely taught foreign language in the country} ----- ☐ 6

{2<sup>nd</sup> most widely taught foreign language in the country} ----- ☐ 7

{3<sup>rd</sup> most widely taught foreign language in the country} ----- ☐ 8

{4<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 9

{5<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 10

Other language(s) ----- ☐ 11

### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

The response categories of TQ14 and TQ24 are identical. The second until the last response categories of TQ14 and TQ24 are identical to the response categories of TQ36.

Please note that the number of response categories can differ between countries.

### NRC note

The questions in this section provide information on the qualifications of the teachers. This question provides information on which subjects a teacher is qualified to teach.



TQ21, TQ22, TQ23 and TQ24 provide information about the specialisation of teachers and allow comparing our findings with the results reported in the Eurydice report on Key Data on Teaching Languages at School in Europe, 2008 Edition. Some teachers are qualified to teach two different subjects, one of which is a foreign language, or qualified solely to teach foreign languages. Other teachers are qualified to teach a group of at least three different subjects, one or more of which is foreign languages and some teachers are qualified to teach all or almost all subjects, including foreign language(s). TQ21 provides information on the number of subjects a teacher is qualified to teach and TQ23 the number of languages. TQ22 provides information on which subjects a teacher is qualified to teach and TQ24 which languages.

The response categories of this question need to be localised identical to TQ14. The response categories 2 through 6 (the most widely spoken “indigenous languages”) should be identical to the first five response categories of TQ8 (and SQ4). The most widely spoken “indigenous” languages presented as response categories should correspond to the languages in the Localisation file (Language Table).

The most widely spoken “indigenous” languages (national and regional) should be included. A maximum number of five “indigenous” languages can be included. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

The response categories 7 and up (the most widely taught languages, including ancient languages) should be identical to the first response categories of TQ11 (and SQ37). The most widely taught languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table).

Please note that the list of languages in TQ11 and TQ14 cannot be exhaustive, given the number of languages in the world.

In total a maximum number of 10 languages can be included.

Please instruct the reconciler which languages should be used as response categories.

### **Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

## About your current employment

25 How are you currently employed at this school?

(Please select only one answer.

Select the answer that applies to most of the subjects you teach)

In full-time temporary employment ----- ☐ 0

In full-time permanent employment ----- ☐ 1

In part-time temporary employment ----- ☐ 2

In part-time permanent employment ----- ☐ 3

### NRC note

This question provides information on the employment status of teachers. This question allows us to compare our findings with the results reported in the Eurydice report about the employment status of teachers in Europe (Key topics in education in Europe, vol. 3, report III, 2003).

26	According to your <u>contract</u> , how many hours a week do you have to devote to the following?	
	(Please write down the number of hours. Please write down 0 if your <u>contract</u> does not specify the hours for one of the following)	
		Hours a week
1)	Teaching -----	<input type="text"/>
2)	Being available at school, apart from your teaching -----	<input type="text"/>
3)	Working in total -----	<input type="text"/>

### NRC note

This question provides information on the working time of teachers and allows comparison with the results reported in the Eurydice report about the employment status of teachers in Europe (Key topics in education in Europe, vol. 3, report III, 2003).

### Translator note

contract – Please underline the appropriate text in the translation.

27 How long does a class period last at your school?

(Please write down the number of minutes) Minutes

#### NRC note

TQ27 and TQ28 allow us to calculate the teaching time and the proportion that is devoted to [target language] and to other languages.

28	In a normal full week at school, how many class periods do you teach the following?	
	(Please write down the number of class periods)	
		Number of class periods
1)	[Target language] .....	<input type="text"/>
2)	Other languages than [target language], including ancient languages .....	<input type="text"/>
3)	All subjects, including languages, together (total) .....	<input type="text"/>

#### WebTrans Note

The same items occur in TQ28 and TQ29.

#### NRC note

TQ27 and TQ28 allow us to calculate the teaching time and the proportion that is devoted to [target language] and to other languages.

#### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

29	By the end of this school year, how many years will you have been teaching the following?	
	<i>(Please write down the number of full years)</i>	
		Years
1)	[Target language] -----	<input type="text"/>
2)	Other languages than [target language], including ancient languages-----	<input type="text"/>
3)	All subjects, including languages, together (total) -----	<input type="text"/>

### WebTrans Note

The same items occur in TQ28 and TQ29.

### NRC note

This question provides information on the teaching experience of teachers in total and their experience with teaching languages and [target language].

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About in-service training

*In this section you will be asked some questions about in-service training.*

*In-service training is training that teachers receive during their career to update, develop and broaden the knowledge acquired during initial training and/or to provide them with new skills and professional understanding which teachers may not have at a given stage in their career. The training might take different forms, such as a course, a conference or a seminar.*

### NRC note

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training.

30 Is participation in in-service training an obligation, a right or an option for you?

*(Please select the answer(s) that describes your situation best)*

Participation in in-service training is an obligation for teachers ----- ☐ 0

Participation in in-service training is a right for teachers ----- ☐ 1

Participation in in-service training is required for promotion ----- ☐ 2

Participation in in-service training is optional ----- ☐ 3

### NRC note

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ30, TQ34 and TQ35 provide information on the organisation and incentives of in-service training. Schools have a growing autonomy and freedom to develop plans for in-service training at school-level.

### Translator note

your – Please underline the appropriate text in the translation.

31	In the past five years, how often have you participated in an in-service training at the following places?	
	<i>(Please write down the number of times you participated in a <u>different</u> in-service training)</i>	
		<i>Number of times</i>
1)	At the school where you teach -----	<input type="text"/>
2)	At another institute in [Educational system] -----	<input type="text"/>
3)	At an institute in a [target language] speaking country -----	<input type="text"/>
4)	At an institute in a non-[target language] speaking country other than [Educational system] -----	<input type="text"/>
5)	Online-----	<input type="text"/>

### NRC note

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ31, TQ32 and TQ33 provide information on the mode and focus of in-service training the teachers have followed. This question allows us to compare our findings with the results reported in the Eurydice report about in-service training of teachers in Europe (Key topics in education in Europe, vol. 3, report III, 2003) and of TALIS (OECD). This question inquires how often the teacher has followed an in-service training, not after the number of course days of a single in-service training.

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by teachers might be needed. Please instruct the reconciler how to adapt the question.

### Translator note

This question inquires how often the teacher has followed an in-service training, not after the number of course days of a single in-service training.

[Educational system] – is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

different - Please underline the appropriate text in the translation.

32	In the past five years, have you, as a teacher, participated in in-service training covering any of the following language related themes?		
		<i>(Please select in each row No or Yes)</i>	
		<i>No</i>	<i>Yes</i>
1)	Study [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
2)	Study other languages than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1
3)	[Target language] literature-----	<input type="radio"/> 0	<input type="radio"/> 1
4)	[Target language] culture -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	Teaching [target language] as a foreign language-----	<input type="radio"/> 0	<input type="radio"/> 1
6)	Didactics and methodology of [target language] teaching-----	<input type="radio"/> 0	<input type="radio"/> 1
7)	ICT for language teaching -----	<input type="radio"/> 0	<input type="radio"/> 1
8)	{Content and Language Integrated Learning (CLIL)} ---	<input type="radio"/> 0	<input type="radio"/> 1
9)	The Common European Framework of Reference -----	<input type="radio"/> 0	<input type="radio"/> 1
10)	The use of a Portfolio, e.g. the European Language Portfolio -----	<input type="radio"/> 0	<input type="radio"/> 1
11)	Development of reflective practice and action research skills -----	<input type="radio"/> 0	<input type="radio"/> 1
12)	Classroom activities for language learning -----	<input type="radio"/> 0	<input type="radio"/> 1

### WebTrans Note

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The items 3) through 12) are identical in TQ15 and TQ32.

### NRC note (similar to Q19)

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ31, TQ32 and TQ33 provide information on the mode and focus of in-service training the teachers have followed. This question allows us to compare our findings with the results reported in the Eurydice report about in-service training of teachers in Europe (Key topics in education in Europe, vol. 3, report III, 2003) and of TALIS (OECD).

This question provides information on the language related subjects teachers studied as part of their in-service training. The items of this question are similar to the items in Q15.

- 15\_5 (and 32\_5) provide information about the training of teachers from other language communities to teach [target language] as a foreign language.





- 15\_10 (and 32\_10) provide information on the received training in using a Portfolio, like the European Language Portfolio.
- 15\_9 (and 32\_9) provide information on the use and purpose of use of CEFR.
- 15\_8 (and 32\_8) provide information on the training in Content and language integrated learning (CLIL).

Some of the items of this question need to be localised.

Content and Language Integrated Learning (CLIL) - involves teaching a curricular subject through the medium of a language other than that normally used. The subject can be entirely unrelated to language learning, such as history lessons being taught in English in a school in Spain”

Please instruct the reconciler which item should be used.

### **Translator note**

Please do not translate the item in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate item.

The Common European Framework of Reference (CEFR) – A guideline developed by the Council of Europe which provides “a practical tool for setting clear standards to be attained at successive stages of learning and for evaluating outcomes in an internationally comparable manner. [...] The Common European Framework of Reference (CEFR) provides a basis for the mutual recognition of language qualifications, thus facilitating educational and occupational mobility. It is increasingly used in the reform of national curricula and by international consortia for the comparison of language certificates” ([http://www.coe.int/t/dg4/linguistic/CADRE\\_EN.asp](http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp)).

Portfolio – “a language portfolio is a document (...) in which individual learners (...) can assemble over a period of time, and display in a systematic way, a record of their qualifications, achievements and experiences in language learning, together with samples of work they have themselves produced” (Trim, 1997, p.3).

Please note that in items 1) 3) and 4) the adapted and translated word for [target language] starts with an uppercase letter.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

33	In the past five years, have you, as a teacher, participated in in-service training treating any of the following themes related to the theory and practice of teaching in general?		
	<i>(Please select in each row No or Yes)</i>		
		<i>No</i>	<i>Yes</i>
1)	Intercultural education -----	<input type="radio"/> 0	<input type="radio"/> 1
2)	Special needs education -----	<input type="radio"/> 0	<input type="radio"/> 1
3)	Dealing with mixed ability in the class -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	Behaviour management and school discipline -----	<input type="radio"/> 0	<input type="radio"/> 1
5)	Development of course materials -----	<input type="radio"/> 0	<input type="radio"/> 1
6)	Tests or student assessment -----	<input type="radio"/> 0	<input type="radio"/> 1
7)	School management -----	<input type="radio"/> 0	<input type="radio"/> 1
8)	Communication skills or public relations -----	<input type="radio"/> 0	<input type="radio"/> 1
9)	Mentoring and coaching -----	<input type="radio"/> 0	<input type="radio"/> 1
10)	Accessing support networks and professional associations -----	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The items of TQ16 and TQ33 are identical.

#### NRC note (similar to Q20)

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ31, TQ32 and TQ33 provide information on the mode and focus of in-service training the teachers have followed. This question allows us to compare our findings with the results reported in the Eurydice report about in-service training of teachers in Europe (Key topics in education in Europe, vol. 3, report III, 2003) and of TALIS (OECD).

The items in this question are identical to the items in TQ16.

#### Translator note

Special needs education – Education and the provision of additional educational resources for students with special educational needs.

34	Which of the following financial compensations can you get for participation in in-service training?	(Please select in each row No or Yes)	
		No	Yes
1)	Payment of enrolment costs of training-----	<input type="radio"/> 0	<input type="radio"/> 1
2)	Payment of other training-related expenditure-----	<input type="radio"/> 0	<input type="radio"/> 1
3)	Paid leave during training with no loss of earnings -----	<input type="radio"/> 0	<input type="radio"/> 1
4)	Increase in salary afterwards -----	<input type="radio"/> 0	<input type="radio"/> 1

#### WebTrans Note

The same response instruction (Please select in each row No or Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

The same response scale (No-Yes) occurs in TQ4, TQ15, TQ16, TQ20, TQ22, TQ32, TQ33, and TQ34.

#### NRC note

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ30, TQ34 and TQ35 provide information on the organisation and incentives of in-service training.

#### Translator note

Financial compensations – all financial compensations whether from the government or employer.

### 35 When are you normally allowed to participate in in-service training?

(Please select only one answer)

During your working hours with a substitute teacher for your classes ----- ☐ 0

During your working hours but not during teaching hours (a substitute teacher for your classes is not organised)----- ☐ 1

Only outside your working hours----- ☐ 2

#### WebTrans Note

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

#### NRC note

The questions in this section provide information on the participation in in-service training and the mode and focus of in-service training. TQ30, TQ34 and TQ35 provide information on the organisation and incentives of in-service training.

**Translator note**

Without - Please underline the appropriate text in the translation.

Only - Please underline the appropriate text in the translation.

## About teaching foreign languages

36 Which of the following languages have you taught during the past five years?

(Please tick as many boxes as applicable)

- {most widely spoken indigenous language1} ----- ☐ 0
- {most widely spoken indigenous language2} ----- ☐ 1
- {most widely spoken indigenous language3} ----- ☐ 2
- {most widely spoken indigenous language4} ----- ☐ 3
- {most widely spoken indigenous language5} ----- ☐ 4
- {1<sup>st</sup> most widely taught foreign language in the country} ----- ☐ 5
- {2<sup>nd</sup> most widely taught foreign language in the country} ----- ☐ 6
- {3<sup>rd</sup> most widely taught foreign language in the country} ----- ☐ 7
- {4<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 8
- {5<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 9
- Other language(s) ----- ☐ 10

### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

The response categories of TQ14 and TQ24 are identical. The second until the last response categories of TQ14 and TQ24 are identical to the response categories of TQ36.

Please note that the number of response categories can differ between countries.

### NRC note

TQ36, TQ37 and TQ38 provide information about the experience of teachers in teaching languages. This question provides information on which languages the teachers have taught.

The response categories of this question need to be localised. The response categories 2 through 6 (the most widely spoken “indigenous languages”) should be identical to the first five response categories of TQ8 (and SQ4). The most widely spoken “indigenous” languages presented as response categories should correspond to the languages in the Localisation file (Language Table).

The most widely spoken “indigenous” languages (national and regional) should be included. A maximum number of five “indigenous” languages can be included. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.



The response categories 7 and up (the most widely taught languages, including ancient languages) should be identical to the first response categories of TQ11 (and SQ37). The most widely taught languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table).

Please note that the list of languages in TQ11 and TQ14 cannot be exhaustive, given the number of languages in the world.

In total a maximum number of 10 languages can be included.

Please instruct the reconciler which languages should be used as response categories.

**Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

37 In which of the following [programmes] have you taught [target language] during the past five years?

(Please tick as many boxes as applicable)

{Programme 1}----- ☐ 0

{Programme 2}----- ☐ 1

{Programme 3}----- ☐ 2

{Programme 4}----- ☐ 3

{Programme 5}----- ☐ 4

{Programme 6}----- ☐ 5

{Programme 7}----- ☐ 6

{Programme 8}----- ☐ 7

{Programme 9}----- ☐ 8

{Programme 10}----- ☐ 9

{Programme 11}----- ☐ 10

{Programme 12}-----

{Programme 13}-----

{Programme 14}-----

{Programme 15}-----

### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

Please note that the number of response categories can differ between countries.

### NRC note

TQ36, TQ37 and TQ38 provide information about the experience of teachers in teaching languages. This question provides information on the different grades in which the teacher has taught [target language]. In which study programs and grades a teacher can teach depends in several countries on the qualifications and experience of the teacher.

The response categories of this question need to be localised (similar to SQ6). The study programmes presented as response categories should correspond to the study programmes at ISCED2 and ISCED3 level in the Localisation file (Study Program Table). It is necessary to phrase study programme labels in such a way that teachers will easily understand.



In some countries there may not be an administrative or structural boundary between some successive ISCED levels (between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels.

In countries where students are grouped in school according to their ability, an adequate adaptation of this question that can be understood by students is advisable.

Please instruct the reconciler which study programs should be used as response categories.

### **Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

[Programmes] – the study programmes the student can follow in secondary education.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



### 38 In which [grades] have you taught [target language] during the past five years?

(Please tick as many boxes as applicable)

- {2<sup>nd</sup> grade of ISCED3} ----- ☐ 0
- {1<sup>st</sup> grade of ISCED3} ----- ☐ 1
- {6<sup>th</sup> grade of ISCED2} ----- ☐ 2
- {5<sup>th</sup> grade of ISCED2} ----- ☐ 3
- {4<sup>th</sup> grade of ISCED2} ----- ☐ 4
- {3<sup>rd</sup> grade of ISCED2} ----- ☐ 5
- {2<sup>nd</sup> grade of ISCED2} ----- ☐ 6
- {1<sup>st</sup> grade of ISCED2} ----- ☐ 7
- {6<sup>th</sup> grade of ISCED1} ----- ☐ 8
- {5<sup>th</sup> grade of ISCED1} ----- ☐ 9
- {4<sup>th</sup> grade of ISCED1} ----- ☐ 10
- {3<sup>rd</sup> grade of ISCED1} ----- ☐ 11
- {2<sup>nd</sup> grade of ISCED1} ----- ☐ 12
- {starting grade of ISCED1} ----- ☐ 13
- Before {the starting grade of ISCED1 } ----- ☐ 14

#### WebTrans Note

The same response instruction (Please tick as many boxes as applicable) occurs in TQ7, TQ11, TQ14, TQ24, TQ36, TQ37, and TQ38.

Please note that the number of response categories can differ between countries.

#### NRC note

TQ36, TQ37 and TQ38 provide information about the experience of teachers in teaching languages. This question provides information on the different grades in which the teacher has taught [target language]. In which study programs and grades a teacher can teach depends in several countries on the qualifications and experience of the teacher.

The response categories of this question need to be localised (identical to SQ39). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3 should be presented as response categories. It is necessary to phrase the response categories in such a way that teachers will easily understand.



The number of response categories equals the total duration of primary education (ISCED1), lower secondary education (ISCED2) and the first two years of upper secondary education (ISCED3) plus one for the pre-ISCED1 period. For example, if in the Educational system both ISCED1 and ISCED2 last four years, then the four grades of ISCED1, the four grades of ISCED2 and two grades of ISCED3 should be presented as response categories (in total 11 response categories).

Please instruct the reconciler which response categories to use.

#### **Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

[Grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

---

39 In general, how many students are there in your classroom during [target language] lessons?

*(Please write down the number of students)*

*Students*

---

#### **NRC note**

This question gives an indication of the class size of the [target language] classes. The general class size for [target language] lessons can differ from the general class size for other subjects.

#### **Translator note**

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

40	How often have you used the Common European Framework of Reference for the following?				
		<i>(Please select only one answer from each row)</i>			
		<i>Never</i>	<i>Sometimes</i>	<i>Quite often</i>	<i>Very often</i>
1)	For curriculum or syllabus development-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	For teacher training-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	For testing or assessment -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	For the development or selection of instructional materials-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	For communication with stakeholders, such as students, other teachers, parents, etc. -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never-Sometimes-Quite often-Very often) occurs in TQ40 and TQ54.

### NRC note

This question assesses the use and purpose of use of the CEFR.

41	During the past <u>three</u> years, how often were you involved in the organisation of the following?					
		<i>(Please select only one answer from each row)</i>				
		<i>Not at all</i>	<i>Once</i>	<i>Twice</i>	<i>Three times</i>	<i>Four times or more</i>
1)	School trips to a [target language] speaking country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	School trips to another (non-[target language] speaking) country -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Visits to your school by a school class from a [target language] speaking country -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Visits to your school by a school class from another (non-[target language] speaking) country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

#### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Not at all - Once - Twice - Three times - Four times or more) occurs in TQ41 and TQ42.

#### NRC note

TQ41 and TQ42 provide information on the teachers' involvement in creating opportunities for exchange visits and school language projects. This question allows comparison with the students' report of the received opportunities (SQ45 and SQ46).

In Educational systems that are not an entire country an adequate adaptation of the items of this questions that can be understood by teachers might be needed. The items should match the items of SQ45. Please instruct the reconciler how to adapt the items.

#### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

three - Please underline the appropriate text in the translation.

42	In the past <u>three</u> years, how often were you involved in the organisation of the following activities at school?	(Please select only one answer from each row)				
		<i>Not at all</i>	<i>Once</i>	<i>Twice</i>	<i>Three times</i>	<i>Four times or more</i>
1)	Collaboration project with schools abroad-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Language clubs -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Language competition -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	European Day of Languages -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Extracurricular language projects-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Pen friends, email or MSN friends for your students -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Excursions and field trips related to foreign language education -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Not at all - Once - Twice - Three times - Four times or more) occurs in TQ41 and TQ42.

### NRC note

TQ41 and TQ42 provide information on the teachers' involvement in creating opportunities for exchange visits and school language projects. This question allows comparison with the students' report of the received opportunities (SQ45 and SQ46). The items of this question should match the items of SQ46.

The items in this question might need to be adapted. If necessary, provide examples appropriate for the [Educational system] of language clubs and language competitions. Please use items that match the items of SQ46.

**Language clubs** – A group of students that meets regularly and organized for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

**A language competition** – an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

**European Day of Languages** – The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.



MSN – is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

### **Translator note**

Language clubs – A group of students that meets regularly for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition – an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

European Day of Languages – The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

MSN – is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

three - Please underline the appropriate text in the translation.

## About the available resources for your [target language] classes

*In this section you will be asked some questions about the resources for teaching [target language]. A general indication of the resources' use across all {grades of ISCED2 and first two years of ISCED3} is adequate for the purposes of the study.*

### **NRC note**

The phrase {grades of ISCED2 and first two years of ISCED3} needs to be localised. The phrase should refer to all the grades from the onset of ISCED2 education until the 2<sup>nd</sup> year of ISCED3. It is necessary to phrase in such a way that teachers will easily understand.

[Grade] – refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

Please instruct the reconciler which phrase to use.

### **Translator note**

Please do not translate the phrase in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate phrase.

[Grade] – refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

43	How often do you use the following devices at school for teaching [target language]?					
		<i>(Please select only one answer from each row)</i>				
		<i>Never, because it is not available</i>	<i>Hardly ever or never</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>(Almost) every week</i>
1)	A teacher PC or laptop in the classroom	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	A projector in the classroom -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Interactive whiteboard in the classroom--	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Multimedia language lab (teacher PC and student PCs <u>with</u> specific language learning software) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Multimedia lab (teacher PC and student PCs <u>without</u> specific language learning software) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	An internet connection in the classroom	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	A virtual learning environment to support language teaching and learning, e.g. Moodle, WebCT, Blackboard, Fronter, Sakai -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

### NRC note

TQ5, TQ6, TQ43, TQ44, TQ45 and TQ47 provide information on the use of ICT in foreign language teaching. This question provides information about the ICT facilities in school.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

with – Please underline the appropriate text in the translation.

without – Please underline the appropriate text in the translation.



44	In general, how often do you or your students use the following resources for a [target language] class you teach?	<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1)	Audio-cassettes, CDs or other audio-material in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Video cassettes, DVDs, video clips from YouTube or other audio-visual material -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Newspapers, magazines, comics or song texts written in [target language] ---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Internet -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Computer programmes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Language laboratory (student PCs with specific language software) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Textbook for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8)	Books written in [target language] for extensive reading e.g. fiction -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9)	Lesson material made by you or your colleagues (e.g. hand-outs, reading texts) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) occurs in TQ44, TQ45, TQ56, and TQ57.

### NRC note

TQ5, TQ6, TQ43, TQ44, TQ45 and TQ47 provide information on the use of ICT in foreign language teaching. This question allows us to compare the use of ICT with the use of other resources.

The question and items are similar (or identical) to SQ51.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

45	In general, how often do you or your students use the following <u>ICT facilities</u> for a [target language] class you teach?	<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1)	Software or websites specifically designed for learning languages-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Online dictionaries and other reference works-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Online news media (TV, radio, newspapers) in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Other websites on life and culture in [target language] speaking country/countries -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Communication tools, e.g. email, chatting, blogging, {MySpace}, {Skype}--	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Custom made tools developed in house for learning and teaching languages-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Online portfolio -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8)	Tools for language assessment-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9)	Language webquest -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) occurs in TQ44, TQ45, TQ56, and TQ57.

### NRC note

TQ5, TQ6, TQ43, TQ44, TQ45 and TQ47 provide information on the use of ICT in foreign language teaching. This question provides information on the use of ICT for the [target language] classes.

Tools – An application program for the computer or an element of a computer program.

### Translator note

Tools – An application program for the computer or an element of a computer program.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

ICT facilities – Please underline the appropriate text in the translation.

46	In your opinion, how useful are the textbooks or instruction materials <u>you use</u> for the following?	(Please select only one answer from each row)			
		<i>Not useful at all</i>	<i>Hardly useful</i>	<i>Quite useful</i>	<i>Very useful</i>
1)	Teaching to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Teaching to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Teaching to understand spoken [target language] --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Teaching [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Teaching to read [target language] texts-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Teaching to pronounce [target language] correctly--	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Teaching [target language] vocabulary-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8)	[target language] culture and literature -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9)	For lesson planning -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question assesses the usefulness of the textbook for learning or teaching the four communicative skills and language content. This question allows us to compare the perception of the teacher with the perception of the students regarding the usefulness of the textbook(s)(SQ52).

The items should match the items of SQ58 and the response scales should be identical. Furthermore the items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

you use - Please underline the appropriate text in the translation.

47	In general, how often do your students have to use a computer for the following?	<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>A few times a week</i>	<i>(Almost) every day</i>
1)	For finding information for [target language] homework or assignments-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	For [target language] homework or assignments -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	For learning to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	For learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	For learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	For learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	For learning to read [target language] texts-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8)	For learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
9)	For learning [target language] vocabulary-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-A few times a month-A few times a week-(Almost) every day) occurs in TQ5, TQ6, and TQ47.



### **NRC note**

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question provides information on the use of ICT for the teaching of [target language] skills and, along with TQ6, TQ7, TQ48, TQ49, and TQ50, provides information on the use of ICT in foreign language teaching.

The items are identical to the items of SQ62 which allows comparing the view of the teacher with the view of the student.

### **Translator note**

homework or assignments – All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About your [target language] classes

*In this section you will be asked some questions about your [target language] classes and your students. A general indication of your lessons and students across all {grades of ISCED2 and first two years of ISCED3} is adequate for the purposes of the study.*

### **NRC note**

The phrase {grades of ISCED2 and first two years of ISCED3} needs to be localised. The phrase should refer to all the grades from the onset of ISCED2 education until the 2<sup>nd</sup> year of ISCED3. It is necessary to phrase in such a way that teachers will easily understand.

[Grade] – refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

Please instruct the reconciler which phrase to use.

### **Translator note**

Please do not translate the phrase in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate phrase.

[Grade] – refers to the administrative level of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

48	In your opinion, how difficult is it in general for your students to learn the following?				
		<i>(Please select only one answer from each row)</i>			
		<i>Very easy</i>	<i>Quite easy</i>	<i>Quite difficult</i>	<i>Very difficult</i>
1)	Learning to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Learning to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Learning [target language] vocabulary -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The items 1) through 7) of TQ48, TQ55, and TQ56 are identical. The item 8) of TQ55 and TQ56 are identical as well.

#### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question assesses the perceived difficulty of learning the four communicative skills and language content. This question allows us to compare the perception of the teacher with the perception of the students (SQ48). The items of this question are identical to the items of SQ48. Please note that the items 1) through 7) of TQ48, TQ55, and TQ56 are identical.

#### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

49	In general, how often do you speak [target language] when you do the following in a [target language] lesson?	<i>(Please select only one answer from each row)</i>				
		<i>Never</i>	<i>Hardly ever</i>	<i>Every now and then</i>	<i>Usually</i>	<i>Always</i>
1)	When you speak to the whole class-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	When you talk with one or two students	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never-Hardly ever-Every now and then-Usually-Always) occurs in TQ49, TQ50, and TQ51.

### NRC note

TQ49 and TQ50 provide information on the use of the target language during foreign language lessons. The questions and items match SQ49 and SQ50.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

50	In general, how often do your students speak [target language] when they do the following in a [target language] lesson?	<i>(Please select only one answer from each row)</i>				
		<i>Never</i>	<i>Hardly ever</i>	<i>Every now and then</i>	<i>Usually</i>	<i>Always</i>
1)	When students speak to you-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	When students work in groups and speak together -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	When students speak in front of the whole class-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never-Hardly ever-Every now and then-Usually-Always) occurs in TQ49, TQ50, and TQ51.



### NRC note

TQ49 and TQ50 provide information on the use of the target language during foreign language lessons. The questions and items match SQ49 and SQ50.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

51	In general, how often do you do the following during a [target language] lesson?	(Please select only one answer from each row)				
		Never	Hardly ever	Every now and then	Usually	Always
1)	Let the students work in mixed-ability groups-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Let the students work in same-ability groups-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Let the students work individually-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Let a group of students speak in front of the whole class-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Let an individual student speak in front of the whole class-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never-Hardly ever-Every now and then-Usually-Always) occurs in TQ49, TQ50, and TQ51.

### NRC note

This question provides information about the within class ability grouping (setting) and allows comparison with the report of students about working individually or in groups (SQ53).

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

52 In general, how much of the time do you give instruction to the whole class during a [target language] lesson?

*(Please select only one answer)*

Not at all ----- ☐ 0

Some of the time ----- ☐ 1

About half of the time ----- ☐ 2

Most of the time ----- ☐ 3

All the time ----- ☐ 4

---

### WebTrans Note

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

### NRC note

This question provides information about whole class teaching.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

53	In general, how often do you teach the following to a [target language] class?					
		<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>A few times a month</i>	<i>A few times a week</i>	<i>(Almost) every lesson</i>
1)	Teaching to write in [target language] ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Teaching to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Teaching to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Teaching [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Teaching to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Teaching to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Teaching [target language] vocabulary -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8)	Teaching about [target language] culture or literature -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

#### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

#### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question assesses the emphasis during lessons and allows comparison with the perception of students (SQ58).

The items should match the items of SQ58 and the response scales should be identical. Furthermore the items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

#### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

54	In general, how often do you point out similarities between [target language] and other languages (including [Questionnaire language]) when teaching the following to one of your classes?				
		<i>(Please select only one answer from each row)</i>			
		<i>Never</i>	<i>Sometimes</i>	<i>Quite often</i>	<i>Very often</i>
1)	Teaching to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Teaching to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Teaching to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Teaching [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Teaching to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Teaching to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Teaching [target language] vocabulary -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never-Sometimes-Quite often-Very often) occurs in TQ40 and TQ54.

The items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

#### NRC note

This question provides information about the emphasis on similarities between known languages and allows comparison with the perception of students (SQ58).

The items should match the items of SQ58 and the response scales should be identical. Furthermore the items 1) through 8) of TQ46, TQ53, and TQ54 are identical.

#### Translator note

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

55	In your opinion, how important is it that your students learn the following?				
		<i>(Please select only one answer from each row)</i>			
		<i>Not important at all</i>	<i>Hardly important</i>	<i>Quite important</i>	<i>Very important</i>
1)	Learning to write in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Learning to read [target language] texts -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Learning [target language] vocabulary -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8)	Learning about [target language] culture and literature -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Not important at all-Hardly important-Quite important-Very important) occurs in TQ55 and TQ59.

The items 1) through 7) of TQ48, TQ55, and TQ56 are identical. The item 8) of TQ55 and TQ56 are identical as well.

### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question assesses the perceived importance of learning the four communicative skills and language content.

Please note that the items 1) through 7) of TQ48, TQ55, and TQ56 are identical. The item 8) of TQ55 and TQ56 are identical as well.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

## About homework and assessment

56	In general, how often do you give a [target language] class homework or assignments aimed at the following?	<i>(Please select only one answer from each row)</i>				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1)	Learning to write in [target language] ----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Learning to speak [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Learning to understand spoken [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Learning [target language] grammar -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Learning to read [target language] texts	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Learning to pronounce [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Learning [target language] vocabulary --	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
8)	Learning about [target language] culture and literature -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) occurs in TQ44, TQ45, TQ56, and TQ57.

The items 1) through 7) of TQ48, TQ55, and TQ56 are identical. The item 8) of TQ55 and TQ56 are identical as well.

### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question provides information on the emphasis in homework.

Please note that the items 1) through 7) of TQ48, TQ55, and TQ56 are identical. The item 8) of TQ55 and TQ56 are identical as well.

### Translator note

homework or assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

57	In general, how often do you give the following to a [target language] class you teach?	(Please select only one answer from each row)				
		<i>Never or hardly ever</i>	<i>A few times a year</i>	<i>About once a month</i>	<i>A few times a month</i>	<i>(Almost) every lesson</i>
1)	Homework or assignments -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Written or oral comments to homework or to an assignment that your students made-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	A [target language] test or assignment that is marked or scored -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Written or oral comments to a test or assignment your students made -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

**Removed apostrophe!**

### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Never or hardly ever-A few times a year-About once a month-A few times a month-(Almost) every lesson) occurs in TQ44, TQ45, TQ56, and TQ57.

### NRC note

This question provides information on the frequency of homework and assignment and can be compared with the students' report of the time spent on homework and study for the subject of [target language] (SQ63\_1).

Items 2) and 3) allow comparison to the students report about the frequency of feedback (SQ59).

### Translator note

homework or assignments – All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

58 In your opinion, how much time should your students spend each week on study and homework for [target language]?

*(Please select only one answer)*

Zero hours----- ☐ 0

Less than one hour a week ----- ☐ 1

About one to two hours a week ----- ☐ 2

About two to three hours a week ----- ☐ 3

More than three hours a week----- ☐ 4

---

### WebTrans Note

The same response instruction (Please select only one answer) occurs in TQ2, TQ3, TQ9, TQ19, TQ21, TQ23, TQ35, TQ52, TQ58, TQ13, and TQ25.

### NRC note

This question provides information about the expected learning time a week and allows comparison with students' reported learning time (SQ59, SQ60 and SQ63).

### Translator note

Homework and assignments - All the school work and tasks that teachers give to students to do outside the lessons, for example preparing an oral presentation, writing a paper or learning words.

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



59	How important are the following when <u>you</u> determine a mark for the final grade of students for the subject of [target language]?				
		<i>(Please select only one answer from each row)</i>			
		<i>Not important at all</i>	<i>Hardly important</i>	<i>Quite important</i>	<i>Very important</i>
1)	Writing [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Speaking [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Understanding spoken [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Knowing [target language] grammar well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Reading [target language] well -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Pronouncing [target language] correctly -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Knowing [target language] vocabulary well---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8)	Knowledge about [target language] culture and/or literature -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans Note

The same response instruction (Please select only one answer from each row) occurs in TQ5, TQ6, TQ40, TQ41, TQ42, TQ43, TQ44, TQ45, TQ46, TQ47, TQ48, TQ49, TQ50, TQ51, TQ53, TQ54, TQ55, TQ56, TQ57, and TQ59.

The same response scale (Not important at all-Hardly important-Quite important-Very important) occurs in TQ55 and TQ59.

#### NRC note

TQ46, TQ47, TQ48, TQ53, TQ55, TQ56 and TQ59 provide information on the emphasis on the four communicative skills and language content within the teaching activities and resources used. This question assesses the emphasis within the final grade and allows comparison with students' perception (SQ61).

#### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

you - Please underline the appropriate text in the translation.



Thank you very much for the time and effort you have put into responding to this questionnaire.



## Principal's Questionnaire for the European Survey on Language Competences 2011



# Main Study

## About this full note version

Accompanying each question are:

- **Notes for WebTrans** indicating the recurring question elements that have to be linked in WebTrans, so they have to be translated only once.
- **Notes for the NRC** clarifying terms and options, noting where localisations should be made, and providing a rationale for the question's inclusion.
- **Notes for the translator** clarifying terms and options, noting where response categories and/or terms should not be translated, because they have to be localised.

## Conventions in this document

Terms in curly brackets { } should not be translated, but localised. The NRC is asked to instruct the reconciler which localised (a term which is appropriate for the country) terms to insert.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.

The curly brackets { } and square brackets [ ] should not appear in the translated text. Please instruct the translators and reconciler to remove these brackets from the translated text.

The term [target language] refers to the language for which the students will be tested prior to filling out the questionnaire, in other words the 1<sup>st</sup> most widely taught language among English, French, German, Spanish and Italian. Please instruct the translator which language to fill in when the term [target language] appears.

The term "SQ" is used as a reference to a question in the Student Questionnaire, "TQ" as a reference to the Teacher Questionnaire and "PQ" as a reference to the Principal Questionnaire.

Please note that all questions should be translated even when it is expected that all respondents in your country will give the same answer. Question order is known to have an effect upon the answers. Removal of questions will compromise the comparability across countries and the comparability with future cycles.

Overall we have tried to prevent questions that require an open-ended text response as much as possible, as the coding of such questions (an NRC task) is very time consuming and costly.

The current lay-out of the questions is not final, as we need to see how the lay-out works for all translated questionnaires (most languages are less concise than English) and in the testing tool. The lay-out and the boxes for answering are added, to allow pre-testing the translated questionnaires.

## Translator note

Throughout the questionnaire the formal address is used.

"You" is singular unless otherwise indicated.

Terms in curly brackets { } should not be translated, but localised. The reconciler and/or NRC should insert a phrase or word which is appropriate for the country.

Terms in square brackets [ ] means that the translator should replace the term with a term which is appropriate. In some cases adaptation is required; for example [Educational system]. In other cases adaptation is optional; for example [grade] may not need adaptation, and may be directly translated.



Your school has agreed to participate in the European Survey on Language Competences, a large European study of student learning of foreign languages, launched by the European Commission.

This survey aims to provide Member States, policy makers, teachers and practitioners with information on the foreign language competence of students enrolled in {secondary education} and to provide insight into the progress towards the objective of improving foreign language learning.

In this questionnaire you will find questions about:

- The school's characteristics
- The school's teaching staff
- In-service training for the school's teaching staff
- The school's curriculum for foreign languages
- The teaching time for foreign languages
- The school's policy and practices to encourage language learning
- The school's resources

This questionnaire is addressed to school principals who are asked to supply information about their schools. Since your school has been selected as part of a nationwide sample, your responses are very important in helping to describe the foreign language education in [Educational system].

Since ESLC is an international study and all countries are using the same questionnaire, you may find that some of the questions seem unusual or are not entirely relevant to you or schools in [Educational system]. Nevertheless, it is important that you do your best to answer all of the questions so comparisons can be made across countries in the study.

Please read the questions carefully and answer each question. If you do not know a precise answer, your best estimate will be adequate for the purposes of the study.

Completing the questionnaire will take between 30 and 45 minutes. If you want to stop and continue filling out the questionnaire at a later time, please press the button "SAVE". The next time you will log on, you can continue where you left off.

If you have completed the entire questionnaire please press "SEND". Please do NOT press the button "SEND" before you have completed the questionnaire, as you will not be able to complete the questionnaire after sending it.



All your answers will be kept confidential and secret. It will be impossible to identify individuals from the combined responses.

**NRC note**

The term {secondary education} in this instruction might need to be localised. Please instruct the reconciler which terms should be used.

**Translator note**

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Educational system] – is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

SAVE - Please underline the appropriate text in the translation.

## About the school's characteristics

1	As of {February 1, 2011}, what was the number of students enrolled in your school?	
	<i>(Please write down the number of students)</i>	
		<i>Number of students</i>
1)	Number of male students -----	<input type="text"/>
2)	Number of female students -----	<input type="text"/>
3)	TOTAL number of students -----	<input type="text"/>

### WebTrans note

The same response instruction (Please write down the number of students) occurs in PQ1, PQ28 and PQ29.

The same response label (Number of students) occurs in PQ1, PQ27, PQ28, and PQ29.

### NRC note

This question provides information about the school size at the time of the survey and can also be used to determine whether a school is co-educational or single sex. It will also be used for calculating the computer-to-student ratio and the teacher-to-student ratio. This ratio can be compared for example to the results of PISA. Furthermore, the question can be used for validation purposes (the proportion of boys and girls in the sample - SQ1).

The date {February 1, 2011}, the start of the testing period in the countries, might need to be localised.

Please instruct the reconciler which date should be used.

### Translator note

Please do not translate terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

2	<u>Approximately</u> what percentage of students in your school has the following characteristics?					
	<i>(Please select only one answer from each row. A rough estimate of the percentage is adequate for the purposes of the study)</i>					
		<i>Less than 10%</i>	<i>11-20%</i>	<i>21-40%</i>	<i>41-60%</i>	<i>More than 60%</i>
1)	Come from economically disadvantaged homes? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Come from economically affluent homes? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Come from a country other than [Educational system]? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Have lived for less than 1 year in [Educational system]? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Do not have [Questionnaire language] as their first language? -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

#### NRC note

This question provides information on the percentage of immigrant students (item 3), 4) and 5)), which can be compared to the results of PISA, TALIS and PIRLS. Items 1) and 2) about the percentage of students from economically advantaged and disadvantaged homes are highly similar to the questions in PIRLS about the students characteristics.

Furthermore, the question can be used for validation purposes: the languages in the students' home environment and wealth.

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by principals is advisable.

#### Translator note

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

[Educational system] – is the country in which the translated questionnaire will be administered. Please replace the term “Educational system” with the name of the country.

Approximately - Please underline the appropriate text in the translation.



3	How often are the following factors considered when students are admitted to your school?					
		<i>(Please select only one answer from each row)</i>				
		<i>Never</i>	<i>Hardly ever</i>	<i>Every now and then</i>	<i>Usually</i>	<i>Always</i>
1)	Residence in a particular area-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Student's record of their language skills (including placement tests) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Student's record of academic performance (including placement tests) in all subjects -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Recommendation of feeder schools-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Parents' endorsement of the instructional or religious philosophy of the school-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
6)	Whether the student requires, or is interested, in a special programme-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
7)	Preference given to family members of current or former students -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

### NRC note

This question provides information on the admission criteria, which is particularly relevant because some schools are specialised in languages and some schools in upper secondary education (ISCED3) use academic performance criteria. This question is highly similar to the questions in PISA and TALIS about admission criteria.

4	What is your school's policy on organising instruction for students with different abilities?			
	<i>(Please select only one answer from each row)</i>			
		<i>Not for any subject</i>	<i>For some subjects</i>	<i>For all subjects</i>
1)	Students are grouped by ability into different study programs -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
2)	Students of the same study program are grouped by ability into different classes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
3)	Students of the same study program are grouped by ability within their classes -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2

### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

### NRC note

This question provides information on the within school streaming based on general ability. In contrast, the sampled teachers and students report on the streaming in the [target language] lessons only (SQ53, TQ51). The question is highly similar to the PISA question about streaming.

## 5 Which of the following [grades] do you have in your school?

*(Please tick as many boxes as applicable)*

- {Lower grades than the starting grade of ISCED1}--- ☐ 0
- {starting grade of ISCED1}----- ☐ 1
- {2<sup>nd</sup> grade of ISCED1}----- ☐ 2
- {3<sup>rd</sup> grade of ISCED1}----- ☐ 3
- {4<sup>th</sup> grade of ISCED1}----- ☐ 4
- {5<sup>th</sup> grade of ISCED1}----- ☐ 5
- {6<sup>th</sup> grade of ISCED1}----- ☐ 6
- {1<sup>st</sup> grade of ISCED2}----- ☐ 7
- {2<sup>nd</sup> grade of ISCED2}----- ☐ 8
- {3<sup>rd</sup> grade of ISCED2}----- ☐ 9
- {4<sup>th</sup> grade of ISCED2}----- ☐ 10
- {5<sup>th</sup> grade of ISCED2}----- ☐ 11
- {6<sup>th</sup> grade of ISCED2}----- ☐ 12
- {1<sup>st</sup> grade of ISCED3}-----
- {2<sup>nd</sup> grade of ISCED3}-----
- {Higher grades than 2<sup>nd</sup> grade of ISCED3}-----

### WebTrans note

The same response instruction (Please tick as many boxes as applicable) occurs in PQ5, PQ6, PQ21, and PQ22.

The response categories of PQ5 are identical to the items of PQ23, PQ24, and PQ25.

Please note that the number of response categories can differ between countries.

### NRC note

PQ5, PQ6, and PQ7 provide information on the organisational structure of the educational system. This question provides information on the grades within the school and is similar to the PISA question about grades within the school. This information is crucial to determine whether the school does not offer foreign languages (see the later questions 29, 30 and 31) in some of the grades the sampled students might have had, simply because the school does not offer these grades.



The response categories of this question need to be localised (similar to SQ39, SQ40, and TQ38). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. The later grades (after the 2<sup>nd</sup> year of ISCED3) do not have to be presented separately. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

**Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

## 6 Which of the following [programmes] does your school offer?

*(Please tick as many boxes as applicable)*

- {Programme 1}----- ☐ 0
- {Programme 2}----- ☐ 1
- {Programme 3}----- ☐ 2
- {Programme 4}----- ☐ 3
- {Programme 5}----- ☐ 4
- {Programme 6}----- ☐ 5
- {Programme 7}----- ☐ 6
- {Programme 8}----- ☐ 7
- {Programme 9}----- ☐ 8
- {Programme 10}----- ☐ 9
- {Programme 11}----- ☐ 10
- {Programme 12}----- ☐ 11
- {Programme 13}----- ☐ 12
- {Programme 14}-----
- {Programme 15}-----

### WebTrans note

The same response instruction (Please tick as many boxes as applicable) occurs in PQ5, PQ6, PQ21, and PQ22.

Please note that the number of response categories can differ between countries.

### NRC note

PQ5, PQ6, and PQ7 provide information on the organisational structure of the educational system. This question provides information on the study program(s) the school offers.

The response categories of this question need to be localised (similar to SQ6 and TQ37). The study programmes presented as response categories should correspond to the study programmes at ISCED2 and ISCED3 level in the Localisation file (Study Program Table). It is necessary to phrase study programme labels in such a way that principals will easily understand.

In some countries there may not be an administrative or structural boundary between some successive ISCED levels (between ISCED 2 and 3) in the educational system. In these cases one should ask about completion of the grade/school year that can be defined as an implicit boundary between the ISCED-levels.



In countries where students are grouped in school according to their ability, an adequate adaptation of this question that can be understood by principals is advisable.

Please instruct the reconciler which study programs should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

[programmes] – the study programmes the student can follow in secondary education.

---

## 7 Is your school a public or a private school?

A {public school} is a school managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise.

A {private school} is a school managed directly or indirectly by a non-government organisation; e.g. a church, trade union, business, or other private institution.

So, the distinction between {public} and {private} does not rely on the sources of funding.

*(Please select only one answer)*

{A public school} ----- ☐ 0

{A private school} ----- ☐ 1

---

### WebTrans note

The same response instruction (Please select only one answer) occurs in PQ7, PQ8, and PQ16.

### NRC note

PQ5, PQ6, and PQ7 provide information on the organisational structure of the educational system. This question is highly similar to the PISA and TALIS questions about the school management.

The terms {public school} and {private school} might need to be localised in some Educational systems.

A {public school} is a school managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise.

A {private school} is a school managed directly or indirectly by a non-government organisation; e.g. a church, trade union, business, or other private institution.

Please instruct the reconciler which term to use.

### Translator note

Please do not translate the terms and response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms and response categories.

8 Which of the following definitions best describes the community in which your school is located?

*(Please select only one answer)*

A village, hamlet or rural area (fewer than three thousand people) ----- ☐ 0

A small town (three thousand to around fifteen thousand people) ----- ☐ 1

A town (fifteen thousand to around hundred thousand people) ----- ☐ 2

{A city (hundred thousand to around one million people)} ----- ☐ 3

{A large city with over one million people} ----- ☐ 4

---

**WebTrans note**

The same response instruction (Please select only one answer) occurs in PQ7, PQ8, and PQ16.

**NRC note**

This question provides information on the school location. A highly similar question is posed in PISA and TALIS.

The response categories are identical to the response categories of SQ3.

Some of the items might need to be localised. Please instruct the reconciler which items should be used.

{A city (hundred thousand to around one million people)} - If a town of this size does not exist in the Educational system, this response option can be omitted.

{A large city with over one million people} - If a town of this size does not exist in the Educational system, this response option can be omitted.

**Translator note**

Community - refers to the village, town or city, not to the larger administrative municipality.

## About the school's teaching staff

9	As of {February 1, 2011}, what was the number of teachers in your school?	
	<i>(Please write down the number of teachers)</i>	
		<i>Number of teachers</i>
1)	Number of male teachers -----	<input type="text"/>
2)	Number of female teachers-----	<input type="text"/>
3)	TOTAL number of teachers -----	<input type="text"/>

### WebTrans note

The same response label (Number of teachers) occurs in PQ9, PQ10, and PQ17.

### NRC note

PQ9 and PQ10 provide information on the teaching staff. This question provides information on the size of the teaching staff at the time of the survey. It will be used for calculating the computer-to-teacher ratio and the teacher-to-student ratio. Furthermore, this question can be used for validation purposes (the proportion of female and male teachers in the sample - TQ1).

The date {February 1, 2011}, the start of the testing period in the countries, might need to be localised. Please instruct the reconciler which date should be used.

### Translator note

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.



10	This school year, how many teachers are scheduled to teach the following?	
	<p>(Please write down the number of teachers. Write 0 (zero) if there are none.</p> <p>The number of teachers scheduled for languages is <u>not</u> the sum of the teachers for [questionnaire language], [target language] and other foreign languages, because teachers can be scheduled for more than one language)</p>	
		Number of teachers
1)	[Questionnaire language] -----	<input type="text"/>
2)	[Target language] -----	<input type="text"/>
3)	Foreign languages (including ancient languages) other than [target language] -----	<input type="text"/>
4)	All languages together (TOTAL) -----	<input type="text"/>

#### WebTrans note

The same response label (Number of teachers) occurs in PQ9, PQ10, and PQ17.

#### NRC note

PQ9 and PQ10 provide information on the teaching staff. This question provides information on the size of the teaching staff for languages.

#### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

not - Please underline the appropriate text in the translation.

11	During the past five years, have you had difficulty in filling teaching vacancies or covering for absent teachers for the following subjects?			
	<i>(Please select in each row No or Yes)</i>			
		<i>No</i>	<i>Yes</i>	
1)	For languages-----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	For [questionnaire language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	For [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	For foreign languages (including ancient languages) other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	For subjects other than languages -----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

### NRC note

PQ11 and PQ12 provide information on teacher shortage, which might be a cause of the employment or scheduling of inadequately qualified teachers (see TQ19). The information obtained can be compared with the teacher shortage reported in TIMSS, PIRLS, TALIS and PISA.

### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

12	In the past five years, have you had to use the following measures for longer than three weeks for the subject of [target language] to fill teaching vacancies or to cover for absent teachers?			
		<i>(Please select in each row No or Yes)</i>		
		No	Yes	
1)	Increase the teaching hours of teachers who are teaching [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Cancel [target language] lessons-----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	Increase class sizes for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	Reassign teachers qualified for other languages to teach [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	Employ teachers qualified for other languages to teach [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
6)	Reassign teachers qualified for subjects other than languages to teach [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
7)	Employ teachers qualified for subjects other than languages to teach [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
8)	Employ teachers qualified for [target language] who are over the retirement age-----	<input type="radio"/> 0	<input type="radio"/> 1	
9)	Employ [target language] teachers who had not as yet completed their teacher training -----	<input type="radio"/> 0	<input type="radio"/> 1	
10)	Employ native speakers who have no teaching qualification to teach [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

#### NRC note

PQ11 and PQ12 provide information on teacher shortage. This question provides information on the measures that can be taken to deal with teacher shortage, see the Eurydice report on the teaching profession in Europe (2003; Volume 2: Supply and demand).

#### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).



## About in-service training for the school's teaching staff

*In this section you will be asked some questions about in-service training.*

*In-service training is training that teachers receive during their career to update, develop and broaden the knowledge acquired during initial training and/or to provide them with new skills and professional understanding which teachers may not have at a given stage in their career. The training might take different forms, such as a course, a conference or a seminar.*

### NRC note

The explanation of in-service training also occurs in the Teacher Questionnaire.

- 
- 13 Is participation in in-service training an obligation, a right or an option for the teachers on your staff?

*(Please select the answers that describe the situation in your school best)*

Participation in in-service training is an obligation for teachers ----- ☐ 0

Participation in in-service training is a right for teachers ----- ☐ 1

Participation in in-service training is required for promotion ----- ☐ 2

Participation in in-service training is optional ----- ☐ 3

---

### NRC note

PQ13, PQ14, PQ15, and PQ16 provide information on the incentives for in-service training. Schools have growing autonomy and freedom to develop plans for in-service training at school-level (see in Eurydice report on the teaching profession in Europe (2003; Volume 3: Working conditions and pay). This question assesses whether within the school in-service training is an obligation, right or option.

The response categories are identical to the response categories of TQ30.

14	Which of the following financial compensations can teachers get from <u>our school</u> for participation in in-service training?			
		<i>(Please select in each row No or Yes)</i>		
		<i>No</i>	<i>Yes</i>	
1)	Payment of enrolment costs of training -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Payment of other training-related expenditure -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	Paid leave during training with no loss of earnings-----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	Increase in salary afterwards-----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The items of PQ14 and PQ15 are identical.

#### NRC note

PQ13, PQ14, PQ15, and PQ16 provide information on the incentives for in-service training. Schools have growing autonomy and freedom to develop plans for in-service training at school-level (see in Eurydice report on the teaching profession in Europe (2003; Volume 3: Working conditions and pay). This question assesses the financial compensations for in-service training teachers can get from the school.

The response categories are identical to the response categories of TQ34.

#### Translator note

your school - Please underline the appropriate text in the translation.

15	Which of the following financial compensations can teachers get for participation in in-service training <u>from sources other</u> than your school, for example, funds or the national or local government?			
	<i>(Please select in each row No or Yes)</i>			
		No	Yes	
1)	Payment of enrolment costs of training -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Payment of other training-related expenditure -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	Paid leave during training with no loss of earnings-----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	Increase in salary afterwards-----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The items of PQ14 and PQ15 are identical.

#### NRC note

PQ13, PQ14, PQ15, and PQ16 provide information on the incentives for in-service training. Schools have growing autonomy and freedom to develop plans for in-service training at school-level (see in Eurydice report on the teaching profession in Europe (2003; Volume 3: Working conditions and pay). This question assesses the financial compensations for in-service training from other sources than the school.

The response categories are identical to the response categories of TQ34.

#### Translator note

from other sources - Please underline the appropriate text in the translation.

16 At your school, when are teachers normally allowed to participate in in-service training?

(Please select only one answer)

During working hours with a substitute teacher for his or her classes ----- ☐ 0

During working hours but not during teaching hours (a substitute teacher for his or her classes is not organised) ----- ☐ 1

Only outside working hours ----- ☐ 2

#### WebTrans note

The same response instruction (Please select only one answer) occurs in PQ7, PQ8, and PQ16.

#### NRC note

PQ13, PQ14, PQ15, and PQ16 provide information on the incentives for in-service training. Schools have growing autonomy and freedom to develop plans for in-service training at school-level (see in Eurydice report on the teaching profession in Europe (2003; Volume 3: Working conditions and pay). This question assesses when teachers are allowed by the school management to participate in in-service training

The response categories are similar to the response categories of TQ35.

#### Translator note

Only - Please underline the appropriate text in the translation.

17	In the previous school year, how many teachers participated in teacher exchange visits to work or study in another country for longer than one month?	
	(Please write down the number of teachers. If none of the teachers on your school's staff took leave for work or study abroad please write 0)	
		Number of teachers
1)	Teachers of [target language] -----	<input type="text"/>
2)	Teachers of languages other than [target language] -----	<input type="text"/>
3)	Teachers of subjects other than languages -----	<input type="text"/>
4)	TOTAL number of teachers -----	<input type="text"/>

#### WebTrans note

The same response label (Number of teachers) occurs in PQ9, PQ10, and PQ17.



### NRC note

PQ17 and PQ18 provide information on the participation in exchange visits and guest teachers from other countries or language communities.

Furthermore, this question can be used for validation purposes (the reported exchange visits of the teachers in the sample – TQ12).

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by principals is advisable.

### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

18	In the previous school year, how many teachers from abroad came to work in your school for longer than one month?	
	(Please write down the number of guest teachers. If no guest teachers from abroad came to your school please write 0)	
		<i>Number of guest teachers</i>
1)	Guest teachers of [target language] -----	<input type="text"/>
2)	Guest teachers of languages other than [target language] -----	<input type="text"/>
3)	Guest teachers of subjects other than languages -----	<input type="text"/>
4)	TOTAL number of guest teachers -----	<input type="text"/>

### NRC note

PQ17 and PQ18 provide information on the participation in exchange visits and guest teachers from other countries or language communities.

In Educational systems that are not an entire country an adequate adaptation of this question that can be understood by principals is advisable.

### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

19	In the previous school year, did any of the teachers or guest teachers receive funding for exchange visits in the following ways?			
		<i>(Please select in each row No or Yes)</i>		
		<i>No</i>	<i>Yes</i>	
1)	Through the European Union, such as a Comenius grant-----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Through the government (including local, regional, state and national government) -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	Through benefactors, donations, bequests, sponsorships, parent fund raising -----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

#### NRC note

This question provides information on the funding of teacher exchange visits.

20	Does your school have a mentoring programme for the following teaching staff?			
		<i>(Please select in each row No or Yes)</i>		
		<i>No</i>	<i>Yes</i>	
1)	For (almost) all teachers-----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	For teachers for whom this is their first teaching job-----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	For teachers who are new to this school -----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

#### NRC note

This question provides information on the school's policy for mentoring. The question is similar to the question about mentoring in TALIS.

## About the school's curriculum for foreign languages

### 21 Which languages are used for giving instruction in subjects other than languages?

*(Please tick as many boxes as applicable)*

{most widely spoken indigenous language1} ----- ☐ 0

{most widely spoken indigenous language2} ----- ☐ 1

{most widely spoken indigenous language3} ----- ☐ 2

{most widely spoken indigenous language4} ----- ☐ 3

{most widely spoken indigenous language5} ----- ☐ 4

{1<sup>st</sup> most widely taught foreign language in the country} ----- ☐ 5

{2<sup>nd</sup> most widely taught foreign language in the country} ----- ☐ 6

{3<sup>rd</sup> most widely taught foreign language in the country} ----- ☐ 7

{4<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 8

{5<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 9

Other language(s) ----- ☐ 10

#### WebTrans note

The same response instruction (Please tick as many boxes as applicable) occurs in PQ5, PQ6, PQ21, and PQ22.

Please note that the number of response categories can differ between countries.

#### NRC note

This question assesses the use of foreign languages and [target language] for the instruction of other subjects (Content and Language integrated learning). "Content and Language Integrated Learning (CLIL) involves teaching a curricular subject through the medium of a language other than that normally used. The subject can be entirely unrelated to language learning, such as history lessons being taught in English in a school in Spain".



The response categories of this question need to be localised similarly to TQ36 (excluding the ancient languages). The most widely spoken indigenous languages presented as response categories should correspond to the languages in the Localisation file (Language Table) and the most widely taught languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table).

The most widely spoken “indigenous” languages (national and regional) should be included. A maximum number of five “indigenous” languages can be included. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

In total a maximum number of 10 languages can be included.

Please instruct the reconciler which languages should be used as response categories.

### Translator note

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

---

## 22 Which of the following languages can students study in your school?

*(Please tick as many boxes as applicable)*

- {most widely spoken indigenous language1} ----- ☐ 0
- {most widely spoken indigenous language2} ----- ☐ 1
- {most widely spoken indigenous language3} ----- ☐ 2
- {most widely spoken indigenous language4} ----- ☐ 3
- {most widely spoken indigenous language5} ----- ☐ 4
- {1<sup>st</sup> most widely taught foreign language in the country} ----- ☐ 5
- {2<sup>nd</sup> most widely taught foreign language in the country} ----- ☐ 6
- {3<sup>rd</sup> most widely taught foreign language in the country} ----- ☐ 7
- {4<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 8
- {5<sup>th</sup> most widely taught foreign language in the country} ----- ☐ 9
- Other language(s) ----- ☐ 10

---

### WebTrans note

The same response instruction (Please tick as many boxes as applicable) occurs in PQ5, PQ6, PQ21, and PQ22.



### **NRC note**

PQ22 and PQ25 provide information on the foreign languages offered in the school. This question provides information on which foreign languages can be studied in the school.

The response categories of this question need to be localised identically to TQ36. The most widely spoken indigenous languages presented as response categories should correspond to the languages in the Localisation file (Language Table) and the most widely taught languages presented as response categories should correspond to the taught languages in the Localisation file (Taught language Table).

The most widely spoken “indigenous” languages (national and regional) should be included. A maximum number of five “indigenous” languages can be included. If a country has less than five “indigenous” languages less response categories can be used. If a country has more than five “indigenous” languages please use a more generic description, such as “One of the Sami languages”.

In total a maximum number of 10 languages can be included.

Please instruct the reconciler which languages should be used as response categories.

### **Translator note**

Please do not translate the response categories in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response categories.

23	In which [grades] are foreign languages (or ancient languages) offered as a subject in your school, either as a compulsory subject or optional subject?			
	Foreign languages (either as a compulsory subject or an optional subject) can be offered to all students or to only students in certain branches of study.			
	<i>(Please select only one answer from each row. If foreign languages cannot be studied in a particular [grade] or your school does not offer a particular [grade], please select "Not on offer")</i>			
		Not on offer	On offer for some	On offer for all
1)	{Lower grades than the starting grade of ISCED1} ---	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
2)	{starting grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
3)	{2 <sup>nd</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
4)	{3 <sup>rd</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
5)	{4 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
6)	{5 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
7)	{6 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
8)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
9)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
10)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
11)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
12)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
13)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
14)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
15)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
16)	{Higher grades than 2 <sup>nd</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2

#### WebTrans note

The response categories of PQ5 are identical to the items of PQ23, PQ24, and PQ25.

The same response instruction (Please select only one answer from each row ... please select "Not on offer") occurs in PQ23 and PQ24.

Please note that the number of items can differ between countries.



### **NRC note**

PQ23 and PQ24 provide information on the onset and duration of foreign language teaching. This question provides information on the grades in which foreign language education is offered. In the student questionnaire (SQ39) the grades in which students have studied foreign languages (in this school or other schools) is assessed.

The response categories of this question need to be localised (identically to PQ5). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. The later grades (after the 2<sup>nd</sup> year of ISCED3) do not have to be presented separately. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

24	In which [grades] are foreign languages (or ancient languages) a <u>compulsory</u> subject for all or some students in your school?				
	Foreign language lessons can be compulsory for all students, only compulsory for students in certain branches of study or not compulsory at all.				
	<i>(Please select only one answer from each row. If foreign languages cannot be studied in a particular [grade] or your school does not offer a particular [grade], please select "Not on offer")</i>				
		Not on offer	Not compulsory at all	Compulsory for some	Compulsory for all
1)	{Lower grades than the starting grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	{starting grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	{2 <sup>nd</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	{3 <sup>rd</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	{4 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	{5 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	{6 <sup>th</sup> grade of ISCED1} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
8)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
9)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
10)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
11)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
12)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
13)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
14)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
15)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
16)	{Higher grades than 2 <sup>nd</sup> grade of ISCED3} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The response categories of PQ5 are identical to the items of PQ23, PQ24, and PQ25.

The same response instruction (Please select only one answer from each row ... please select "Not on offer") occurs in PQ23 and PQ24.

Please note that the number of items can differ between countries.





### **NRC note**

PQ23 and PQ24 provide information on the onset and duration of foreign language teaching. This question provides information on the grades in which foreign language education is compulsory in the school.

The response categories of this question need to be localised (identically to PQ5). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. The later grades (after the 2<sup>nd</sup> year of ISCED3) do not have to be presented separately. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

compulsory - Please underline the appropriate text in the translation.

25	How many foreign languages (including ancient languages) can students study in each [grade]?		
<i>(Please write down the minimum and the maximum number of foreign languages a student can study. If foreign languages cannot be studied in a particular [grade] or your school does not offer a particular [grade], please write 0)</i>			
		Minimum number of foreign languages	Maximum number of foreign languages
1)	{Lower grades than the starting grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
2)	{starting grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
3)	{2 <sup>nd</sup> grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
4)	{3 <sup>rd</sup> grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
5)	{4 <sup>th</sup> grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
6)	{5 <sup>th</sup> grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
7)	{6 <sup>th</sup> grade of ISCED1} -----	<input type="text"/>	<input type="text"/>
8)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
9)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
10)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
11)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
12)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
13)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
14)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
15)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
16)	{Higher grades than 2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>

**WebTrans note**



The response categories of PQ5 are identical to the items of PQ23, PQ24, and PQ25.

Please note that the number of items can differ between countries.

### NRC note

PQ22 and PQ25 provide information on the offered foreign languages in the school. This question provides information on the number of foreign languages that can be studied in the school in each grade.

The response categories of this question need to be localised (identically to PQ5). All the grades from the onset of ISCED1 education until the 2<sup>nd</sup> year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. The later grades (after the 2<sup>nd</sup> year of ISCED3) do not have to be presented separately. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### Translator note

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[grade] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

How many - Please underline the appropriate text in the translation.

26	What is the status of [target language] in the school's curriculum?			
	<i>(Please select in each row No or Yes)</i>			
		No	Yes	
1)	[Target language] is a compulsory subject in the curriculum of <u>all</u> students-----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	[Target language] is a compulsory subject in the curriculum of <u>some</u> students -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	[Target language] is the first taught foreign language for <u>all</u> students -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	[Target language] is the first taught foreign language for <u>some</u> students -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	[Target language] is the most widely taught foreign language in school-----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

## NRC note

This question provides information on the status of [target language] in the school (compulsory or not, most widely taught) and the order of learning (1<sup>st</sup> foreign language or not).

## Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

all - Please underline the appropriate text in the translation.

some - Please underline the appropriate text in the translation.

27	How many students in your school study [target language] either as a compulsory subject or an optional subject?
	<i>(Please write down the number of students. If foreign languages cannot be studied in a particular [grade] or your school does not offer a particular [grade], please write 0)</i>
	<i>Number of students</i>
1)	{1 <sup>st</sup> grade of ISCED2} ----- <input type="text"/>
2)	{2 <sup>nd</sup> grade of ISCED2} ----- <input type="text"/>
3)	{3 <sup>rd</sup> grade of ISCED2} ----- <input type="text"/>
4)	{4 <sup>th</sup> grade of ISCED2} ----- <input type="text"/>
5)	{5 <sup>th</sup> grade of ISCED2} ----- <input type="text"/>
6)	{6 <sup>th</sup> grade of ISCED2} ----- <input type="text"/>
7)	{7 <sup>th</sup> grade of ISCED2} ----- <input type="text"/>
8)	{8 <sup>th</sup> grade of ISCED2} ----- <input type="text"/>
9)	{1 <sup>st</sup> grade of ISCED3} ----- <input type="text"/>
10)	{2 <sup>nd</sup> grade of ISCED3} ----- <input type="text"/>

## WebTrans note

The items of PQ27, PQ32, PQ33, PQ34, and PQ35 are identical (do NOT link these items to PQ5, PQ23, PQ24 and PQ25).

The same response label (Number of students) occurs in PQ1, PQ27, PQ28, and PQ29.



Please note that the number of items can differ between countries.

### NRC note

This question provides information on the number of students studying [target language] in the school.

The response categories of this question need to be localised. All the grades from the onset of ISCED2 education until the 2nd year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### Translator note

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[grade] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

- 28 What is the average size of classes in {all grades of ISCED2 and the first two years of ISCED3} in your school?

(as far as these grades are present in your school)

*(Please write down the number of students)*

*Number of  
students*

### WebTrans note

The same response instruction (Please write down the number of students) occurs in PQ1, PQ28 and PQ29.

The same response label (Number of students) occurs in PQ1, PQ27, PQ28, and PQ29.

### NRC note

PQ28 and PQ29 provide information on the average class size. The information on the average class size in the whole school (PQ28) can be compared with the class size for [target language] lessons (PQ29 and TQ39) and the size of the classes of the sampled students (SQ42).

The phrase {all grades of ISCED2 and the first two years of ISCED3} needs to be localised. Please instruct the reconciler which phrase to use.

### Translator note

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

29 What is the average size of [target language] classes in {all grades of ISCED2 and the first two years of ISCED3} in your school?

(as far as these grades are present in your school)

*(Please write down the number of students)*

*Number of  
students*

#### **WebTrans note**

The same response instruction (Please write down the number of students) occurs in PQ1, PQ28 and PQ29.

The same response label (Number of students) occurs in PQ1, PQ27, PQ28, and PQ29.

#### **NRC note**

PQ28 and PQ29 provide information on the average class size. The information on the average class size in the whole school (PQ28) can be compared with the class size for [target language] lessons (PQ29 and TQ39) and the size of the classes of the sampled students (SQ42).

The phrase {all grades of ISCED2 and the first two years of ISCED3} needs to be localised. Please instruct the reconciler which phrase to use.

#### **Translator note**

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

30	Generally, in your school, how often are students in {all grades of ISCED2 and the first two years of ISCED3} assessed using the following methods?  (as far as these grades are present in your school)	(Please select only one answer from each row)				
		Never	Once or a few times a year	About once a month	A few times a month	(Almost) every week
1)	Standardised tests, e.g. {a national examination} -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
2)	Teacher-developed tests -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
3)	Teachers' judgmental ratings -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
4)	Student portfolios -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4
5)	Student assignments/projects/homework -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4

#### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

#### NRC note

This question provides information on the frequency and mode of assessment in the school and is highly similar to PISA question about assessment.

The phrase {all grades of ISCED2 and the first two years of ISCED3} needs to be localised. Please instruct the reconciler which phrase to use.

#### Translator note

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

## About the teaching time for foreign languages

*In this section you will be asked some questions about the number of class periods per year in which languages are taught.*

31 How long does a class period last at your school?

*(Please write down the number of minutes)*

*Minutes*

---

### **NRC note**

This question allows us to calculate the teaching time per year.



32	According to the school curriculum what is the number of class periods per year in each of the following [grades] for the subject of [target language]?		
	<i>(Please write down the minimum and maximum number of class periods. If the number of hours is equal for all students irrespective of the branch of study, please write down the same number twice.)</i> <i>If your school does not offer a particular [grade], please write 0)</i>		
		Minimum number of class periods	Maximum number of class periods
1)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
2)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
3)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
4)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
5)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
6)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
7)	{7 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
8)	{8 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
9)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
10)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>

### WebTrans note

The items of PQ27, PQ32, PQ33, PQ34, and PQ35 are identical (do NOT link these items to PQ5, PQ23, PQ24 and PQ25).

The same response instruction (Please write down the minimum number of teaching hours ...please write 0) occurs in PQ32, PQ33, PQ34, and PQ35.

The same response scale (Minimum number of class periods-Maximum number of class periods) occurs in PQ32, PQ33, PQ34, and PQ35.

Please note that the number of items can differ between countries.

### NRC note

PQ32, PQ33, PQ34 and PQ35 provide information on the teaching time in each grade: for [target language] (PQ32), for all foreign language (PQ33), for all languages (PQ34) and for all subjects (PQ35).



The response categories of this question need to be localised (identically to PQ27). All the grades from the onset of ISCED2 education until the 2nd year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

33	According to the school curriculum what is the number of class periods per year in each of the following [grades] for foreign languages (including ancient languages and [target language])?		
<p><i>(Please write down the minimum and maximum number of class periods. If the number of hours is equal for all students irrespective of the branch of study, please write down the same number twice.)</i></p> <p><i>If your school does not offer a particular [grade], please write 0)</i></p>			
		<i>Minimum number of class periods</i>	<i>Maximum number of class periods</i>
1)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
2)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
3)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
4)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
5)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
6)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
7)	{7 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
8)	{8 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
9)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
10)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>

### WebTrans note

The items of PQ27, PQ32, PQ33, PQ34, and PQ35 are identical (do NOT link these items to PQ5, PQ23, PQ24 and PQ25).

The same response instruction (Please write down the minimum number of teaching hours ...please write 0) occurs in PQ32, PQ33, PQ34, and PQ35.

The same response scale (Minimum number of class periods-Maximum number of class periods) occurs in PQ32, PQ33, PQ34, and PQ35.

Please note that the number of items can differ between countries.

### NRC note

PQ32, PQ33, PQ34 and PQ35 provide information on the teaching time in each grade: for [target language] (PQ32), for all foreign language (PQ33), for all languages (PQ34) and for all subjects (PQ35).



The response categories of this question need to be localised (identically to PQ27). All the grades from the onset of ISCED2 education until the 2nd year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

34	According to the school curriculum what is the number of class periods per year in each of the following [grades] for languages (including all foreign languages, ancient languages and [questionnaire language])?		
<p><i>(Please write down the minimum and maximum number of class periods. If the number of hours is equal for all students irrespective of the branch of study, please write down the same number twice.)</i></p> <p><i>If your school does not offer a particular [grade], please write 0)</i></p>			
		Minimum number of class periods	Maximum number of class periods
1)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
2)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
3)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
4)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
5)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
6)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
7)	{7 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
8)	{8 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
9)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
10)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>

### WebTrans note

The items of PQ27, PQ32, PQ33, PQ34, and PQ35 are identical (do NOT link these items to PQ5, PQ23, PQ24 and PQ25).

The same response instruction (Please write down the minimum number of teaching hours ...please write 0) occurs in PQ32, PQ33, PQ34, and PQ35.

The same response scale (Minimum number of class periods-Maximum number of class periods) occurs in PQ32, PQ33, PQ34, and PQ35.

Please note that the number of items can differ between countries.

### NRC note

PQ32, PQ33, PQ34 and PQ35 provide information on the teaching time in each grade: for [target language] (PQ32), for all foreign language (PQ33), for all languages (PQ34) and for all subjects (PQ35).



The response categories of this question need to be localised (identically to PQ27). All the grades from the onset of ISCED2 education until the 2nd year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

### **Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.

35	According to the school curriculum what is the teaching time per year in each of the following [grades] (for all subjects together)?		
<p><i>(Please write down the minimum and maximum number of class periods. If the number of hours is equal for all students irrespective of the branch of study, please write down the same number twice.</i></p> <p><i>If your school does not offer a particular [grade], please write 0)</i></p>			
		<i>Minimum number of class periods</i>	<i>Maximum number of class periods</i>
1)	{1 <sup>st</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
2)	{2 <sup>nd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
3)	{3 <sup>rd</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
4)	{4 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
5)	{5 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
6)	{6 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
7)	{7 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
8)	{8 <sup>th</sup> grade of ISCED2} -----	<input type="text"/>	<input type="text"/>
9)	{1 <sup>st</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>
10)	{2 <sup>nd</sup> grade of ISCED3} -----	<input type="text"/>	<input type="text"/>

### WebTrans note

The items of PQ27, PQ32, PQ33, PQ34, and PQ35 are identical (do NOT link these items to PQ5, PQ23, PQ24 and PQ25).

The same response instruction (Please write down the minimum number of teaching hours ...please write 0) occurs in PQ32, PQ33, PQ34, and PQ35.

The same response scale (Minimum number of class periods-Maximum number of class periods) occurs in PQ32, PQ33, PQ34, and PQ35.

Please note that the number of items can differ between countries.

### NRC note

PQ32, PQ33, PQ34 and PQ35 provide information on the teaching time in each grade: for [target language] (PQ32), for all foreign language (PQ33), for all languages (PQ34) and for all subjects (PQ35).



The response categories of this question need to be localised (identically to PQ27). All the grades from the onset of ISCED2 education until the 2nd year of ISCED3, that sampled schools offer, should be presented as response categories. Grades that none of the sampled school offer do not have to be included. It is necessary to phrase the response categories in such a way that principals will easily understand.

Please instruct the reconciler which response categories to use.

**Translator note**

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

[grades] – refers to the administrative levels of the student in the school. The number of years in schooling is the usual measure of grade. It does not mean the name of a class.



## About the school's policy and practices to encourage language learning

36	Does your school offer the following to encourage language learning?			
	<i>(Please select in each row No or Yes)</i>			
		<i>No</i>	<i>Yes</i>	
1)	{Content and Language Integrated Learning (CLIL)} -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	The classes for foreign language lessons are smaller than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	A wider choice of languages is offered than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	More teaching hours are devoted to foreign language learning than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	Students can study <u>more</u> languages than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	
6)	More extra curricular activities related to language education are organised than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	
7)	Foreign language lessons are offered in earlier grades than is common or required -----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

### NRC note

This question provides information on the school's foreign language specialisation.

Item 1) in this question needs to be localised identical to item 8) of TQ15 and TQ32.

{Content and Language Integrated Learning (CLIL)} – involves teaching a curricular subject through the medium of a language other than that normally used. The subject can be entirely unrelated to language learning, such as history lessons being taught in English in a school in Spain. Please use a term that is commonly used in your [Educational system] and is easily understood by teachers.

Please instruct the reconciler which item should be used.

### Translator note

Please do not translate the items in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate items.

more - Please underline the appropriate text in the translation.

37	In the past three school years, have the following activities been organised at your school?	(Please select only one answer from each row)			
		Not at all	In one school year	In two school years	In three school years
1)	Collaboration project with schools abroad -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Language clubs -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Language competition -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	European Day of Languages-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Extracurricular language projects -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
6)	Pen friends, email or MSN friends for students -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
7)	Excursions and field trips related to language education -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

#### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

The same response scale (Not at all-In one school year-In two school years-In three school years) occurs in PQ37 and PQ38.

#### NRC note

PQ37 and PQ38 provide information on the opportunities created in school for exchange visits (PQ37) and school language projects (PQ38).

The items in this question might need to be adapted (identical to SQ46 and TQ42). If necessary, provide examples appropriate for the [Educational system] of language clubs and language competitions. Please use items that match the items of SQ46.

Language clubs – A group of students that meets regularly and organized for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition – an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

European Day of Languages – The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

MSN – is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

### Translator note

Language clubs – A group of students that meets regularly for the common purpose of learning and using a foreign language, e.g. a [target language] conversation club, a [target language] debating club or a club for learning a foreign language that is not part of the regular curriculum.

A language competition – an event in which students compete with other students to see who has the best language skills, e.g. a debating competition in which students have to debate in a foreign language or a spelling competition.

European Day of Languages – The first European Day of Languages took place on 26 September 2001. On this day, 26 September, activities are organised throughout Europe to celebrate linguistic diversity in Europe and to promote language learning.

MSN – is an internet service for instant messaging, a form of real-time communication between two or more people based on typed text. The term MSN might need to be replaced with a suitable term that is understood by students in your country.

38	In the past three school years, have the following trips or visits been organised at your school?	(Please select only one answer from each row)			
			<i>In one school year</i>	<i>In two school years</i>	<i>In three school years</i>
		<i>Not at all</i>			
1)	School trips to a [target language] speaking country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	School trips to another (non-[target language] speaking) country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Visits to your school by a school class from a [target language] speaking country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Visits to your school by a school class from another (non-[target language] speaking) country-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

The same response scale (Not at all-In one school year-In two school years-In three school years) occurs in PQ37 and PQ38.

### NRC note

PQ37 and PQ38 provide information on the opportunities created in school for exchange visits (PQ37) and school language projects (PQ38).

In Educational systems that are not an entire country an adequate adaptation of the items of this questions that can be understood by principals might be needed. The items in this question are identical to the items of TQ41 and similar to the items 1), 2), 5) and 6) of SQ46 (parallel questions).

### Translator note

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

39	To what extent are intercultural exchanges for students (such as exchange visits) funded in the following ways?	<i>(Please select only one answer from each row)</i>			
		<i>Not at all</i>	<i>To a small extent</i>	<i>To a large extent</i>	<i>Completely</i>
1)	Through the European Union, such as a Comenius grant -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
2)	Through the government (including local, regional, state and national government) -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
3)	Through regular student fees or school charges paid by parents-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
4)	Through specific contributions paid by parents -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3
5)	Through benefactors, donations, bequests, sponsorships, parent fund raising-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3

### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

### NRC note

This question provides information on the funding of intercultural exchanges for students.

40	What type of <u>extra</u> lessons does your school offer to students?			
	<i>(Please select in each row No or Yes)</i>			
		<i>No</i>	<i>Yes</i>	
1)	{Enrichment lessons} for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	{Enrichment lessons} for other foreign languages (including for Latin and ancient Greek) -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	{Remedial lessons} for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	{Remedial lessons} for other foreign languages (including for Latin and ancient Greek) -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	Extra [questionnaire language] lessons for students with a different home language to [Questionnaire language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
6)	Extra lessons in students' home language if this is a different language to [questionnaire language] -----	<input type="radio"/> 0	<input type="radio"/> 1	

#### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

#### NRC note

This question provides information on the extra lessons the school provides for foreign languages and [target language], and the provision of extra lessons for immigrant students (extra lessons in the host language and in the language of origin). Schools with a foreign language specialisation usually offer more opportunities for out-of-school-time lessons, than schools without such a specialisation. The information can be compared to the findings in other international surveys (PISA, PIRLS, TIMSS, TALIS) and the participation of the samples students in the offered extra lessons (SQ64).

The items of this question might need to be localised (identical to SQ64). It is necessary to use terms that principals will easily understand. Please instruct the reconciler which terms should be used.

{Remedial lessons} — Any lessons in addition to regular lessons designed to help students with learning difficulties.

{Enrichment lessons} — Any lessons in addition to regular lessons designed to extend abilities of more able students.

#### Translator note

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.



Extra - Please underline the appropriate text in the translation.

## About the school's resources

41	Is your school's capacity to provide instruction hindered by any of the following issues?			
	<i>(Please select in each row No or Yes)</i>			
		<i>No</i>	<i>Yes</i>	
1)	A lack of qualified language teachers-----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	A lack of qualified [questionnaire language] teachers-----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	A lack of qualified [target language] teachers-----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	A lack of qualified teachers of foreign languages (including ancient languages) other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	A lack of qualified teachers of subjects other than languages ---	<input type="radio"/> 0	<input type="radio"/> 1	
6)	A lack of library staff-----	<input type="radio"/> 0	<input type="radio"/> 1	
7)	A lack of other support personnel-----	<input type="radio"/> 0	<input type="radio"/> 1	
8)	Shortage or inadequacy of instructional materials (e.g. textbooks) -----	<input type="radio"/> 0	<input type="radio"/> 1	
9)	Shortage or inadequacy of computers for instruction -----	<input type="radio"/> 0	<input type="radio"/> 1	
10)	Lack or inadequacy of Internet connectivity -----	<input type="radio"/> 0	<input type="radio"/> 1	
11)	Shortage or inadequacy of computer software for instruction----	<input type="radio"/> 0	<input type="radio"/> 1	
12)	Shortage or inadequacy of library materials-----	<input type="radio"/> 0	<input type="radio"/> 1	
13)	Shortage or inadequacy of audio-visual resources-----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

### NRC note

PQ41 and PQ42 provide information on the available resources (other than ICT) in the school. This question is highly similar to the question about resources in PISA, PIRLS, TIMSS and TALIS.

**Translator note**

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

[Questionnaire language] – Please substitute this term with the name of the language into which the questionnaire is being translated.



42	In your opinion, are the following resources in your school sufficient to support the instruction <u>in foreign languages</u> ?			
		<i>(Please select in each row No or Yes)</i>		
		No	Yes	
1)	Audio cassettes, CDs or other audio-material spoken in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Audio cassettes, CDs or other audio-material spoken in foreign languages other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	Video cassettes, DVDs, video clips from YouTube or other audio-visual material spoken in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	Video cassettes, DVDs, video clips from YouTube or other audio-visual material spoken in foreign languages other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	Newspapers, magazines, comics or song texts written in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
6)	Newspapers, magazines, comics or song texts written in foreign languages other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
7)	Textbook(s) for [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
8)	Textbook(s) for foreign languages other than [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
9)	Books written in [target language] for extensive reading, e.g. fiction -----	<input type="radio"/> 0	<input type="radio"/> 1	
10)	Books written in foreign languages other than [target language] for extensive reading e.g. fiction -----	<input type="radio"/> 0	<input type="radio"/> 1	
11)	Lesson materials prepared by the teacher(s) of [target language] (e.g. hand-outs, reading texts) -----	<input type="radio"/> 0	<input type="radio"/> 1	
12)	Lesson materials prepared by the teacher(s) of other foreign languages than [target language] (e.g. hand-outs, reading texts) -----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

### NRC note

PQ41 and PQ42 provide information on the available resources (other than ICT) in the school. This question provides information on the specific resources for language education.

The items are highly similar to the items of SQ51 and TQ44.

### Translator note

[Target language] – Please replace the term “target language” according to the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

in foreign languages - Please underline the appropriate text in the translation.

43	At your school, how many computers do you have for educational purposes?	
	<i>(Please write 0 if there are none)</i>	
		<i>Number of computers</i>
1)	Approximately, how many computers are available for students for educational purposes? -----	<input type="text"/>
2)	Approximately, how many of these computers for students are connected to the Internet/World Wide Web? -----	<input type="text"/>
3)	Approximately, how many computers are available specifically for teachers? -----	<input type="text"/>
4)	Approximately, how many of these computers for teachers are connected to the Internet/World Wide Web? -----	<input type="text"/>

### NRC note

PQ43, PQ44, and PQ45 provide information on the ICT-facilities in the school. This question provides information on the number of computers (with and without internet connection) available for students and teachers.

44	Are the following devices available in the classrooms?			
	<i>(Please select only one answer from each row)</i>			
		<i>No</i>	<i>In some classrooms</i>	<i>In (almost) all classrooms</i>
1)	A teacher PC or laptop-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
2)	A projector -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
3)	Interactive whiteboard-----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2
4)	An internet connection -----	<input type="radio"/> 0	<input type="radio"/> 1	<input type="radio"/> 2

#### WebTrans note

The same response instruction (Please select only one answer from each row) occurs in PQ3, PQ4, PQ30, PQ37, PQ38, PQ39, and PQ44.

#### NRC note

PQ43, PQ44, and PQ45 provide information on the ICT-facilities in the school. This question provides information on the ICT-facilities in the classrooms.

The items of this question are identical to the items 1), 2), 3) and 6) of TQ43 (question about the ICT facilities the teachers use).

45	Does your school have the following ICT facilities?			
	<i>(Please select in each row No or Yes)</i>			
		No	Yes	
1)	Multimedia language lab (teacher PC and student PCs <u>with</u> specific language learning software) -----	<input type="radio"/> 0	<input type="radio"/> 1	
2)	Multimedia lab (teacher PC and student PCs <u>without</u> specific language learning software) -----	<input type="radio"/> 0	<input type="radio"/> 1	
3)	A virtual learning environment to support teaching and learning, e.g. Moodle, WebCT, Blackboard, Fronter, Sakai -----	<input type="radio"/> 0	<input type="radio"/> 1	
4)	Software or access to websites specifically designed for learning languages -----	<input type="radio"/> 0	<input type="radio"/> 1	
5)	Software for communication tools, e.g. email, chatting, blogging, {MySpace}, {Skype} -----	<input type="radio"/> 0	<input type="radio"/> 1	
6)	Software or tools developed in house for learning and teaching languages -----	<input type="radio"/> 0	<input type="radio"/> 1	
7)	Digital student portfolio -----	<input type="radio"/> 0	<input type="radio"/> 1	
8)	Software for language assessment-----	<input type="radio"/> 0	<input type="radio"/> 1	
9)	Access to online dictionaries and other reference works-----	<input type="radio"/> 0	<input type="radio"/> 1	
10)	Access to online news media (TV, radio, newspapers) in [target language] -----	<input type="radio"/> 0	<input type="radio"/> 1	
11)	Access to other websites on life and culture in [target language] speaking country/countries -----	<input type="radio"/> 0	<input type="radio"/> 1	

### WebTrans note

The same response instruction (Please select in each row No or Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

The same response scale (No-Yes) occurs in PQ11, PQ12, PQ14, PQ15, PQ19, PQ20, PQ26, PQ36, PQ40, PQ41, PQ42, and PQ45.

### NRC note

PQ43, PQ44, and PQ45 provide information on the ICT-facilities in the school. This question provides information on the available language specific ICT-facilities (software and access to websites).

Items 1), 2) and 3) are identical to the items 4), 5) and 7) of TQ43.

The items 4) through 11) are similar to the items of TQ45.

Some of the terms in this question might need to be localised. Please instruct the reconciler which terms should be used.



Skype - Skype is a software application that allows users to make voice calls over the Internet.

MySpace – is an example of a social networking website with an interactive, user-submitted network of friends, personal profiles, blogs, groups, photos, music, and videos for teenagers and adults. Please use an example of a social networking website that is most widely known in your country (e.g. Friendster, Twitter, e.d.).

Tools – An application program for the computer or an element of a computer program.

### **Translator note**

Please do not translate the terms in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate terms.

[Target language] – Please replace the term “target language” according the instructions of the National Research Coordinator (the name of the language for which the students will be tested prior to filling out the questionnaire).

with - Please underline the appropriate text in the translation.

without - Please underline the appropriate text in the translation.

46	For your school, who has considerable responsibility for the following tasks?					
		<i>(Please tick in each row as many boxes as applicable)</i>				
		<i>Principals</i>	<i>Teachers</i>	<i>{School governing board}</i>	<i>{Regional or local education authority}</i>	<i>National education authority</i>
1)	Selecting teachers for hire -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
2)	Firing teachers -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
3)	Establishing teachers' starting salaries -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
4)	Determining teachers' salary increases -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
5)	Formulating the school budget -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
6)	Deciding on budget allocations within the school -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
7)	Establishing student disciplinary policies -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
8)	Establishing student assessment policies -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
9)	Approving students for admission to the school -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
10)	Choosing which textbooks are used -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
11)	Determining course content -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
12)	Determining which foreign languages are offered as a subject -	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
13)	Determining in which order foreign languages are offered -----	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4

### NRC note

This question provides information on the autonomy of schools in several areas. This question is identical to the PISA question about school autonomy and highly similar to the TALIS question about school autonomy).

Some of the terms in this question might need to be localised. Please instruct the reconciler which terms should be used.



{School governing board} - a board directly responsible for the governance of the school. This board may be totally external to the school or may have staff and student representation. The school's governing board is usually the governing board of that school only (i.e. it is not a district board).

{Regional or local education authority} - an authority that is not a national authority and does not directly govern the school. For example, a provincial ministry.

**Translator note**

Please do not translate the response labels in curly brackets { }. The reconciler and/or NRC is asked to add the appropriate response labels.

Thank you very much for the time and effort you have put  
into responding to this questionnaire.

## Appendix 3: Sampling Forms



## 14.3 *Sampling forms*

### 14.3.1 Form 1. Organisation, logistics

Educational system:	<b>PRE-FILLED</b>
1.2 Educational system code (ISO) :	<b>XX (PRE-FILLED)</b>
1.3 National Research Centre:	<b>NAME OF ORGANISATION</b>
National Research Coordinator:	1.4 Name: <b>FILL IN</b>
	1.5 Email: <b>FILL IN</b>
	1.6 Direct telephone: <b>FILL IN</b>
	1.7 Fax: <b>FILL IN</b>
	1.8 Postal address: <b>FILL IN</b>
Expert responsible for sampling information: (fill in only if different from National Research Coordinator, otherwise leave blank)	1.9 Name: <b>FILL IN</b>
	1.10 Email: <b>FILL IN</b>
	1.11 Direct telephone: <b>FILL IN</b>
	1.12 Fax: <b>FILL IN</b>
	1.13 Postal address: <b>FILL IN</b>
1.14 Educational system formally committed to participate in ELSC (Y/N):	

### 14.3.2 Form 2. Language and Grade Definition

2.1 The most widely taught Foreign Language tested (LANGUAGE1) is:	LANGUAGE1 (PRE-FILLED)	YES/NO
2.2 Students who will be assessed for the most widely taught foreign language are in the following grade:	GRADE1 <sup>6</sup> (PRE-FILLED)	YES/NO
2.3 Total number of schools teaching LANGUAGE1 in GRADE1	FILL IN THE NUMBER	PROVISIONAL / FINAL
2.3a Source of information in 2.3, if exact figures are given (describe and insert an URL where available):	SOURCE URL ( <a href="http://inserturl">http://inserturl</a> ) or: ESTIMATED	
2.4 Total national population of students studying LANGUAGE1 in GRADE1	FILL IN THE NUMBER	PROVISIONAL / FINAL
2.4a Source of information in 2.4, if exact figures are given (describe and insert an URL where available):	SOURCE URL ( <a href="http://inserturl">http://inserturl</a> ) or: ESTIMATED	

2.5 The second most widely taught Foreign Language tested (LANGUAGE2) is:	LANGUAGE2 (PRE-FILLED)	YES/NO
2.6 Students who will be assessed for the second most widely taught Foreign language are in the following grade:	GRADE2 (PRE-FILLED)	YES/NO
2.7 Total number of schools teaching LANGUAGE2 in GRADE2	FILL IN THE NUMBER	PROVISIONAL / FINAL
2.7a Source of information in 2.7, if exact figures are given (describe and insert an URL where available):	SOURCE URL ( <a href="http://inserturl">http://inserturl</a> ) or: ESTIMATED	
2.8 Total national population of students studying LANGUAGE2 in GRADE2	FILL IN THE NUMBER	PROVISIONAL / FINAL
2.8a Source of information in 2.8, if exact figures are given (describe and insert an URL where available):	SOURCE URL ( <a href="http://inserturl">http://inserturl</a> ) or: ESTIMATED	

<sup>6</sup> Select one: Last grade of ISCED2 or Second grade of ISCED3

### 14.3.3 Form 3. School-level exclusion types

ESTIMATED % OF ELIGIBLE STUDENTS CONCERNED		
REASON (DESCRIBE IN SUFFICIENT DETAIL)	LANGUAGE 1.	LANGUAGE 2.
3.0 EXTREMELY SMALL SCHOOLS (WITH LESS THAN 10 ELIGIBLE STUDENTS)	FILL IN	FILL IN
3.1	FILL IN	FILL IN
3.2	FILL IN	FILL IN
3.3	FILL IN	FILL IN
3.4	FILL IN	FILL IN
3.5	FILL IN	FILL IN
3.6	FILL IN	FILL IN
3.7	FILL IN	FILL IN
3.8	FILL IN	FILL IN
3.9	FILL IN	FILL IN
3.10	FILL IN	FILL IN
TOTAL :	FILL IN	FILL IN

#### 14.3.4 Form 4. Student-level exclusion types

ESTIMATED % OF ELIGIBLE STUDENTS CONCERNED,		
REASON (DESCRIBE IN SUFFICIENT DETAIL)	LANGUAGE 1.	LANGUAGE 2.
4.1	FILL IN	FILL IN
4.2	FILL IN	FILL IN
4.3	FILL IN	FILL IN
4.4	FILL IN	FILL IN
4.5	FILL IN	FILL IN
4.6	FILL IN	FILL IN
4.7	FILL IN	FILL IN
4.8	FILL IN	FILL IN
4.9	FILL IN	FILL IN
4.10	FILL IN	FILL IN
4.11	FILL IN	FILL IN
4.12	FILL IN	FILL IN
TOTAL :	FILL IN	FILL IN

### 14.3.5 Form 5A. Explicit stratification, Language<sup>1</sup>

5A.1 Do you intend to use explicit stratification in your sample? **YES/NO** (If not, the rest of the form should be left blank)

5A.2: Allocation of cases across strata: **PROPORTIONAL/DISPROPORTIONAL**

5A.3 Details: (d) only to be filled in if disproportional allocation is foreseen)

(a) CODE	(b) STRATUM	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE STRATUM	(d) % OF STUDENT SAMPLE TO BE ALLOCATED IN THE STRATUM
1	FILL IN	FILL IN	FILL IN
2	FILL IN	FILL IN	FILL IN
3	FILL IN	FILL IN	FILL IN
4	FILL IN	FILL IN	FILL IN
5	FILL IN	FILL IN	FILL IN
6	FILL IN	FILL IN	FILL IN
7	FILL IN	FILL IN	FILL IN
8	FILL IN	FILL IN	FILL IN
9	FILL IN	FILL IN	FILL IN
10	FILL IN	FILL IN	FILL IN
11	FILL IN	FILL IN	FILL IN
12	FILL IN	FILL IN	FILL IN

### 14.3.6 Form 5B. Explicit stratification, Language2

**FILL IN ONLY IF YOU REQUIRE DIFFERENT STRATIFICATION FOR THE SECOND LANGUAGE!**

5B.1 Do you intend to use explicit stratification in your sample? **YES/NO** (If not, the rest of the form should be left blank)

5B.2: Allocation of cases across strata: **PROPORTIONAL/DISPROPORTIONAL**

5B.3 Details: (d) only to be filled in if disproportional allocation is foreseen)

(a) CODE	(b) STRATUM	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE STRATUM	(d) % OF STUDENT SAMPLE TO BE ALLOCATED IN THE STRATUM
1	FILL IN	FILL IN	FILL IN
2	FILL IN	FILL IN	FILL IN
3	FILL IN	FILL IN	FILL IN
4	FILL IN	FILL IN	FILL IN
5	FILL IN	FILL IN	FILL IN
6	FILL IN	FILL IN	FILL IN
7	FILL IN	FILL IN	FILL IN
8	FILL IN	FILL IN	FILL IN
9	FILL IN	FILL IN	FILL IN
10	FILL IN	FILL IN	FILL IN
11	FILL IN	FILL IN	FILL IN
12	FILL IN	FILL IN	FILL IN

### 14.3.7 Form 6A. Implicit stratification, Language1

6A.1 Do you intend to use implicit stratification in your sample? **YES/NO** (If not, the rest of the form should be left blank)

6A.2 Variables to be used (e.g. 'Region'), up to three (FILL IN):

1	<b>VARIABLE A1 (FILL IN)</b>
2	<b>VARIABLE A2 (FILL IN)</b>
3	<b>VARIABLE A3 (FILL IN)</b>

6A.3 **VARIABLE A1 (FILL IN)** categories:

6A.3.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

(a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
2	<b>FILL IN</b>	<b>FILL IN</b>
3	<b>FILL IN</b>	<b>FILL IN</b>
4	<b>FILL IN</b>	<b>FILL IN</b>
5	<b>FILL IN</b>	<b>FILL IN</b>
6	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

6A.4 **VARIABLE A2 (FILL IN)** categories:

6A.4.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

(a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
...	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

6A.5 **VARIABLE A3 (FILL IN)** categories:

6A.5.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

(a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
...	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

### 14.3.8 Form 6B. Implicit stratification, Language2

**FILL IN ONLY IF YOU REQUIRE DIFFERENT STRATIFICATION FOR THE SECOND LANGUAGE!**

6B.1 Do you intend to use implicit stratification in your sample? **YES/NO** (If not, the rest of the form should be left blank)

6B.2 Variables to be used (e.g. 'Region'), up to three, FILL IN:

1	<b>VARIABLE B1 (FILL IN)</b>
2	<b>VARIABLE B2 (FILL IN)</b>
3	<b>VARIABLE B3 (FILL IN)</b>

6B.3 **VARIABLE B1 (FILL IN)** categories:

6B.3.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
2	<b>FILL IN</b>	<b>FILL IN</b>
3	<b>FILL IN</b>	<b>FILL IN</b>
4	<b>FILL IN</b>	<b>FILL IN</b>
5	<b>FILL IN</b>	<b>FILL IN</b>
6	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

6B.4 **VARIABLE B2 (FILL IN)** categories:

6B.4.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

(a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
...	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

6B.5 **VARIABLE B3 (FILL IN)** categories:

6B.5.0 Is this a continuous variable? **YES/NO** (If yes, rows 1-7 should be left blank)

(a) CODE	(b) CATEGORY	(c) % OF ALL ELIGIBLE STUDENTS FALLING IN THE CATEGORY
1	<b>FILL IN</b>	<b>FILL IN</b>
...	<b>FILL IN</b>	<b>FILL IN</b>
7	<b>FILL IN</b>	<b>FILL IN</b>

**NO FORM 7**



### 14.3.9 Form 8. School Master List (metadata)

The information below is only a reference, the information has to be provided in the spreadsheet provided by *SurveyLang*.

CODE	FIELD	TYPE7:	REMARK:	FIXED8	DATA EXAMPLE:
8.1	ASSIGNED SLANG ID	INT	<b>NUMERIC UNIQUE ID KEY, assigned by database, by SurveyLang</b>	y	1210001
8.2	EDUCATIONAL SYSTEM CODE	INT	ISO code pre-defined for the NRCs, refer to Sampling Manual	y	12
8.3	Nat'l School ID	STR	UNIQUE KEY, according to national nomenclature	n	10001
8.4	Name	STR	The full name of the school	n	XYZ District School
8.5	Postal code	STR	AS IS	n	1111
8.6	City	STR	AS IS	n	Aalborg
8.7	Address	STR	Including street and number	n	1. Main square
8.8	Education level	INT	Specify if the school provides education on ISCED2 level (1), on ISCED3 level (2), or on both levels (3)	n	1
8.10a	Grade corresponding to test level for <b>LANGUAGE1</b>	INT	Either the last grade of ISCED2 or the second grade of ISCED3 level should be specified in a way that corresponds to the grade system used in the given school	y	3rd grade
8.11	Students enrolled in eligible grade of LANGUAGE1	INT	Number of all students in the eligible grade, as specified on Form 2, Field 2.2.	n	98

CODE	FIELD	TYPE7:	REMARK:	FIXED8	DATA EXAMPLE:
8.12	Students eligible for <b>LANGUAGE1</b> testing in the eligible grade	INT	Number of those in the eligible grade who had prior instruction in the given language for at least one full academic year	y	78
8.13	Students enrolled in the grade below the eligible grade of <b>LANGUAGE1</b>	INT	Number of all students in the grade below the eligible grade. E.g. if testing is among those in the final year of ISCED2, the number of those in the pre-final year of ISCED2 should be specified here.	n	94
8.14	Students learning <b>LANGUAGE1</b> in the grade below the eligible grade	INT	Number of all students currently learning the 1st foreign language in the grade below the testing grade, i.e. those who will be eligible for testing in the next academic year.	y	71
8.15a	Grade corresponding to test level for <b>LANGUAGE2</b>	INT	Either the last grade of ISCED2 or the second grade of ISCED3 level should be specified in a way that is relevant locally. If the two languages are tested on the same ISCED level, this field equals 8.10a.	y	3rd grade
8.17	Students enrolled in eligible grade of <b>LANGUAGE2</b>	INT	Number of all students in the eligible grade, as specified on Form 2, Field 2.6.	n	98
8.18	Students eligible for <b>LANGUAGE2</b> testing in the eligible grade	INT	Number of those in the eligible grade who had prior instruction in the given language for at least one full academic year	y	45



CODE	FIELD	TYPE7:	REMARK:	FIXED8	DATA EXAMPLE:
8.19	Students enrolled in the grade below the eligible grade of <b>LANGUAGE2</b>	INT	Number of all students in the grade below the eligible grade. E.g. if testing is among those in the final year of ISCED2, the number of those in the pre-final year of ISCED2 should be specified here.	n	98
8.20	Students learning <b>LANGUAGE2</b> in the grade below the eligible grade	INT	Number of all students currently learning the 2nd foreign language in the grade below the testing grade, i.e. those who will be eligible for testing in the next academic year.	y	45
8.21	Excluded	INT	This field specifies if the school is excluded or not. A code should be used where '0' means that the school is NOT excluded, and codes 1-n (corresponding to broad exclusion categories) mean that the school is not included in the sampling frame. Codes will be provided in Lookup 8.21, by <i>SurveyLang</i> , on basis of information collected from Form 3.) Note that this field will be provisional, subject of <i>SurveyLang</i> evaluation.	y	0
8.22	Detailed justification for exclusion	STR	Beyond categorising the reason of exclusion in 8.21, NRCs are requested to provide qualitative explanation as well, in sufficient detail that allows <i>SurveyLang</i> to understand the specific situation.	n	"Due to significant security problems stemming from the hostility between the local minority and the majority the access...etc."
8.23	Explicit stratum, <b>LANGUAGE1</b>	INT	If explicit stratification is applied, for each school a category (stratum) has to be specified, using the codes in Form 5A, from the field 5A.3(a)	y	1

CODE	FIELD	TYPE7:	REMARK:	FIXED8	DATA EXAMPLE:
8.24-8.26	Implicit strata <b>LANGUAGE1</b>	INT	If implicit stratification is applied, for each school, for each variable (up to three) a stratum has to be specified. As many columns as many stratification criteria Stratum codes should be used as specified in Form 6A, under points 6A.3 – 6A.5. In the spreadsheet, three columns are reserved for implicit strata. 8.24 corresponds to field 6A.2.1, 8.25 to 6A.2.2 and 6A.2.3 to 6A.2.3 on Form 6A.	y	1
8.27	Explicit stratum, <b>LANGUAGE2</b>  <b>FIELD TO BE REMOVED IF NO DIFFERENCE BETWEEN LANGUAGES</b>	INT	If explicit stratification is applied, for each school a category (stratum) has to be specified, using the codes in Form 5B, from the field 5B.3(a)	y	1
8.28-8.30	Implicit strata <b>LANGUAGE2</b>  <b>FIELD TO BE REMOVED IF NO DIFFERENCE BETWEEN LANGUAGES</b>	INT	If implicit stratification is applied, for each school, for each variable (up to three) a stratum has to be specified. As many columns as many stratification criteria Stratum codes should be used as specified in Form 6B, under points 6B.3 – 6B.5. Three columns are reserved for implicit strata. 8.28 corresponds to field 6B.2.1, 8.29 to 6B.2.2 and 8.30 to 6B.2.3 on Form 6B.	y	1



### 14.3.10 Form 9A. Student listing form A – Tab 1

#### Form 9A. Student listing form A

Form 9A is for schools sampled for only one test language, or two test languages but in separate grades. If the latter is the case, the school has to fill in two 9A Forms, separately providing the student lists for the two languages tested (in different grades).

9A.1 SL School ID:	
9A.2 School name:	
School address	9A.3 Street and number:
	9A.4 City:
	9A.5 Postal Code:
9A.6 Educational system:	
9A.7 Prepared by:	
9A.8 Telephone number:	
9A.9 Date:	
9A.10 Test language:	
9A.11 Number of students listed in the eligible grade :	
9A.12 Grade of testing:	
9A.13 Preferred testing mode:	



### 14.3.11 Form 9A. Student listing form A – Tab 2

ID	A. Student name			B. Gender	C1. Year of birth (YYYY)	C2. Month of birth (MM)	D. At least one year of training Yes=1 No=2	E. Exclusion (use codes)	F. Mark on the Routing test	G. Questionnaire Language*	H. Study programme
	First	Middle	Last	M=1 F=2							
1	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	1/2	FILL IN	FILL IN	FILL IN	FILL IN
2	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	1/2	FILL IN	FILL IN	FILL IN	FILL IN
3	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	1/2	FILL IN	FILL IN	FILL IN	FILL IN
4											
5											
6											



## 14.3.12 Form 9B. Student listing form B – Tab 1

### Form 9B. Student listing form B

Form 9B is for schools sampled for both test languages in the same grade.

9B.1 SL School ID:

9B.2 School name:

School address

9B.3 Street and number:

9B.4 City:

9B.5 Postal Code:

9B.6 Educational system:

9B.7 Prepared by:

9B.8 Telephone number:

9B.9 Date:

9B.10A Test language A:

9B.10B Test language B:

9B.11A: Number of students listed for French only (as in 9B.14):

9B.11B: Number of students listed for German only (as in 9B.14):

9B.11C: Number of students listed for French and German (as in 9B.14):

9B.11: Number of all students listed (A+B+C) (as in 9B.14):

9B.12: Grade of testing:

9B.13: Preferred testing mode:



### 14.3.13 Form 9B. Student listing form B – Tab 2

A. Student name			B. Gender	C1. Year of birth (YYYY)	C2. Month of birth (MM)	E. Exclusion (use codes)	G. Questionnaire Language	H. Study Programme	LANGUAGE 1 (pre-filled by SurveyLang)			LANGUAGE2 (pre-filled by SurveyLang)			
ID	First	Middle	M = 1 F = 2						L1. YES/ NO yes = 1 no = 2	D1. At least one year of training Yes=1 No=2	F1. Mark on the Routing Test	L2. YES/NO yes = 1 no = 2	D2. At least one year of training Yes=1 No=2	F2. Mark on the Routing Test	
1	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	FILL IN	FILL IN	FILL IN	1/2	1/2	FILL IN	1/2	1/2	FILL IN
2	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	FILL IN	FILL IN	FILL IN	1/2	1/2	FILL IN	1/2	1/2	FILL IN
3	FILL IN	FILL IN	FILL IN	1/2	YYYY	MM	FILL IN	FILL IN	FILL IN	1/2	1/2	FILL IN	1/2	1/2	FILL IN
...															



### 14.3.14 Form 10A. Teacher listing form A – Tab 1

10A.1 SL School ID:	PRE-FILLED
10A.2 School name:	PRE-FILLED
10A.3 School address:	STREET & NUMBER, PRE-FILLED
10A.4	CITY, PRE-FILLED
10A.5	POSTAL CODE, PRE-FILLED
10A.6 Educational system:	PRE-FILLED
10A.7 Prepared by:	FILL IN YOUR NAME
10A.8 Telephone number:	FILL IN (+EDUCATIONAL SYSTEMCODE-AREACODE-NUMBER)
10A.9 Date:	FILL IN (DD-MM-YYYY)
10A.10 Test language:	PRE-FILLED
10A.11 Number of teachers listed (as in 10A.13):	FILL IN
10A.12 Level of testing:	ISCED2/ISCED3, PRE-FILLED

### 14.3.15 Form 10A. Teacher listing form A – Tab 2

10A.13 English Teacher list (use as many rows as needed)

A .Teacher name			
ID	First	Middle	Last
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

### 14.3.16 Form 10B. Teacher listing form B – Tab 1

<b>Form 10B. Teacher listing form B</b>	
10B.1 SL School ID:	
10B.2 School name:	
School address:	10B.3. Street & number :
	10B.4. City :
	10B.5. Postal code :
10B.6 Educational system :	
10B.7 Prepared by :	
10B.8 Telephone number :	
10B.9 Date :	
10B.10A Test language A :	
10B.10B Test language B :	
10B.11A Number of teachers listed for Language 1 (as in 10B.13 – D.1) :	
10B.11B Number of teachers listed for Language 2 (as in 10B.13 – D.2) :	
10B.12A Level of testing, Language A :	
10B.12B Level of testing, Language B :	

### 14.3.17 Form 10B. Teacher listing form B – Tab 2

#### 10B.13 Teacher list (use as many rows as needed)

ID	A .Teacher name			D .Teaches (at the target level)	
	First	Middle	Last	1. Test	2. Test
				language A	language B
				yes = 1	yes = 1
				no = 2	no = 2
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					



### 14.3.18 Form T1. School tracking form

Form T1. School tracking form

	Stratum ID:	School Sample ID:	Educational system:
Target language:		Test administrator:	

Only one school from the list below can be sampled, in the order which is presented (the next school can only be approached if the previous one has refused to participate)

	SCHOOL INFO															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	SL School ID	National school ID	School ENR	Name, Address, Phone #	Name and phone # of school coordinator	Status	PQ yes = 1 no = 2	Issued student sample	Test mode (CB/PB)	(IF CB) Number of test stations available	(IF PB) Date test materials sent	Test date, first session	Test date, second session	(IF PB) Date test materials returned	Achieved student sample - tests	Achieved student sample - questionnaire
Sampled School																
First replacement school																
Second replacement school																



### 14.3.19 Form T2. Student tracking form

School name:	PRE-FILLED	Coordinator:	PRE-FILLED	Stratum ID:	PRE-FILLED	School ID:	PRE-FILLED	Educational system	PRE-FILLED			
Test administrator:						FILL IN NAME						
Outcome use codes 0-8												
1. SL Student ID	2. Name First, Middle Initial, Last)	3. Gender (M=1 F=2)	4. Year of birth (YYYY)	5. Month of birth (MM)	6. Exclusion code	7. Skills tested L – Listening R – Reading W- Writing	Test difficulty (1/2/3) 1 - low 2 - medium 3 - high	9a. R (Reading test completed)	9b. L (Listening test completed)	9c. W (Writing test completed)	10. Student Questionnaire	11. #USB
PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	FILL-IN	PRE-FILLED	PRE-FILLED	FILL-IN	FILL-IN	FILL-IN	FILL-IN	FILL-IN
PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	FILL-IN	PRE-FILLED	PRE-FILLED	FILL-IN	FILL-IN	FILL-IN	FILL-IN	FILL-IN
PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	FILL-IN	PRE-FILLED	PRE-FILLED	FILL-IN	FILL-IN	FILL-IN	FILL-IN	FILL-IN
PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	PRE-FILLED	FILL-IN	PRE-FILLED	PRE-FILLED	FILL-IN	FILL-IN	FILL-IN	FILL-IN	FILL-IN

## **14.4 ESLC Technical Standards**

This document is an adaptation of the PISA technical standards of the 2006 cycle. SurveyLang would like to thank the PISA consortium for kindly making their standards available.

### **14.4.1 Purpose of document**

The purpose of this document is to list the set of standards upon which the ESLC 2010 and 2011 data collection activities will be based.

The standards described in this document have been developed for the purpose of creating an international dataset of a quality that will enable valid cross-national inferences to be made.

There are three types of standards; each type has a specific purpose. Data Standards ensure that all collected data can be added to the final ESLC 2011 dataset that will be released by the Commission. Management Standards ensure that all ESLC operational objectives are met in a timely and coordinated manner. National Involvement Standards ensure that the internationally developed instruments meet the highest standards of cross-national, cross-cultural and cross-linguistic validity and equivalence and that the ESLC results have the greatest possible meaning for national stakeholders.

### **14.4.2 Format of document**

The standards are grouped into sections that relate to specific tasks in the ESLC data collection process. The standards in each section consist of two distinct elements. First, there are the Standards themselves that are numbered and are shown in **shaded boxes**. Second, there are Notes that provide additional information pertaining directly to the standards. The notes are listed after the standards in each section. In addition, the standards contain words that have a defined meaning in the context of the standards. These words are shown in italics throughout the document and are defined in the Definitions section at the end of the document, where the terms are listed alphabetically.

### **14.4.3 Data Standards**

#### **Target population and sampling**

Standard 1.1

The ESLC Desired Target Population must be agreed upon.

Standard 1.2

Students participating in the test must be ESLC-Eligible students.

#### Standard 1.3

The testing period must be no longer than six consecutive weeks, and be inside the ESLC testing window, unless otherwise agreed upon.

#### Standard 1.4

Schools must be sampled using agreed upon, established and professionally recognised principles of scientific sampling.

#### Standard 1.5

National Research Coordinators must identify appropriate stratification variables to reduce sampling variance when appropriate.

#### Standard 1.6

Students must be sampled using *agreed upon*, established and professionally recognised principles of scientific sampling and in a way that represents the full population of *ESLC-Eligible students*.

#### Standard 1.7

The *ESLC Defined Target Population* must cover 95 percent or more of the *ESLC Desired Target Population*. Acceptable exclusions are as follows:

school level exclusions that are exclusions due to geographical inaccessibility, extremely small school size, where administration of ESLC would be not feasible

within the school, where all students would be within-school exclusions and other agreed upon reasons and that total to less than 2.5 percent of the ESLC Desired Target Population,

School-level and within-school exclusions must total less than 5 percent of the ESLC Desired Target Population.

#### Standard 1.8

School response rates should be above 85 percent of sampled schools. If a response rate is below 85 percent then an acceptable response rate can still be achieved through *agreed upon* use of replacement schools.

#### Standard 1.9

The student sample size must be a minimum of 1,500 assessed students per language for each *educational system*, or the *ESLC Defined Target Population* where the *ESLC Defined Target Population* is below 1,500, unless otherwise agreed.

#### Standard 1.10

The school sample size must be a minimum of 71 schools per language for each *educational system*, or all schools that have students in the *ESLC Defined Target Population* where the number of schools with students in the *ESLC Defined Target Population* is below 71, unless otherwise agreed.

#### Standard 1.11

The *target cluster size* is typically a national average of 25 *ESLC-Eligible students* which *upon agreement* can be increased, or reduced to a number not less than 10.

#### Standard 1.12

More than 80 percent of all sampled students must respond to the test.

#### Note 1.1

The Target Population and Sampling standards apply to the Main Study but not the Field Trial.

#### Note 1.2

Data from schools where the student response rate is greater than 25 percent will be included in the SurveyLang dataset.

#### Note 1.3

A participating school is defined as a sampled school in which more than 50 percent of sampled students respond.

### Language of tests and questionnaires

#### Standard 2.1

The language of the ESLC language test (including the rubrics) is exclusively the *target language*. The language of the student background questionnaire must be a *questionnaire language*.

The language of the teacher background questionnaire is the *questionnaire language* at the school unless otherwise *agreed upon*.

The language of the school background questionnaire is the *questionnaire language* at the school unless otherwise *agreed upon*.

The language of the national background questionnaire is English.

#### Note 2.1

Although the language test rubrics will be in the target language, guidelines on how to respond to the test tasks will be made available by the National Research Coordinator to students, teachers and test administrators in the *questionnaire language*.

### Field Trial participation

#### Standard 3.1

*ESLC participants* participating in the ESLC 2011 Main Study must successfully implement the Field Trial in 2010.

The number of schools participating in the Field Trial is set to a minimum 40 per tested language unless otherwise *agreed upon*.

The *target cluster size* is typically a national average of 25 *ESLC-Eligible students* which *upon agreement* can be increased, or reduced to a number not less than 10.

#### Note 3.1

The ESLC Technical Standards for the Main Study generally apply to the Field Trial, except for the Target Population and Sampling Standard. For the Field Trial a sampling plan needs to be *agreed upon*. More information is given in the ELSC sampling manual.

### Adaptation of test rubrics, questionnaires, test administrator manual and school coordinator manual

#### Standard 4.1

The questionnaire instruments must be equivalent to the *source version*. They must be adapted to the local context only if needed and *agreed upon*.

#### Standard 4.2

The Test Administrator Manual and the School Coordinator guidelines must be equivalent to the *source version* and must be adapted to the local context only if needed and *agreed upon*. Adaptations must not alter the intent of the manual.

### Translation of tests, questionnaires, manuals and guidelines

#### Standard 5.1

The following documents are classified as *type A documents*.

- Testing tool navigation details
- The student, teacher and principal questionnaires

#### Standard 5.2

The following documents are classified as *type B documents*.



- Brief instructions for test administrators
- Language test familiarisation materials
- Instructions for marking paper-based language tests
- Routing test instruction sheet
- Sampling forms and tracking forms (student and teacher level)
- Sampling FAQ
- Short sampling guidelines
- Testing tool guidelines
- Testing tool training manual for the students
- Test administration FAQ
- Test administration manual
- School Coordinator guidelines

**Note 5.1**

Documents not listed in standard 5 are so-called type C documents, and do not need to be translated. Technical documentation on sampling and testing tool and the translation manuals are type C documents. The national questionnaire (to be filled out by the National Research Coordinator) is also type C, as well as all documents made available by SurveyLang for use by the National Research Coordinator.

**Note 5.2**

The marking guidelines for the Writing tests will be made available by SurveyLang in the five target languages, and therefore they are to be considered as type C documents.

## Test administration

### Standard 6.1

All test sessions must follow procedures as specified in the *ESLC operations manuals*, particularly the procedures that are:

- related to test session timing,
- for maintaining test conditions,
- for student tracking, and
- for assigning booklets.

### Standard 6.2

Test Administrators must be trained according to *agreed procedures*.

### Standard 6.3

Test Administrators must be trained in person.

### Standard 6.4

The relationship between the Test Administrator and the participating student must not compromise the credibility of the test session. In particular, the Test Administrator should not be the teacher of the language any of the participating students is tested in.

### Note 6.1

Generally it is preferred that Test Administrators are not staff of the school.

## Background questionnaires

### Standard 7.1

Adjudicated entries may contribute up to five additional questions to each contextual questionnaire.

## Security of the material

### Standard 8.1

ESLC materials designated as secure must be kept confidential at all times. Secure materials include all test materials, data, and draft materials. In particular:

- no person other than approved project staff and students participating in the test session are able to access the room during the session or view the test material,

### Note 8.1

As per the SurveyLang Confidentiality Agreement, no person other than approved project staff will have access to secure ESLC data and embargoed material. The National Research Coordinator is responsible for ensuring confidentiality amongst all approved project staff.

## Quality monitoring

### Standard 9.1

The National Research Coordinator must nominate an *agreed upon* number of *ESLC Quality Monitors* a minimum of six weeks prior to the Field Trial and Main Study *testing period*.

### Standard 9.2

The nominees for *ESLC Quality Monitors* must:

- be knowledgeable about testing procedures,
- fluently speak the questionnaire language and be sufficiently fluent in English to communicate effectively with SurveyLang and to effectively implement required procedures,
- have a background in education, assessment or research,
- be sensitive to the needs of schools and students and feel comfortable in a school environment,
- be able to represent ESLC in schools diplomatically and positively,
- be able to conduct their work in such a way as to ensure that it is independent of the National Research Coordinators within the countries where they are working,
- not be a member of the same institution where the National Research Coordinator works or where the National Research Coordinators is based,
- not be an immediate relative of an employee at the National Research Coordinators,
- not report to the National Research Coordinator in their day-to-day work, and
- be able and have the capacity to independently and effectively communicate with SurveyLang using email.

### Standard 9.3

The ESLC Quality Monitoring visits must be unannounced.

### Standard 9.4

The National Research Coordinator must assist the *ESLC Quality Monitors* in performing their duties by:

- providing a list of all participating schools that includes the school name, complete address, the School Coordinator's name and phone number, the Test Administrator's name for each testing session, and the dates and times of all test sessions,
- forwarding all amendments to the above information as soon as the National Research Coordinator becomes aware of a change, and
- establishing protocols to enable ESLC Quality Monitors to monitor testing sessions unannounced.

**Standard 9.5**

The National Research Coordinator must train the *ESLC Quality Monitors* according to the quality auditor logistics sheet to be provided by SurveyLang.

**Note 9.1**

The Quality Monitoring standards apply to both the Main Study and the Field Trial.

**Printing of material****Standard 10.1**

All student assessment material must be printed using an *agreed upon* paper quality, print quality and according to the standard specified by SurveyLang. National Research Coordinators will submit a sample of printed material to SurveyLang for agreement. A sample test and questionnaire booklet per language will be provided by SurveyLang. National Research Coordinators must demonstrate the print quality and the correct collation and stapling of the sample booklets as specified in the *ESLC operations manuals*.

**Note 10.1**

Standard 9 only applies to paper based testing. Tests will be provided as ready-to-print .PDF files to the National Research Coordinators.

**Note 10.2**

The high degree of individualisation of test booklets requires professional printing facilities. Booklets are individualised at student level, printed black-and-white double-sided on A3 paper and centre-stapled.

**Note 10.3**

Printing should thus be undertaken centrally in each educational system.

## Response marking, coding and data entry

### **Standard 11.1**

The questionnaire coding scheme as described in the coding guidelines will be implemented in the manner described by SurveyLang test developers.

### **Standard 11.2**

One or more representatives from each National Research Coordinators will attend the international ESLC marking training session for both the Field Trial and the Main Study.

### **Standard 11.3**

Markers, coders and data entry staff will be recruited and trained by National Research Coordinators following *agreed upon* procedures.

### **Standard 11.4**

National Research Coordinators will provide a sample of marked writing scripts to SurveyLang as specified in the *ESLC operations manuals*.

### **Note 11.1**

Marking of Writing applies to paper-based as well as computer-based responses.

### **Note 11.2**

Marking for the Reading and Listening tests is done automatically in Computer based testing; for Paper based testing Reading and Listening requires only entry of response data.

### **Note 11.3**

Coding and data entry of open questionnaire responses applies to paper-based as well as computer-based questionnaires. Data entry of all other questionnaire responses applies only to paper-based questionnaires.

#### **14.4.3.1.1 Data submission**

### **Standard 12.1**

Data from computer-based administration must be retrieved from PCs and submitted as specified in the *ESLC operations manuals*.

### **Standard 12.2**

Data from paper-based administration must be submitted using the data entry applications provided by SurveyLang as described in the *ESLC operations manuals*. This includes the test data, questionnaire data, and tracking data.

#### 14.4.4 Management Standards

##### Schedule for completion of tasks

**Standard 13.1**

Unless otherwise *agreed upon*, National Research Coordinators will complete tasks and submit materials according to the schedule of tasks maintained on the ESLC Basecamp website.

**Standard 13.2**

Any variations to the agreed upon Schedule of National Research Coordinator tasks on the ESLC Basecamp website must be further *agreed upon*.

**Standard 13.3**

Procedures laid down in the *ESLC operations manuals* will be followed for all tasks.

##### Drawing samples

**Standard 14.1**

SurveyLang will be responsible for drawing both the school sample and the student sample for the Main Study, and only the student sample for the Field Trial.

##### Management of data

**Standard 15.1**

National response data must be forwarded to SurveyLang within six weeks of the last day of testing for the Field Trial and within eight weeks of the last day of testing for the Main Study, unless otherwise *agreed upon*.

**Standard 15.2**

National Centres must execute data checking procedures as specified in the *ESLC Operation Manuals* before submitting data to SurveyLang.

**Standard 15.3**

National Centres must make a data manager available upon submission of the database. The data manager must be:

- authorised to respond to Consortium data queries,
- available for a three-month period immediately after the database is submitted unless otherwise agreed upon,
- able to respond to Consortium queries within three working days, and

- able to resolve data discrepancies.

### Archiving of materials

#### Standard 16.1

National Centres must develop and maintain an electronic archive of all *type A* and *type B* documents as listed in Standard 5.

#### Standard 16.2

Unless otherwise requested, National Centres will archive all Field Trial materials until the commencement of the Main Study, and all Main Study materials until the publication of the international report. Materials to be archived must include:

- student booklets and Questionnaires,
- sampling forms,
- student lists,
- student tracking forms, and
- all data submitted to SurveyLang.

### 14.4.5 National involvement standards

#### National report

##### Standard 17.1

National Research Coordinators are responsible for writing a national report for their educational system.

#### National feedback

##### Standard 18.1

National Research Coordinators are responsible for developing a Communications Plan. National Research Coordinators are responsible for planning communication with educational authorities on every level in their educational system and for finding the most suitable tools and means to achieve this.

##### Standard 18.2

National Research Coordinators are responsible for promoting participation, effective implementation, and dissemination of results amongst all relevant national stakeholders.

##### Standard 18.3

National Research Coordinators should take all available measures to convince

schools to participate in the survey. National Research Coordinators need to obtain permission, where required, from the appropriate authorities to approach schools. In some systems, it is not enough to notify the appropriate authorities; permission must be obtained. This includes permission from one governmental level to contact another and/or permission from a governing board to contact individual schools.

**Standard 18.4**

National Research Coordinators are responsible for providing information about legislation in their educational system concerning privacy rules, accessibility etc.

**Standard 18.5**

National Research Coordinators are responsible for providing feedback to SurveyLang on the development of instruments, the adaptation of instruments, and other domain related matters that represent the perspectives of the key national stakeholders.

## 14.5 Definitions

**educational system**

A educational system, geographic region, or similarly defined population, for which SurveyLang fully implements quality assurance and quality control mechanisms and endorses, or otherwise, the publication of separate ESLC results.

**agreed procedures**

Procedures that are specified in the *ESLC operations manuals*, or variations that are *agreed upon* between the National Research Coordinator and SurveyLang.

**agreed upon**

Variations and definitions agreed upon between the National Research Coordinator and SurveyLang.

**ESLC Defined Target Population**

All *ESLC-Eligible students* in the schools that are listed on the school sampling frame. That is, the *ESLC Desired Target Population* minus exclusions.

**ESLC Desired Target Population**

The *ESLC Target Population* defined for a specific *educational system*. It provides the most exhaustive coverage of *ESLC-Eligible students* in the *educational system* as is feasible.



## **ESLC-eligible students**

Students who are in the *ESLC Target Population*.

**ESLC operations manuals** - manuals provided by SurveyLang, that is the following:

- Test Administrator Manual including security and receipt control protocols,
- Sampling Manual,
- Testing tool manual,
- NRC Field Operations manual,
- Translation role manuals.

The following guidelines will be provided:

- School Coordinator guidelines,
- Marking guidelines,
- Marking of Writing guidelines,
- Translation guidelines,
- Testing tool guidelines,
- Sampling guidelines,
- Instructions for marking paper-based language tests and questionnaires,
- Brief instructions for test administrators,
- Language test familiarisation materials.

The preparation of the *ESLC operations manuals* will be carried out by SurveyLang and will describe procedures developed by SurveyLang.

## **ESLC participant**

An administration centre, commonly called a National Research Centre, that is managed by a person, commonly called a National Research Coordinator, who is responsible for administering the ESLC in an *educational system*. The National Research Coordinator must be authorised to communicate with SurveyLang on all operational matters relating to the *educational system* for which the National Research Coordinator is responsible.

## **ESLC Quality Monitor**

A person nominated by the National Research Coordinator to monitor test administration quality in an *educational system*.

## **ESLC target population**

The target population for each language in an *educational system* consists of students enrolled in ISCED2 level (final year) or after the first completed year of ISCED3 level. What we call the ESLC international target population in each *educational system* corresponds to the total number of students in eligible grades (ISCED2 or ISCED3) that are:

- attending educational institutions located within the educational system, #
- studying the specific language to be tested for a defined minimum period of one academic year prior to the testing year (further referred to as “eligible students”).

Note that students at ISCED3 level are to be sampled only when the language is not taught at ISCED2 level, unless otherwise agreed upon.

## **ESLC testing window**

A two-month interval in which the *testing period* must be contained.

## **questionnaire language**

The language that the questionnaires, testing tool navigation details, sampling forms, guidelines and manuals will be administered and available in. This language, to be agreed upon with SurveyLang, must be one of the official languages within the Educational system which is used in most or most important communicative situations (for work, life in society, etc.) in the region where the school is located and that is the language of instruction in the school's region.

The decision about questionnaire languages within an educational system is taken at educational system level and *agreed upon* with SurveyLang. However, implementation is at school level. Thus, all documentation available within a school will be in one questionnaire language, unless otherwise *agreed upon*.

### **school level exclusions**

Exclusion of schools from the sampling frame because:

- of geographical inaccessibility (but not part of a region that is omitted from the ESLC Desired Target Population),
- of extremely small size,
- administration of the ESLC assessment within the school would not be feasible,
- all students in the school would be within-school exclusions, or
- of other reasons as agreed upon.

### **source version**

Documents provided in English by SurveyLang.

### **target cluster size**

The number of students that are to be sampled from schools where not all students are to be included in the sample.

### **target language**

The languages to be tested in the survey, which are English, French, German, Italian and Spanish.

### **testing period**

Six-week period of time during which data is collected in an *educational system*.

### **Type A documents**

The translation process comprises double translation from the English source version, reconciliation, back translation and sign-off by SurveyLang, unless otherwise *agreed upon*.

### **Type B documents**

The translation process comprises single translation from the English source version, reconciliation and quality control sign-off by the National Research Coordinator through content checklists, unless otherwise *agreed upon*.

### **Type C documents**

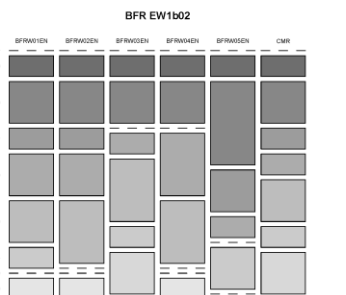
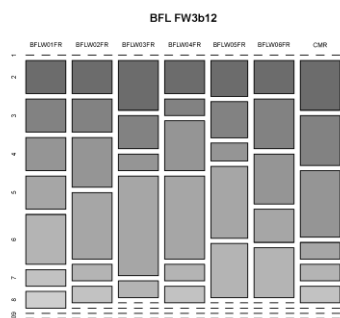
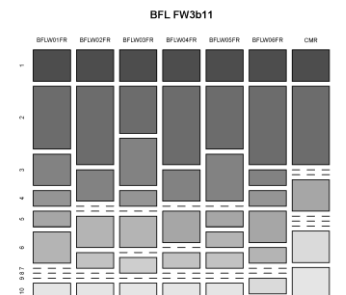
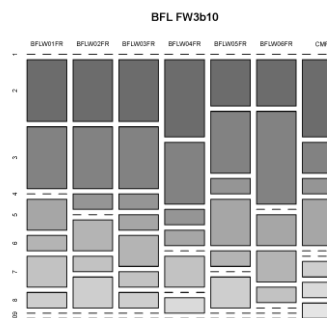
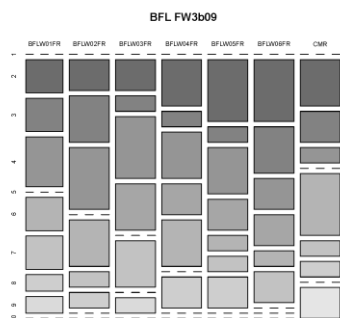
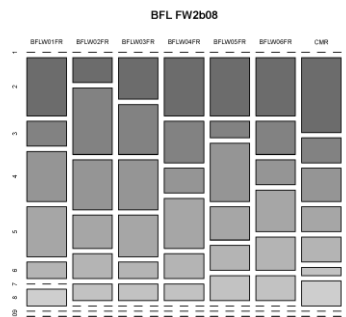
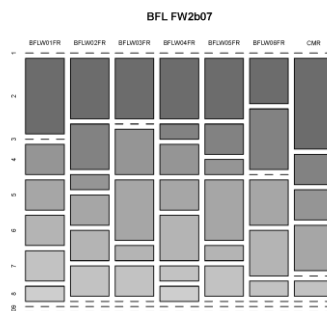
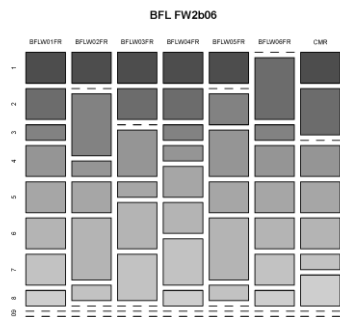
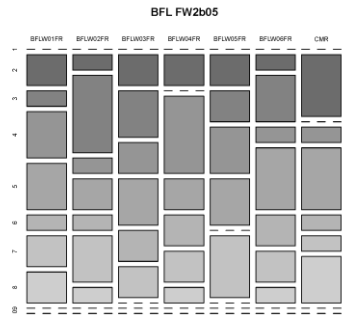
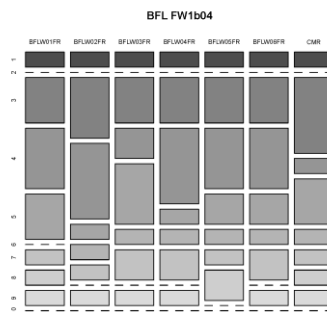
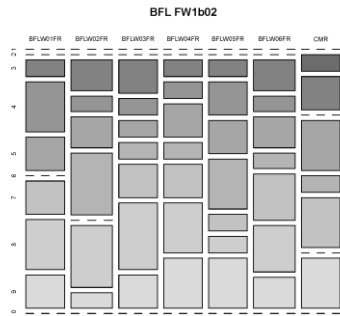
Documents which do not need to be translated.

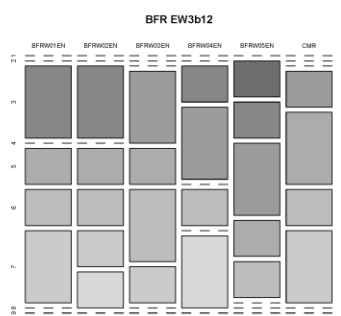
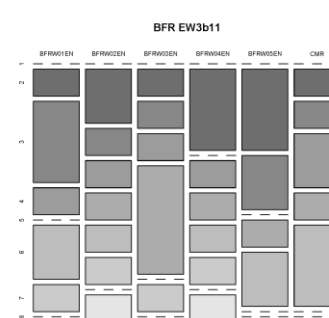
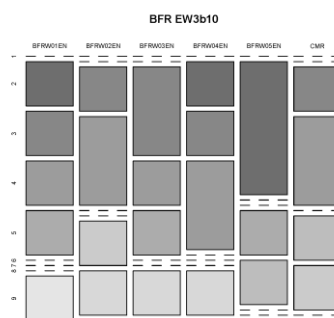
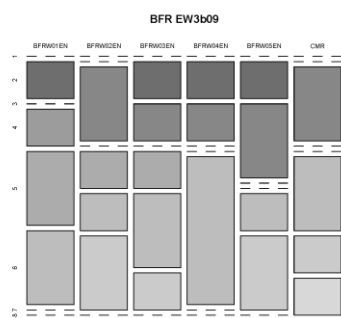
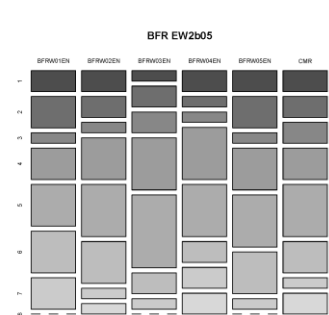
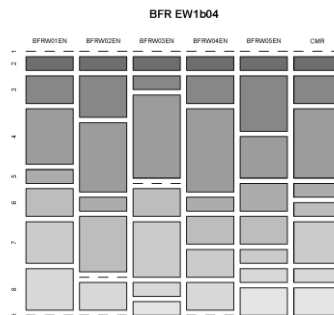
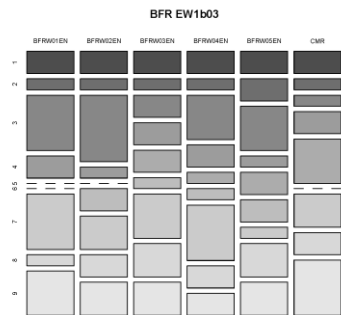
### **within-school exclusions**

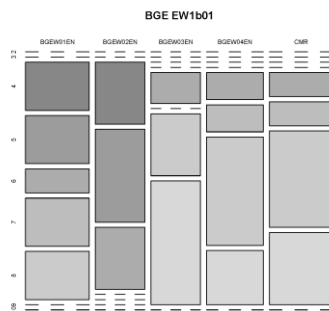
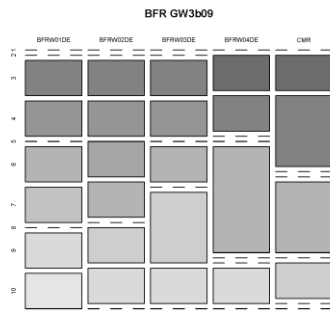
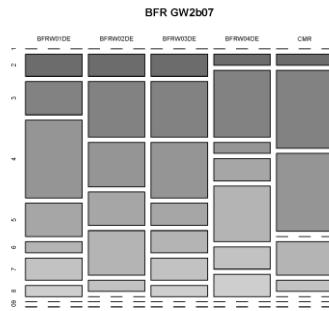
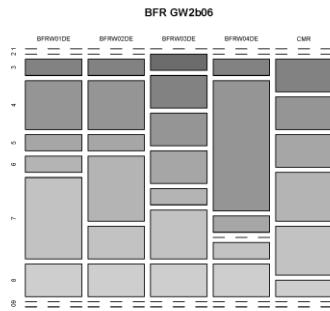
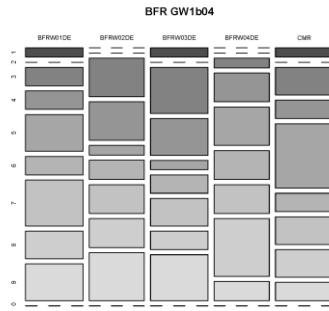
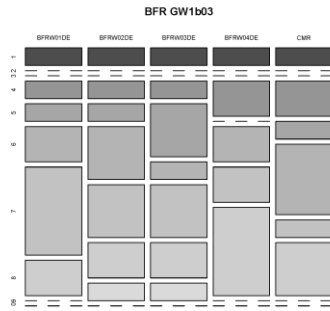
Exclusion of students from potential assessment because of one of the following:

- Functionally disabled students – The students suffer from a permanent disability that prevents them from taking part in the ESLC test. The exclusion will not apply to functionally disabled students who actually have the physical ability to participate.
- Intellectually disabled students – This intellectual disability should have been previously diagnosed by professionals such as the school principal, qualified staff members or psychologists. Students who are emotionally or mentally not capable to follow even the general instructions of the test will be included in this group. However, students who do not do well academically or have standard discipline problems will not be considered to fall under that category.
- Students with limited command of the official language of the educational system where they will be taking the test, and in which the instructions will be provided.
- Students who have received less than one year of instruction in the assessment language.
- Students who are not assessable for some other reason as agreed upon.

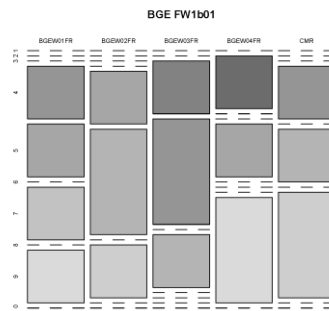
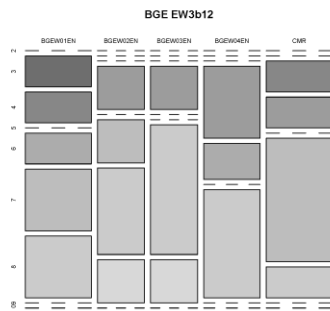
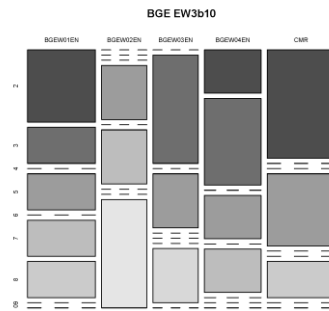
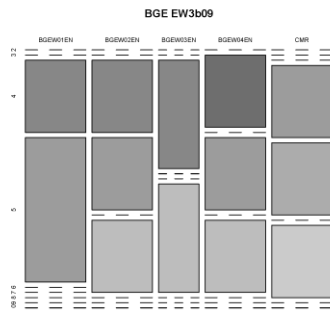
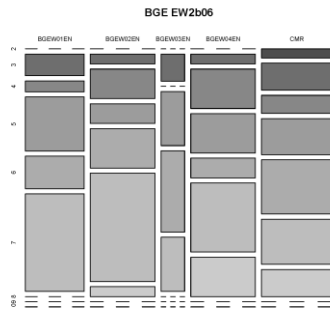
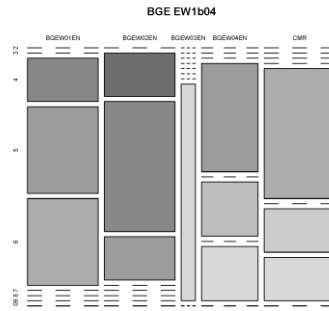
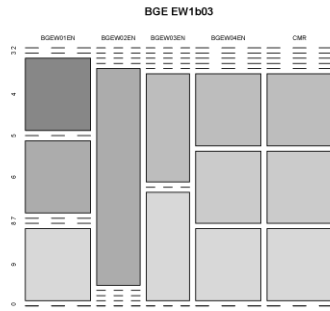


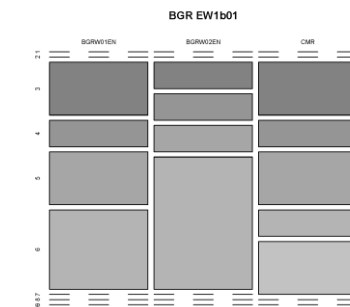
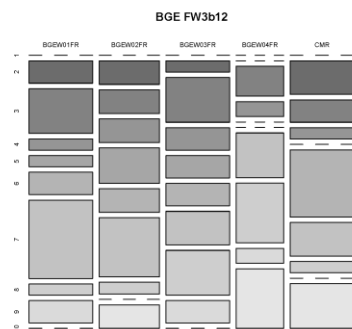
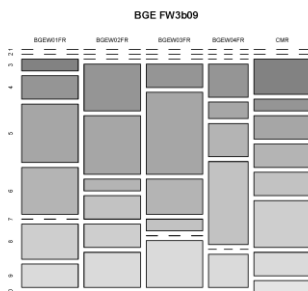
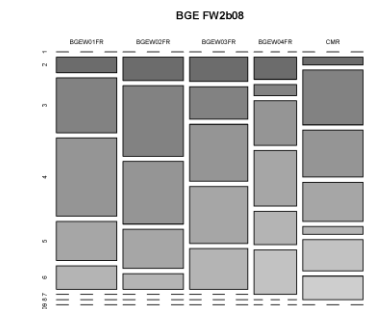
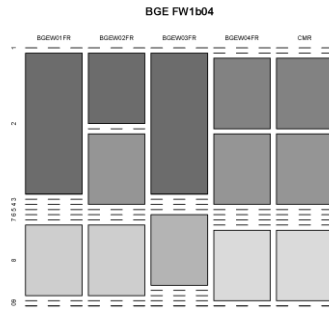
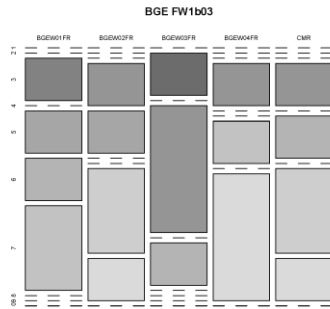


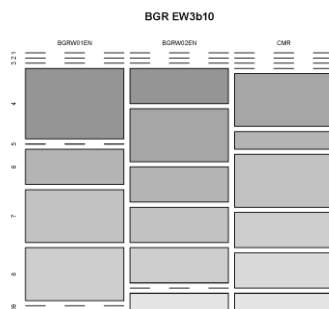
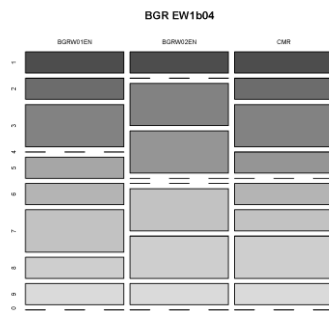


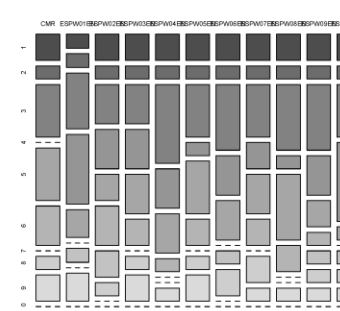
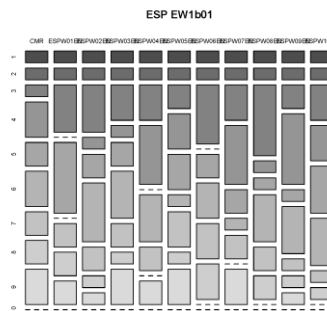
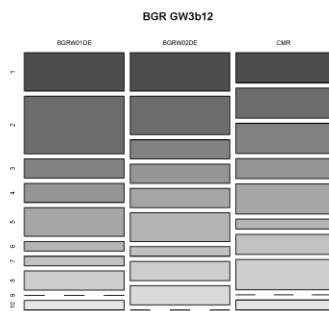
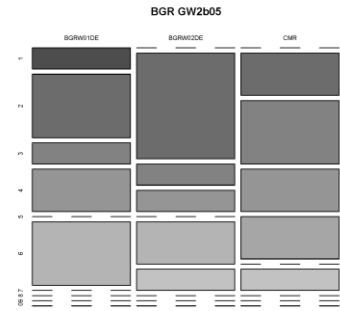
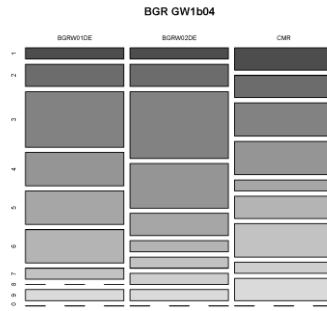
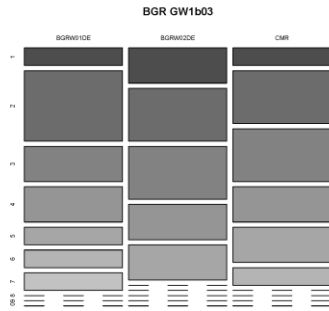




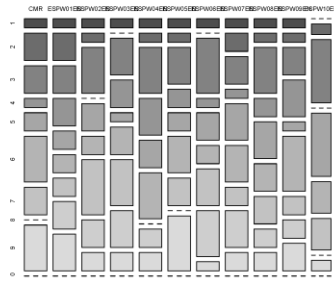




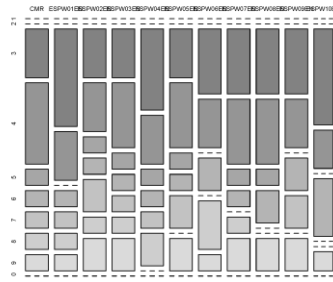




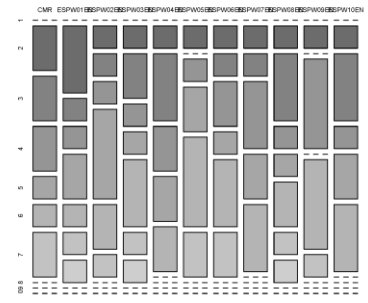
ESP EW1b03



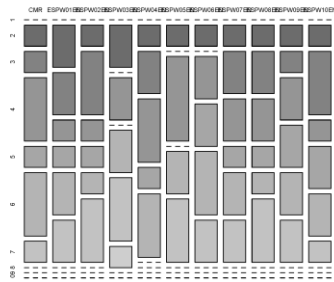
ESP EW1b04



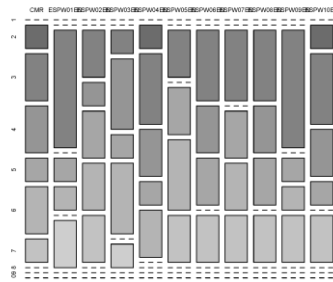
ESP EW2b05



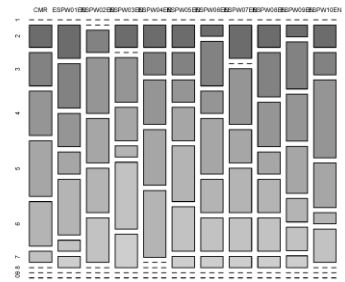
ESP EW2b06



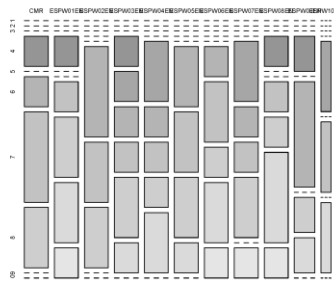
ESP EW2b07



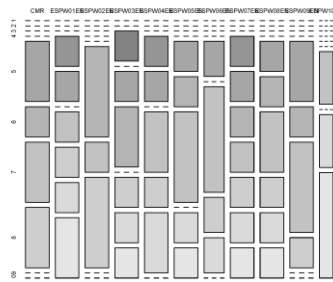
ESP EW2b08



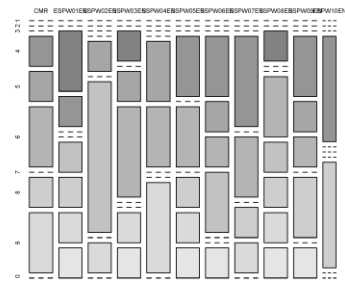
ESP EW3b09



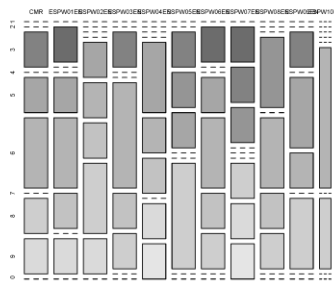
ESP EW3b10



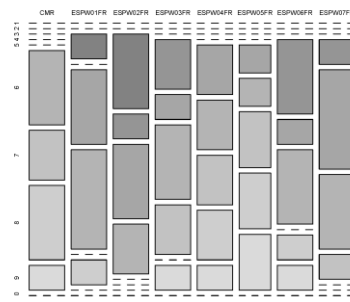
ESP EW3b11



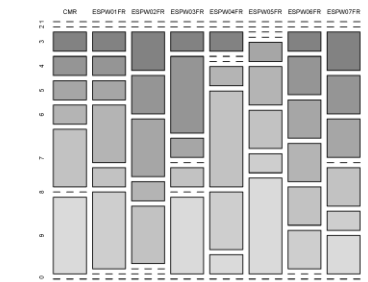
ESP EW3b12

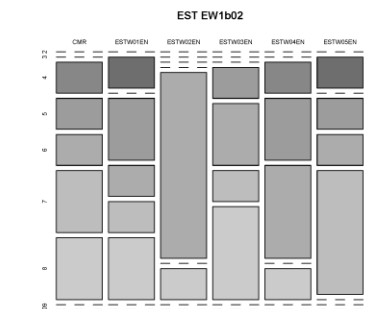
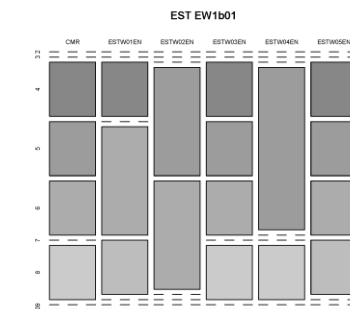
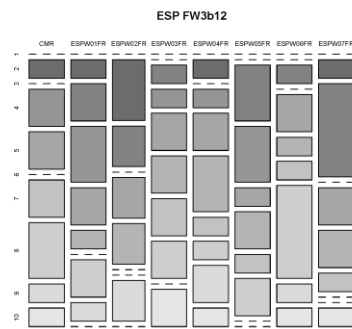
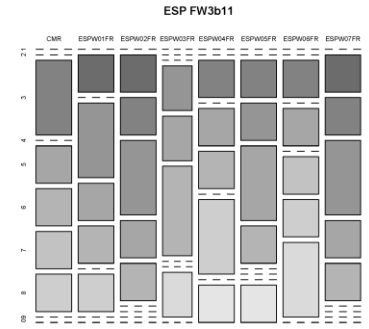
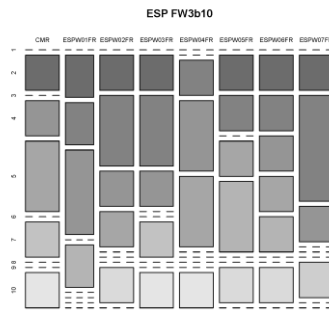
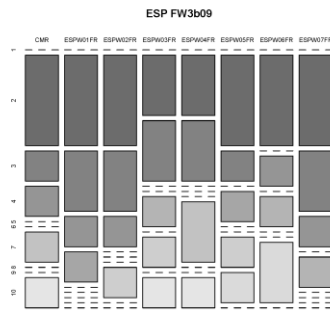
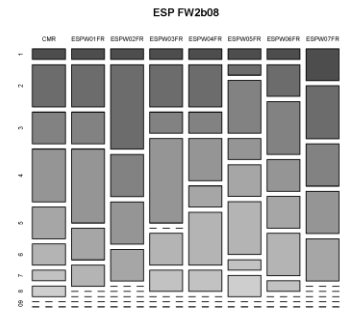
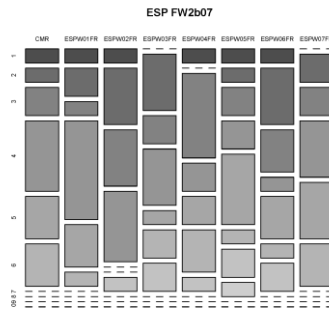
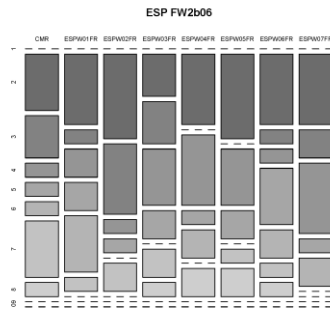
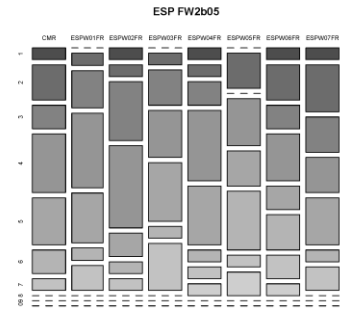
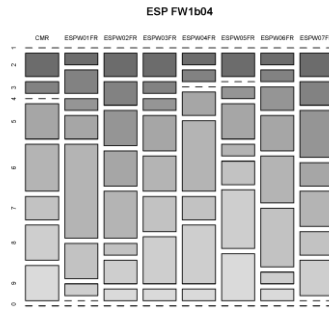
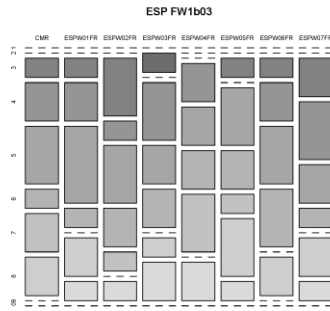


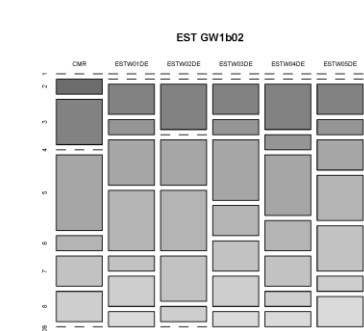
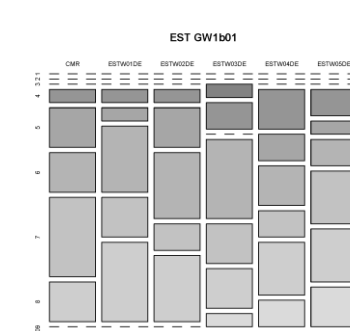
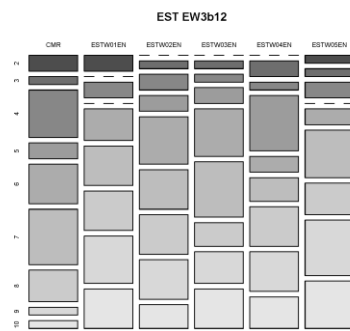
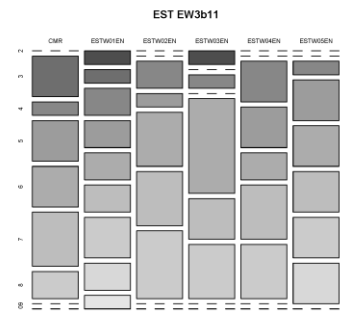
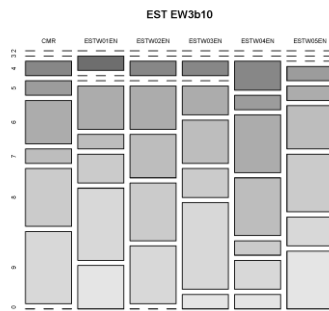
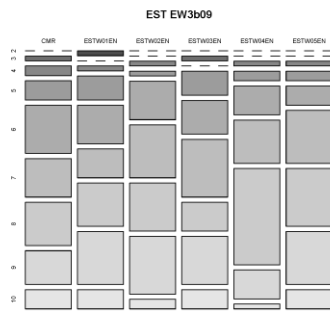
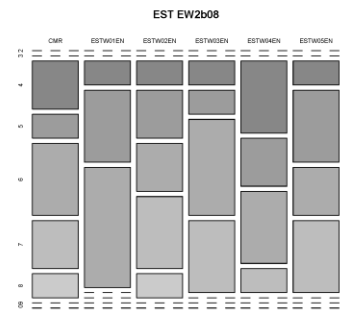
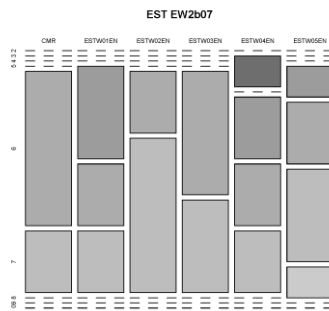
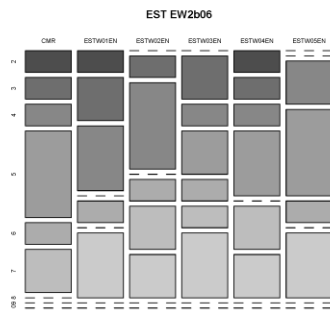
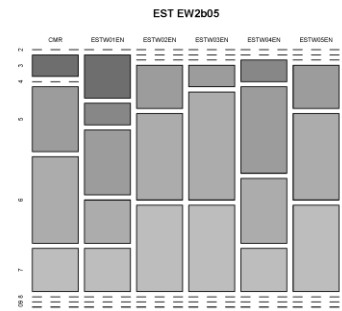
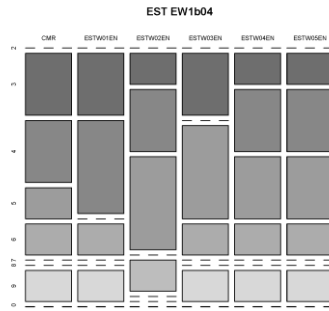
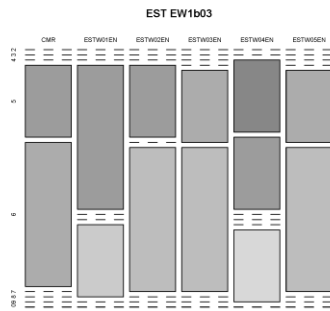
ESP FW1b01

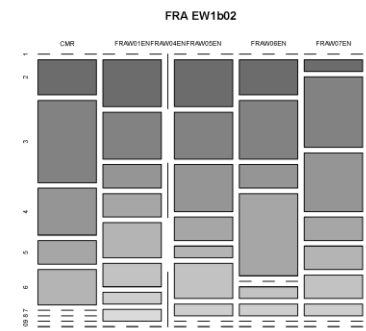
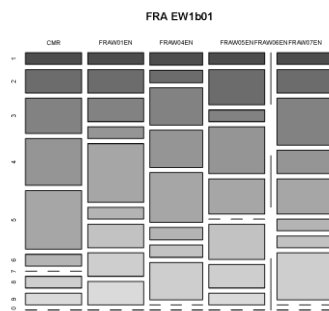
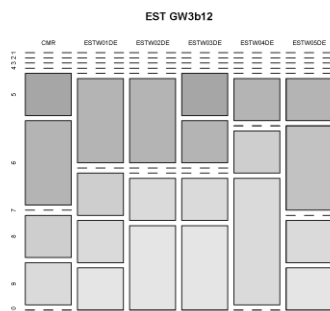
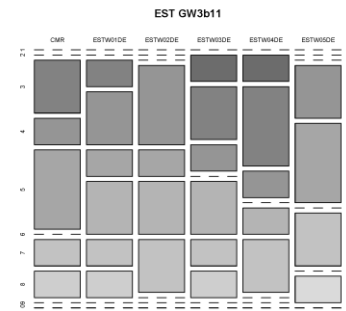
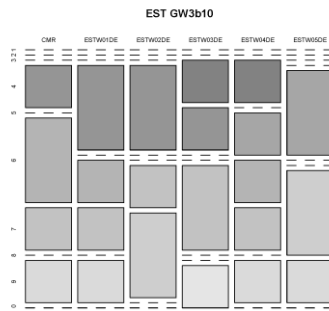
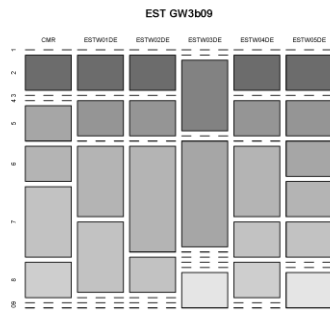
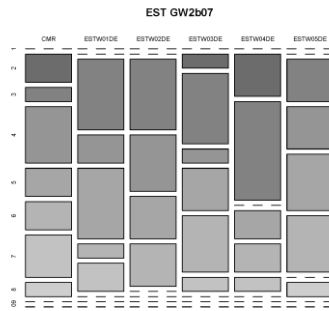
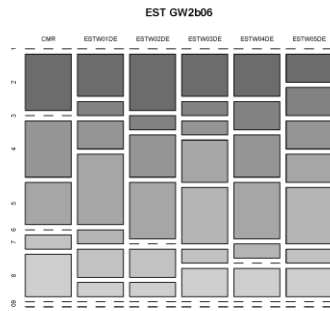
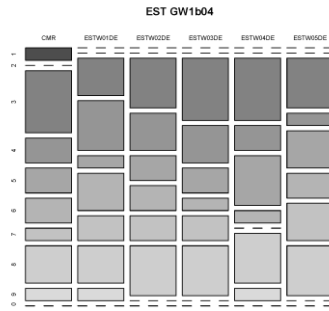
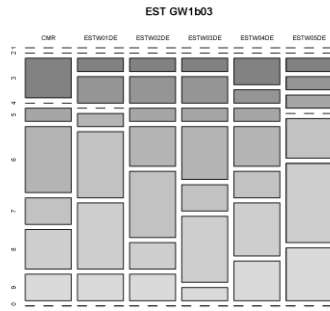


ESP FW1b02

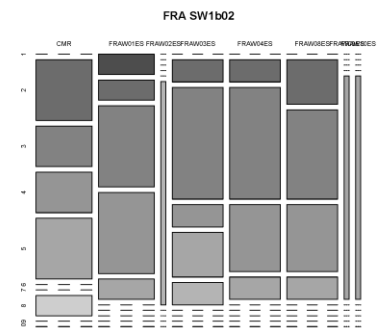
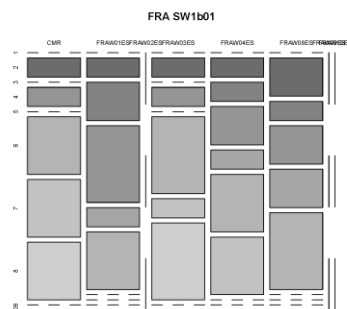
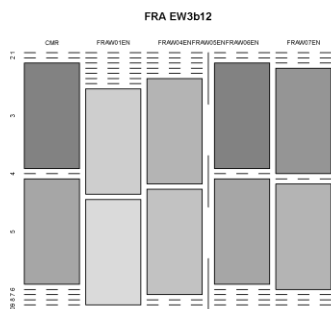
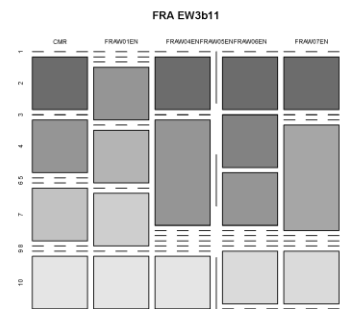
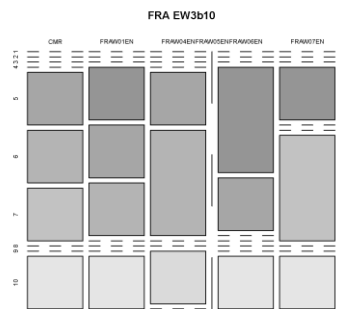
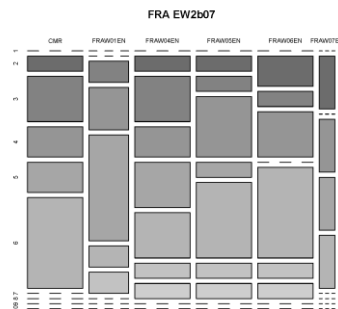
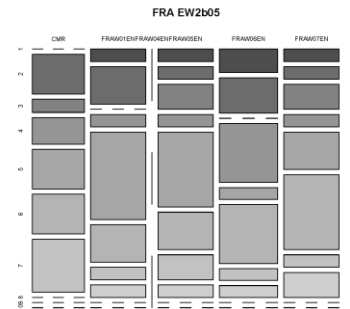
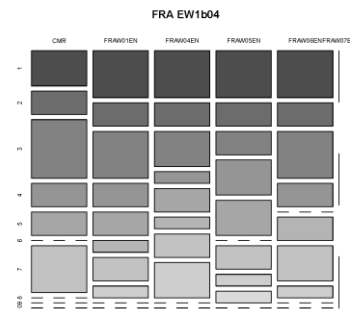
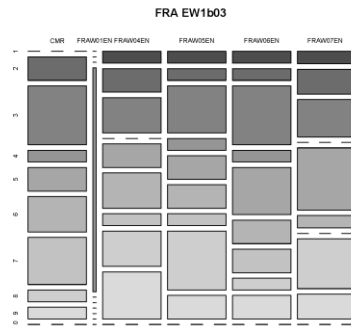


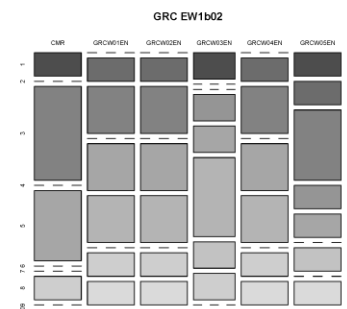
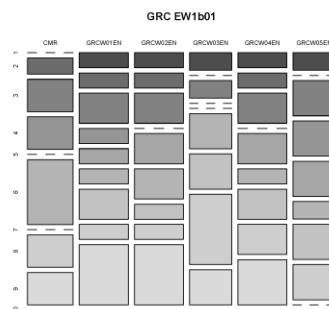
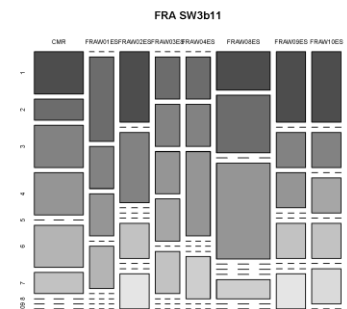
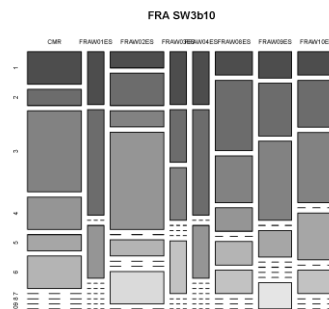
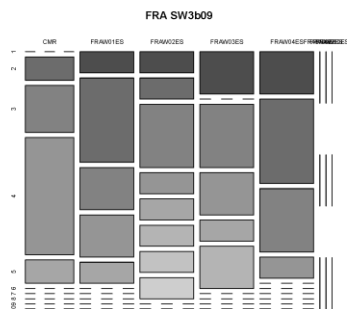
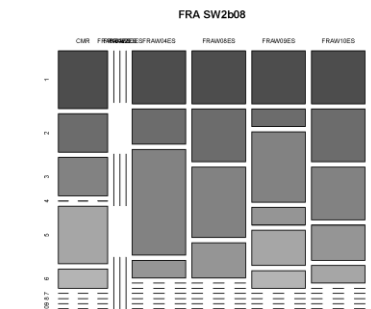
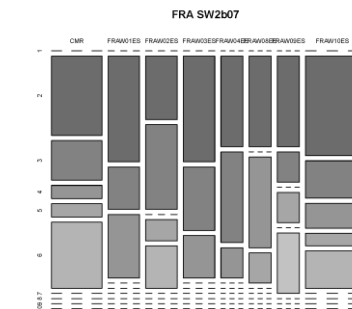
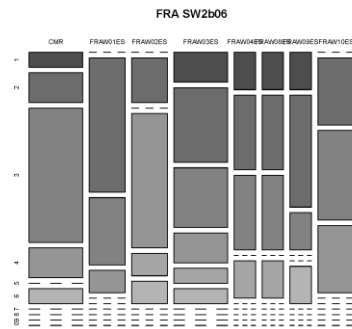
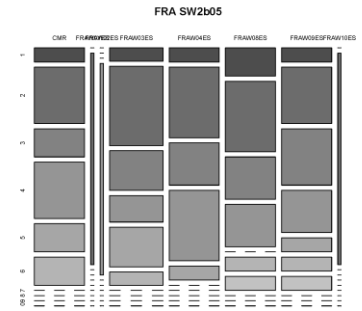
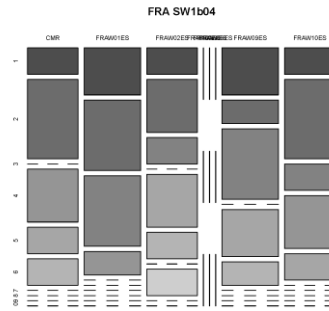
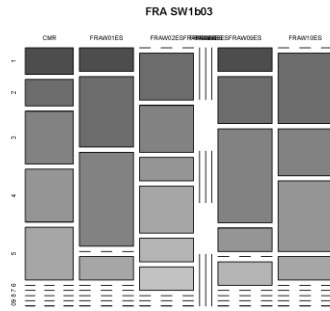


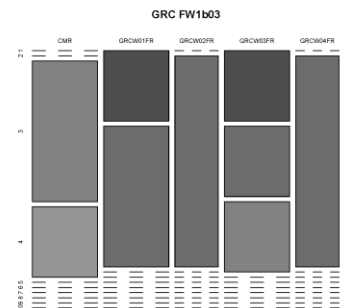
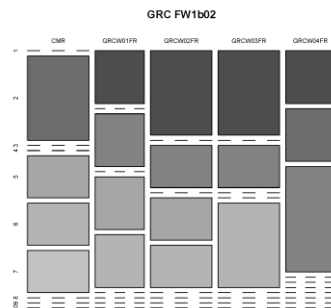
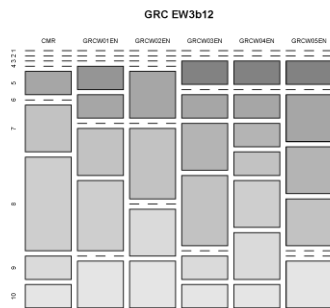
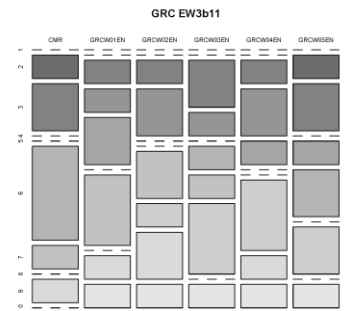
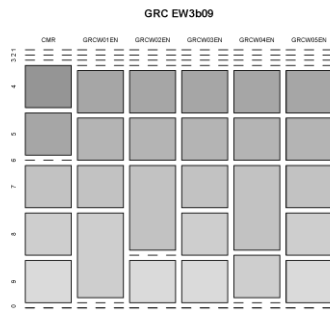
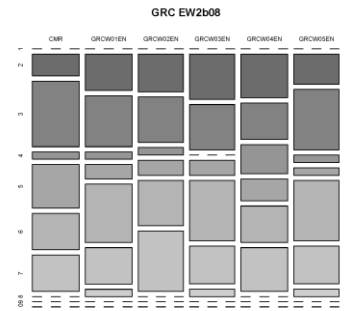
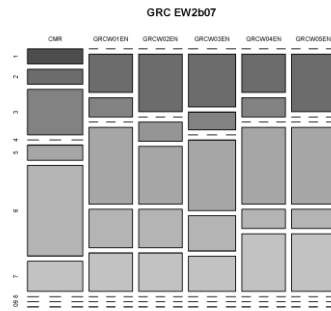
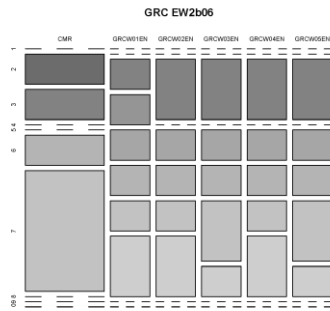
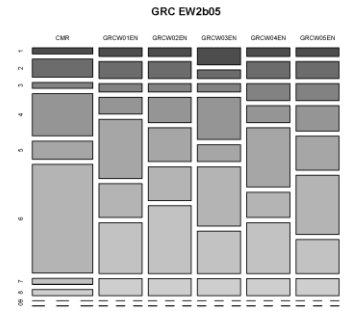
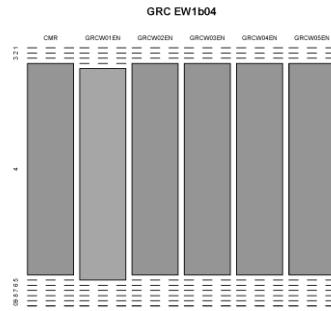
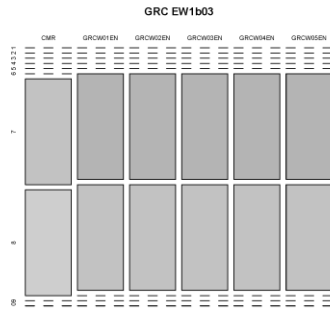


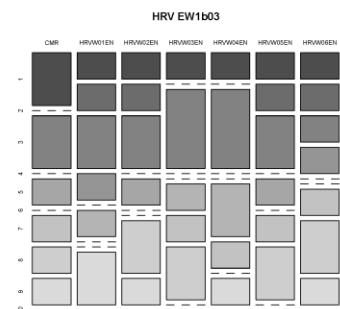
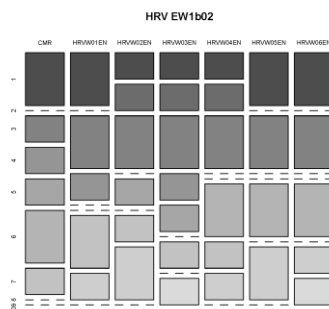
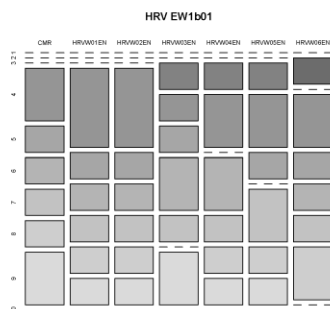
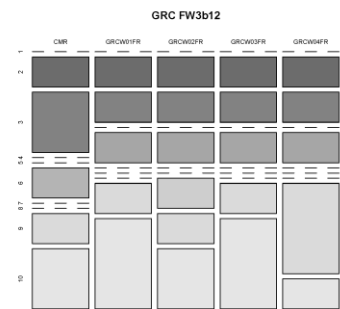
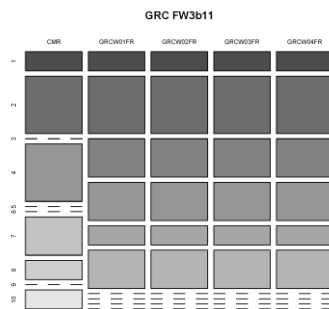
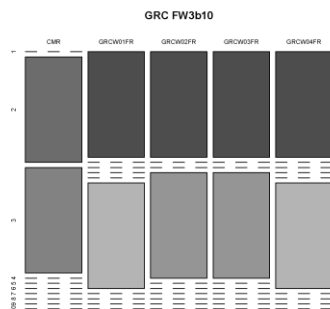
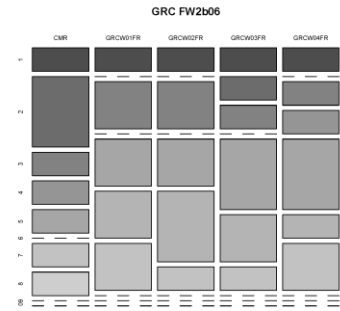
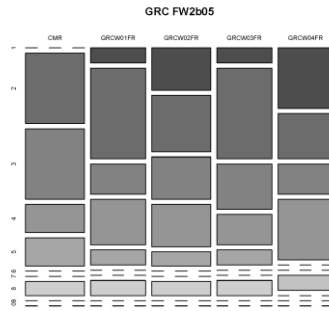
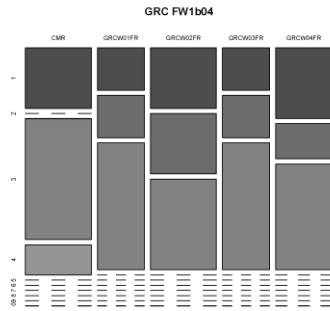


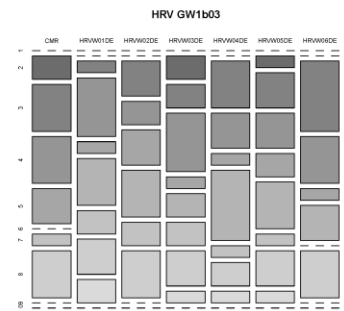
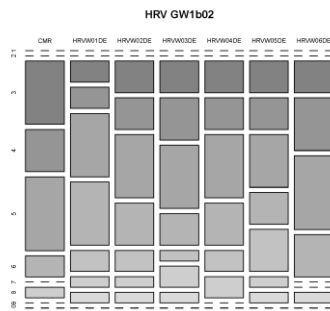
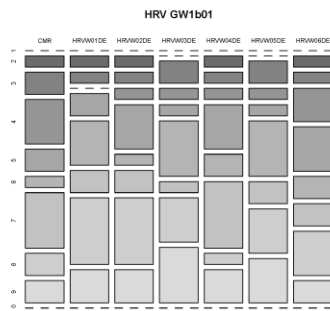
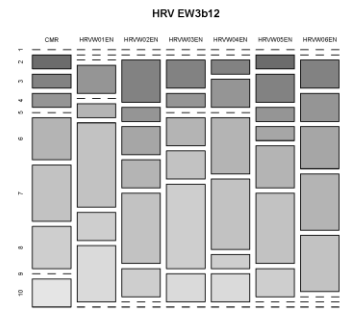
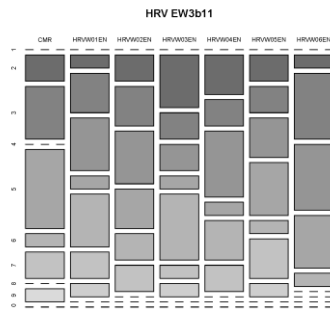
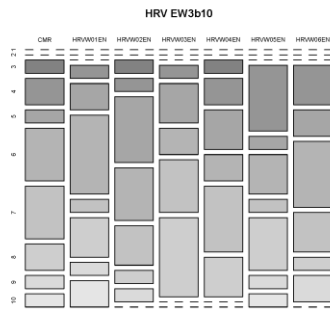
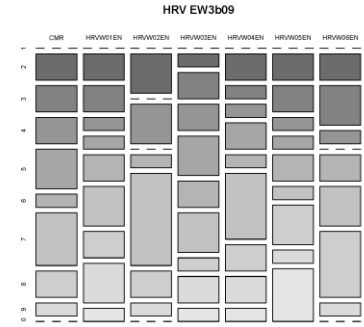
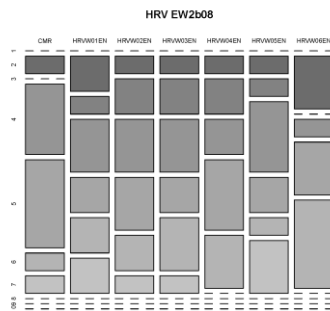
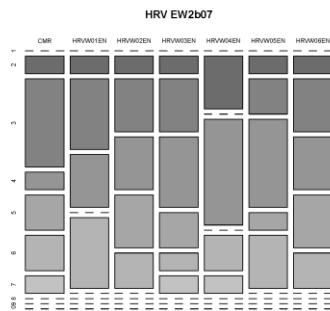
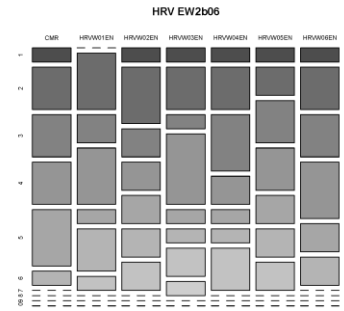
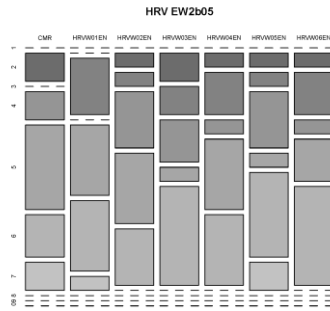
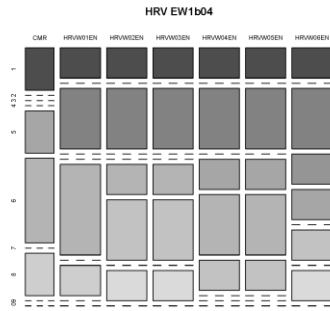


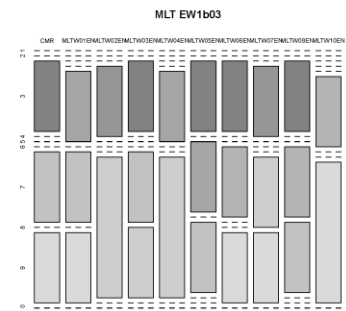
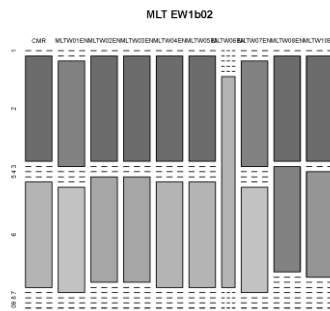
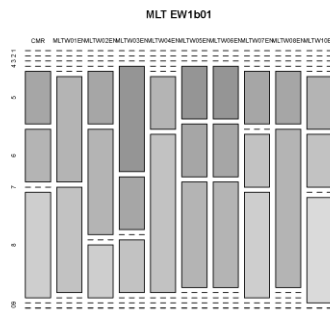
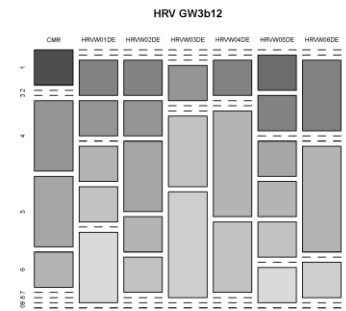
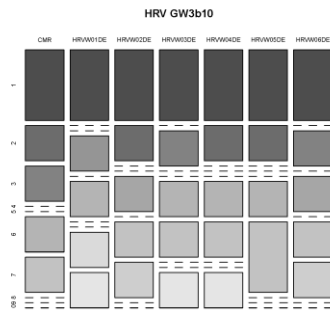
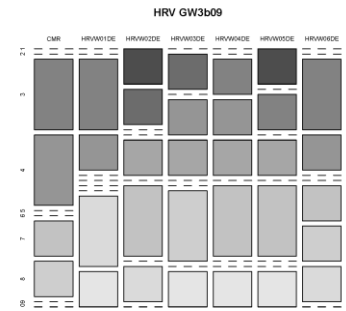
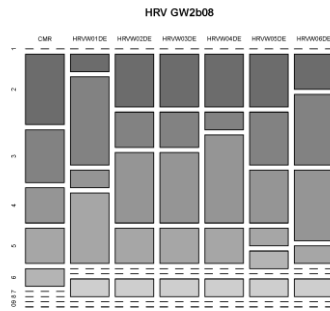
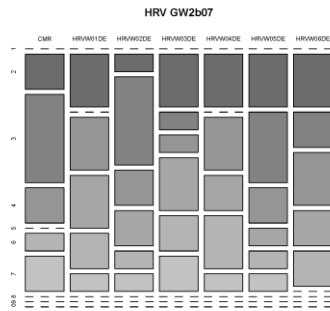
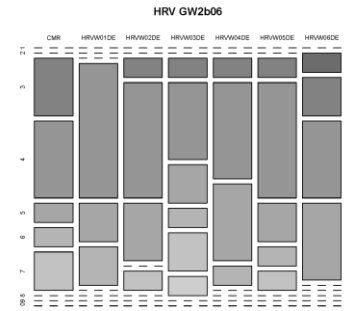
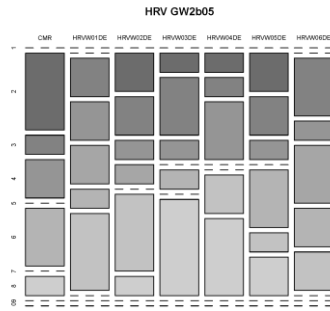
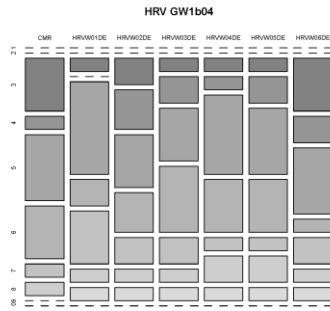


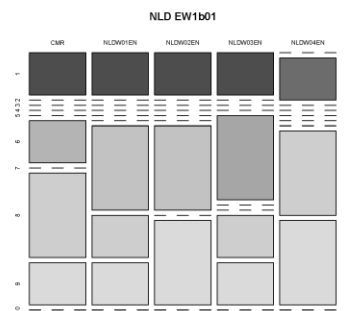
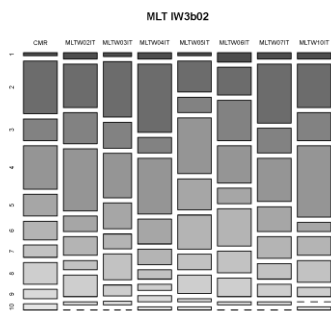
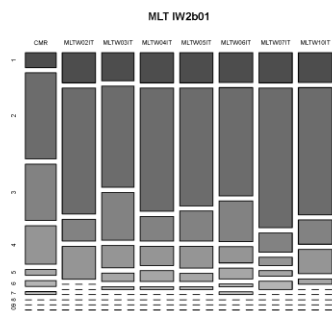
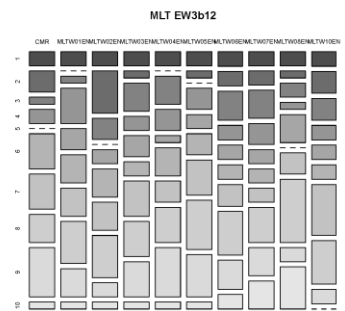
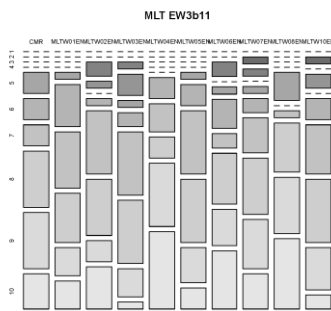
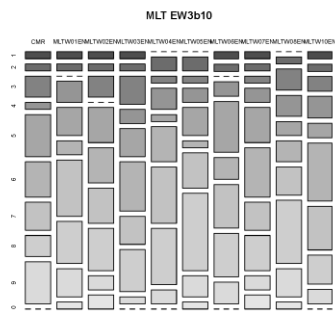
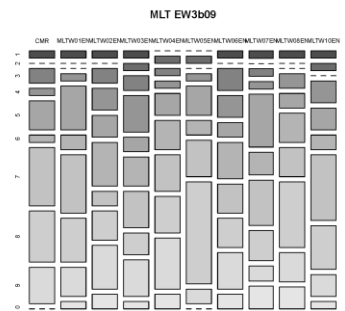
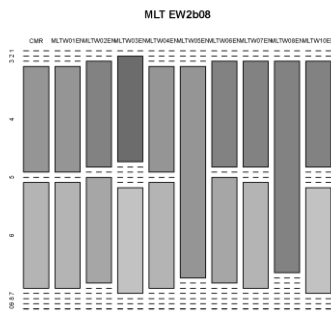
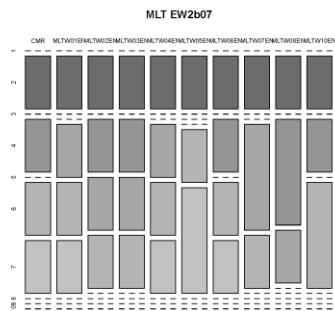
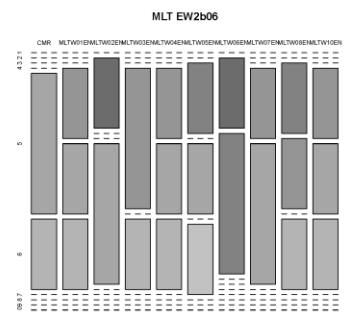
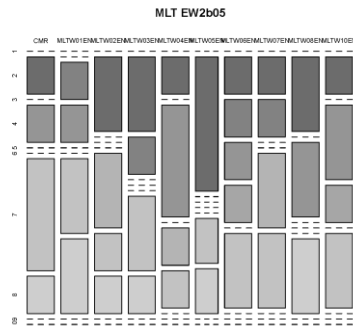
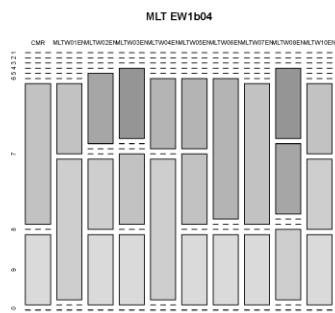


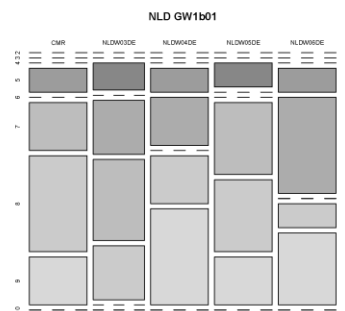
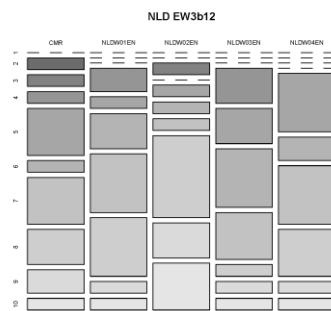
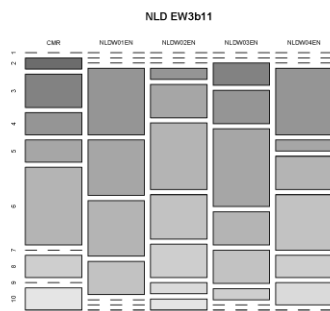
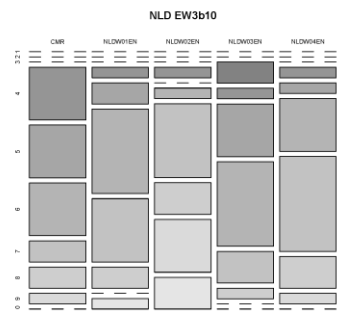
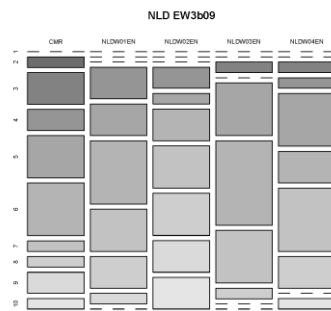
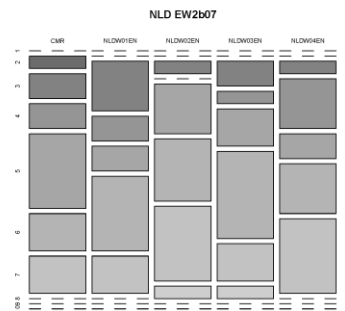
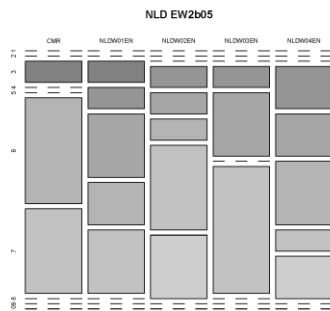
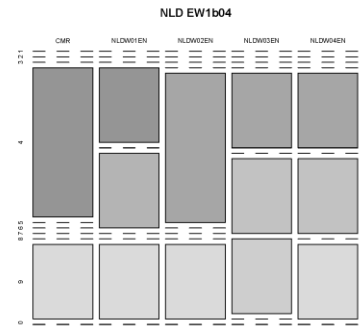
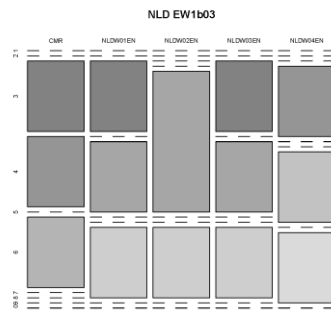
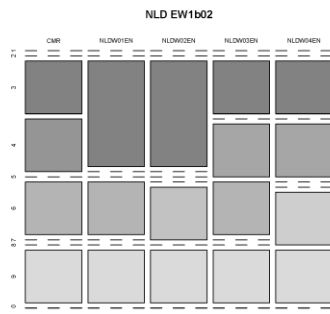




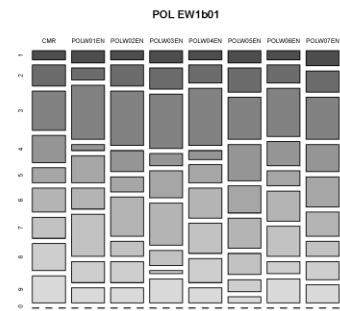
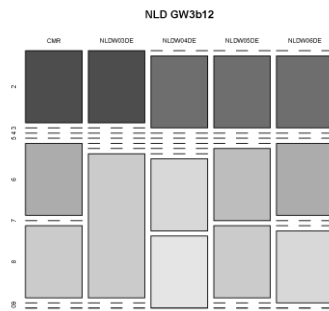
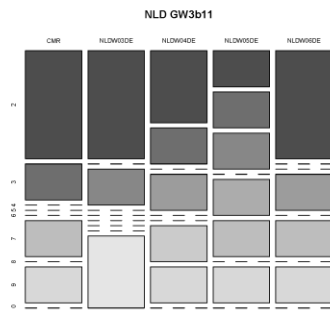
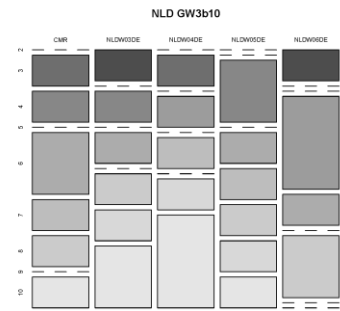
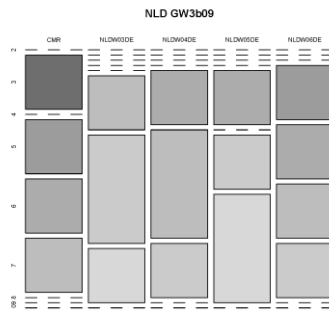
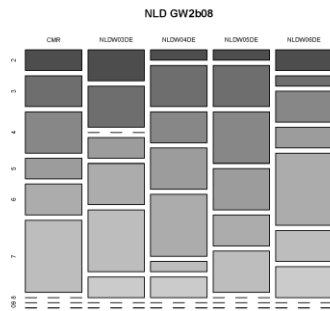
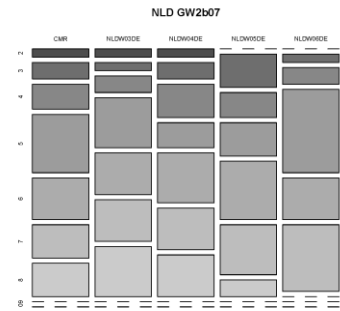
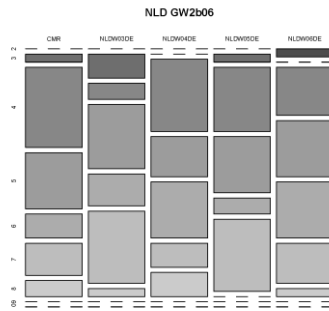
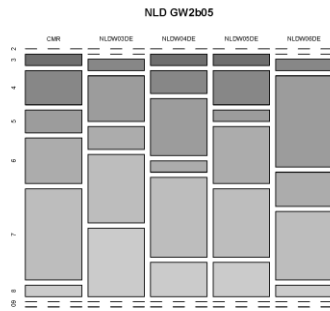
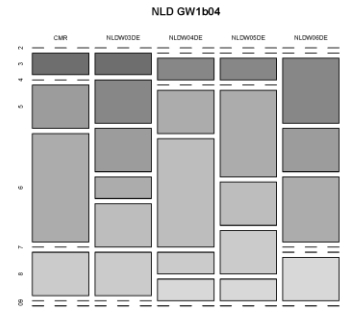
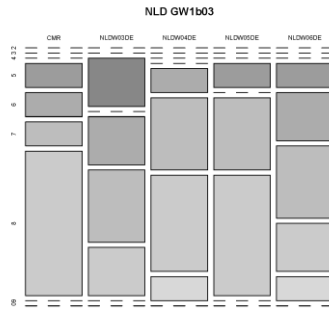
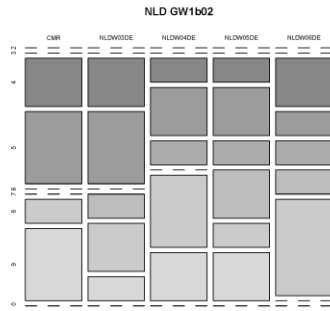


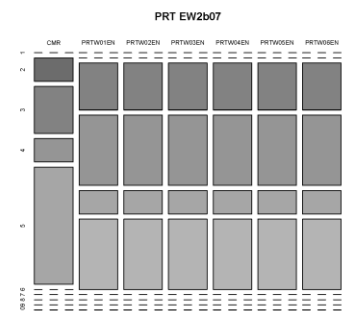
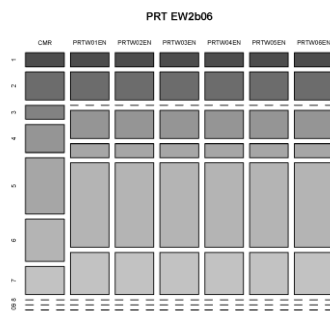
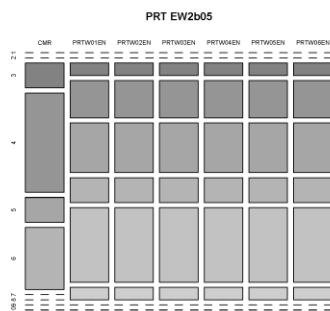
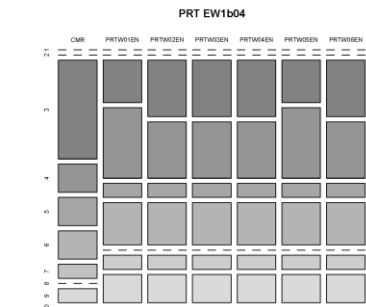
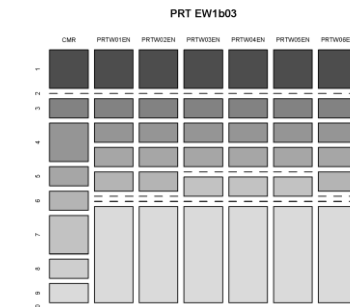
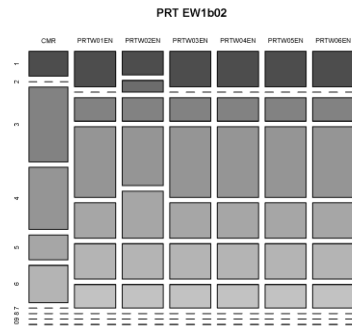
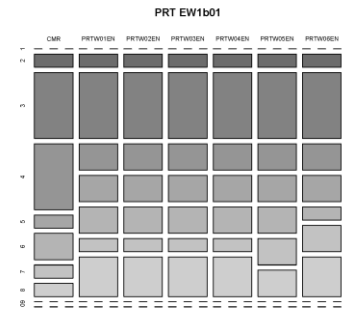
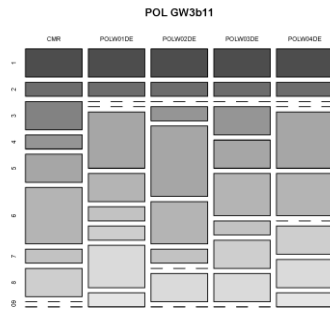
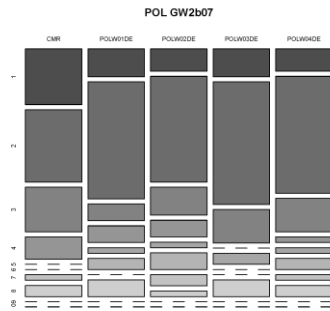
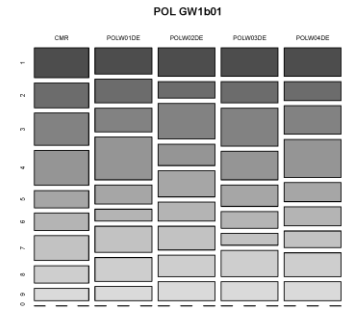
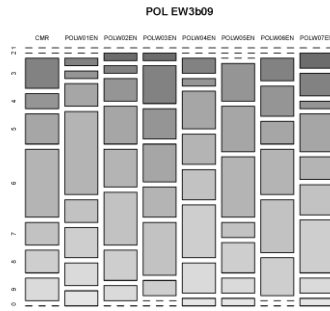
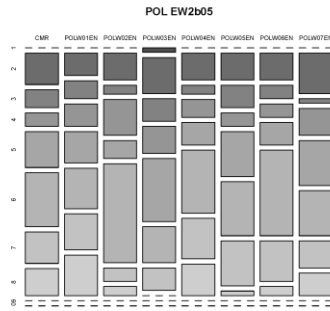


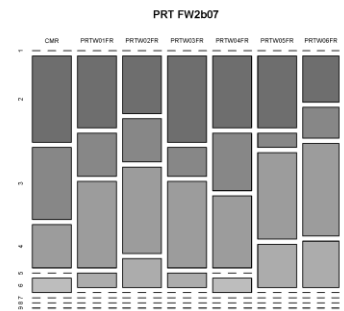
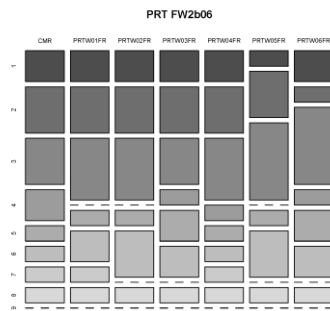
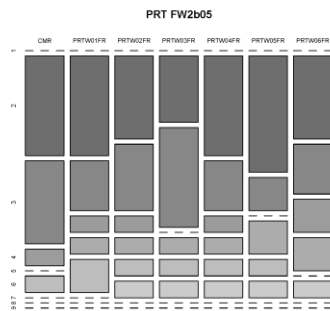
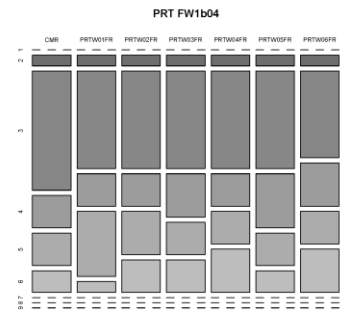
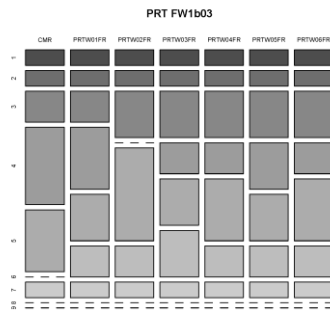
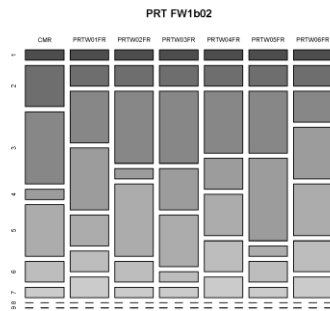
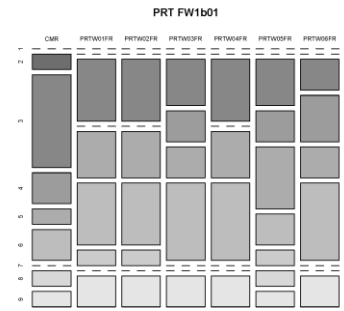
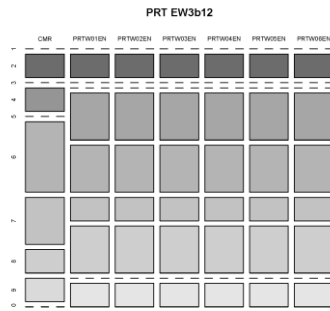
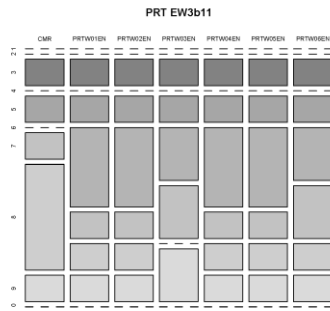
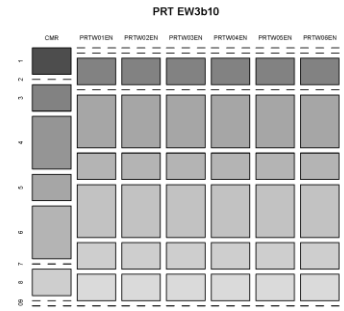
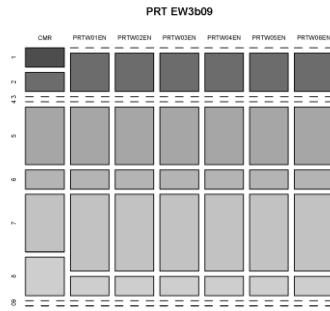
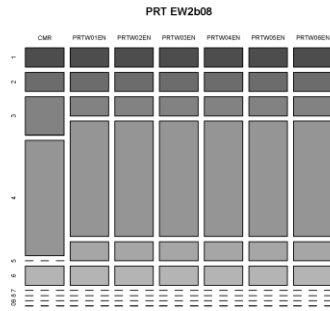


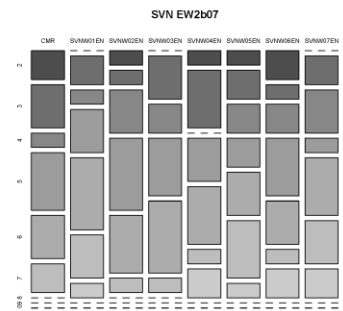
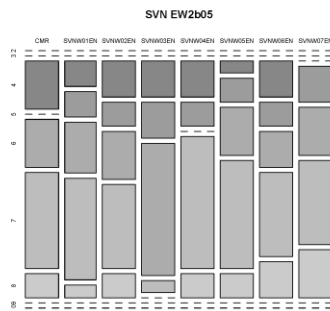
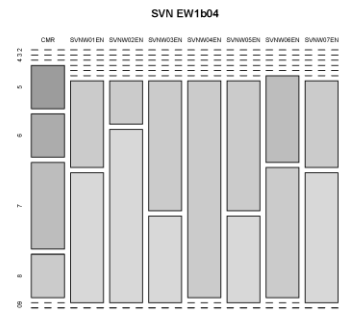
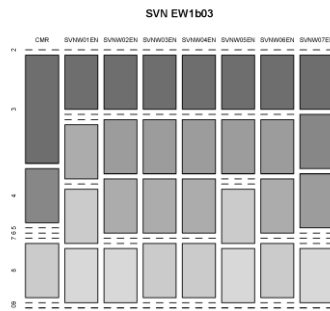
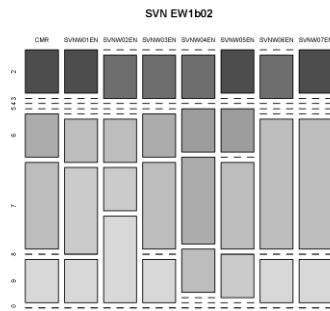
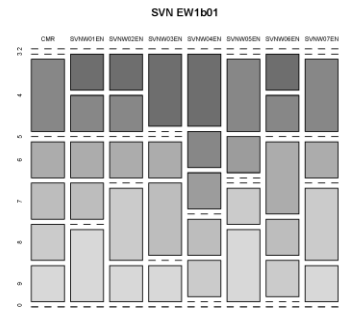
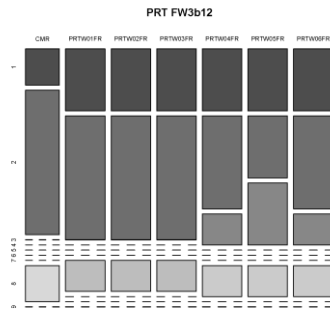
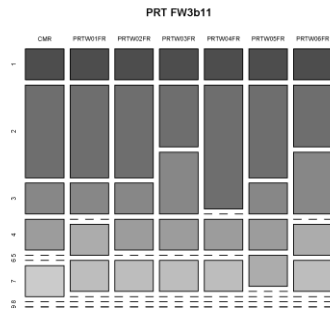
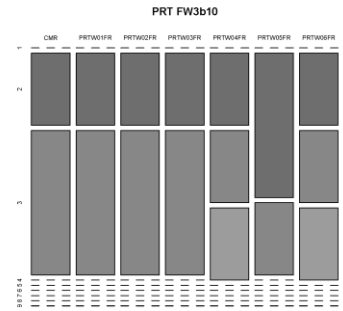
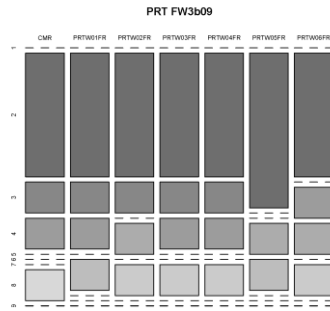
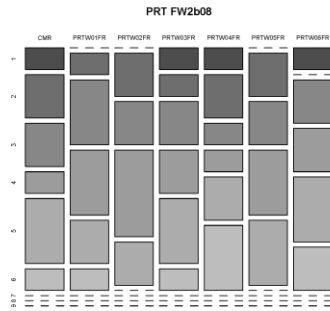


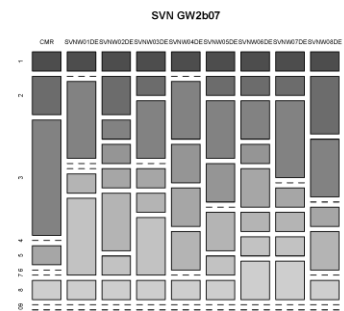
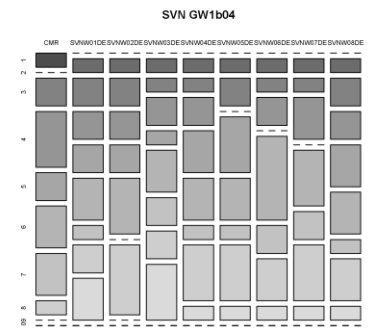
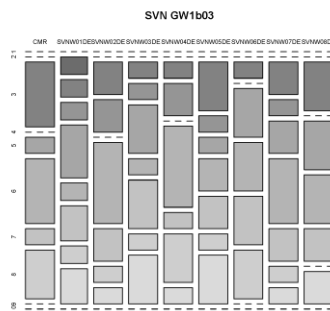
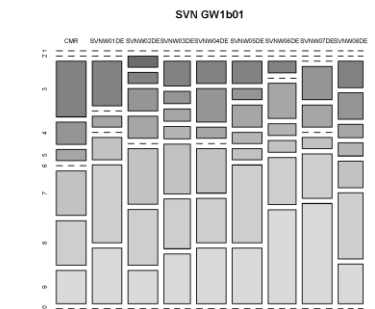
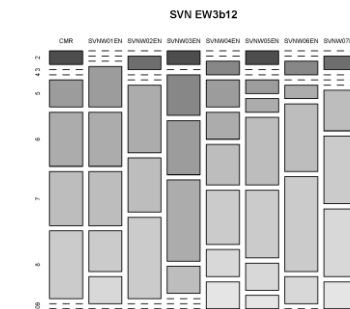
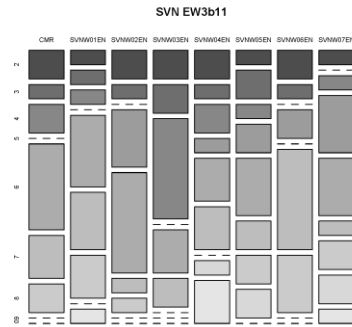
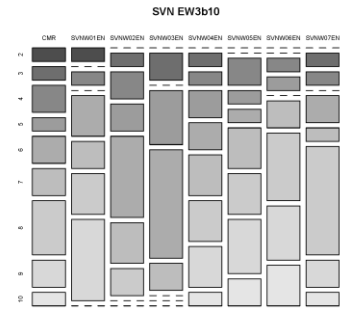
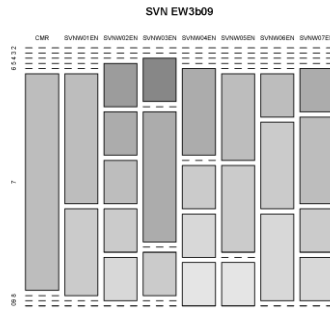
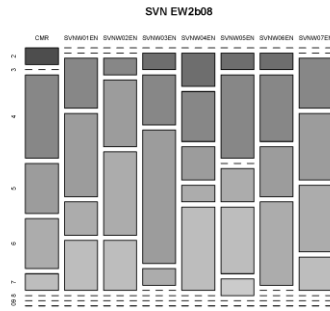


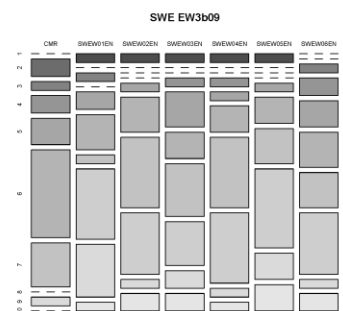
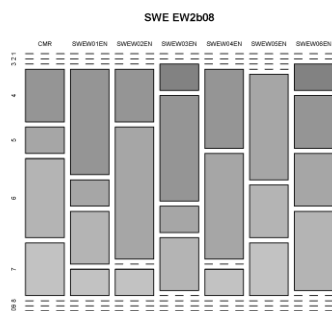
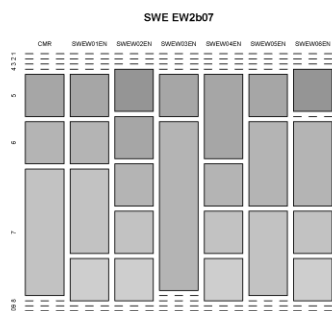
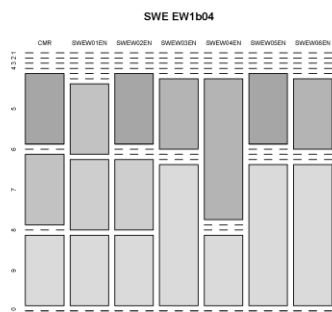
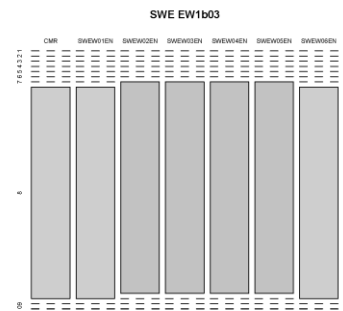
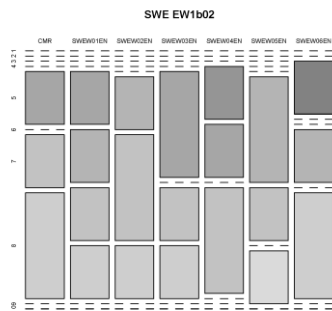
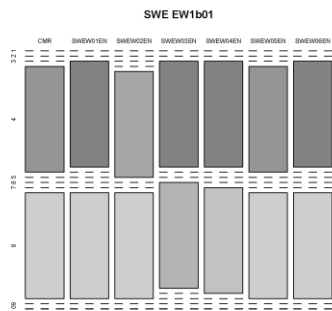
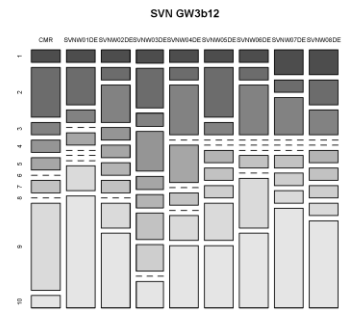
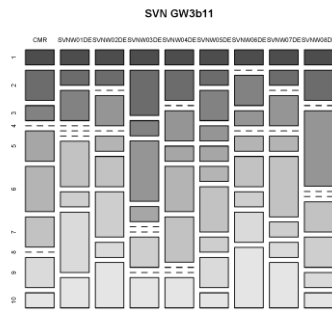
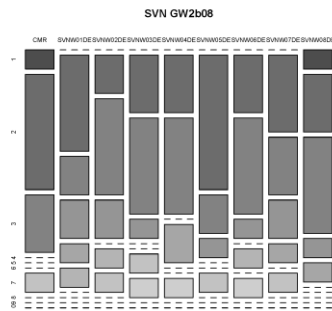


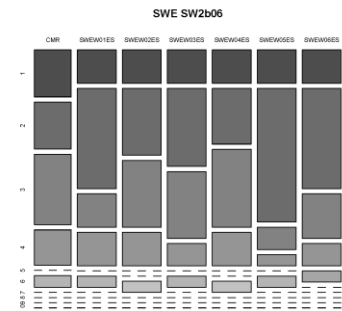
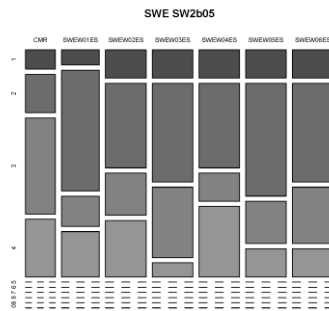
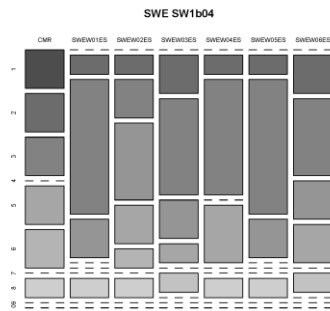
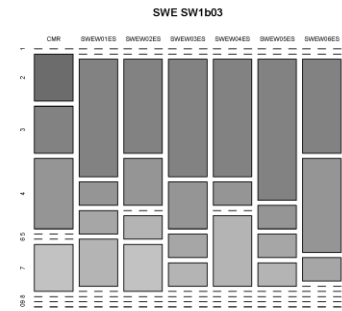
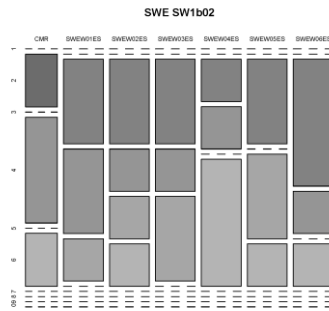
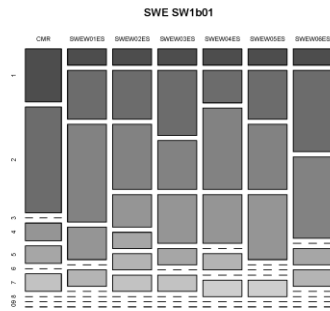
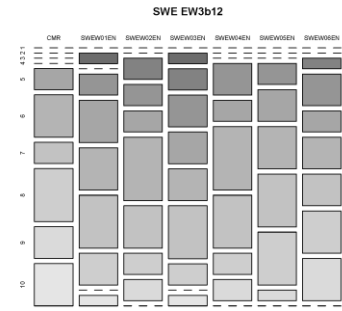
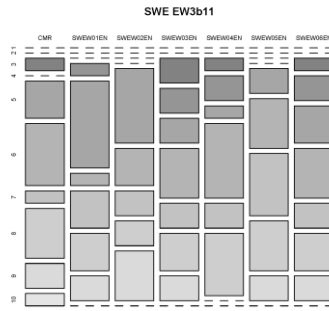
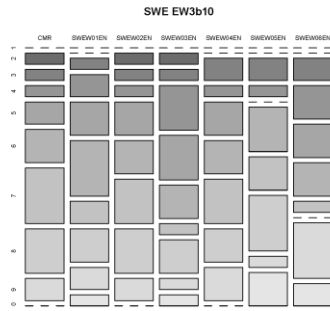


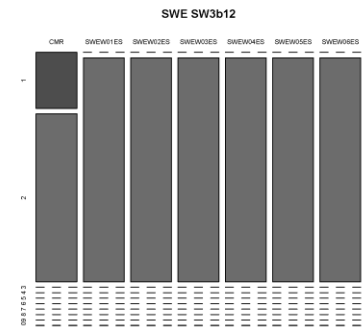
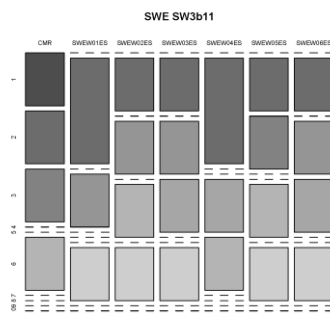
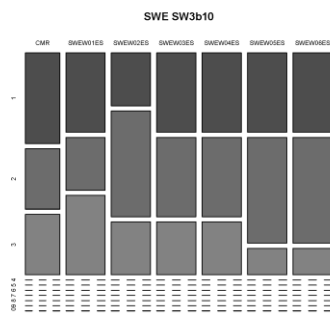
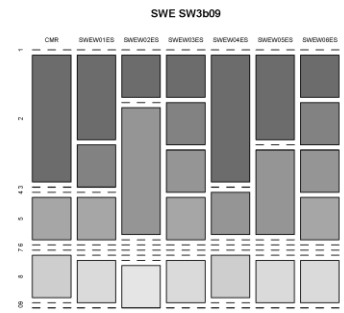
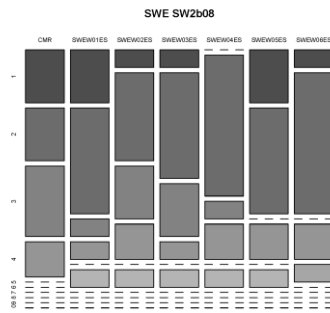
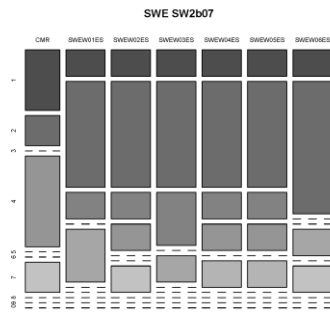
















## 14.7 Managing and implementing the ESLC

The design and implementation of the European Survey on Language Competences (ESLC) was the responsibility of an international consortium, SurveyLang, led by Cambridge ESOL. The partners and key people in the consortium are outlined below.

### 14.7.1 SurveyLang partners

SurveyLang Institutions	Work areas	Website
Centre international d'études pédagogiques (CIEP)	Language testing (French)	<a href="http://www.ciep.fr">www.ciep.fr</a>
Gallup Europe	Sampling including base weights, testing tool development, translation	<a href="http://www.gallup-europe.be">www.gallup-europe.be</a>
Goethe Institut	Language testing (German)	<a href="http://www.goethe.de">www.goethe.de</a>
Instituto Cervantes	language testing (Spanish)	<a href="http://www.cervantes.es">www.cervantes.es</a>
National Institute for Educational Measurement (Cito)	Psychometric analysis, questionnaires, sampling weights, data sets	<a href="http://www.cito.nl">www.cito.nl</a>
University of Cambridge ESOL Examinations	Language testing (English), project management, field operations	<a href="http://www.cambridgeesol.org">www.cambridgeesol.org</a>
Universidad de Salamanca	Language testing (Spanish)	<a href="http://www.usal.es">www.usal.es</a>
Università per Stranieri di Perugia; Centre for Assessment and Language Certification (CVCL)	Language testing (Italian)	<a href="http://www.cvcl.it">www.cvcl.it</a>

### 14.7.2 SurveyLang team members

SurveyLang has had a large team and is grateful to the many people involved in the ESLC over the course of the project, more than can be mentioned here. The table below presents the key SurveyLang contributors to their project together with their role.

SurveyLang member	Institutions	Position
Mike Milanovic	ESOL	Chair of Executive (February 2008 - December 2010), Vice Chair of Executive (December 2010 - present)
Jan Wiegers	Cito	Vice Chair of Executive (February 2008 - December 2010), Chair of Executive (December 2010 - present)
Neil Jones	ESOL	Project Director (December 2010 - present), Vice Project Director (February 2008 – December 2010)
Norman Verhelst	Cito	Project Director (February 2008 - December 2010), Testing design Lead, Standard Setting Listening and Reading
Erna Gille	Cito	Vice Project Director (December 2010 - present)
Johanna Kordes	Cito	Assistant Vice Director (December 2010-present), Reporting Lead on the results of the questionnaires, Questionnaires Verification Coordinator
Karen Ashton	ESOL	Project Manager, Field Operations Lead
Rebecca Stevens	ESOL	Project Assistant, Field Operations Assistant
Robert Manchin	Gallup	Program Board Representative
Michaela Perlmann-Balme	Goethe	Program Board representative,

SurveyLang member	Institutions	Position
		German Language Testing Lead
Inma Borrego	USAL	Program Board representative, Spanish Language Testing Lead
Guiliana Bolli	CVLC	Program Board representative,
Sylvie Lepage	CIEP	Program Board representative, French Language Testing Lead
Victoria Rubini	Cervantes	Program Board representative
Sanneke Schouwstra	Cito	Questionnaires development Lead including framework and indices, Field Trial Questionnaire analyses Lead
Gunter Maris	Cito	Data Analysis Lead
Ivailo Partchev	Cito	Sampling weights Lead, Data Management Lead
Remco Feskens	Cito	Questionnaires analysis support
Jesse Koops	Cito	Data Management Co-Lead
Joost Schotten	Cito	Data Entry systems specialist
Roselyne Marty	CIEP	French Language Testing Lead
Francesca Pelliccia	CVLC	Italian Language Testing specialist (March 2009 – October 2009)
Danilo Rini	CVLC	Italian Language Testing specialist (from October 2009 to present)
Francesca Parizzi	CVLC	Italian Language Testing specialist (February 2008 to March 2009)
Barbara Spinelli	CVLC	Italian Language Testing Lead
Martin Robinson	ESOL	Language Testing Team Lead
Glyn Hughes	ESOL	English Language Testing Team Lead (July 2009 – April 2010)
Heidi Endres	ESOL	English Language Testing Team Lead (from April 2010 to present)
Andrew Balch	ESOL	English Language Testing Team Lead (February 2008 – July 2009)

SurveyLang member	Institutions	Position
Agnes Illyes	Gallup	Translation Lead
Peter Husztik	Gallup	Software systems specialist (WebTrans, sampling portal, Testing Tool) Translation Lead
Jostein Ryssevik	Gallup	Software systems Lead
Anna Chan	Gallup	Sampling specialist
Gergely Hideg	Gallup	Sampling Lead
Manas Chattopadhyay	Gallup	Sampling Lead
Claudia Schulze	Goethe	German Language Testing specialist
Julia Guess	Goethe	German Language Testing specialist
Marian Villoria	USAL	Spanish Language Testing specialist

### 14.7.3 Special thanks

SurveyLang would also like to thank the following people, institutions and boards for their contribution to the ESLC:

The European Commission, their experts and the Advisory Board of the European Indicator of Language Competence have provided invaluable, constructive feedback throughout the project and have contributed to key decisions on technical issues.

NRCs had a fundamental role in the ESLC. This project would simply not have been possible without their dedication and efforts.

The SurveyLang Program Board, consisting of one member of each project partner, have met regularly throughout the project to review the project's progress and have taken key decisions on work processes and technical issues. Michael Milanovic, the CEO of Cambridge ESOL, and Jan Wieggers, a Director at Cito, have shared the role of Chair of this group.

SurveyLang would like to offer special thanks to Norman Verhelst, Project Director from February 2008 - December 2010, for his contribution and input to the ESLC. His work was fundamental in shaping the project and in developing and contributing to many of the technical elements outlined in this report. SurveyLang would like to pay special thanks to Norman for his work in constructing the test designs used in the ESLC (see section 2.5).