Matura GSE Alignment Summary Report

September 2016

David Booth
Katarzyna Wojciechowska
Mike Mayor

Global Scale of English

Fast-track your progress

—

Contents

Executive Summary

The Polish Matura is a high-school exit exam taken by young adults (usually aged around 18) at the end of their secondary education in order to matriculate. Pearson was interested to investigate how attainment on the Matura related to the Global Scale of English (GSE), a granular proficiency scale built on the Common European Framework of Reference for Languages (CEFR). Given that Matura scores are equated with CEFR levels, the study would also provide the opportunity to validate these CEFR levels. The study found a strong relationship between the Polish Matura objectives and the GSE Learning Objectives. Furthermore, whilst Matura is aimed at a B1/B1+ level, students can still study effectively for the test at different levels and be successful.

GLOBAL SCALE OF ENGLISH

For more information on the Global Scale of English, visit English.com/gse

Background

Two assessment instruments were used in this study: the Polish Matura and *Progress*.

Polish Matura is a high-stakes matriculation examination taken by students at the end of their secondary school studies and is a requirement for those wishing to move on to tertiary education. The exam has two levels: a basic level and an advanced (or extended) level. In 2015, 159,996 students took basic Matura and 88,727 students took the advanced level. The exam is a paper-based test followed by a face-to-face oral exam. The basic level paper exam takes 120 minutes, the advanced level takes 150 minutes. The oral exam for both levels takes 15 minutes.

Progress is a low-stakes proficiency examination developed by Pearson that reports on the Global Scale of English. It is a computer-based test, administered by teachers in the classroom. It tests all four skills and scores all skills, including speaking and writing, automatically. The test is sold as a package of three tests for one of six levels (from below A1 to C2). It is divided into six levels so that an accurate assessment of proficiency can be obtained in under one hour. For this study, *Progress* Level 4 was selected (reporting proficiency levels from GSE 45–60, B1 to low B2, but able to measure from GSE 40–65, A2+ to mid B2).

PROGRESS

For more information on Progress, visit pearsonelt.com/progress

The study was divided into three sections: an analysis of the Polish National Curriculum; a comparison of test-taker scores on Matura and *Progress*; a standard setting exercise.

In the first phase to compare the Polish Curriculum to the GSE Learning Objectives for Adults, a good relationship was found. This was unsurprising as both the GSE Learning Objectives and the Matura are based on the CEFR.

The second phase was to look at mapping student outcomes on the Matura test to the GSE. To investigate this, a two stage methodology was implemented. Firstly a group of students' scores on Matura were correlated with those on *Progress*, a Pearson assessment product which reports scores on the GSE. The second stage involved relating the standard of the Matura students and the Matura examination paper to the GSE. To do this, an experienced panel of teachers estimated the difficulty of Matura test items from the 2015 exam and made judgements about the level of ability a student needed to have to be successful on the examination.

Key outcomes of the study

- This was the first study of its kind relating Matura to an externally calibrated scale, confirming its level

- Through the mapping of the curriculum and the test papers, a positive relationship between the Polish curriculum objectives and the GSE Learning Objectives was found, confirming the Matura level

- The panel was also able to arrive at clear conclusions regarding the level of Matura based on the Borderline Person Profiles. This showed that whilst Matura is aimed at a B1/B1+ level, students can still study effectively for the test at different levels and be successful. To pass Matura at the 30% level, the panel estimated the level of the student to be at the borderline of A1 and A2. This indicates a GSE Score of 29. The table below summarises the Borderline Person scores.

Borderline Person Profile at Matura score	CEFR Level	GSE minimum score
30	A1/A2	29
50	B1	43
80	B1+/B2	59

The correlation studies and the mapping of Matura to the GSE are detailed further in the following pages.

TABLE 1

GSE minimum score for the Borderline Person Profile on Matura

Study **1** Correlation

The correlation study looked at two groups of students. Just over three hundred students were recruited and took *Progress* within 3 weeks of sitting the Matura test and provided Pearson with their results on Matura. The students came from towns and cities all over Poland with the biggest groups of 30 from Warsaw and Gdansk. The population was therefore a fair reflection of students across Polish cities. The two groups were the Matura Basic group and the Matura Advanced (Extended) group. Matura Basic is the test taken by all graduating students in Poland. Matura Advanced is an extended level which the student elects to take. They select from a range of subjects. Since Matura Advanced is elective we would expect a higher level of ability for students taking this examination.

Matura Basic

The Matura pass boundary is set at 30% of the available marks. This is a very low bar. In the sample of students used for this study almost all students passed.
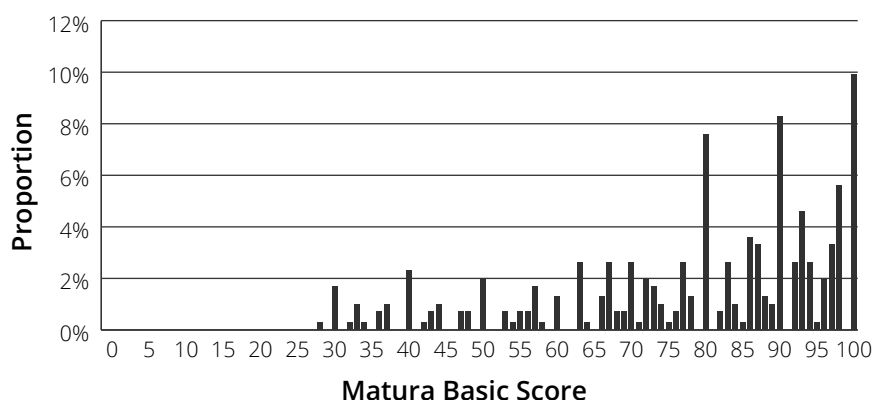

FIGURE 1

Distribution of scores obtained in Matura Basic (Mean 78.3 SD 19.1 n=300)

The students in this study almost all reached the minimum score required to pass Matura. From the data therefore we cannot identify a minimum GSE score required. Instead we correlated the Matura scores with the *Progress* scores. Using a complete data set of students we were able to predict a positive relationship between GSE scores and Matura scores. This gives an indication based on GSE scores of the minimum Matura Basic score and had a positive correlation of .4. This is not a strong correlation but this is to be expected given the different types of test used here. For Matura Advanced a similar picture was found.

Matura Advanced (Extended)

As mentioned, this test is set for students who have elected to sit the extended paper in this area. The paper is more difficult and there is no pass boundary.
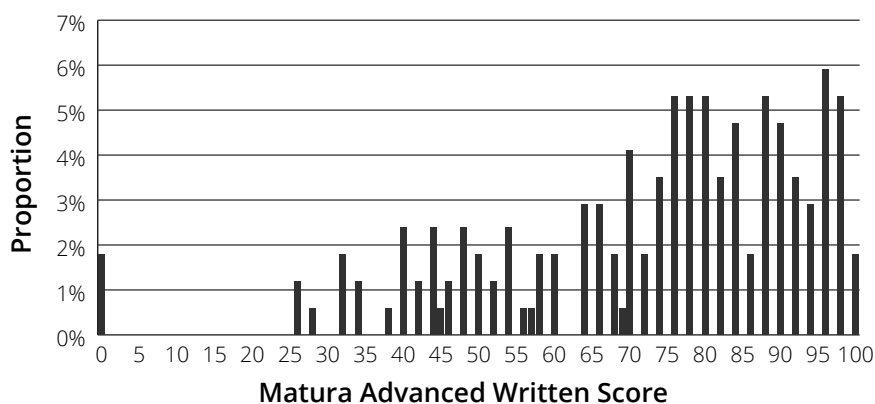

FIGURE 2

Distribution of scores obtained in Matura Advanced Written (Mean 72 SD 21.3 n=170)

The same caveats apply to Matura Advanced as to Basic. Again we can correlate the scores using the sample of 170 students who completed the Matura Advanced written paper. Again a positive relationship can be seen between *Progress* scores and the Matura Advanced scores with a correlation of .45.
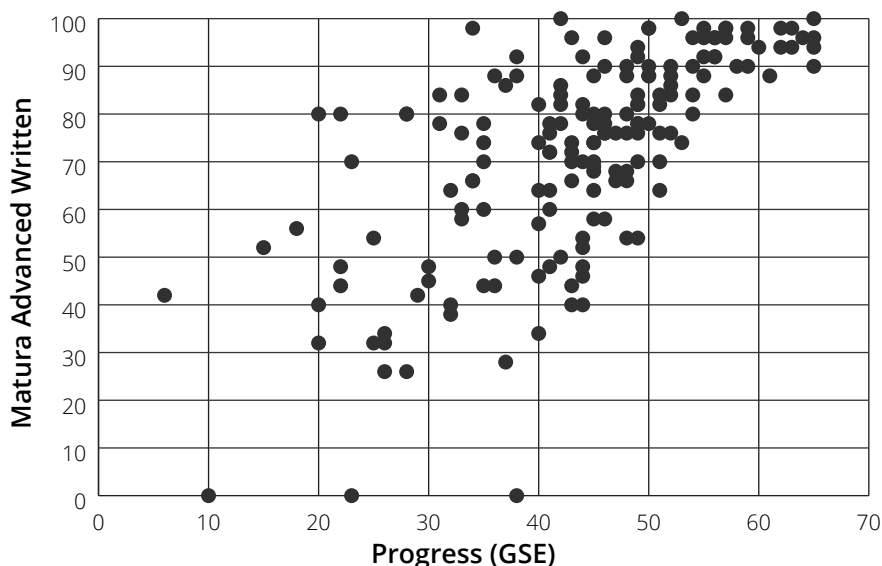
FIGURE 3

Scatter graph of *Progress* scores plotted against Matura Advanced Written scores.

Conclusions

There is a positive relationship between GSE scores as surfaced by *Progress* and Matura scores as shown in the scatter plot above. Success on Matura is almost guaranteed for matriculating students in Poland with the pass mark set at 30% of the available marks. *Progress* predicts that higher levels of ability will lead in general to a higher score in the Matura examination as reflected in the correlations above. This is even more true for Matura Extended students.

Study **2** Mapping Matura to the GSE

The second part of the study involved a mapping exercise to confirm the relationship between the Matura exam and the GSE. The mapping involved two parts, firstly understanding the level targeted in the English Matura Basic test, specifically the level of difficulty of the test items in relation to the GSE. And secondly, an understanding of the language ability level of students, specifically the probability of the success of candidates with different levels of ability on the test. As a result of these two measures an attempt was made to relate test difficulty to the ability of the students.

The study is set out in detail below:

Method and Process

For both parts of the study, a panel of experts (referred to as judges) was established. These experts were teachers based in Poland who understood in detail the requirements of the Matura exam. For the first part of the study a ratings exercise took place with the panel estimating test item difficulty. For the second part of the study and given the affinity between the aims and the principled approach to Standard Setting outlined by the Council of Europe (2009, chapter 6), it was decided to implement an adapted version of Tucker-Angoff (T–F) methodology.

The study was limited in scope by the following:

- Test performance data was not available for the Matura test as this is retained by the local testing authorities. This limits what can be said about student and item performance

- Matura has not been formally mapped to the CEFR so there is no a priori statements about expected levels of attainment

- The majority of students for Matura are expected to achieve the Basic test with the Advanced or Extended version aimed at students with an interest in the subject. This means the data around the Basic examination is limited in terms of discrimination between different student abilities

Twelve judges were selected based on their teaching experience. All the judges had 10+ years of experience teaching English. All had experience of teaching at the Matura level. All but two judges had experience as a Matura examiner.

—

The Workshop

Introduction

At the beginning of the workshop judges were asked for their informal impressions of the new Matura exam (revised in 2015). This was in order to:

- Establish group dynamics and/or commonality
- Highlight any issues which may need to be discussed or clarified during the workshop
- Discuss the panels opinions relating to the new exam

The whole group expressed positive reactions to the revised Matura test. The judges felt that the test was more closely related to real life needs and, therefore, more motivating for the students. They also felt that the test items were better aligned with what happens in the classroom. These factors demonstrated a strong argument for the increased validity of the test. The group indicated that, in achieving this, the test was potentially harder, although overall they were unsure how level had been affected, either intrinsically or in relation to the CEFR.

Following this discussion, the objectives and the context of the workshop were clarified for the judges i.e. that other research was taking place centrally and that this workshop was to gain expert qualitative data to support and illuminate the statistical studies.

The process to be followed throughout the workshop was then explained to the judges together with a brief rationale for each stage.

The approach to making judgements was outlined, i.e. that the judges were not obliged to agree. They were encouraged to use their expertise and to trust, and argue for, their own opinions. Judges were reminded the process would take a learner-centred approach and so it was not about testing the judges but about the relationship between the test and the learner.

CEFR Familiarisation

This section focused on A2 – B2 aspects of the CEFR as these levels are the target of the Matura exam. It was conducted without allowing judges to refer to the pre-assigned material.

The purpose of this section was to re-familiarise judges with the detail of the relevant CEFR levels and use discussion to standardise judges on the CEFR levels.

GSE Familiarisation

This section focused on introducing judges to GSE, its development history, role and purpose.

The purpose of the presentation was to introduce judges to the GSE and to clarify its link to the CEFR, as well as to promote the usefulness of a granular approach to the very broad bands of the CEFR in measuring achievement and progression.

The purpose of the interactive activities was to build confidence and familiarisation in participants in upcoming tasks and to indicate how standardised the group were on the levels of GSE/CEFR. Any outliers were flagged in order that the judges could make their own adjustments to level.

Mapping Matura to GSE: Process

This section required judges to use the agreed GSE Learning Objectives to confirm the construct of the Matura exam.

The purpose of this section was to agree the test focus of each task as well as the level targeted in each task. Judges were reminded to focus on the learning objective of a task rather than the method by which it is tested.

There were two rounds of discussion at each stage and agreement was reached as a group.

Tucker-Angoff Process Familiarisation

The group was introduced to the methodology for deciding candidates who sit at borderline levels in a test. In the case of the Matura, the borderlines were pre-set at 30%, 50% and 80%. 30% is required to pass the Matura.

The agreed process was that judges would decide on a borderline person's profile based on the broad CEFR bands. They would then decide which GSE Learning Objectives within this selected band were relevant for each skill. Judges were asked as individuals to follow this process and there were 2 rounds of discussions.

General Conclusions

It was felt the reading section of the Matura Basic test had more variety than the previous test and thus targeted more of the reading construct. The listening section tended to test the main idea only i.e. it has a restricted construct. The use of English section was not covered as a construct because it did not relate directly to a performance skill. The writing section favoured description and personal experience but in fact demanded far more.

—

The Matura Learner: Borderline Person Profiles

In making decisions about borderline profiles it should be noted that most of the judges commented that their experience was with learners who generally did the more advanced level of the Matura so they were not especially familiar with the judgement levels required for the lower levels.

However the group were able to agree the following:

- 30% borderline person overall CEFR level A1/A2
- 50% borderline person overall CEFR level B1
- 80% borderline person overall CEFR level B1+/B2

Psychometric Analysis

The psychometric analysis stage had two aims. Firstly to establish GSE values for the Matura Basic 2015 exam and secondly to establish a Matura to GSE concordance table. To establish the GSE values, the CEFR judgements provided from the workshop were used. The table below shows where the GSE values were generated based on the most frequent selected CEFR ratings for each item and their corresponding GSE values.

Item	GSE Value Final
1	48
2	45
3	46
4	47
5	51
6	56
7	57
8	46
9	55
10	49

TABLE 2

GSE Values for Matura Basic Test Items
(June 2015 paper)

—

In order to establish a Matura and GSE concordance table, the raters used the adapted Tucker Angoff method to determine probability estimates of candidate ability.
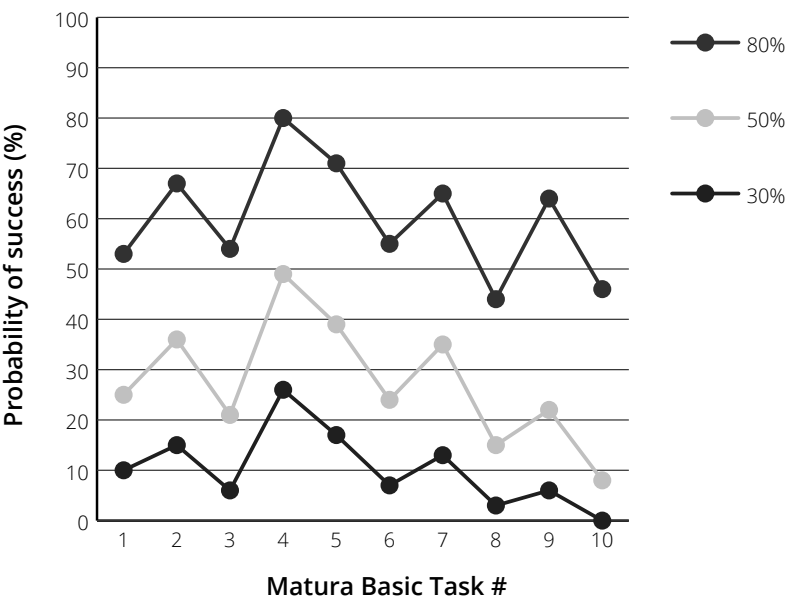The results from these steps were collated, including 30% pass, 50%, 80%. .

FIGURE 4

Estimates of GSE level required to be successful on each Matura task

The graph demonstrates the three collective estimates of candidate ability required to be successful on the Matura tasks at different Matura performance levels. A nice progression can be observed together with variations in the estimated item difficulties. Included in the table below are also the estimates of item difficulty from the CEFR mapping activity.

	item 1	item 2	item 3	item 4	item 5	item 6	item 7	item 8	item 9	item 10
Panel Estimate 30%	9.73	14.73	6.00	26.18	17.45	6.55	13.09	3.00	5.82	0.00
Panel Estimate 50%	25.00	35.80	20.80	49.30	38.80	24.20	34.50	15.10	22.30	7.60
Panel Estimate 80%	52.50	66.50	54.30	79.60	71.30	54.70	64.60	43.50	63.80	45.50
GSE value from table 2	48	45	46	47	51	56	57	46	55	49

The estimates of candidate ability were then correlated with the estimates of the item difficulty. However, the correlations between the rater-probability estimates were not in line with the GSE values of item difficulty obtained from the stage one analysis. This may reflect some unfamiliarity of the panel with the GSE Learning Objectives and the limited range of difficulty of the Can Do statements which were selected. In this instance it would seem better to rely on the panel's CEFR judgements regarding the level required by borderline students as opposed to their judgement of task difficulty, given their greater familiarity with the student population and the CEFR.

TABLE 3

Summary of panel estimates of candidate and item ability

—

Discussion and summary

The study was successful in confirming the CEFR and GSE level of the Matura examination and showing that students who are below the B1 level can still work towards it and be successful. It is important to give students encouragement and support in learning and *Progress* is designed to do this by giving a clear and accurate picture of a student's language ability in real life contexts.

The correlation study showed a clear relationship between the results on the two tests, which is encouraging given they are designed for different purposes; one to test achievement and one to test progress. The study demonstrated the value of expert judgement in relation to systems such as the Global Scale of English and the value it can bring to understanding examinations in a local context, from a measurement and quality perspective.

The panel judgements were helpful in defining the minimum requirements for students working towards the Matura and there are opportunities here to define clearer pathways in terms of study plans and clearer measurement of progress.

References and further reading

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

Council of Europe (2009) *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: CUP.

Be yourself
in English.

Pearson