Technical Report

Implementation of Speakout with MyEnglishLab and perceptions of impact on student outcomes: Examining three Universities in Poland

Authors
Elpida Ahtaridou
Katya Saville
Iasonas Lamprianou

Pearson Global Product Organization
Efficacy & Research
Impact & Evaluation

Pearson

Table of Contents

Acknowledgements **4**

About the authors **5**
- **Elpida Ahtaridou** **5**
- **Professor Iasonas Lamprianou** **5**
- **Katya Saville** **5**

Introduction **6**
- **Background** **6**

Description of Speakout with MEL **7**

The present study **9**
- **The structure of the report** **10**

Sample **11**

Methods **14**
- **Student Questionnaire** **15**
  - Development of the questionnaire 15
  - Questionnaire structure and content 17
  - Administration, collection and analysis of the student questionnaire 17
- **Coordinator's pre-questionnaire** **22**
- **Student focus groups and staff interviews** **23**
  - Student focus groups conducted 23
  - Teacher interviews conducted 25
- **Analysis of interview and focus-group data** **26**
- **Classroom observations and post-observation debriefs** **28**
- **MyEnglishLab (MEL) student data** **29**
  - Analysis of assignments/practices and tests 31
  - Assignments/practices 32
  - Student progress analysis on assignments/practices 33
  - Analysis of student test scores 34
  - Reliability and validity of MEL practices / assignments / tests 34
  - Method for deriving findings on perceptions of impact from all data 35

Discussion of findings **37**
- **Students' attitudes towards Speakout with MEL** **38**
- **Teachers' attitudes towards Speakout with MEL** **38**
- **Implementation of Speakout with MEL** **38**
- **Student performance on MEL** **44**
- **Perceived impact of Speakout with MEL on students** **45**
- **Perceived impact of Speakout with MEL on teachers** **50**

Recommendations **51**

Generalisability of findings, limitations and future research **52**

References **54**

Appendix A: Consolidated Framework of Implementation Research (CFIR) **56**

Appendix B: Results **59**
Students motivation to learn English **60**
Teacher attitudes towards Speakout with MEL **64**

Implementation of Speakout with MEL **66**
Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie Filia w Białej Podlaskiej (AWF) **66**
Reasons for selecting Speakout with MEL 66
Training in MEL 67
Teaching and learning 67
Assessment and MEL 67
Wyższa Szkoła Logistyki **69**
Reasons for selecting Speakout with MEL 69
Training in MEL 69
Teaching and learning 70
Assessment and MEL 71
Wyzsza Szkoła in Warsaw (WS) **72**
Reasons for selecting Speakout with MEL 73
Training in MEL 73
Teaching and learning 74
Assessment and MEL 74

Student use of MEL **76**
Student weekly use of MEL **76**
Hours / minutes spent on MEL weekly 77
Student frequency of checking the gradebook 79
Number of practice, assignment, and test completions 81
Number of attempts on assignments/practices 83

Student performance on MEL **86**
Overall student performance on MEL assignments/practices **86**
Student progress on practices and assignments **89**
Correlation between students' average practice/assignment scores and percentage of practice/assignments completed **90**
Correlation between the average unit practice or assignment **93**
Correlation between average practice score and average test score **94**
Correlation between test scores **95**

Student and teacher perceptions of Speakout with MEL **96**
Infrastructure **96**
Training **96**
Teaching and learning **96**

The Students' Book **99**

MyEnglishLab **101**

Teacher perceptions of the impact of Speakout with MEL on teachers and teaching 104

Appendix C: Rasch fit statistics for the five questions of the motivation scale 106

Appendix D: Percentage of students achieving within different score bands in practices per institution when 0% scores are included in the analysis. 107

Appendix E: Student progress on practices and assignments per institution and course, MEL data 109

Pearson

Acknowledgements

Pearson and the authors of this report, Elpida Ahtaridou, Professor Iasonas Lamprianou and Katya Saville, are grateful to a number of people who have made this research possible.

We would like to thank the Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie Filia w Białej Podlaskiej (AWF) in Biała Podlaska, Wyższa Szkoła Logistyki in Poznan, Lingua Nova in Warsaw and Wyzsza Szkoła in Warsaw (WS, pseudonym) for agreeing to take part in this study. In particular, we would like to thank Grażyna Szyke, Director of Studies at Lingua Nova, Mariusz Buszta, head of the Foreign Languages Department in AWF and Witosław Awedyk, Director of Academic Studies for Wyższa Szkoła Logistyki for their support with the data collection and their continuous feedback, their openness and their patience in answering all our queries.

We would also like to thank Pietro Alongi, Portfolio Director for the adult sector at Pearson, for his support throughout the process of generating this report. We are grateful for his invaluable feedback on the report's findings and advice on how the findings could be translated into realistic and actionable recommendations for institutions and different Pearson teams.

Furthermore, we owe a big thank you to Susan Trory, English Portfolio Manager for the adult sector and responsible for Speakout, and Alicja Tarkowaska, Market Planning and Research Manager for Central Europe. We are grateful for their hard work, great spirits and continuous support — from arranging the case study visits and other data collection processes to providing feedback during all the phases of the research; from conceptualisation to delivery.

We are also most grateful to SRI Education and Nicole Arshan, Katherine Nagle, Jared Boyce, Elise Levin-Guracar and Allison Milby for their valuable feedback on the report. Lastly, we would like to acknowledge Emily Lai, Pearson's Vice President for the Impact Evaluation, Efficacy and Research team; Vikki Weston, Director of Programme Management, also part of the Efficacy and Research team; Liz Hudson of Little Red Pen; editor, Hilary Duplantis; and Redhouse Lane Communications for their continuous support during the lifespan of the research — from providing valuable feedback to checking the data for accuracy and proofreading the report.

About the authors

Elpida Ahtaridou

Elpida Ahtaridou is Director for Pearson's English products globally as part of the Impact Evaluation, Efficacy and Research team within Pearson. Elpida leads an array of efficacy and effectiveness studies across different countries. Prior to becoming director for English, she worked at Sir Michael Barber's global strategy and research team within Pearson and she was also UK Head of Efficacy Studies and Research, leading a number of studies, including a randomised control trial (RCT) of 1500 pupils on the efficacy of Bug Club, Pearson's reading programme for primary schools. Prior to that, Elpida was a researcher to Professor David Hopkins at the London Centre for Leadership in Learning (LCLL), Institute of Education (IoE), University College London (UCL), a lecturer in Education Studies and a teacher trainer for post-compulsory education. Projects she has worked on included the *Global Evaluation of the Policy Impact of PISA* and the *Reflections of the Performance on the Education System in Mexico* for the Organization for Economic Co-operation and Development (OECD), and a three year study on the *Impact of School Leadership on Pupil Outcomes*. She has published papers and book chapters including *System Leadership: A Response to the Challenges Facing Urban Schools in England* and *Research Essentials: An Introduction to Designs and Practices*.

Professor Iasonas Lamprianou

Iasonas Lamprianou is an Assistant Professor of Quantitative Methods at the University of Cyprus. He has a special interest in Educational Assessment and Educational Measurement and considerable experience with large-scale assessment. Iasonas has been involved in various projects for Examination Boards and Testing Services across the world, especially in the UK. He has experience in using summative assessment results for formative assessment purposes in the context of the School Sampling Project in England. Iasonas has published articles in a number of journals, including the Journal of Educational Measurement and the Journal of Educational and Psychological Measurement.

Katya Saville

Katya Saville has completed her PhD in bilingual education in 2017 at the Institute of Education (IoE), University College London (UCL), where she also works on the MA in Education programme. Her research experience is in qualitative and mixed methods case studies. Katya has also been teaching in primary schools and training teachers for nearly 20 years in five different country education systems.

Introduction

In this section, we give a brief background to the initiative behind this study. Before we outline the study's research questions, we provide a description of Speakout with MEL and its components, the sample, and the methods used to collect and analyse data.

Background

This report is part of Pearson's commitment to efficacy, launched in 2013. In the drive to improve student outcomes, we committed to reporting publicly on the impact of our products on student outcomes. Part of this commitment was, by 2018, to publish research that has been audited by a third party; PricewaterhouseCoopers LLP (PwC).
In 2017, the Global Impact Evaluation team, part of the Global Efficacy and Research team at Pearson, designed and embarked on a series of studies on Speakout and its counterpart, Top Notch, both of which are used by tertiary institutions and private language schools (PLSs) worldwide. The studies aimed to examine the implementation, perceived impact and relationship between each product and its intended student outcomes, across different countries and multiple sites.

Pearson

Description of Speakout with MEL

Speakout is a communicative course for adults and young adults. It is currently in its 3rd edition (2015), with the 2nd edition (2012) still in use in some countries. Speakout is a course with six levels:

1. Speakout Fundamentals (Starter)
2. Speakout Level 1 (Elementary)
3. Speakout Level 2 (Pre-intermediate)
4. Speakout Level 3 (Intermediate)
5. Summit Level 1 (Upper Intermediate)
6. Summit Level 2 (Advanced)

The course components include:

- **Students' Book** — 90–120 hours of learning material available in split or full editions (the split editions come with a split workbook or a split MyEnglishLab access code)
- **Classroom Audio CDs** — audio materials to use in class
- **Workbook** — additional exercises to consolidate learning in print
- **MyEnglishLab (MEL)** — student and teacher versions — a platform that contains an array of exercises to consolidate learning; meaningful feedback on wrong answers; remedial grammar exercises; grammar and pronunciation coaching videos and auto-graded achievement tests
- **Teacher's Book** — including detailed, interleaved lesson plans, language culture notes and answer keys
- **ActiveTeach** — a disc for front-of-class use, which includes a digital version of the Students' Book; digital grammar exercises; videos, photocopiable activities for every unit; and unit, mid-course and end-of-course tests
- **audio and extra activities on English.com** — online grammar, vocabulary, reading and listening practice activities; and downloadable classroom audio files
- **full-course placement tests** — printable or online versions

MEL is an optional component and is designed to support Speakout by:

1. providing students with the opportunity to work whenever they want, using the resources most likely to enhance their learning of course material
2. helping students develop the skills to become responsible and autonomous students
3. allowing students to do work at their own pace and to track their progress

MEL content can be assigned for the whole class, groups or individuals (Vymetalkova, 2016; Vasbieva & Klimova, 2015; Pearson, 2014a; 2014b). The use of MEL allows for blending classroom learning with

synchronous and/or asynchronous learning outside of the classroom. It has the potential to build a bridge, whereby teaching and technology support learning and inform each other.

Overall, Speakout with MEL aims to support students to:

- engage with the materials and have a positive learning experience
- develop positive learning behaviors when using the materials
- progress in learning English
- be ready for the next stage of their learning in English
- achieve their goal(s)

The present study

The present study is an exploratory study, which set out to answer the following questions:

- What is students' motivation to learn English?
- What are students' attitudes towards Speakout with MEL when compared to other instructional experiences?
- What are teachers' attitudes towards Speakout with MEL and its use?
- How are courses implemented when using Speakout with MEL and why do institutions engage with it in the first place?
- Why and how do students and teachers use MEL?
- What can we learn about teaching and learning from the MEL data?
- What is the perceived impact of Speakout with MEL on students, teachers and institutions?

The present study is not an evaluation of practice. Pearson aims to keep Speakout with MEL flexible so that teachers are able to make their own choices on how to implement it. The flexibility in implementation, together with the fact that Speakout with MEL is less widely used across countries and institutions made it imperative, as a first step, to explore institutions' approach to implementing Speakout with MEL, and teachers' and students' experiences and perceived impact when using the materials.

Implementation studies have the potential to improve learner outcomes by allowing us to know what works where, when and why. They are key to informing teaching and learning by providing evidence on which products and services are likely to 'work' within a particular context, institution, and classroom. Implementation research is rooted in capturing the real-life experiences and insights of students and teachers to develop an evidence-informed understanding of the factors that can enable or impede intended and unintended outcomes. Variability across settings, cultures, institutional preferences or priorities, professional development, and infrastructure can all affect the implementation of, and outcomes associated with, products and services.

This study examines the implementation of Speakout with MEL in three Polish institutions: Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie Filia w Białej Podlaskiej (AWF) in Biała Podlaska, Wyższa Szkoła Logistyki in Poznan and Wyzsza Szkoła (WS) in Warsaw.

The structure of the report

For ease of navigation and readability, the next section outlines the methods used in this study, including details on the data collection and analyses, and then moves directly to the discussion of the findings. Before presenting the full results in the appendices, we make recommendations and explain the limitations of the study, as well as listing the references.

Sample

Given resources and timelines, a maximum of three institutions were sampled for this study. Stratified sampling was used to account for potential variations in practice. To derive a sample, a number of criteria were developed, as seen in Table 1. In hierarchical order, the non-negotiable criteria for selection were that:

- institutions used the most up-to-date version of Speakout
- institutions used the 'full Speakout package', i.e., the Students' Book, the Teacher's Book, MEL and ActiveTeach
- at least two institutions would be using the full package for more than one year

To ensure triangulation of the data collected for the sampling and to make sure that the data was accurate and current, four approaches were used:

1. Using the non-negotiable criteria, targeted searches were conducted of the sales databases that hold institutions' information related to Speakout with MEL.
2. Searches identified institutions using the MEL platform. This information was corroborated with the sales data to point us to institutions with higher student activity using MEL.
3. Direct conversations were held between our Poland efficacy lead and sales representatives, who have an in-depth and up-to-date knowledge of institutional practices and issues.
4. Institutions were identified by our Poland efficacy lead, who also has in-depth knowledge of the institutions through research and the marketing activities that she leads.

Information collected was recorded in an Excel spreadsheet for each institution and for each of the criteria. Approximately 20 institutions were found to use the latest Speakout version with MEL. No institution was found to use ActiveTeach. From these 20 institutions, intelligence suggested that they were either not using MEL in practice, they did not fit the length of time of using MEL, or they did not fit other of our criteria (i.e. they were providing in-company courses to small numbers of students).

After applying the sampling criteria in February 2017, four institutions were deemed suitable, three of which took part in the research. Table 1 provides a summary of the initial criteria drawn, the final criteria used and the reasons behind the sampling decisions leading to the final sample.

Pearson

Table 1: Original and final sampling criteria

Criteria type	Original inclusion criteria	Inclusions and criteria met
Institutional criteria per a country		
Type of adoption[1] & experience in using Speakout with MEL (non-negotiable criteria)	• The most recent edition is used by the institution • The 'full' package is used, including the: Students' Book; Teacher's Book (whether using the print version or PDFs in ActiveTeach); MEL and ActiveTeach • Institutions using the Workbook will be included in the study, but the Workbook is not an essential component for selection • At least two institutions with more than one year experience of using the Speakout package	• Four institutions met these criteria, apart from using ActiveTeach, for which no institution did. From the four institutions, three were selected[2]. One of the institutions initially agreed, but was unable to gain permission from senior leaders to take part in the research within the timeframe available. • Two of the selected institutions had more than one year of experience using the Speakout package.
Size of adoption per institution	• Institutions with the highest number of students (e.g., serving university students and not in-service company courses, which have small numbers of students and are also hard to access) • At least one institution with one of the highest numbers of students using Speakout with MEL, one medium and one small	• Due to the small number of available institutions to choose from, this criterion was abandoned
Levels taught	• Ideally, focus should be placed on courses that have the largest number of students • If possible, one institution whose program includes all Speakout with MEL levels (preferable criterion)	• In the institutions selected, the large majority of students concentrated in intermediate classes. Intermediate classes were selected for observations, and students at that level were interviewed.
Type of institution	• Institutional variation is preferable	• One institution was a private language school (PLS) and two were public universities.[2]
Institutional effectiveness	• Rating of effectiveness by an external regulatory body • If the above criterion cannot be applied, we should consider the following: two institutions that are perceived to have influence on others; their reputation in the country; institutions' success in national exams; their support of government initiatives and associations; the number of accreditations they are	• No regulatory body operates in Poland.

[1] Adoption is defined as the number of books sold per institution.
[2] To reach one of the two universities, a teaching agency which supplied the Speakout with MEL implementation model and the teachers teaching it was initially contacted. The teaching agency connected us to the university and also provided relevant data for the study.

Criteria type	Original inclusion criteria	Inclusions and criteria met
	awarded; the number of students that complete their courses; the level students reach when they complete their courses.	
Geographical spread	• Institutions are located in different regions	• Institutions were located in three different regions: Greater Poland, Masovia and Lublin.
Length of courses	• At least one or two institutions teach full-length and short courses	• Two institutions offered both long and short courses.
Teacher		
Experience	• New and experienced teachers (experienced teachers should have more than one year working with Speakout with MEL)	• All teacher criteria were abandoned given the very small number of teachers teaching Speakout with MEL within the institutions (e.g. two or three teachers per institution). • However, information on the criteria was collected through the interviews: • All teachers (bar one) had been using Speakout with MEL for more than one year • All teachers reported that, overall, they were comfortable using technology • All teachers (bar one) had experience in teaching courses before their institution switched to Speakout with MEL
Comfort with technology	• At least one teacher who the institution believes is comfortable with technology and has used MEL effectively to deliver Speakout lessons	
Courses experience	• At least one teacher per institution who has experience of how the course was taught before Speakout with MEL	
Courses taught	• Teachers who teach long and short courses	

Pearson

Methods

We used a multiple-case-study design to conduct the study. To conceptualise the study design and manage the development of the research instruments, we were guided by the Consolidated Framework of Implementation Research (CFIR) devised in 2009. The CFIR appeared relatively exhaustive. We undertook an exercise to reduce the number areas to be examined in order to better fit with the study's aims, and to allow themes to be explored in depth.

A brief outline of the three overarching categories of investigation and a small sample of the sub-categories is provided below. The first two categories were part of the CFIR framework, details of which can be found in Appendix A. We added the third category referring to perceptions of impact, as it was an important part of our research questions:

- the 'what' of the intervention — components used to teach English, structural and processual, and their organisation
- the 'why' and 'how' of the intervention — the intervention characteristics; inner settings, such as the priority placed on the use of Speakout with MEL; institutional goals and whether and how they were enacted, and, the usefulness of the 'intervention', etc.
- the perceived impact of the intervention — on the department/institution as a whole; on students' and teachers' access and engagement, and, on student achievement, etc.
- Methods used to collect the data included:
- a student questionnaire
- a pre-questionnaire filled in by the language departments' coordinators
- student focus groups
- staff interviews
- observations
- student MEL data analysis
-

In the next section, we provide more detail on each of the methods used.

Student Questionnaire

Development of the questionnaire

The questionnaire was developed with two aims in mind:

1. That they could be used in different countries (Poland and Turkey for Speakout and Mexico, Peru, and Colombia for a previous study on a similar product)
2. That they would elicit robust evidence to address the aims of the study.

To achieve these aims we were guided by the conceptual framework used for the study. We:

- acknowledged, evaluated and, where appropriate, adapted or revised questions already used in questionnaires developed within Pearson for similar research purposes
- were guided by questionnaires used in other research studies (e.g., Dorneyei with Taguchi 2010; national student questionnaires used by universities, etc.)
- consulted the Pearson product development teams and our local Pearson efficacy leads, who have a deep understanding of both the products and the local context in which the products operate
- took advice from other researchers within our Impact Evaluation team in terms of the design
- elicited feedback from co-ordinators or teachers and students through a pilot.

The development of the student questionnaire was supported by a group of key stakeholders who had the knowledge and experience to support the robustness of the instruments. Table 2 outlines the individuals involved in the process, their knowledge and expertise and the tasks they performed. After feedback was sought and contextual information was adapted, the questionnaire was translated into Polish and Spanish by the Poland, Mexico, and Colombia efficacy leads, and was piloted. Pilots involved one round in Poland and two rounds in Colombia, including three institutions and 18 students in total.[3] Following the pilots, relevant adjustments were made. Draft and final questionnaires were completed by the Impact Evaluation team in English and were then transferred to the final student questionnaire for Poland by the efficacy lead in Polish.

[3] Pilots included 15 students from Colombia (5 in the first round and 10 in the second round of piloting) and 3 students from Poland.

Table 2: Stakeholders involved in the development of the questionnaire, experience and task performed

Stakeholder	Experience	Task performed
English, Global Impact Evaluation team: Elpida Ahtaridou	• Expertise in methodological and quantitative and qualitative instrument design • Expertise in product research for different purposes and for the different stages of product development	• Designed the questionnaire • Reviewed previously developed instruments used in internal and external to Pearson research and their related reports • Decided on which questions might be appropriate for reuse • Drafted new questions
Speakout portfolio manager: Susan Trory	• In-depth knowledge of Speakout with MEL and of the different country needs and issues in relation to its use and delivery	• Provided feedback on previously used questions from internal to Pearson questionnaires • Reviewed questionnaire items and drafted new items, when appropriate
Efficacy geography lead: Alicja Tarkowaska	• Experience in developing and conducting research • Deep understanding of the study aims and of individual institutions' needs, and issues related to the delivery of Speakout with MEL	• Provided feedback on previously used questions from internal to Pearson questionnaires • Reviewed questionnaire items • Translated, culturally adapted and piloted the student questionnaire • Created the questionnaire online
Coordinators, teachers and students	• In-depth knowledge of Speakout with MEL in their institution and related issues • Understanding of terminology used in the institution and by their students related to Speakout with MEL	• Piloted the student questionnaire • Provided feedback on potential issues with questions and wording

Questionnaire structure and content

The first page of the questionnaire contained elements of informed consent, a brief explanation of the study purpose and use of data, confidentiality, data-protection procedures and the voluntary nature of participation.

In brief, the student questionnaire aimed to draw information about:

- the course students were attending
- their motivation related to learning English
- their comfort with the use of technology
- the use and frequency of different Speakout with MEL components in class
- the usefulness of different components and features
- perceptions of the impact of Speakout with MEL and of its components on access, student learning behaviours and attitudes, performance, preparation for the next stage in their learning and for their ultimate goal
- the likelihood of recommending Speakout and MEL to other students

The student questionnaire included a total of 25 questions. The majority were closed questions, using a four-point Likert scale that included a number of sub-questions. Open-ended questions were also used to gather qualitative responses related to students' views on the most useful features, benefits, challenges and areas for improvement.

Administration, collection and analysis of the student questionnaire

Print copies of the student questionnaire were distributed to students by their teachers in May 2017. Students were informed of the research and its purpose in advance, and their participation was voluntary. Based on past experience, institutions felt that distributing questionnaires in print was less burdensome and more likely to get higher returns when compared to administering the questionnaire online. Distribution was opportunistic.

After teachers had collected the completed questionnaires, they were either handed to the researchers on the day of their visit, or posted to the Pearson Poland office using secure mail. Then Polish native speakers, who also spoke English at a proficient level, inputted each of the print questionnaires in Google Forms in English.

Three datasets corresponding to the three institutions were received in Microsoft Excel. Analysis was performed by three Pearson researchers independently, using Microsoft Excel and the R platform.

Regarding data cleaning, the following processes were applied:

1. The number of columns, the number of rows and the type of each of the variables were confirmed for each dataset.
2. The data were scanned for suspiciously large numbers of missing data, unexpected values and other possible irregularities.
3. For each dataset, the online questionnaires were checked to confirm that there were no missing questions.
4. The value labels for each of the variables were confirmed.

A total of 201 questionnaires were analysed out of 202 returned (one questionnaire was blank and was removed from the analysis). The majority of students who responded to the questionnaire were from AWF (153). 39 student responses were from Wyższa Szkoła Logistyki and nine from Uniwersytet Centralny[4]. This means that AWF represents 76% (153/201) of all of the respondents in the student questionnaire. Although representation of views might seem imbalanced as regards the overall sample, questionnaire results from AWF and Wyższa Szkoła Logistyki represent the majority of student views that were studying with Speakout with MEL in their own institutions at the time of the research. The majority of students who responded to the questionnaire were from AWF (153/201). They were all studying at the Intermediate B1+ level (167/195) and represented approximately 85% of the total number of students studying using Speakout with MEL in their institution at the time of the research (153/180 approximately). 39 student responses were from Wyższa Szkoła Logistyki, representing approximately 65% of the student population studying with Speakout with MEL at their institution at the time of the research (39/60 approximately). Students from the Uniwersytet Centralny, who were using Speakout with MEL as part of their studies, were also given questionnaires to fill in. This was arranged by their lecturer, who also taught at Wyższa Szkoła Logistyki. Despite the small sample size and the fact that this university was not part of the research per se, the decision to use the results in this study was threefold:

a) the results remain the same if one includes or excludes the data collected from Uniwersytet Centralny;
b) the university used the same approach to using Speakout with MEL, thus, more data had the potential to increase the sample size and the learnings from the research;
c) for completeness and transparency. Permission to distribute questionnaires to WS students was not granted. (Students who completed a questionnaire may or may not have participated in the focus groups.)

Table 3 provides a summary of respondents per institution and level. Respondents came from four different levels, with the large majority studying across the three intermediate levels and the intermediate level itself. 86% (167/195) of the total number of respondents were studying at intermediate level. Representation of students from the intermediate level was large from within the

[4] Please note that a pseudonym is used to represent this university.

institutions too. 97% (144/148) of all AWF respondents and 61%, 23/38 of Wyższa Szkoła Logistyki respondents were studying at the intermediate level. All students from Public University (9/9) were studying at an advanced level.

Table 3: Total number of students studying with Speakout with MEL, and questionnaire responses per level

Institution	Pre-intermediate B1	Intermediate B1+	Upper-intermediate B2	Advanced C1	Total
AWF	3 (2%)	144 (97%)	1 (1%)	0 (0%)	148
Wyższa Szkoła Logistyki	13 (34%)	23 (61%)	2 (5%)	0 (0%)	38
Uniwersytet Centralny	0 (0%)	0 (0%)	0 (0%)	9 (100%)	9
Total	16 (8%)	167 (86%)	3 (2%)	9 (4%)	195[5]

Note: Five students from AWF and one student from Wyższa Szkoła Logistyki did not provide the level of study and, therefore, could not be included in the total number of students per level.

Furthermore, 90% (180/201) of students in the sample were studying in courses lasting 120 hours across two years. Most students (77%, 146/189) were at the middle of their course and 18% (34/189) were at the end. No student was at the beginning of their course

[5] Percentages have been rounded. In cases where the total was 101% after more than one of the percentages were rounded, one of the percentages was rounded down. This made no difference to the results.

Table 4: Number of student responses analysed by institution and level, student questionnaire

	AWF Poznan	Logistyki	Public University	Total
Course duration				
120 hours	153	26	1	180
Between 20 and 35 hours	0	13	0	13
60 hours	0	0	8	8
Total	153 (76%)	39 (19%)	9 (5%)	201
Length of time in course				
Beginning of the course	0	0	0	0
Middle of the course	85	13	8	106
End of the course	61	21	1	83
Total	146 (77%)	34 (18%)	9 (5%)	189

For each question with discrete answers, responses were analysed using descriptive statistics (frequencies and percentages). Psychometric analysis was used for a set of questions related to students' motivation levels.[6] The student questionnaire included five questions to ascertain their level of motivation. They were asked how important they felt each of the following was for them:

- using English in my life, work and study
- learning English
- succeeding in learning English
- successfully completing the course level
- achieving my goal by learning English

Responses were recorded as 'not important', 'somewhat important', 'important' and 'very important', and were aggregated to derive a single measure of motivation for every student. This aggregate measure of motivation is more reliable and useful for further analyses than using each one of the five individual questions separately, especially when comparing sub-groups.

Aggregating responses to groups of questions is common practice in academic research. In this case, to aggregate the responses of the students to the five motivation questions, we used a standard psychometric technique — the Rasch model (see Bond and Fox, 2001, for more details on the use of Rasch models.) An important characteristic of the Rasch model is that (in the case where there are no missing responses) there is a one-to-one correspondence between the raw score (i.e. the sum or the mean) and the Rasch measure (in technical parlance, one might say that the raw score is a 'sufficient statistic'). Due to the polytomous nature of the items, we chose to use the rating scale Rasch model.

One respondent was removed from the analysis because she/he failed to respond to four of the questions. Using the remaining data, most of the correlations between the responses of the students to the five questions were high, suggesting that this was a rather homogenous group of questions. Also, all the correlation coefficients were statistically significant at the 0.001 level. Establishing that we have a homogenous group of questions is a prerequisite when applying the Rasch model on such data. It is also interesting to observe that the scale demonstrated a high internal consistency, with a Cronbach's alpha = 0.81. The Cronbach's alpha reliability index was estimated using the Psych package on R platform (Revelle, 2017).

Before using the outcome of the Rasch analysis operationally, we investigated the model–data fit. In the context of Rasch analysis, if the model–data fit is not appropriate (within specific limits), then using the aggregated motivation measures of the students for further analyses is not appropriate. The model–data fit statistics of the Rasch model for this particular dataset were satisfactory for all the

[6] Frequency tables, cross-tabulations, visualizations and descriptive statistics were produced using standard commands of the sjPlot package (Lüdecke, 2017) and of the tidyverse package (Wickham, 2017). Inferential statistics (e.g. t-tests, correlations) were produced using standard R base functions, but more advanced analyses, e.g. reliability analysis for Cronbach's alpha and Rasch analysis, were run using the psych package (Revelle, 2017) and the TAM package (Robitzsch et al., 2017).

practical intents and purposes of this study (see Appendix 2 for details). The Infit Mean Squares were within the usual rule of thumb of 0.7–1.3 (see, for example, Lamprianou & Boyle, 2004).

The Rasch measures were transformed linearly to a scale from 0 to 10 to facilitate comprehension by less technically oriented readers. On this scale, 0 means less motivation and 10 means more motivation. To compare groups of students on numeric variables (e.g., to compare the aggregated Rasch-estimated motivation of students from different levels or institutes), we used graphical methods (e.g., boxplots), but we also used the appropriate t-tests and analysis of variance or their non-parametric equivalents, when necessary (e.g., in the case of very skewed distributions).

Due to time constraints, filled in responses to the open-ended questions in the students' questionnaires were not analysed.

Coordinator's pre-questionnaire

The co-ordinator's pre-questionnaire aimed to gain an understanding of the context, course structure and implementation model used in the different institutions before the day of the visit. It also supported the development of more targeted questions to be asked during the visit. In more detail, data collected through the pre-questionnaire focused on:

-
- **context** — factual information on the student and teacher numbers studying using Speakout with MEL, the number of classes offered and the length of the courses offered
- **classroom time** — details on how classroom time was structured, including the number of face-to-face lessons weekly, the length of class time, the structure of the lessons, preparation required by teachers and students before the lesson, etc.
- **assessment** — how assessment, both formal and informal, was structured, including an understanding of what type of tests students took and the conditions under which they completed them
- **homework** — whether or not students were required to complete homework; the place and suggested time for completing homework; whether homework was obligatory; whether students were allowed one or more attempts in MEL; and other details on how MEL was used
- **tracking student usage of MEL** — how teachers tracked student usage of MEL and what they did with the data they collected
- **tracking student performance** — how student performance was tracked, who was responsible for tracking it and how the performance data collected was used

The coordinator's pre-questionnaire followed the design processes and quality-control mechanisms outlined in the student and teacher questionnaire section above. The pre-questionnaire included both closed and open-ended questions. Coordinators were sent the pre-questionnaire at least a month before the day of the visit, and were asked to return it two weeks before the visit. To reduce the

burden on coordinators, they were advised to leave the open-ended questions blank if they felt that they had no time to fill them in.

Three pre-questionnaires were collected, one from each of the three institutions.

Student focus groups and staff interviews

Focus group and interview schedules followed the process outlined in the student questionnaire section as regards their design and the quality-assurance mechanisms followed for their finalisation. Student focus groups and one-to-one interviews with teachers were held on the universities' premises in March 2017 and were conducted by our Pearson researcher, who led the research and the design of the schedules. No student or teacher was excluded from participation. All students and teachers chosen by the coordinator were interviewed.

The purpose of the study and the procedures were explained to all potential participants by their institution, and teachers and students had the opportunity to decline taking part in the study. No incentives for participating were offered, and both the institution and the research participants could withdraw from the study at any point during the lifespan of the research project. Those who agreed to take part were asked to sign a consent form and agreed to the focus groups/interviews to be recorded.

Student focus groups conducted

Institutions were asked to arrange two student focus group with five to 10 students in each institution. Students were to be selected randomly from those students who agreed to take part in the study and from classes held on the day of the visit. The visit was arranged based on the availability of classes and a convenient time for the institutions.

Twenty-nine students were reached through a total of four focus groups. Students were studying in Wyższa Szkoła Logistyki and AWF. As per the student questionnaire, a focus group was also arranged with Uniwersytet Centralny by one of the lecturers who also taught in Wyższa Szkoła Logistyki. Given that they experienced very similar, if not the same, approach to the implementation of Speakout with MEL to students from Wyższa Szkoła Logistyki, their views were seen as equally valid. Permission was not granted to interview WS's students.

Table 5 shows the number of students who took part in the four focus groups conducted, and the course they attended at the time of the research. (Students who participated in the focus groups may or may not have filled in a student questionnaire.)

Table 5: Number of students who attended the focus groups, course they attended and timeframe within the course

Role			Number of focus-group students	
	Course attended	Timeframe within the course	Per course attended	Total institution
AWF	Bachelor students, Speakout Intermediate	End of first academic year (part of two years of study)[1]	5	11
	Masters' students, Speakout Intermediate		6	
Wyższa Szkoła Logistyki	Bachelor students, Speakout Intermediate	As above	3	3
Uniwersytet Centralny	Masters' students, Speakout Intermediate	As above	15	15
Total number of students in focus groups				29

Note: An academic year starts in September and ends in July.

Mature students usually attend evening sessions and were purposefully not interviewed, given that young adults, who study mainly during the day, constitute the majority of students attending these universities. Interviewing only young adults also allowed for more compatible views to be recorded. Participants' ages ranged from 18 to 25, studying full time in courses lasting two years for a total of 120 hours towards a Bachelors or a Masters degree. Focus groups were conducted in English, with a translator at hand, who supported students when they felt they needed it. The translator's help was required very few times.

To further support the validity of the focus-group interviews, a more informal conversational interview style was followed. Although key areas of the schedule were covered most times, the researcher allowed more questions to emerge from individuals' answers, asked in the natural course of things. This approach aimed to make participants feel more at ease. It also aimed to reduce item order effect, i.e. when the order in which the items are presented affects people's responses. In addition to the conversation naturally 're-ordering' the questions and themes (as outlined in the interview schedules),

the researcher also specifically asked questions in a different order for each focus group. To reduce the potential of dominant and shy bias, the researcher invited individuals to offer their views one by one, or randomly, when they were seen not to take part in the discussion. The same approach was taken if a 'declaration' was made by a participant.[7] Overall, participants were forthcoming in refuting a pattern or a view made by a colleague of theirs, if they indeed had a different view. At times, salient questions were asked to each one of the members of the group, either in random order or starting from whoever sat on the right or left side of the table.

Teacher interviews conducted

A total of 11 one-to-one staff interviews were conducted, lasting 45 minutes to one hour each. All interviews were conducted in English. For the three institutions in our study, 100% representation was achieved in terms of the number of teachers using Speakout with MEL in AWF and Wyższa Szkoła Logistyki (3/3 and 2/2 teachers respectively), and 20% (2/10) for WS.

Teachers from Wyższa Szkoła Logistyki who were teaching using MEL with a different Pearson book at the time of the visit were also interviewed. These teachers offered their views of MEL only, and reported practices they used in their own classes. These practices are transferable when teaching with Speakout with MEL. Lingua Nova's director and its teacher training coordinator were also interviewed, given that their perspective on the implementation of Speakout with MEL in WS was imperative.[8] Table 6 outlines a summary of the interviews conducted.

[7] 'Declaration' refers to a pattern mentioned by a participant. The researcher then seeks to confirm or refute that pattern by collecting data from the rest of the group.

[8] Lingua Nova is a school of foreign languages with clients all over Poland, who teach as Lingua Nova in many Polish institutions. Lingua Nova does not supply teachers; rather, teachers work on their contracts using the Lingua Nova curriculum and materials. The decision to use Speakout with MEL, the implementation model for its delivery and the training of the WS teachers are the responsibility of Lingua Nova.

Table 6: Summary of staff interviews

Role	No. of staff interviews				Representation per institution
	AWF	Wyższa Szkoła Logistyki	WS and Lingua Nova	Total	Total
Director	N/A	N/A	1	1	100%
Coordinator	1	1	1	3	100%
Teacher	2	3	2	7	100%
Total	3	4	4	11	100%

To further support reliability of the teacher interviews, only slight differences between the interview schedules for each role (director, coordinator and teacher) were made, with the interviewer asking follow-up questions for clarification when necessary.

Analysis of interview and focus-group data

Audio recordings from the focus groups and interviews were transcribed in full and transferred into the NVivo qualitative analysis software. Two researchers supported the coding of the data. The first researcher provided the initial coding based on the CFRI framework used to guide the study, initially deductively (see Braun & Clarke, 2006). The second researcher completed the coding across all transcripts.

The initial coding included 12 primary codes and 42 secondary codes. After the coding structure was created in NVivo, the two researchers coded the same three transcripts, separately. The three transcripts were selected randomly by the first researcher and represented 10% of the total interview transcripts available (Hodson, 1999), and were between 10 to 20 pages long each (Miles and

Huberman, 1884). The aim of the exercise was twofold: a) to derive final coding structure and assess the degree of agreement; and b) to support the consistency in the approach to coding between the two researchers.

The deductive coding structure was treated as an overall structure rather than a strict outline within which the researchers needed to fit the data. Thus, as the researchers read the transcripts, they adapted the structure to better suit the research questions and to include first and second order codes that had risen inductively from the data. Discussion between the two researchers adjudicated any coding disagreements, involving codes being moved, deleted, merged, split or renamed within the hierarchy. This process resulted in a final coding structure involving a total of seven first codes and 30 secondary and third order codes.

At the outset, the researchers agreed that a unitisation strategy that focused on meaning units would be pursued. This strategy meant that codeable units of text varied in length from a few lines to a whole paragraph. Key to the coding was that contextual information was included to further support researchers reach common ground in their coding decisions. Thus, paragraphs before and after the coded text were included most times for contextualisation. At times, the data within the code related to more than one first and/or secondary codes. In this way, a balance was sought between condensing data for analysis and retaining the uniqueness of meaning.

The proportion agreement method was used to understand the agreement between the two researchers. Although this method comes with limitations, e.g. it does not take into account that researchers might agree occasionally by chance (Bernard 2000), more complex methods were not deemed appropriate given that:

a) not all of our codes had equal probability of being used;
b) multiple codings on a text unit, which we followed, creates problems for calculating agreement between researchers with some statistics because they require that only one code is applied to a unit of text (Burla et al. 2008);
c) Our aim, from the outset, was not to generate variables for use in statistical analysis. Our aim was to systematically classify and retrieve text using clear, distinct definitions to do so;
d) this was an exploratory study for which other researchers have argued that the simple proportion agreement method is an acceptable approach (Kurasaki 2000).

Overall, the two researchers were relatively aligned in their assigning of codes and in the process followed to derive codes. This meant that they were satisfied in proceeding with the coding of the full set of transcripts without replicating the activity. The relative agreement between the researchers could have been due to:

a) the knowledge and experience they built through using the same coding system in three similar research projects;

b) their good knowledge of the research and of its aims and objectives from the outset;
c) the fitness of purposes of the CFRI framework sections used;
d) the clear structure of the product components and features examined;
e) the simplicity of the coding system and the clear and explicit definitions given to the secondary codes, so that meaning was not sacrificed in favour of simplicity. (Despite the relatively large number of codes, they were relatively distinct in their definition.)

Patterns were identified not only by looking at repeated occurrences, but also by similarity, 'declaration' and confirmation, missing patterns expected to be present and co-occurrences. Data collected from the other sources used in this study and their findings also supported the development of patterns. In addition, findings from four similar studies conducted at the same time as this research on Speakout and its counterpart, Top Notch, supported the development and our understanding of different patterns. These studies sought to answer the same research questions as this research project and, overall, used the same research instruments to collect their data. Finally, consideration was placed on whether emerging patterns appeared to be congruent with prior hypotheses and relevant literature (Hopkins & Ahtaridou, 2009; Quartaroli, 2009).

Classroom observations and post-observation debriefs

A structured classroom observation schedule was used to understand how Speakout with MEL materials were implemented in practice. Classroom observation did not aim to evaluate practice. Observations focused on the following areas of exploration:

- **lesson preparation** — what that entailed and why the lessons were set in a particular way
- **exposure to materials** — what materials were used during the lessons by students and teachers
- **how materials were used** — e.g., if materials were used to support individual activities or pair activities, etc.
- **implementation** — whether the teacher guidance was implemented; if teachers differentiated well for students' different needs and levels and how they did it in practice; whether guidance was given to students on how to use the materials; the pace of the lesson; and the types of activities used in the classroom based on the topic learnt
- **navigation of materials** — whether they were easy or hard to navigate and whether support was requested by students during the lesson
- **feedback** — from teachers to students or from peer to peer
- **technology integration** — whether teachers and students had access to technology in the classroom and how it was used
- **engagement** — whether students were on task during the lessons, and how they exemplified it, e.g., by gesturing or asking questions, etc.

The observations also included three sections that required researchers to describe what students and teachers did at the beginning, middle and end of the lesson, as well as record whether the lesson followed the Teacher's Book and what, if anything, was done differently.
Detailed descriptors were used for each category of exploration outlined in the observation schedule. Descriptors aimed to support researchers to record observations systematically and consistently. The observation schedule and associated descriptors were used in a previous Pearson study published by What Works Clearinghouse, and were slightly modified to better meet the needs of this study.

After the observation, debriefs with teachers took place, each lasting 20 to 30 minutes. The debriefs aimed to understand teacher perceptions on how the session went, the use of materials and its usefulness when teaching the specific lesson, barriers and suggestions for improvement.

MyEnglishLab (MEL) student data

A framework was developed in order to extract and report on the MEL student data. The framework aimed to collect usage and performance data (as outlined in Table 7). The analyses outlined below were not always possible or appropriate based on the data received. We adapted the framework when this was the case.

Pearson

Table 7: Summary of framework that guided the MEL data extraction

Item	Data collection, analysis and presentation
Assignments and practices[9]	• Number of assignments/practices and tests assigned • Percentage of completions by institution, level and class • Average number of completions by institution, level, class and individual student from the total number assigned
Time on task[10]	• Average of students' time on task on assignments/practices and tests per institution, level, class and student
Attempts	• Number of attempts on assignments/practices per institution and level
Assignments/practices scores	• Average percentage achieved in first, last and highest attempt per institution, level and class • Percentage of students achieving within different score boundaries per institution
Test scores	• Average percentage achieved on test scores per institution, level and class • Percentage of students achieving within different score boundaries
Progress	• Improvement between scores on first and highest attempts at assignments/practices • Relationship between individuals' average assignment/practice scores and % of assignments/practices completed • Relationship between individuals' test scores and percentage of tests completed • Relationship between individuals' average test scores and percentage of assignments/practices completed
Reliability/validity	• Correlation between average student scores on assignments/practices • Correlation between average student practice scores per unit. (In the case where many correlations were computed, the results were presented in the form of quartiles to avoid the presentation of long tables

Data was extracted in July 2017. Due to the manual extraction of the MEL data and the short timeframe in which to extract it, a total of 20 classes per institution was agreed for analysis. The sampling strategy involved two steps:

1. Stratified sampling to include all level courses offered in the institutions, most recently completed courses and courses with more than five students.
2. Randomly selected classes from the remaining sample.

[9] MEL records student scores on assignments and practices. Exercises available in the system can be assigned by teachers. These exercises automatically show as assignments. Exercises not assigned show as practices. Although some exercises might show as practices, it does not necessarily mean that they have not been assigned by a teacher, as, they might have assigned them verbally in class, something that seems to happen often. It is not possible to know which practices have been assigned by teachers and which might have been undertaken by students on their own accord.

[10] Time on task is purely the time spent on assignments/practices and tests, and not log-in time.

A total of 41 classes' MEL data was extracted into .csv files: 15 classes from AWF; eight from Wyższa Szkoła Logistyki and 18 from WS across the Elementary, Intermediate, Pre-Intermediate and Advanced levels. In some cases, it was not possible to sample a total of 20 classes per institution because of the small number of classes or students. The 41 classes included data for a total of 750 students: 328 from AWF, 113 for Wyższa Szkoła Logistyki and 309 for WS.

MEL data was extracted from classes from 2014 to 2017, the majority of which ran for two years (38 classes). Three classes from one institution ran for one year. Table 8 provides a summary of the number of classes extracted per institution and level.

Table 8: Summary of the number of classes extracted per institution and level

	No. of classes				
	Elementary	**Intermediate**	**Pre-intermediate**	**Advanced**	**Total**
AWF	–	15	–	–	15
Wyższa Szkoła Logistyki	–	–	5	3	8
WS	4	9	5	–	18
Total	4	24	10	3	41

Analysis of assignments/practices and tests

We conducted an analysis of students' performance on practice/assignments and tests. Statistical means and other descriptive statistics were calculated for each analysis. Student performance was analysed using the score boundaries set by the platform: 90–100% (A); 80–89% (B); 70–79% (C); 60–69% (D); 50–59% (E); and 0–49% (F).

All scores of 0% were excluded from the analysis. The reasons for this exclusion is that:

- students were given multiple attempts to answer assignment questions, thus it is highly unlikely that a 0% score represents a student who had tried to complete an assignment multiple times and had failed each time
- a student who received a 0% score because she/he had not submitted an assignment/practice or had not submitted it on time, had made no attempt to complete an assignment/practice
- a teacher might not have marked a student's response to open ended questions

In all of the above cases, we are confident that 0% does not correspond to the real ability of students. Therefore, including 0% scores in the overall analysis would have artificially skewed the distribution of the scores. For transparency, analysis including the 0% scores is included in the report's appendices.

For each assignment/practice, the database reported one overall score per student. There was no information in the data regarding individual items.

Assignments/practices

Assignment scores are reported based on the highest attempt (MEL captures the first, last and highest attempts.) We decided to use the highest attempt because we wanted to understand the progress students made by the end of the course, instead of recording their average performance from the beginning to the end of the course. In addition, a comparison of the highest and last attempt scores showed that there was no difference between them in terms of student performance.

A total of 736 out of the original 750 students were included in the practices score analysis — 14 students were excluded; four had completed no practices and ten only had zero scores. Only 59 students had completed assignments and they were all included in the analysis.
A total of 106,526 practices were analysed: 5% (4,805) were excluded as they scored 0%. For the assignments, out of the 5,402 assignments, 3,194 were analysed; 2,208 (41%) were excluded as they scored 0%.

Results on assignments are presented separately from practices in the report, wherever meaningful. Overall, little data on assignments were found, and therefore they are less reliable compared to practice data. However, the analysis on assignments is included in the report for completeness

Table 9: Practice analysis — number of students whose data was collected and analysed and number of practices completed and excluded

Institution	Total number of students		Total number of practices	
	Collected	Analysed	Completed	Excluded for having a 0% score
AWF (Pre-intermediate level)	328	326	48,875	3% (N=1,548)
Wyższa Szkoła Logistyki (Pre-intermediate and Advanced levels)	113	103	10,199	23% (N= 2,364)
WS (Elementary, Pre-intermediate and Intermediate levels)	309	307	58,048	2% (N=893)
Total	750	736	117,122	5% (4,805)

Student progress analysis on assignments/practices

To gain an insight into student progress on assignments/practices, we analysed data for 733 students in total (from the .cvs files received). For the progress analysis, we removed all practices or assignments included in the files that were not attempted at least twice, or that had a 0% or 100% score.

Exclusions included three AWF students, who only had 0% or 100% scores in all assignments/ practices; five Wyższa Szkoła Logistyki, seven AWF students, and two WS students, who had only one attempt in all the assignments/practices they completed; 9,672 assignments/practices, which had a 0% score; and 1,981 assignments/practices, which had a 100% score. A summary of the number of student data collected and analysed and of the assignments/practices completed and excluded per institution is provided in Table 10.

Table 10: Progress analysis — number of students whose data was collected and analysed, and assignments/practices completed and excluded

Level	Total number of students		Total number of assignments/practices	
	MEL data collected	MEL data analysed	Completed	Excluded having a 0% or 100% score
AWF (Pre-intermediate level)	328	318	15,616	15% (2,326) (1,560 0% scores and 766 100% scores)
WS (Elementary, Pre-intermediate and Intermediate levels)	309	307	30,949	29% (8,835) (7,768 0% scores and 1067 100% score)
Wyższa Szkoła Logistyki (Pre-intermediate and Advanced levels)	113	108	2,815	17% (492) (344 0% scores and 148 100% score)
Total	750	733	49,380	24% (11,653)

Analysis of student test scores

Out of 328 students from AWF (all at the intermediate level), thirteen students were removed because they only had zero scores, so test data from 315 students were analysed. Overall, students completed 3,411 tests. As per the assignments analysis, zero scores were removed. This meant that we excluded 1,145 tests and analysed a total of 2,266.

Reliability and validity of MEL practices / assignments / tests

To investigate the reliability and validity of student scores from assignments/practices and tests, we estimated the following:

- correlation between average practice score and average test score
- correlation between test scores
- correlation between the average unit practice/assignment

To gauge the internal consistency of unit scores as a measurement of student performance, we computed the correlations between the students' average practice/assignment score on different units. The first step was to compute, for each student, the average practice/assignment score for all of the exercises/assignments within each unit. Then, these average practice/assignment scores for each student and for each unit were correlated between them across all students.

When appropriate, both Spearman and Pearson correlations were computed and presented to accommodate for skewed distributions of the data. The effect of outliers was also investigated wherever appropriate. Removing a limited number of outliers did not change the values of the correlations significantly.

Method for deriving findings on perceptions of impact from all data

To derive findings on the perceptual impact statements the following method was used:

- Student questionnaire results took prominence. This was because we believe that students are the most important voice in the learning process and because the student questionnaire data have the largest samples. Everything above 50% was seen as the majority. If most of the questionnaire questions on the same outcome, e.g. access, showed above 50% of respondents to have a positive view, then the outcome was also considered to be positive. 70% and above was seen as a 'very positive' outcome. When results were between 50% and 70% the outcome or the component(s) 'assessed' were seen to need some level of improvement. When the results of two out of the three main Speakout components (the Students' Book, ActiveTeach or MEL) or the majority of features and/or skills for each of the component were positive, then the overall finding about the outcome category was also seen to be positive.

- In the same realm, if less than half of the areas under investigation relating to the components' features and/or skills were below 50% in the student questionnaire, the outcome was considered not positive. When questionnaire results were skewed by one institution, this is mentioned in the report and is taken into account to derive overall findings.

- The approach to the teacher data was the same as that used for the student data. When students and teacher data agreed, the outcome was seen as positive. There was never a disagreement between teacher and student data because at no point were any two or more of the Speakout components, or half of the components' features/skills, negative. In cases where individual components, features or skills had negative results, this was clearly stated in the relevant section(s) of the report.

- When data from the student focus groups and teacher interviews agree with the student questionnaire results, the outcome is considered positive. When there is disagreement

between the two, if the finding from the student focus groups and teacher interviews does not derive from a strong pattern, then the outcome is still considered positive. If the finding from the student focus groups and teacher interviews derives from a strong pattern, then results are considered mixed or the outcome negative.

- If student focus groups and teacher interviews point to what one would consider a 'deal breaker' (the organisation switching to another product or students found to not use the product and able to provide a number of reasons that point to their dissatisfaction for doing so), even if the questionnaire results are positive, the outcome or the feature/skill is seen to not have had a positive impact, and this is mentioned clearly in the findings.

- When observational data disagreed with the student focus groups and teacher interviews, and quantitative data, it was thought that a judgement would need to be made on which would take precedence, based on the strength of the evidence collected. This was not necessary, however, as no disagreement was found between observational and interview or focus group data. In all cases, they were either in agreement or complimented each other.

- Lastly, four researchers who worked on the Speakout and Top Notch research studies conducted in 2017 (which, had the same research questions and used the same research instruments) read the results multiple times and independently applied the method outlined above. When there were disagreements, a reapplication of the method took place by each of the researchers. Subsequent meetings took place to reach a consensus.

Pearson

Discussion of findings

In this section, we discuss the study results. The discussion is based on the data outlined in the results section in Appendix B, and includes data from the student questionnaire, the coordinators' pre-questionnaires, the student focus groups and staff interviews, and the MEL student data analyses.

Please note that students who responded to the questionnaire these represent the views of AWF and Wyższa Szkoła Logistyki students. Students' views from AWF and Wyższa Szkoła Logistyki were relatively representative of the views of all the students studying in their institution using Speakout with MEL (approximately 85% representation; 153/180 students) and 65% representation (39/60 students). Thus, respondents' views from the questionnaire represent the views of students in two out of the three institutions that took part in the research.

In terms of its characteristics, the majority of the student sample in the questionnaire was relatively homogeneous. All students who responded to the questionnaire from AWF and Wyższa Szkoła Logistyki were studying at the intermediate level, representing 95% (186/195) of the total sample. Furthermore, most students were studying in courses lasting 120 hours in total (90%, 180/201) over two years of study and were mid-way through their courses, meaning they were at the end of their first academic year of study (56%, 106/189).

The 29 students who took part in the focus groups were all young adults between 18 and 25 years old, studying full time at intermediate or advanced levels in courses lasting 120 hours in total over the course of two years. Students who took part in the focus groups may or may not have filled in a questionnaire.

From the 11 teachers interviewed, 100% representation was achieved in terms of the number of teachers using Speakout with MEL at AWF and Wyższa Szkoła Logistyki (3/3 and 2/2 teachers taught using Speakout with MEL, respectively), and 20% (2/10 teachers) at WS. Two teachers from Wyższa Szkoła Logistyki, who were teaching using MEL with a different Pearson book at the time of the visit were also interviewed and offered their views of MEL only. Lingua Nova's director and teacher training coordinator were also interviewed, given that their perspective on the implementation of Speakout with MEL at WS was imperative.

Findings are presented thematically and follow the order below:

- Students attitudes towards Speakout with MEL
- Teachers attitudes towards Speakout with MEL
- Implementation of Speakout with MEL, including the reasons for engaging with Speakout with MEL; institutions' readiness for implementation; the teaching and learning and assessment approach implemented, especially relating to MEL; and the training and support for teachers and students
- Student and teacher usage of MEL
- Learnings for teaching and learning from the MEL data

- Student and teacher perceptions of the impact of Speakout with MEL on students, teachers, teaching and the institution

Students' attitudes towards Speakout with MEL

Finding 1: The majority of students report to prefer Speakout with MEL to classes that do not use it.

Student questionnaire results, student focus groups and teacher interviews agree that, overall, students like learning using Speakout with MEL. Student questionnaire results in particular show that 78% (154/198) of students liked the combination of teacher instruction and independent practice using MEL; 68% (134/197) reported that they prefer to complete practices in MEL than in the print workbook; 67% (127/189) enjoyed learning with Speakout with MEL more than other classes without Speakout with MEL; 61% (117/192) were more engaged and 58% (111/190) were motivated in classes that used MEL, compared to those that did not use Speakout with MEL.

Teachers' attitudes towards Speakout with MEL

Finding 2: Evidence indicates that teachers are in favour of using MEL, but not all make it an integral part of their teaching. Teacher implementation of Speakout with MEL is influenced by teachers' beliefs, attitudes and competence, interviews suggest.

Student focus groups and teacher interview data (at least eight out of ten teachers), and students' questionnaire results (79%, 155/197) agreed that their teacher was in favour of the use of MEL. However, 40% (78/197) disagreed that their teacher made MEL an integral part of their learning. The implementation models examined also showed that, although MEL was one of the main parts of the learning model, its enforcement allowed for variation in usage by both teachers and students, which could lead, in some cases, to MEL becoming less of an important tool for learning.

Teachers' use of MEL seemed to be intertwined with their beliefs, attitudes and competence in using technology and/or MEL. Interviews suggested that the use of MEL was challenging for some teachers because of their beliefs about the usefulness of technology in learning, their confidence in using technology, and concerns about the time needed to master MEL.

Implementation of Speakout with MEL

Finding 3: Evidence suggests that Speakout with MEL has the potential to respond to institutional challenges, such as student motivation, and to bridge classroom and home learning. Speakout with MEL can also raise the profile of the university, giving it a competitive edge.

Evidence from this study suggests that Speakout with MEL has the potential to positively impact student motivation when compared to other modes of instruction. Interviewees noted that students were familiar with technology and used it in their everyday lives. They indicated that such familiarity can further support student engagement with learning. This is particularly useful given that the obligatory nature of English courses for students entering university who operate below B2+ level has a negative impact on students' motivation, according to teachers. Student questionnaire data suggests that the majority of students showed moderate levels of motivation (45%, 91/200), while 30% were found to have low levels of motivation. Teacher interviews suggest that student motivation was low.

MEL can also bridge the gap between home and classroom and allow for continuity in learning English, according to interviewees. Given the limited class-contact time, especially with extramural students and absenteeism,[11] MEL is a key solution to supporting learning beyond the classroom. Additionally, learning with MEL can be monitored by teachers at any point.

Speakout with MEL is also seen as a type of learning suitable for today's world. Given the prominence of technology in every aspect of daily life, teachers felt that learning should be no different. One teacher noted that they could not envisage teaching without a platform in the future.

Lastly, Speakout with MEL has the potential to raise an institution's profile and give them a competitive advantage, interviewees suggested. Teachers reported that the use of Speakout with MEL made students gravitate towards their institutions, instead of others in the area that did not use technology to support learning.

Finding 4: Data indicates that the Students' Book, the Teacher's Book and MEL are the main Speakout components used in the three institutions visited. The use of MEL features beyond the assigning of practices is not widespread, teacher interview data suggests. As regards students' use of MEL, the majority used it to complete assigned tasks. Although the majority of students reported that they checked the MEL gradebook weekly or monthly, 28% (56/197) said that they did not check it at all. The use of ActiveTeach was rare.

[11] Class attendance is mandatory in Poland; only one session is allowed to be missed per semester without a valid reason (e.g., illness). Any other absence must be accompanied with a doctor's or employer's note. In theory, this policy should promote high levels of attendance across classes. In practice, however, teachers noted that the minimum attendance for extramural students could be as low as 50%. They reported that many students produced notes from employers about having to work on given weekends.

Student questionnaire, student focus group, teacher interview and observational data suggests that the Students' Book, the Teacher's Book and MEL were the core components used for the delivery of Speakout across the three institutions. The Students' Book provided the guide for the content to be covered, while MEL was used as the main homework tool for re-enforcing learning. The Teacher's Book was used to support teaching and was supplemented by other materials as teachers saw fit. Extra grammar and extra vocabulary were frequently used, while ActiveTeach was used rarely.

The use of MEL was perceived to be compulsory by nearly all students, according to both focus groups and questionnaire respondents (94%, 188/201). Assignments and practices were the core MEL feature used across all three institutions, and the majority of students indicated that the main reason for using MEL was to complete assigned work. A further 20% (37/186) reported in the questionnaire that they also did extra unassigned practice.

Checking one's progress in learning is key to success. It seems that there is more to be done in order to encourage more students to understand the usefulness of the gradebook and encourage its use further, as only 55%, (108/197) responding to the questionnaire suggested that they checked the MEL gradebook at least once a month or more, and 28% (56/197) of students reported not to use it all. Students in one focus group told us that they viewed their performance on the gradebook nearly every day, and several noted that self-monitoring through the gradebook motivated them to constantly improve.

Data also suggest that teachers are less likely to use the diagnostic and error reports. Lastly, tests in the assessment pack were used relatively infrequently, as teachers were concerned that students could purchase the answers online. One teacher, discussing the tests in the assessment pack noted that tests would not be beneficial to students if they could use 'cheat sheets' to copy the answers from an answer key.

The Table below provides an overview of Speakout's mostly used and less used components, and those not used at all.

Pearson

Table 11: Speakout non-core components, all data, student questionnaire, student focus groups and teacher interviews

Non-essential components		
Mostly used[1]	**Less used[2]**	**Not used**
• Extra grammar (81%, 151/186) • Extra vocabulary resources (83%, 154/185).	• **English.com**: it seems that less than half of students used English.com, given that only 47% (78/167) reported that they found the extra practices useful, and 43% (72/167) that they found the audio programmes to be useful. However, student focus groups and teacher interview data suggests that even this percentage is over-inflated as all teachers and students who were asked about English.com in interviews were not aware of it. In addition, 17% of those answering the questionnaire chose not to answer the question about usefulness of English.com, implying either no awareness or limited to no use. • **Extra BBC clips worksheets:** (55%, 87/157) reported using them • Extra Pronunciation practice: (27%, 44/162) • Extra Skills Development: (29%, 46/161)	• Only one teacher across the three institutions appears to have used **ActiveTeach.** Although teachers identified a number of benefits, they reported two key barriers that prohibited them from using ActiveTeach: o the lack of infrastructure in their institutions, e.g., lack of interactive whiteboards or Wi-Fi in classrooms o its location on a CD-ROM. This was troublesome for an instructor, who suggested that it was time-consuming to download the software, and because of this she abandoned using it • Overall, teachers agreed that they would be more likely to use ActiveTeach if it was accessed through MEL.

Notes: (1) Mostly used means > 50% of students agreed or strongly agreed that they were used. (2) Less used means < 50% of students agreed/strongly agreed that they were used.

Finding 5: Overall, implementation of Speakout with MEL between and within institutions varies, leading to different student learning experiences. In two of the institutions, individual teachers deployed Speakout with MEL as they saw fit. In one institution, an overall framework for implementation was provided. Variation is especially evident in: a) the use of MEL scores as part of institutions' assessment system; b) their weighting towards students' scores; and c) the monitoring of students' progress in MEL assignments by teachers.

Teachers working for Lingua Nova deployed the Speakout resources based on an overall framework provided to them. Beyond the framework, teachers were given guidance on the different MEL functions so that they used them, or not, based on their students' needs. The framework included:

- application: MEL is used as homework to help with the consolidation of learning
- assessment: MEL scores count for 40% of students' final score
- monitoring of student progress in MEL: teachers have to produce a report on student progress before every lesson and send it to the central administration office

Although some uses of MEL seem to be more commonplace than others across the three institutions, variation in implementation was also observed. Variation was especially evident in the use of MEL scores as part of institutions' assessment system and their weighting towards students' scores, as well as the monitoring of students' progress in MEL assignments by teachers. In more detail:

Majority of teachers

Class vs home use: two teachers used MEL in the classroom, while the majority did not.
Number of practices assigned: most teachers seemed to assign all assignments at once (ranging from 320 to 445 as observed in the MEL database), with the exception of one group of students who were only assigned, on average, around 90 practices.

- Deadlines for completion: in most cases assignments were open for a whole semester.
- Mixed number of teachers
- How practices were assigned: assignments could be assigned formally through the system, or verbally by teachers in class.
- Number of attempts: some teachers allowed for three attempts, some for unlimited attempts, and others for one attempt. (Limiting attempts was part of teachers' efforts to gain what they saw as a more accurate interpretation of students' performance, and to gauge their scores in tests.)

- Use of different MEL features: some teachers used the MEL assignments/practices only, and others used the tests too.

- Assessment and weighting: MEL scores counted formally towards students' final score in one institution. Students had to score >50% on assignments to be considered a pass, although the quantity of assignments completed was not a factor taken into account. In the other two institutions, practices and assignments did not count towards students' final scores, although a few teachers used them as a way to differentiate between, for example, a B1 and B1+ student. Other teachers 'rewarded' students with a + in their final score, looking both at the quantity of assignments they completed and whether their scores in MEL were above the class average.

- Monitoring: in one institution, teachers had to submit a weekly student progress report from MEL to central administration. Other teachers monitored student progress in MEL less frequently.

As regards time spent on MEL, the majority of students in the questionnaire reported to use MEL once to twice weekly for half an hour to an hour. Results from the MEL data analysis show that, on average, students seemed to spend less than 13 hours on task during the two years of their course, equating to less than 15 minutes a week. When we looked at student time on task in terms of quartiles, for two year courses, 25% of the total in the 1st lowest Quartile spent an average of 7 hours or less, and 25% in the 4th highest Quartile spent 16.3 hours or more. For one year courses, the equivalent hours were about 5.5. and 7.7.

However, the results from the MEL data analysis outlined above, need to be treated with caution. This is because the classes were randomly selected for the analysis and dated back to 2014. As such, the teachers we interviewed might be different to those who were teaching the classes for which student MEL data were analysed. Furthermore, it is likely that variation in implementation has influenced the results. However, although the teachers who taught the sampled classes might not be the same, the variation in assignments and practice completions seem to agree with the variation in the implementation of MEL and of its usage by students and teachers recorded through the feedback students offered in the focus groups and teachers in the interviews.

Finding 6: Pearson's training on MEL concentrates on technical aspects and is considered sufficient and useful, given the platform's intuitive nature, teachers suggest. Top-up training focusing on how MEL can best support pedagogy is welcomed, as is additional technical training, especially for teachers who are less confident with the use of technology.

Training by Pearson concentrates on the technical aspects of how to access and control several MEL functions, such as how to log in, assign assignments and view the gradebook, according to teacher interviews. Given its intuitive nature, in-depth training on MEL's technical aspects is not a requirement

for most teachers, interviewees suggest. However, for teachers whose confidence with technology is not as high, more elaborate top-up training might be useful. Teachers also welcomed top-up training on how MEL could further support teaching in a pedagogical way.

In the institutions visited, the cascade model of training was employed. Training on MEL was usually provided by a Pearson representative to coordinators. In turn, coordinators trained other teachers, or, in some cases, teachers trained their colleagues. One teacher, however, reported to be self-trained as they found MEL intuitive to use.

Student performance on MEL

Finding 7: There are indications that teachers can use students' scores from MEL assignments and tests confidently for formative assessment purposes. Further investigation however would be helpful.

Given that only 20% of practices were attempted more than two times, one could assume that, in general, most students did not necessarily look up the answers after the third attempt.[12]

Furthermore, practices and tests seem to be relatively reliable indicators of students' performance. So, a students' performance in one practice could be used as an indication of their future performance in another practice. For each student, the average practice score was computed for all the exercises of each unit. The average practice scores for each student and for each unit were then correlated between them, in the same way one might correlate the item scores of students in a test. In most cases, the correlations ranged from moderate to high (average correlation of 0.45 for AWF, 0.53 for Wyższa Szkoła Logistyki and 0.41 for WS. Most of the correlations were statistically significant, at the $p = 0.05$ level). As regards tests, the correlations between the test scores, across students, mainly ranged between 0.38 (1st Quartile) and 0.57 (3rd Quartile), with a mean of 0.46. Seventy-six out of 81 correlations, around 94%, were statistically significant at the $p = 0.05$ level.

Lastly, practice scores could indicate a student's performance in MEL tests. We have strong indications of convergence validity when practice and test scores are considered, meaning that practices and tests tend to yield consistent information about a students' ability. There was a strong association between the average practice score and the average test score. The correlations were moderate to high (0.40 to 0.69) and statistically significant in all cases at the $p < 0.001$ level.

Finding 8: The more assignments and practices students complete, the better their scores, data suggests.

[12] Students have the opportunity to see the answer to a question after their third attempt.

There is evidence that students should be encouraged to complete MEL assignments/practices for learning purposes. In nine out of 11 cohort cases, a high correlation was found between the percentage of practices and assignments completed and the average score. Coefficients ranged between 0.44 to 0.83 with $p < 0.001$ for eight of the correlations and $p=0.01$ for the ninth; two of the eleven correlations had non-statistically significant results. (Completions did not include repeated attempts on the same practice/assignments.)

Perceived impact of Speakout with MEL on students

Finding 9: Overall, Speakout with MEL is accessible, data indicates. Above 82% (163/199) of students in most of the access related questions find the Students' Book and MEL accessible and easy to navigate. Student access to ActiveTeach is limited due to the lack of infrastructure and because ActiveTeach is currently only accessible by CD-ROM.

All data collected suggests that Speakout with MEL is accessible to students. Between 82% (163/199) and 99% (198/199) of students reported that the Students' Book was up to date, interesting, relevant to real life, culturally appropriate and at the appropriate level. Students and teachers also pointed to the book's ease of navigation and flexibility to be adapted based on need. During the observations, minimal guidance was given on how to use the materials, and no navigation issues were raised.

Data suggest that MEL is also accessible. The large majority of students were able to access MEL on their computer (95%, 190/199) and smartphone (81%, 157/193), and access their assignments easily (94%, 188/200). They also found MEL easy to navigate (95%, 188/198). Student and teacher data from the focus groups and interviews support these findings and also point to MEL being perceived to be intuitive and relatively trouble-free. Issues related to access referred to wrong codes distributed, internet issues and other institutional-level problems.

ActiveTeach however, was deployed by only one teacher, who used her/his computer to project content. Although interviewees acknowledged the potential benefits of ActiveTeach, they mainly noted the lack of institutional infrastructure, e.g., the lack of availability of whiteboards in classrooms, as a key barrier to its use. Its usability was also an issue for a few teachers interviewed. One teacher in particular, pointed to how time-consuming it was to download ActiveTeach from the CD-ROM and for that reason she/he did not use it.

Other issues identified that related to Speakout and MEL's accessibility were:

- It was hard to cover the book's content fully. In two institutions, teachers adapted the Speakout content so that they covered the necessary, especially with extramural students. Choosing units, however, was a challenge — so much so that one institution decided to use Speakout with MEL only in its one long course (of 120 hours) and not in its shorter courses (of 60 and 80 hours).

- Speakout with MEL would benefit from more challenging grammar content and the inclusion of more academic, as well as technical, language.

- British humour was reported to be hard to understand, e.g., comedies such as Blackadder, which were also seen as 'dated'.

Finding 10: For the majority of students, their confidence, enjoyment and motivation to learn English has increased since using Speakout with MEL, as well as their confidence across all the English language skills, data suggests. The use of MEL further encourages student motivation as well as nurtures students' independent and self-monitoring skills, according to teachers and students. Attention on how Speakout can further support students' confidence in speaking and writing is required.

The majority of students reported that their confidence in learning English (74%, 142/192), reading (76%, 145/192) and listening (75%, 143/191) had increased since using Speakout with MEL. For writing, 59% (113/191) of students reported that their confidence had increased since using Speakout with MEL and for speaking, 55% (106/192). However, the overall percentage for speaking was skewed due to one institution's results. 84% (31/37) of Wyższa Szkoła Logistyki students agreed or strongly agreed that their confidence in speaking had increased, but the equivalent for AWF was 49% (71/146).

As regards motivation and enjoyment, according to 70% (135/193) of students, Speakout with MEL significantly or very significantly improved their motivation to learn English. In particular, interviewees pointed to the video clips and their diverse topics, which, according to teachers, also broadened students' cultural horizons. Additionally, 66% (128/194) of students reported that Speakout with MEL significantly or very significantly helped them enjoy learning English.

Data suggests that the interesting stories and variety of spoken exercises (e.g., the dialogues, discussions, role plays and the activities where students plan and produce a topic for discussion themselves) in the Students' Book are major contributors to the encouragement of positive learning behaviours. One coordinator noted that they liked the revision exercises at the end of each unit in Speakout, and that almost all exercises included an option to practice speaking in pairs or groups.

Student motivation was further enhanced through the use of MEL, as well as skills such as self-monitoring. Particularly, students who tracked their own progress reported that it motivated them to continually improve. Giving students the responsibility to complete MEL practices nurtured skills such as autonomy, independence and taking control of one's learning, teachers suggested.

Finding 11: Speakout with MEL is engaging for the majority of students, especially the Students' Book topics, which encourage discussion. However, data indicates that a sizeable group of students needs further convincing as regards MEL. This is both because of perceived issues with the product itself, as well as the way it is implemented.

The large majority of students (77%, 149/194) reported that Speakout with MEL helped them engage with learning English. Overall, teachers suggested that Speakout with MEL met the needs of students of all ages well, although younger audiences were not necessarily attracted to the same topics as more mature audiences. Student engagement was also evident during the observations. The majority of students across the three institutions were relatively engaged in the tasks, contributing their thoughts and asking questions. In the speaking and listening session observed in one institution in particular, all students were speaking for the majority of the session, laughing and giggling whilst discussing one of the book's topics. The students' NPS score, however, was -11.9, suggesting that they are less likely to recommend Speakout with MEL to others. Although students have pointed to Speakout with MEL's misgivings, overall, this finding does not seem to match well with the findings from across the sources used in this study. Further investigation is required to fully understand whether the negative NPS score is a reflection of the issues related to Speakout with MEL, or if this is an issue with the interpretation of the scale. This investigation could assess the suitability of NPS as a measure to gather students' views, given that they are not the 'direct customer'.

Although the majority of students suggested that MEL was engaging (64%, 124/195) and that they enjoyed learning by completing assignments on it (57%, 113/198), a relatively sizeable 36% (71/196) and 43% (85/198) disagreed. Questionnaire results also showed that the majority of students viewed a number of MEL features positively: 78% (153/196) suggested that seeing the completion dates for assignments was useful or very useful, and between 89% (175/197) and 96% (190/198) of students thought that there was a good variety of practices and tests in MEL. Furthermore, the majority thought that checking answers immediately, accessing MEL at any time and place, and seeing a summary of their score and progress were useful or very useful to their learning. Lastly, 57% (110/193) of students reported that communicating with their teachers online was useful, although this finding might reflect more the under-usage of the function, rather than the usefulness attributed to it by students. It might also be the case that full-time day students, who saw their tutors more often, did not see this function as being as useful as extramural students did. Lastly, students' NPS score for MEL was negative at −50.4. Even though data suggests that a sizeable group of learners need convincing about MEL, this is a very low NPS score. Data across sources does not seem to agree with the level of dissatisfaction suggested by the NPS score. Thus, further investigation is required, including examining the validity and reliability of the scale used for the purposes of this type of research.

Data from the focus groups, however, shows that students were relatively divided in their views of MEL, which could have an impact on engagement. Students who viewed MEL as modern and helpful pointed to a number of aspects that brought value to their learning and were relatively keen to use it. Reported criticisms of MEL included that:

- students saw MEL repetitive over time
- the instructions to the open-ended questions were not specific enough regarding the length of the answer required
- the exercises were not challenging enough
- the large number of assignments given to them at once was overwhelming, which was more related to the teaching approach rather than MEL itself.

Some students in the focus groups offered different views on the purpose of MEL to their learning and used it for different purposes. For example, a few students made a direct comparison between MEL and the Students' Book, which might suggest that the different purposes of the two components were not necessarily well understood. Also, some students reported that they only used practices on MEL, while others used it as a revision or test-preparation tool.

Finding 12: Overall, students believe that Speakout with MEL supports them to improve their English, as well as their skills in speaking, listening, vocabulary, grammar and writing. MEL positively contributes to the development of different English-language skills too, although some exercises and the auto-scoring system need to be reviewed, according to the data.

The majority of students indicated that Speakout with MEL had significantly or very significantly supported them in improving their English (61%, 115/189), as well as helped them in improving their skills in speaking, listening, vocabulary, grammar and writing (between 65%, 126/194 and 87%, 169/194 of students). Although 65% (126/194) of students reported that Speakout with MEL has supported their speaking skills significantly or very significantly, responses from one institution relatively skewed the overall percentage downwards — 60% (87/146) of students from AWF reported that Speakout with MEL helped them to improve their speaking significantly or very significantly, compared to 87% (34/39) from Wyższa Szkoła Logistyki.

Teachers found that, since other skills were practiced at home through MEL, more time was freed up to practice speaking in class. Furthermore, the majority of students found MEL exercises and the feedback it provided to be supportive of learning:

- Between 84% (167/198) and 96% (190/198) of students agreed or strongly agreed that feedback from MEL helped students understand how to improve their English. Students also found the immediacy of the feedback they received from MEL; the grammar and vocabulary exercises; and repeating exercises and seeing their scores as a summary, helpful or very helpful to their learning.

- Between 71% (140/198) and 78% (155/198) of students agreed or strongly agreed that MEL helped students understand content learned in class; and found the writing exercises helpful or very helpful. Teachers asked for more support with writing.

- 64% (124/195) of students found the pronunciation coach videos useful. A sizeable 36% (71/195), however, disagreed. Teachers asked for additional exercises in both vocabulary and pronunciation.

MEL's auto-scoring system was a point of frustration for both students and teachers. Students suggested that they got penalised for what they saw as minor mistakes in punctuation or spelling, such as non-conformity to capitalisation rules. Teachers agreed and pointed to the increase in queries

they had to deal with, as well as, for a few teachers, the increase in their assessment workload as they felt obliged to alter student scores.

Furthermore, some MEL exercises were seen as not challenging enough, especially those that required one-word answers, or those that students saw as having obvious answers (e.g., choosing between 'I am' and 'I'm'). Reported issues with MEL activities not being challenging enough match student and teacher feedback on the difficulty level of Speakout's grammar, which were also noted.

As regards students' performance on MEL practices and assignments, this is relatively high; the mean scores range between 84% to 97% for different institutions and levels. Findings need to be treated with caution, however, as they involve a sample of classes for which we have no information regarding the identity of the teachers who taught these classes or of the implementation approach followed. It would be useful to have further data on student performance. This is because, amongst other reasons, MEL assignments/practices comprise one piece of a wider assessment system; students also take other formative and summative assessments, but we only had access to their MEL scores. It is important that we cross reference student scores from MEL with more standardised assessments.

Finding 13: Speakout with MEL prepares students well for the next stage of their learning and to achieve their goal, data suggests. Progression in grammar difficulty, especially from intermediate to upper-intermediate, however, needs reviewing.

84% (161/192) of students reported that Speakout with MEL prepared them well for the next level of their English studies. 79% (152/192) of students suggested that it helped them achieve their goal, with 78% (151/194) saying that it did so significantly or very significantly.
However, according to interviews, grammar in particular posed issues to progression. Grammar content at Intermediate level was seen generally to be relatively 'easy', and grammar difficulty from Intermediate to Upper-intermediate a 'jump'. One teacher noted that she/he used complementary materials to compensate for the difference in difficulty. The B1+ level added as part of the Speakout materials will respond to this need.

Pearson

Perceived impact of Speakout with MEL on teachers

Finding 14: Data suggest that the perceived impact of Speakout with MEL on teachers was positive. They suggested that the Students' Book was accessible, engaging and helpful in supporting teaching, especially the teaching of speaking. They also reported that the Teacher's Book supported lesson planning and that MEL reduced their assessment workload and freed up time for more speaking activities and 'fun' teaching methods to be introduced. Monitoring individual student performance in MEL however, was seen as time-consuming.

Teachers reported that the Students' Book was accessible, engaging and helpful in supporting the teaching of speaking. Teachers referred to the ease of navigation and logical structure of the Students' Book, and found its topics interesting to teach. They also praised the book's focus on speaking. They thought that it offered an array of opportunities for spoken interaction and prompted teachers to monitor and assess speaking more regularly.

The Teacher's Book was used by all teachers in the study. Overall, it was seen to support lesson planning and the implementation of a variety of teaching approaches. Deviations from the Teacher's Book aimed to accommodate group and individual students' needs, as well as to include necessary content, such as academic and technical English.

As for MEL, teachers thought it was accessible, and that it:

- freed up classroom time to develop speaking and deliver other necessary content — for some teachers, freeing time also meant including a variety of teaching approaches, including those considered more fun, such as games
- reduced assessment workload and teacher feedback — despite the increase in queries due to the issues with the auto-scoring system, overall teachers thought that by using MEL they spent less time marking homework and providing feedback
- reduced administrative workload — the time needed to print out homework exercises was reduced
- supported monitoring and assessment — teachers could quickly see an overview of class and student scores. However, when they needed to drill down to an individual student's data, e.g., to see a student's score in a particular assignment, this was seen as time consuming.

Recommendations

Recommendations deriving from this report will be provided to the relevant Pearson teams and to the institutions directly.

Generalisability of findings, limitations and future research

The section below provides the key limitations of the study and outlines potential areas for future research:

- Findings are based on specific implementation models carried out with specific samples of students and teachers in specific cultural context and settings. Further research should aim to replicate findings in similar contexts and with a similar sample, to identify whether their results are consistent with the results of this study. Further research should also aim to expand our understanding of the repertoire of implementation models, so that using Speakout is better understood in different contexts, settings and countries. A better representation of students and teachers from all the institutions to be studied should be also considered.

- In the students' questionnaire, students' views from AWF and Wyższa Szkoła Logistyki were relatively representative of the views of students studying within the same institutions and who were using Speakout with MEL (approximately 85% representation, 153/180 students approximately) and 65% representation (39/ 60 students approximately). However, no permission was given to survey the students from WS. Thus, respondents' views for the questionnaire represent the views of students in two out of the three institutions that took part in the research.

- Findings are based on triangulating inferences across different evidence sources. The aim is to allow Pearson to use the data to screen major occurrences and major trends across institutions for decision making and to develop an understanding of the different implementation models, not to provide precise answers (Ewell, 2009; McCormick & McClenney, 2012; Pike, 2013). Furthermore, self-report methods are known to be vulnerable to both unreliability and bias. Self-reported perceptions of impact on achievement and progression do not provide objective evidence of impact. This is more of a limitation for evidencing student achievement and progression than for student access and experience, where self-reported perceptions are extremely valuable. Further research should seek to incorporate objective external measures of achievement and progression, to compare outcomes for users and non-users and to control for potentially confounding factors, such as prior achievement. Additionally, individual student level analysis will allow us to understand the relationship between student usage and performance. Moreover, we were unable to control for prior achievement for all samples in the study. The addition of an indicator for socioeconomic status as a covariate would also strengthen the study.

- At times, data from the focus groups and from the interviews are presented from the students' perspective only, or from the teachers' perspective only. At times these findings are corroborated by the students' questionnaires. When findings from the student focus

groups and from the teacher interviews are not corroborated with each other, or with the student questionnaire data, they should be treated with caution and become items for follow-up questioning in future research.

- Due to time constraints, analysis that allowed the 'bundling' of answers referring to the same category, i.e. usefulness of MEL or student confidence, to derive a standardised index was not performed. This type of analysis will provide us with a more reliable estimate of the construct and will also allow for comparisons across different institutions and countries. Future research could complete further analysis by looking at different variables in the student and teacher questionnaires, i.e. student motivation to learn or teacher confidence in teaching English and usage. Lastly, responses to open ended questions in the student and teacher questionnaires were also not analysed. Analysing these would further enrich our understanding of the themes explored in this study.

- Between and within classroom variation in terms of the usage and performance of MEL should be further investigated.

- Students' NPS scores with regards to MEL contradicts results from the data collected across the different research instruments. Further research should aim to understand this difference, including to qualitatively investigate the reliability of the scale used for this type of research.

- A more rigorous design would compare the performance of students using Speakout with MEL to students not using Speakout with MEL. Students would either be randomly assigned to treatment condition or would be matched to students in the other group on important background characteristics, such as prior achievement and demographic factors.

- Future research should aim to either control for implementation or understand implementation first, before analysing MEL data.

Pearson

References

Bond, T., & Fox, C. (2001). Applying the Rasch Model: Fundamentals Measurement in the Human Sciences, Mahwah, NJ: Lawrence Erlbaum Associates.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology, Qualitative Research in Psychology, 3 (2). pp. 77-101. http://www.informaworld.com/smpp/content~db=all~content=a795127197~frm=titlelink.

Cohen, J. (1988). Statistical Power Analysis for the Behavioural Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dorneyei, Z., with Taguchi, T. (2010). Questionnaires in Second Language: Construction, Administration and Processing, London and New York: Routledge.

Ewell, P.T. (2009). Assessment, Accountability, and Improvement: Revisiting the Tension. National Institute for Learning Outcomes Assessment, Occasional Paper #1. http://www.learningoutcomeassessment.org/documents/PeterEwell_005.pdf

Hopkins, D., & Ahtaridou, E. (2009) Applying research methods to professional practice. In S. Lapan & M. T. Quartaroli (Eds.), Research Essentials: An Introduction to Designs and Practices, San Francisco, CA: Jossey-Bass.

Lamprianou, I., & Boyle, B. (2004). Accuracy of measurement in the context of mathematics National Curriculum tests in England for ethnic minority pupils and pupils who speak English as an additional language. Journal of Educational Measurement, 41(3): 239–60.

Lüdecke D (2017). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.4.0, <URL: https://CRAN.R-project.org/package=sjPlot>

McCormick. A.C. and McClenney, K. (2012) Will These Trees Ever Bear Fruit?: A Response to the Special Issue on Student Engagement. The Review of Higher Education. Volume 35, Number 2, Winter 2012: 307-333.

Pearson (2014a) Efficacy Results: Studies illustrating the efficacy, impact and motivation of MyEnglishLabs for adult students of English. http://myenglishlab.com/pdf/Efficacy%20Report_Adult%20final_screen%20online%20version.pdf

Pearson (2014b) Efficacy Results: Studies illustrating the efficacy, impact and motivation of MyEnglishLab for Secondary school students of English. https://support.pearsonelt.com/content/dam/professional/english/knowledge-help/KB_Images/Efficacy%20Report%20Secondary_final_online.pdf

Pike, G.R. (2013). NSSE benchmarks and institutional outcomes: A note on the importance of considering the intended uses of a measure in validity studies. Research in Higher Education. March 2013, Volume 54, Issue 2, pp 149–17.

Quartaroli, M. T. (2009) Qualitative data analysis. In S. Lapan & M. T. Quartaroli (Eds.), Research Essentials: An Introduction to Designs and Practices, San Francisco, CA: Jossey-Bass.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Revelle, W. (2017) Psych: Procedures for Personality and Psychological Research, Northwestern University.

Revelle, W. (2017) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Versio 1.7.8.

Robitzsch, A., Kiefer, T., & Wu, M. (2017). TAM: Test analysis modules. R package version 2.8-21. https://CRAN.R-project.org/package=TAM

Vasbieva, D. G., & Klimova, I. I. (2015) Transformational potential of blended learning to personalize foreign language teaching in a non-linguistic higher school. XLinguae, 8 (1), pp 2-10.

Vymetalkova, D. (2016) Blended Learning and MyEnglishLab Application as a Powerful Tool Leading Towards Autonomy in the Process of Learning English Language Targeting and Activating Both Brain Hemispheres. New Trends and Issues Proceedings on Humanities and Social Sciences, 10, pp 75-81.

Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse

Wickham, H. and Bryan, J. (2017). readxl: Read Excel Files. R package version 1.0.0. https://CRAN.R-project.org/package=readxl

Appendix A: Consolidated Framework of Implementation Research (CFIR)

Table A1: Consolidated Framework of Implementation Research (CFIR)

Construct		Short Description
I. Intervention characteristics		
A	Intervention Source	Perception of key stakeholders about whether the intervention is externally or internally developed.
B	Evidence Strength and Quality	Stakeholders' perceptions of the quality and validity of evidence, supporting the belief that the intervention will have desired outcomes.
C	Relative Advantage	Stakeholders' perception of the advantage of implementing the intervention versus an alternative solution.
D	Adaptability	The degree to which an intervention can be adapted, tailored, refined or reinvented to meet local needs.
E	Trialability	The ability to test the intervention on a small scale in the organisation and to be able to reverse course (undo implementation) if warranted.
F	Complexity	Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality and intricacy, and number of steps required to implement.
G	Design Quality and Packaging	Perceived excellence in how the intervention is bundled, presented and assembled.
H	Cost	Costs of the intervention and costs associated with implementing the intervention, including investment, supply and opportunity costs.
II. Outer setting		
A	Patient needs and resources	The extent to which patient needs, as well as barriers and facilitators to meet those needs, are accurately known and prioritised by the organisation.
B	Cosmopolitanism	The degree to which an organisation is networked with other external organisations.
C	Peer pressure	Mimetic or competitive pressure to implement an intervention, typically because most, or other, key peer or competing organisations have already implemented, or are in a bid for, a competitive edge.
D	External policy and incentives	A broad construct that includes external strategies to spread interventions, including policy and regulations (governmental or other central entity), external mandates, recommendations and guidelines, pay-for-performance, collaboratives, and public or benchmark reporting.
III. Inner setting		
A	Structural characteristics	The social architecture, age, maturity and size of an organisation.
B	Networks and communications	The nature and quality of webs of social networks and the nature and quality of formal and informal communications within an organisation.
C	Culture	Norms, values and basic assumptions of a given organisation.
D	Implementation climate	The absorptive capacity for change, shared receptivity of involved individuals to an intervention and the extent to which use of that intervention will be rewarded, supported and expected within their organisation.

Construct		Short Description
1	Tension for change	The degree to which stakeholders perceive the current situation as intolerable or needing change.
2	Compatibility	The degree of tangible fit between meaning and values attached to the intervention by involved individuals, how those align with individuals' own norms, values and perceived risks and needs, and how the intervention fits with existing workflows and systems.
3	Relative priority	Individuals' shared perception of the importance of the implementation within the organisation.
4	Organisational incentives and rewards	Extrinsic incentives, such as goal-sharing awards, performance reviews, promotions and raises in salary, and less tangible incentives, such as increased stature or respect.
5	Goals and feedback	The degree to which goals are clearly communicated, acted upon and fed back to staff, and alignment of that feedback with goals.
6	Learning climate	A climate in which (a) leaders express their own fallibility and need for team members' assistance and input; (b) team members feel that they are essential, valued and knowledgeable partners in the change process; (c) individuals feel psychologically safe to try new methods; and (d) there is sufficient time and space for reflective thinking and evaluation.
E	Readiness for implementation	Tangible and immediate indicators of organisational commitment to its decision to implement an intervention.
1	Leadership engagement	Commitment, involvement and accountability of leaders and managers with the implementation.
2	Available resources	The level of resources dedicated for implementation and ongoing operations, including money, training, education, physical space and time.
3	Access to knowledge and information	Ease of access to digestible information and knowledge about the intervention and how to incorporate it into work tasks.
IV. Characteristics of individuals		
A	Knowledge and beliefs about the intervention	Individuals' attitudes toward and value placed on the intervention, as well as familiarity with facts, truths and principles related to the intervention.
B	Self-efficacy	Individual belief in their own capabilities to execute courses of action to achieve implementation goals.
C	Individual stage of change	Characterisation of the phase an individual is in as they progress toward skilled, enthusiastic and sustained use of the intervention.
D	Individual identification with organisation	A broad construct related to how individuals perceive the organisation and their relationship and degree of commitment with that organisation.
E	Other personal attributes	A broad construct to include other personal traits such as tolerance of ambiguity, intellectual ability, motivation, values, competence, capacity and learning style.
V. Process		
A	Planning	The degree to which a scheme or method of behaviour and tasks for implementing an intervention are developed in advance, and the quality of those schemes or methods.
B	Engaging	Attracting and involving appropriate individuals in the implementation and use of the intervention through a combined strategy of social marketing, education, role modeling, training and other similar activities.
1	Opinion leaders	Individuals in an organisation who have formal or informal influence on the attitudes and beliefs of their colleagues with respect to implementing the

Construct		Short Description
		intervention.
2	Formally appointed internal implementation leaders	Individuals from within the organisation who have been formally appointed with responsibility for implementing an intervention as co-ordinator, project manager, team leader or other similar role.
3	Champions	Individuals who dedicate themselves to supporting, marketing, overcoming indifference or resistance that the intervention may provoke in an organisation.
4	External change agents	Individuals who are affiliated with an outside entity, who formally influence or facilitate intervention decisions in a desirable direction.
C	Executing	Carrying out or accomplishing the implementation according to plan.
D	Reflecting and evaluating	Quantitative and qualitative feedback about the progress and quality of implementation accompanied with regular personal and team debriefing about progress and experience.

Pearson

Appendix B: Results

This section provides the results from the analysis of the data collected for this study, including from the student questionnaires, the coordinators' pre-questionnaires, student focus groups, teacher interviews, classroom observations and MEL student data.
Out of the total 201 student questionnaire respondents, results represent the majority of students who were studying in AWF (153). Thirty-nine of the respondents were studying in Wyższa Szkoła Logistyki and nine at Uniwersytet Centralny. No permission was given to survey the students from WS, and results would not be different if data collected from Uniwersytet Centralny were excluded from the analysis.

Students' views from AWF and Wyższa Szkoła Logistyki were relatively representative of the views of all the students studying in their institution using Speakout with MEL (approximately 85% representation, 153/180 students) and 65% representation (39/60 students). Thus, respondents' views from the questionnaire represent the views of students in two out of the three institutions that took part in the research.

In terms of its characteristics, the majority of the student sample in the questionnaire was relatively homogeneous. All students who responded to the questionnaire from AWF and Wyższa Szkoła Logistyki were studying at the Intermediate level, representing 95% (186/195) of the total sample. Furthermore, most students were studying in courses lasting 120 hours in total (90%, 180/201) over two years of study and were mid-way through their courses, meaning they were at the end of their first academic year of study (56%, 106/189).

The 29 students who took part in the focus groups were all young adults, between 18 and 25 years old, studying full time at Intermediate or Advanced levels in courses lasting 120 hours in total, over the course of two years. Students who took part in the focus groups may or may not have filled in a questionnaire.

From the 11 teachers interviewed, 100% representation was achieved in terms of the number of teachers using Speakout with MEL at AWF and Wyższa Szkoła Logistyki (3/3 and 2/2 teachers, respectively). For WS, representation was 20% (2/10 teachers). Two teachers from Wyższa Szkoła Logistyki, who were teaching using MEL with a different Pearson book at the time of the visit, were also interviewed and offered their views on MEL only. Lingua Nova's director and teacher training coordinator were also interviewed, given that their perspective on the implementation of Speakout with MEL at WS was imperative.

Findings are presented thematically, and refer to:

- students' motivation to learn English
- students' attitudes towards Speakout with MEL

- teachers' attitudes towards Speakout with MEL
- implementation of Speakout with MEL, including the reasons for engaging with Speakout with MEL; institutions' readiness for implementation; the teaching and learning, and assessment approach implemented, especially relating to MEL and the training and support for teachers and students
- student usage of MEL
- findings for teaching and learning from the MEL data
- student and teacher perceptions of the impact of Speakout with MEL on students, teachers, teaching and the institution

Students motivation to learn English

Overall, the level of student motivation to learn English among students across the three institutions in Poland was moderate; about a third of students reported low levels of motivation, and a quarter of students reported high levels of motivation.
Questionnaire participants were asked to respond to five questions related to motivation.

Most of the respondents selected the options 'important' or 'very important' to the following statements:

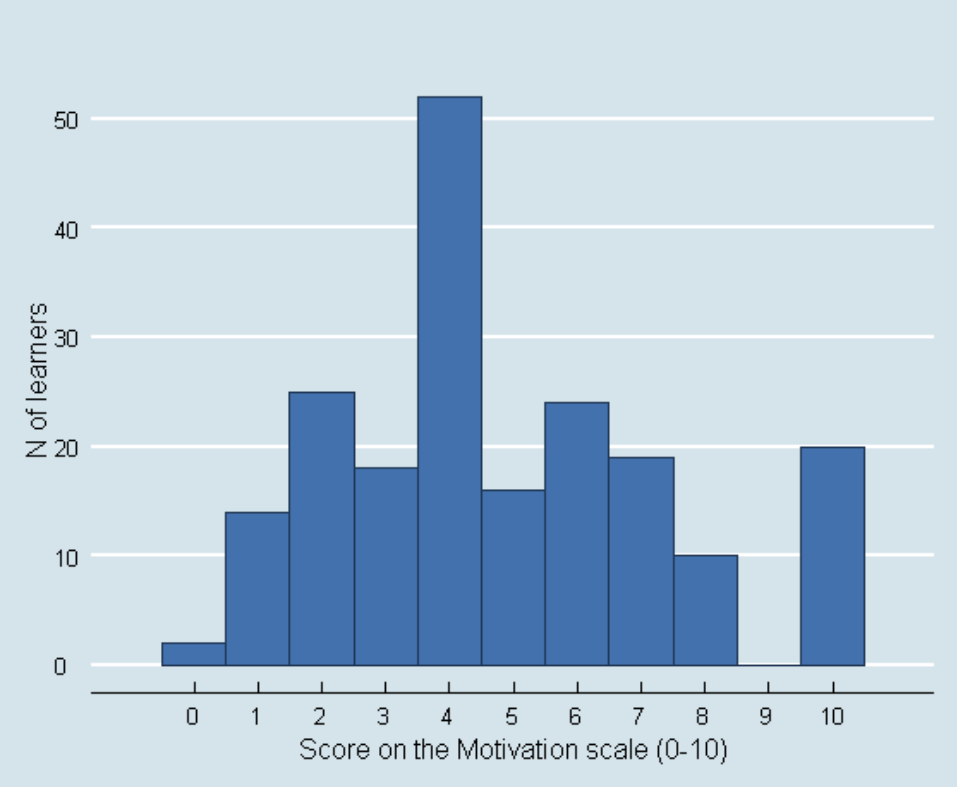
- 91% (181/199) said that learning English was important/very important
- 86% (171/200) said that achieving their goal by learning English was important/very important
- 86% (172/200) said that using English in their life, work and study was important/very important
- 77% (152/197) said that succeeding in learning English was important/very important
- 77% (153/200) said that successfully completing a level was important/very important

To derive an overall measure of motivation for each student, we used a scale from 0 to 10.[13] Students were then divided into three groups based on their motivation score:

- low motivation — 30% of respondents (59/200)
- moderate motivation — 45% of respondents (91/200)
- high motivation — 25% of students (50/200) [14]

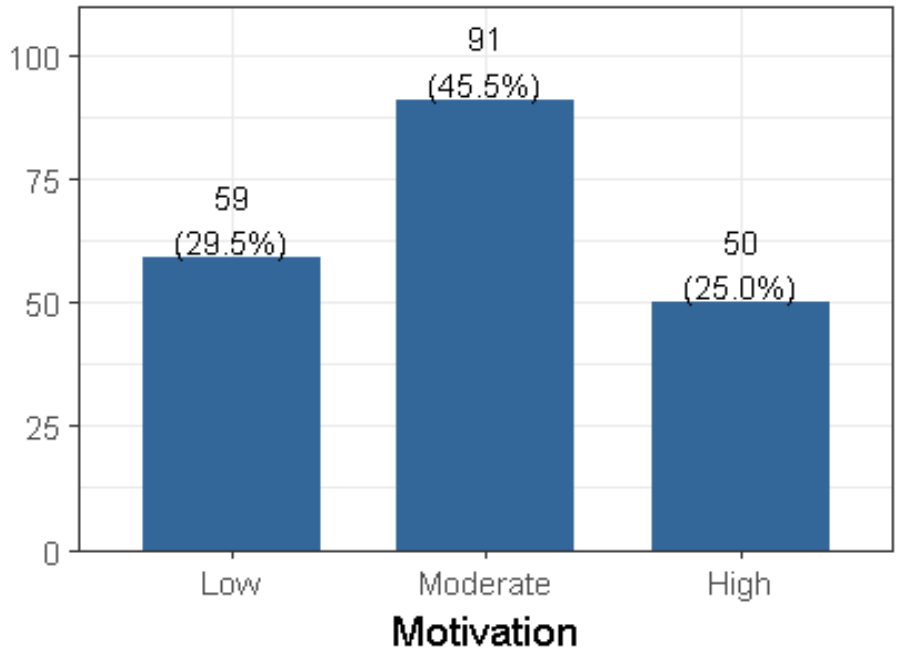
[13] The five questions related to motivation were Rasch-analysed in order to produce one overall measure of motivation for each student. The motivation measures that derived the Rasch analysis were transformed linearly to a scale from 0 to 10 to facilitate comprehension by less technically oriented readers. On this scale, 0 means less motivation and 10 means more motivation. The students achieved a mean motivation score of 4.6, with a standard deviation of 2.7.
[14] Students with a motivation score less than 4.0 were classified as the 'low motivation' group. This group includes all the students who failed to at least respond 'important' to all five questions related to motivation (e.g., a student in the low motivation group may have responded 'important' to four questions and 'somewhat important' to one question). Students with a motivation score between 8 to 10 were classified as the 'high motivation' group. This corresponds to a student responding 'important' to all five

Figure B1: Student combined motivation scores — student questionnaire



questions, or to more positive responses. Those students with a motivation score between 4 and 7 were grouped in the 'moderate motivation' group.

Figure B2: Distribution of student motivation — student questionnaire

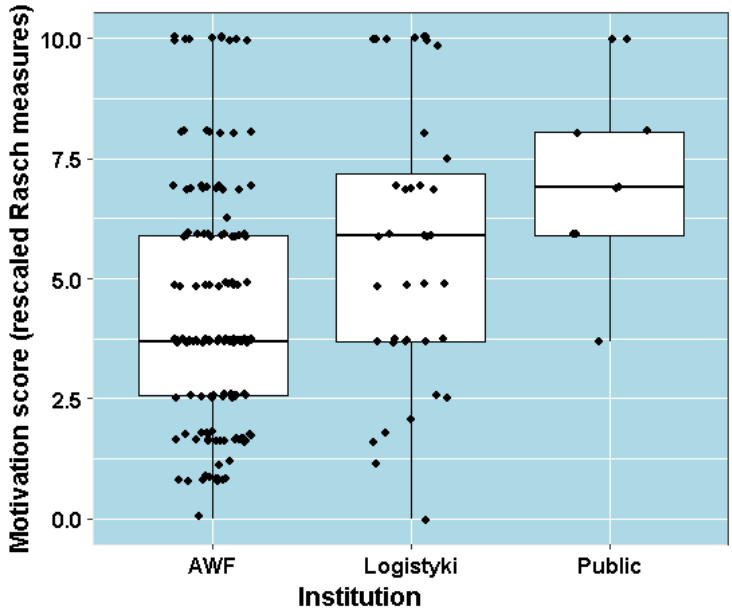


There were large and statistically significant differences ($\chi 2[4] = 17.144$, $p = 0.003$) between the percentage of students in each institution who were classified in each of the three motivation groups:

- AWF — 34% of students (52) were in the low motivation group; 47% (71) were in the moderate motivation group; and 19% (29) were in the high motivation group (mean is 5.1[SD = 2.2]).
- Wyższa Szkoła Logistyki — 18% (7) were in the low motivation group; 44% (17) were in the moderate motivation group; and 39% (15) were in the high motivation group (mean is 6.3[SD = 2.7]).
- Uniwersytet Centralny: we found no students in the low motivation group; 33% (3) were in the moderate motivation group; and 67% (6) were in the high motivation group (mean is 7.8 [SD = 1.7]).

The distribution of motivation scores per institution are provided in Figure B3 below.

Figure B3: Distribution of motivation scores per institution—student questionnaire



Uniwersytet Centralny

According to teachers, the biggest challenge they were faced with was student motivation to learn English, which they thought was relatively low. One teacher suggested this may be due to students' ages, while another pointed to the obligatory nature of the English courses for these students, as well as their personal circumstances.

A further challenge identified was what teachers saw as limited available classroom time- to learn a language, a result of university funding cuts. For extramural students in particular, limited contact time, coupled with absenteeism, was seen as a significant problem. Teachers suggested that sometimes students ended up attending one session per month due to personal and work commitments.

Teaching mixed-ability groups was an additional challenge; teachers had to accommodate students at beginner and advanced levels in the same classroom. Some teachers suggested that such extremes in the level of ability within classrooms was also a result of students intentionally scoring lower in their diagnostic tests.

Overall, focus groups and interview data indicated that students preferred and enjoyed Speakout with MEL more than other instructional approaches they had experienced previously, as well as learning using MEL instead of the print Workbook. When asked in interviews if they would continue to use MEL even if it was not compulsory, most suggested that they would. Teachers also pointed to students being familiar with technology in their everyday lives and that using it to learn feels natural to them. Additionally, student questionnaire data suggested that:

- 78% (154/198) of students agreed or strongly agreed that they liked the combination of teacher instruction and independent MEL practice
- 68% (134/197) of students agreed or strongly agreed that they would rather complete exercises in MEL than in the print Workbook
- 67% (127/189) of students appeared to enjoy their classes using Speakout with MEL more than other classes without Speakout with MEL
- 61% (117/192) of students reported being more engaged in classes using Speakout with MEL than other classes without Speakout with MEL
- 58% (111/190) of students indicated that they were motivated by classes using Speakout with MEL more than other classes without the package

Teacher attitudes towards Speakout with MEL

The large majority of students in the questionnaire (79%, 155/197) suggested that, overall, their teachers were in favour of MEL. However, 40% (78/197) disagreed or strongly disagreed that their teacher had made MEL an integral part of their learning.

Student focus group and teacher interview data support students' views recorded in the questionnaire that their teachers were overall in favour of using MEL. From the administrators and teachers interviewed, two out of the ten appeared less enthusiastic about the use of MEL. In particular, they seemed less enthusiastic about the use of technology in learning, or thought that students did not enjoy using it.

Indeed, some teachers found the use of MEL challenging, according to interviewees, due to: their views about technology and its use in learning and/or the time required to learn how to use MEL. However, there were teachers not as willing to implement MEL according to interviewees, because they felt that it would be a big effort to learn how to use it.

The variation in teachers' use, and especially their motivation to use MEL, had an impact on the whole department and its approach to teaching and learning. According to one teacher, imposing or encouraging standardisation across teaching staff is difficult if not all teachers have 'bought into' the use of technology and of MEL. Two other teachers, who were using MEL and who were enthusiastic about its use, agreed. They suggested that they would welcome standardisation in the use of MEL (and of assessment practices in general) across the department. They believed that standardising assessment practices and the use of MEL would encourage dialogue and the exchange of practices

among teachers. They suggested that standardisation of practices would also benefit students, as there would be one 'rule' across different classes. A more uniform approach, it was suggested, would benefit both teachers and students as sharing a common language and approach to using the platform would allow them to discuss and solve easier issues between themselves.

Implementation of Speakout with MEL

Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie Filia w Białej Podlaskiej (AWF)

AWF is a state university specialising in physical education, with majors in various disciplines, including sports and physiotherapy, as well as tourism and leisure. It offers bachelor's, master's and doctoral degrees.

About 180 students are currently taking Speakout with MEL courses, and a total of four teachers teach at the university's English-language department, two of whom use Speakout with MEL.

Speakout with MEL is used in one field only. Courses in that field last 120 hours in total, while other courses last from 60 to 80 hours. Speakout with MEL is not used in shorter courses because it is felt that there is no time to complete the whole course within that time frame, especially as teaching technical vocabulary can take much of the limited time available. Instead, a different half-book/half-workbook is used.

Reasons for selecting Speakout with MEL

AWF staff noticed that many students entered the university with low levels of spoken English compared to their other skills in English, e.g., writing or reading. Thus, Speakout's focus on speaking seemed the ideal solution to the problem.
According to interviews, the decision to use MEL alongside Speakout was the result of a combination of factors:

- The administrator was enthusiastic and eager to try new things, especially technology-related.
- MEL would engage students with learning because it is novel and because the use of technology comes naturally to them.
- There was a need to keep teaching 'alive' and interesting, and to improve teaching and learning by using new and innovative resources.
- Using MEL had the potential to raise the profile of English teachers within the university, as well as to give AWF English courses a competitive advantage.
- After trying MEL on a small scale, the coordinator was convinced of its learning value and its ability to reduce assessment workload.

Training in MEL

Training in MEL was provided to the coordinator by a Pearson representative at the time of purchase, more than three years before use. During the training, the coordinator was shown the basic features of MEL. In turn, the coordinator trained other teachers.

Teaching and learning

Students studying with Speakout with MEL attended one 90-minute class per week and usually covered one unit within three to five weeks. Every seven weeks, one week before their test, students attended a revision class. A typical lesson involved the following steps:

1. revision of learning, e.g., checking homework, correcting a written assignment, doing a short quiz or using a conversation or discussion point
2. a presentation (of varying length, depending on level and unit) using the Students' Book. Both students and teachers were active in explaining the learning points. This may also have been reading a text together and checking comprehension or discussing new vocabulary. Lesson time was also dedicated to professional language related to the students' field of study
3. students completed exercises related to the Students' Book with scaffolding and support from the teacher and their peers
4. homework was assigned

To support speaking and listening, paired activities were preferred, with students being paired randomly or as instructed by their teacher. Pair work was also used so that students could mark each other's written work. Group work was not used as often, as teachers suggested it allowed more opportunity for students to start speaking in Polish.

Assessment and MEL

Teachers had the flexibility to decide on the assessment approach that was most appropriate for their students. Despite this, teachers employed a relatively similar approach. Table B1 below provides a summary of AWF's approach to assessment.

Pearson

Table B1: AWF's approach to assessment, interviews

Type of assessment	Teacher 1	Teacher 2
Formal	• End-of-unit tests on paper and unannounced five-minute spot tests (85%–90% of the overall score) — on grammar and vocabulary to check previous material about four or five times a semester • Classroom participation (10% of the overall score) — included spoken competence. However, classroom participation only counted when a student was on the borderline between two scores. Teacher observations of a student's participation were used to make relevant judgements	• Two written tests per semester (90% of the overall score) • Classroom participation (10% of the overall score) — included students' spoken ability
	• Final end of two years exam — for bachelor students. 70% of the material in the final exam was general English and 30% professional vocabulary. On the masters, the split was roughly 50/50. Speaking was not part of the assessment. This may be due to time constraints, according to one teacher. Another thought that using a presentation in front of the class as an assessment would be stressful for students	
Informal[1] (MEL)	• MEL was used as homework and it was something that students 'must do', according to teachers. • MEL did not count towards students' overall score and no formal 'penalty' was enforced if students did not complete the assigned practices. Teachers sometimes spoke individually to students who did not complete practices, but, in general, they encouraged students to complete the MEL assignments by highlighting the benefits of doing so. • Students were usually assigned all unit assignments at once. • The time frame for completion of assignments was usually a semester (4 months), but students were encouraged to complete the exercises weekly. Teacher feedback suggested that practices are usually completed at the last minute. Using the MEL gradebook, one teacher told us that most students completed the bulk of their practices close to the final deadline, and very few used MEL on a weekly basis. The same teacher thought that incentivising students to complete tasks in shorter timescales would probably support more regular engagement. • Usually, students were allowed unlimited attempts at assignments. • Monitoring of student work in MEL happened periodically, with one teacher using the MEL gradebook data and examining the common errors report to inform revision classes. Teachers suggested that time constraints prohibited them from closely monitoring individual student performance.	

Note: (1) Different informal assessment took place at the university. However, given the focus of the research, in this section, we focus on MEL only.

Wyższa Szkoła Logistyki

Wyższa Szkoła Logistyki is a private university offering both Bachelors and Masters degrees. Bachelors degree students are a mixture of daytime students (of around 20 years of age) and extramural students (normally in their 20s to 40s), who work full-time during the week and study during weekends.

Some students pay their fees themselves while others' expenses are covered by employers. Approximately a total of 60 students use Speakout with MEL, of which 19 are B1+ level, and the remainder are enrolled in the A2/B1 level classes. Class sizes are typically 18 to 20 students. There is a total of six teachers, two of whom teach with Speakout with MEL. All teachers in the department, however, use MEL alongside other Pearson products.

Improving students' spoken ability in English is the focus of courses. Students' experience with the language is generally seen as more passive before they start their course, and is gained through reading and watching content on the Internet or on TV. Therefore, focus is placed on students expressing themselves at the level they need for their field of study. One teacher mentioned that by improving their speaking, students' understanding of the English culture could also be improved.

Reasons for selecting Speakout with MEL

MEL was seen to support the university-wide move to blended learning, a decision taken approximately six years ago. Other reasons included that it:

- reflects real-life use of English better than its predecessor scheme
- focuses on communication and so would better prepare students for the world of work, where they will be required to speak in English frequently
- would better serve the needs of native and international students, who access university, even at master's level, with low levels of English

Training in MEL

Teacher training on MEL was run about five or six years ago by the central Pearson Poland office. Overall, teachers felt that the training was sufficient, given the intuitive nature of MEL. Despite this, teachers said that they would welcome additional training as it would:

- help them use MEL more effectively in their instruction, if it focused on pedagogical aspects
- allow them to keep up to date with new features and functions
- help them understand how to encourage students to use MEL
- be particularly important for new teachers and for those not as comfortable with using technology

According to the focus groups, students' training on MEL was also adequate, given that the platform is largely intuitive. Teachers either booked the computer suite or asked students to bring their smartphones to class, and everyone practiced registering and logging in. This often took time. Teachers suggested that training students involved:

- about 1.5 hours on why MEL should be used and how it replaces the print workbook
- some time spent exploring the different components of MEL, navigating the platform and even doing some practice activities so students become familiar with it
- explaining which components are compulsory and which are optional
- allowing time for students to raise questions about the features and settings, which can be quickly resolved

Teaching and learning

Overall, teachers followed the Students' Book, with bachelor's students taking two semesters (one year) to complete one level and master's students taking one semester to complete one level. Both groups of students covered most of the Students' Book for their level. Extramural students met every two weeks for a total of ten sessions per semester, each lasting 90 minutes. It was estimated that within this time they managed to cover 50%–70% of the Students' Book.

Wyższa Szkoła Logistyki teachers suggested that speaking was a key focus. Grammar was also prioritised as it was perceived to be harder than vocabulary, and could be practiced independently on MEL.

One teacher said that, typically, a lesson involved:

- A student presentation lasting five to ten minutes (which students prepared in advance on a topic of their choice); or students would ask questions about previous material, e.g., MEL activities or a written assignment
- Conversational warm-up
- Presentation of new material
- Reading the text for the lesson in the book, discussing new vocabulary, checking for comprehension, and examining new speaking and writing structures
- Group, pair or individual practice (sometimes a student pair would present a role play)
- Reflecting and reviewing what had been achieved at the end of the lesson
- Assigning homework

During the observation, it was noted that this lesson structure was followed. The students, a group of five, were also observed speaking for most of the lesson. They were discussing the topic referring to how they could become millionaires. The majority of students were engaged throughout the lesson

and a lot of laughter was observed. Students read the questions and answered them in turn. They listened to the audio and asked for clarifications on vocabulary that was unknown to them.

Assessment and MEL

Interviewees suggested that teachers design the assessment approach to be followed in their classes, to allow for different class and student needs to be accommodated. MEL was used by both teachers teaching with Speakout as the main homework tool.

A further two teachers, who used a different Pearson book with MEL, used MEL in class when possible. According to one of these teachers, using MEL in class was helpful because it could provide an immediate link to class learning. It allowed teachers to view students' progress 'live' through successive attempts, and to intervene quickly when they saw repeated errors. The teacher described how they refreshed the page when using MEL in class, so they could see when students finished their exercises almost straight away. The other teacher experimented with completing MEL exercises on an interactive whiteboard in class, but highlighted how the limited classroom time prohibited her/him from using this practice frequently, as it could become time consuming.

Table B2 outlines the assessment approach adopted by the two teachers who used Speakout with MEL, as described in the interviews.

Pearson

Table B2: Wyższa Szkoła Logistyki's approach to assessment — interviews

	Formal	Informal
Teacher 3	• **Three written tests per semester (60%–70% of final score)** — tests were internally devised. Per semester,[15] students had to pass at least one test the first time and were allowed one retake • **Three MEL unit tests per semester (20–30% of final score)** — tests were open for a week and could be repeated two or three times • One oral presentation per semester (10% of final score) — on a topic of the student's choice • Speakout exams were considered less challenging and were used with struggling students or for re-sits.	• MEL practices • Mel Practices were compulsory and were open throughout the semester. • Teachers suggested that, despite being encouraged to use MEL weekly, students usually completed the assignments a few days before the end of the semester. • Students usually had two or three attempts. This was because, otherwise, some students would overly rely on getting good scores in MEL and coast in class, according to one teacher. According to the other teacher, if an unlimited number of attempts was allowed, the overall class average would be inflated and the scores would also not reflect students' performance in test situations. • For one teacher, students whose scores were above the class average could get a few more 'points' towards their final score. For the other teacher, a final score could increase, e.g., from B to a B+, if a student had completed all or most of the assigned work. • Students who were absent were assigned more MEL assignments. • Short quizzes were used as an additional assessment by one teacher, which could also '*nudge*' the final score up.
Teacher 4	• **Two or three written tests per semester (majority of the score) —** tests were internally devised and included similar questions to those on MEL and in the Students' Book • Speakout exams were considered less challenging and were used with struggling students or for re-sits.	
End of two years exam	Skills assessed included speaking, writing and grammar.	

Wyzsza Szkoła in Warsaw (WS)

[15] There are two semesters in a year.

WS's model of implementation and its teachers are supplied by Lingua Nova. Lingua Nova is a school of foreign languages with clients all over Poland, and they teach as Lingua Nova in many institutions all over Poland. Teachers are hired on a freelance basis and are qualified to at least master's level, if they are Polish, or hold a TEFL qualification, if they are native speakers. The decision to use Speakout with MEL, the implementation model for its delivery, and the training of the teachers, are the responsibility of Lingua Nova.

About 700 students in WS are taught using Speakout with MEL, mainly at Pre-intermediate level and sometimes at Intermediate level. Most of the students are in their twenties or thirties, attend extramural classes, and study towards a bachelor's or master's degree in their chosen discipline. Class sizes vary from 15 to 17 students.

Class contact time is every fortnight on Friday evenings, and Saturday and Sunday, for a total of 45 minutes at a time. This amounts to 1.5 hours a week, 3 hours a month and 16 hours a term. Students cover one Speakout level in four semesters, with two semesters of 4.5 months a year. For most courses, this means classes cover two units a semester and four within a year.

During class time, an emphasis is placed on speaking English. A key goal is increasing student confidence in communicative English, as well as fostering an appreciation for the real-life importance of knowing English. Learning the necessary vocabulary required by students' specific field of study and providing them with interesting opportunities to practise were also important instructional goals.

Reasons for selecting Speakout with MEL

Lingua Nova's decision to use Speakout with MEL was based on the belief that it focuses on speaking, it supports general English well and it is user-friendly for both teachers and students. The decision to use MEL was also deliberate. Reasons included:

- textbooks supported by technology and learning platforms becoming standard
- the need to increase practice outside the classroom and create learning links between class and home — given the limited class time, MEL was seen as useful for students to continue being in 'contact' with the language. Part of the appeal of MEL was the large number of practices available, which allowed for the two-week gap between classes to be filled
- support for the development of independent learning. Students can take control of their learning by deciding when, where and how they would complete their assignments, as well as whether they would do additional practices
- the need to track student progress outside the class — the ability to continue monitoring student engagement with the language and student performance beyond the classroom was important in deciding to adopt MEL

Training in MEL

Training on how to teach English and Speakout with MEL is offered by Lingua Nova to all new teachers. Training topics include:

- how to balance teacher and student talk in class
- how peer observations for new teachers are conducted
- training on MEL and a video on how to use it

Teachers are also given a manual that covers topics, such as:

- how to teach reading, writing, listening and speaking
- how to assign homework
- expectations from teacher presentations and use of multimedia in class
- rules for teachers (e.g., 'Never speak in Polish')

Overall, teachers felt that it would be useful to have top-up training on MEL's technical aspects, and understand how it could better support teaching, as well as how it could be implemented in order to improve student engagement.

Teaching and learning

In the classroom, focus was placed on speaking and on teaching technical vocabulary. This was, in part, because the other skills could be fostered independently through the use of MEL.

Assessment and MEL

Apart from MEL, classroom participation and formal assessment was comprised of end-of-semester tests and an end of year exam. Table B3 outlines how MEL was implemented.

Pearson

Table B3: Lingua Nova's approach to assessment — interviews

Assessment type	% of overall score	Assessment Tasks	Implementation
MEL	40%	Assignments in MEL	• **Number of assignments and due dates** — students were assigned all unit assignments. End dates were set for one week before the end of the semester. Thus, students completed assignments for two units within 4.5 months. The time frame allowed students to fit the completion of the assignments alongside their other commitments, according to interviewees. Teachers could set different completion dates if they wished to. One teacher suggested that they once changed the due date to bi-weekly, but had to deal with a number of students asking for extensions, so she/he reverted to leaving the submission date open for the whole semester • **MEL features used** — in the main, MEL assignments were used. One teacher also used the Review tests on MEL for revision and exam preparation purposes. Students completed the Review tests just before the end of the semester, and after their teacher alerted them that the relevant content was fully covered. Review tests could be attempted once and were completed outside the classroom • **Scoring** — students had to score >50% in the MEL assignments in order to sit the end-of-semester test, which was part of their assessment. The number of assignments students completed was not taken into account • **Number of attempts** — guidance suggests that students should be allowed multiple attempts, a practice followed by one teacher. Another teacher allowed for two attempts but no more, given that students could see the answer after the third attempt. She/he felt that three attempts could 'inflate' scores. She/he also used the assignments as test preparation rather than purely for practice • **Monitoring** — teachers were expected to monitor student performance before every class, attend a bi-weekly meeting and write a relevant monthly report, which was then sent to LINGUA NOVA's coordinator. One teacher mentioned that she/he checked the gradebook before every bi-weekly meeting, made notes of students who had not completed the assignments and of particular areas with which students seemed to struggle. Teachers spoke to students individually about their work on MEL or commented on the importance of completing the assignments in class.

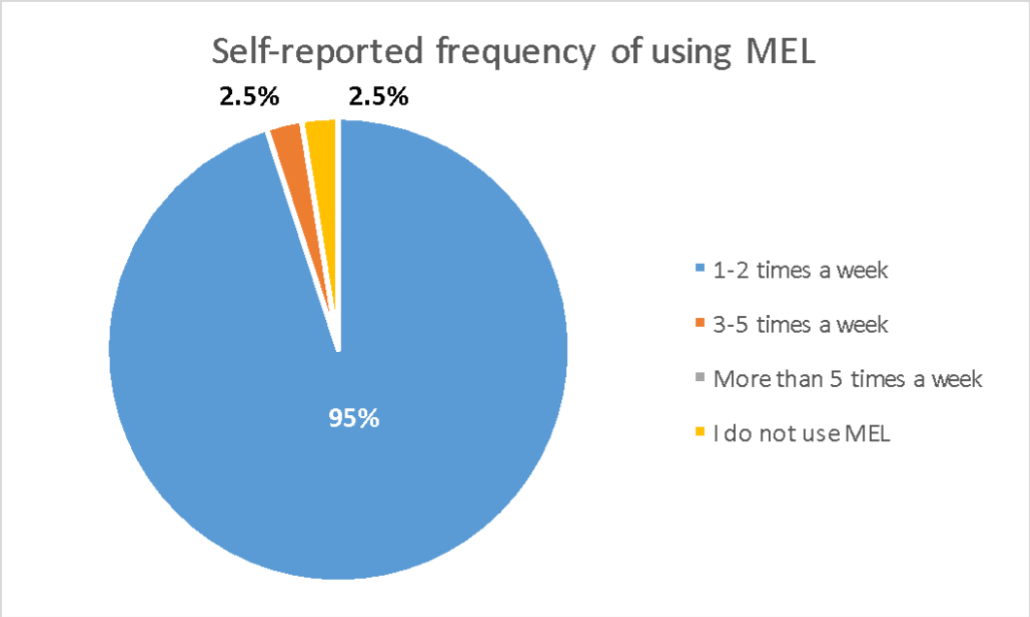
Student use of MEL

The use of MEL by students can be used to make inferences about student access, experience, and engagement. In addition, patterns of use and the reasons given for using MEL provides a more nuanced portrait of how learners are enacting blended instruction. The following section describes the analysis of usage data collected through the student questionnaire and data extracted from the MEL platform. Analysis of usage data focuses on students' weekly use of MEL; time spent on MEL weekly; student frequency of using the gradebook; the number of practice/assignment set and completed; and the number of attempts on assignments/practices.

Student weekly use of MEL[16]

The large majority of students reported that they used MEL once or twice a week (95%, 188/198). Very few indicated using MEL three to five times a week or not using it at all. No students reported using it more than five times a week.

Figure B4: Students' self-reported weekly use of MEL — student questionnaire

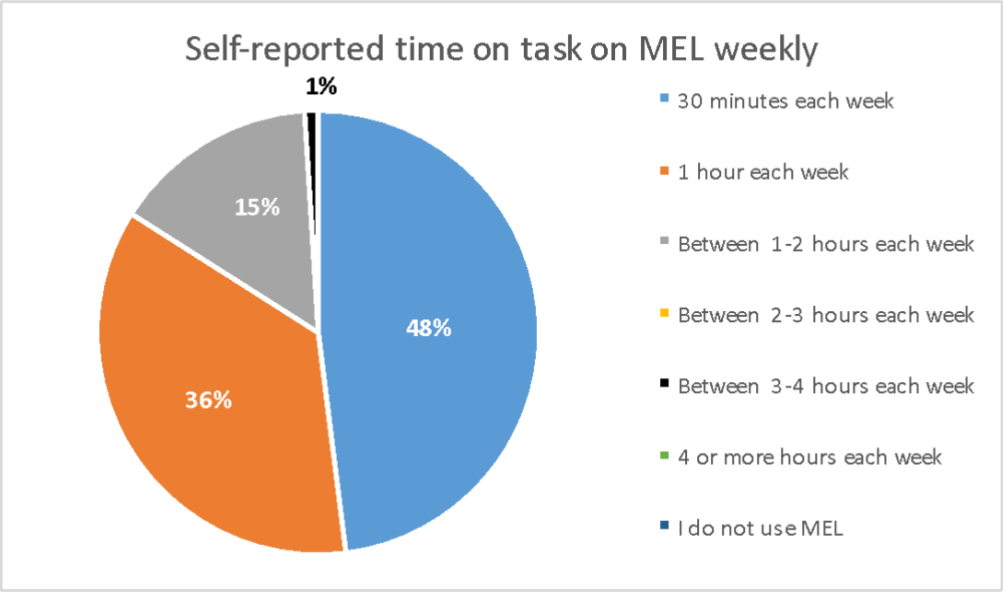


[16] Please note that if we were to remove the nine Uniwersytet Centralny students from the calculation, due to the small sample number, there would be no significant change in the result.

Hours / minutes spent on MEL weekly

Questionnaire data shows the majority of students across the three institutions (48%, 90/189) used MEL for 30 minutes per week, while a sizeable 36% (68/189) spent one hour. Nearly 15% (28/189) reported spending between one and two hours per week, while very few students spent more time on MEL than that.

Figure B5: Students' self-reported hours/minutes of using MEL weekly — student questionnaire



A more even split is observed between AWF students who suggest that they used MEL for 30 minutes weekly and those who suggested they used it for one hour weekly. There is some variation between the responses of the students from different institutions, however, these differences are not substantial. Students from the three different institutions and across levels did not seem to spend more than 12.5 hours over the two years of their course on practices and assignments, and 5.8 hours in one-year courses (see Table B4).

Table B4: Student self-reported frequency of using MEL weekly per institution —student questionnaire

Institute	% of students							
	30 minutes	1 hour	1–2 hours	2–3 hours	3–4 hours	4 or more hours	I do not use MEL	Total
AWF	63 (42%)	58 (38%)	21 (14%)	5 (3%)	2 (1%)	1 (1%)	1 (1%)	151 (100%)
Wyższa Szkoła Logistyki	19 (50%)	9 24%)	7 (19%)	0 (0%)	1 (2%)	0 (0%)	2 (5%)	38 (100%)
Uniwersytet Centralny	8 (89%)	1 (11%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	9 (100%)
Total	90 (45%)	68 (34%)	28 (14%)	5 (3%)	3 (2%)	1 (1%)	3 (1%)	198 (100%)

Notes: $\chi 2 = 16.278$ · df = 12 · Cramer's V = 0.203 · Fisher's p = 0.142.

It seems that students' self-reports about the time they spent on the platform and the time they spent on task as recorded in the MEL platform, show that the former fall under the 'Engaged' category. Individual students' time spent on practices during their courses shows substantial variability, as seen in Table B5 below.

Pearson

Table B5: Time on task by students per quartile for two-year courses—MEL data

Least engaged (1st Quartile)	Less engaged	Engaged	Most engaged
7 hours or less	7 – 11.2 hours	11.2 – 16.3 hours	16.3 hours or more
1st Quartile	2nd Quartile	3rd Quartile	4th Quartile

One-year courses show similar results to the two-year courses (see Table B6).

Table B6: Time on task by students per quartile for one-year courses — MEL data

Least engaged	Less engaged	Engaged	Most engaged
36 hours or less	3.6 – 5.6 hours	5.6 – 7.7 hours	7.7 hours or more
1st Quartile	2nd Quartile	3rd Quartile	4th Quartile

Student frequency of checking the gradebook

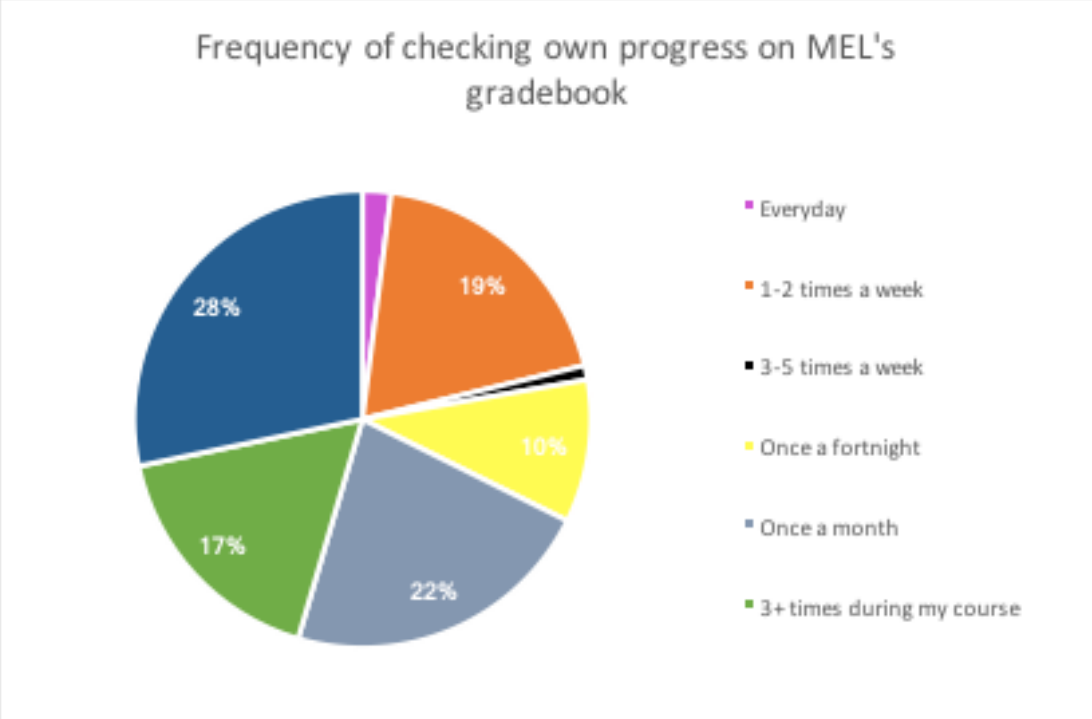
The gradebook feature helped students self-assess, an important practice for improving one's learning. For some learners, the gradebook can also be motivating. This is known in the literature and was also stated in the student focus groups.

Although the majority of students (72%, 141/197) reported using the gradebook, 28% (56/197) suggested that they did not use it at all. 22% (44/197) of students reported that they checked their progress on the gradebook once a month, and just below 30% (58/197) of students reported that they checked their progress on the gradebook one or two times a week, or once a fortnight. In more detail:

- 22% (44/197) of students checked their progress once a month
- 19% (37/197) of students checked their progress one or two times a week
- 17% (33/197) of students checked their progress more than three times during their course

- 11% (21/197) of students checked their progress once every two weeks
- 2% (4/197)) of students checked their progress every day
- 1% (2/197) of students checked their progress three to five times a week
- 28% (56/197) of students did not check their progress on MEL

Figure B6: Students' self-reported frequency of checking their progress in the gradebook — student questionnaire



There is some variation between and within institutions in students' response related to the frequency of checking their progress in the gradebook, but this variation is not substantial. Around a quarter of the students in the AWF and Wyższa Szkoła Logistyki institutions reported checking their progress in the gradebook once a month. The corresponding percentage for Uniwersytet Centralny was 0%. However, the sample size is too small to yield reliable results. Moreover, 21% of AWF's students checked the gradebook once or twice per a week. The corresponding percentage for Wyższa Szkoła Logistyki was 10%. 29% of Wyższa Szkoła Logistyki 's students checked their progress more than three times during their course. The corresponding percentage for AWF's students was 13%.

Table B7: Student self-reported frequency of checking own progress in the gradebook per institution — student questionnaire

Institute		No of students (%)						
	Every day	3–5 times a week	1–2 times a week	Once every two weeks	Once a month	3+ times during my course	I do not check my gradebook	Total
AWF	0 (0%)	2 (1%)	32 (21%)	17 (11%)	35 (23%)	20 (13%)	44 (29%)	150 (100%)
Wyższa Szkoła Logistyki	3 (8%)	0 (0%)	4 (10.5%)	4 (10.5%)	9 (24%)	11 (29%)	7 (18%)	38 (100%)
Uniwersytet Centralny	1 (11%)	0 (0%)	1 (11%)	0 (0%)	0 (0%)	2 (22%)	5 (56%)	9 (100%)
Total	4 (2%)	2 (1%)	37 (19%)	21 (11%)	44 (22%)	33 (17%)	56 (28%)	197 (100%)

Notes: $\chi 2$ = 16.278; df = 12; Cramer's V = 0.203; Fisher's p = 0.142.

Number of practice, assignment, and test completions

The number of students who completed practices and assignments varies between institutions and, at times, within institutions. Differences in completion rates in one-year courses appear substantial when compared to those of two-year courses. Overall, completions of practices appear relatively low with an average completion of 51% for the two-year courses and 13% for the one-year courses. In more detail:

- AWF — students completed between 36% and 52% of all practices assigned.
- Wyższa Szkoła Logistyki — students completed a significantly variable number of practices.
- WS — students of all levels completed about 50% of all assignments/practices assigned across levels.

Pearson

Table B8: Assignments/practices/tests assigned and average number completed — MEL data

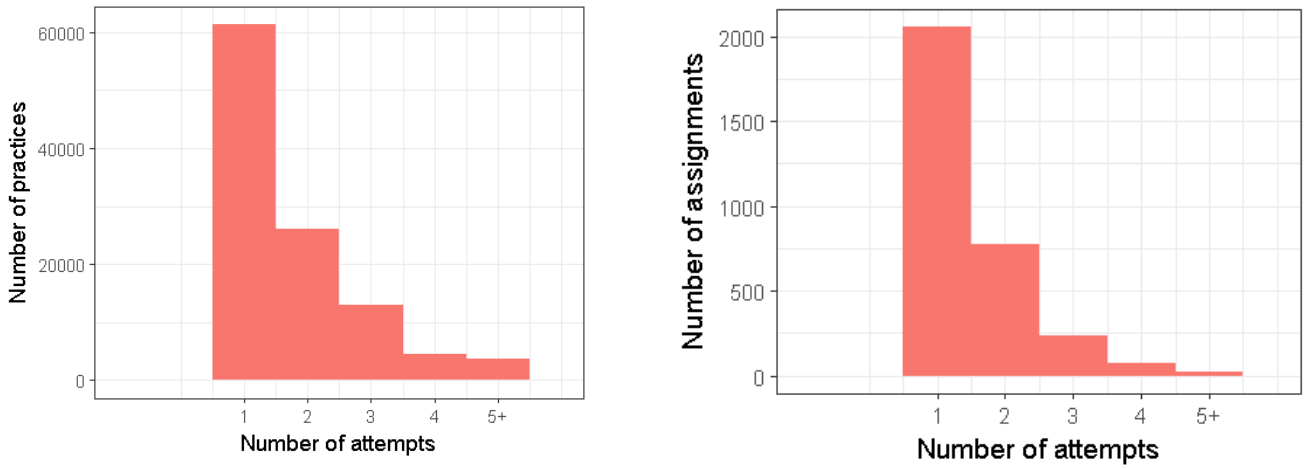
Class	Practices		Assignments		Tests	
	Assigned	Completed	Assigned	Completed	Assigned	Completed
AWF						
Intermediate (2014–16)	366	190 (52%)	–	–	10	7 (70%)
Intermediate (2015–17)	366	160 (44%)	–	–	12	6 (50%)
Intermediate (2016–18, data	445	160[1] (36%)[1]	–	–	10	7[1] (70%)[1]
Wyższa Szkoła Logistyki						
Pre-intermediate	392	103 (26%)	–	–	–	–
Pre-intermediate	392	16 (4%)	88	51 (58%)	–	–
Advanced (2016–17)	387	79 (20%)	–	–	–	–
Advanced (2014–15)	320	41 (13%)	–	–	–	–
WS						
Elementary (2015–17)	377	187 (50%)	–	–	–	–
Pre-intermediate	392	185 (47%)	–	–	–	–
Intermediate (2015–17)	366	184 (50%)	–	–	–	–

Note: (1) Completions are for one year. Due to the very small number of tests completed by some Wyższa Szkoła Logistyki students, we did not report them.

Number of attempts on assignments/practices

Across institutions, the majority of practices were completed using one attempt (56%, 61,349/108,653), followed by two attempts (24%, 26,034/108,653). Twenty per cent of practices were attempted more than two times (21,270/108,653). As regards the assignments, 65% (2,059/3,194) were attempted once, 24% (773/3,194) twice, and the remaining 11% (362/3,194) more than two times.

Figure B7: Number of attempts on practices and assignments, all institutions (> 49%), followed by two attempts — MEL data



The number of attempts per institution were as follows:

- **AWF —** between 66% and 68% of tasks were attempted once; about a quarter twice; and the minority (< 10%) more than three times
- **Wyższa Szkoła Logistyki** — although the majority of tasks were completed using one attempt, around one quarter of them were attempted twice. In most cases, less than 10% were attempted more than two times.
- **WS** — in the Elementary class we observed more of a spread in terms of the attempts on practices, when compared to Pre-intermediate and Intermediate classes. The latter two levels showed half of the practices had been attempted once, around a quarter, twice, and between 24% and 27% using three or more attempts. However, for the Elementary students, around 43% of the tasks were attempted three or more times.

Pearson

Table B9: Percentage of number of attempts of assignments/practices per institution and class—MEL data

Level	Number of attempts				
	1	2	3	4	5+	Total
AWF						
Intermediate (2014–2016)	14803 (68 %)	5392 (25%)	1146 (5%)	322 (1%)	178 (1%)	21841 (100%)
Intermediate (2015–2017)	10979 (68%)	3829 (24%)	993 (6%)	229 (1%)	86 (1%)	16116 (100%)
Intermediate (2016)	5909 (66%)	2226 (25%)	572 (6%)	132 (2%)	61 (1%)	8900 (100%)
Wyższa Szkoła Logistyki						
Pre-intermediate (2014)	1114 (89%)	120 (10%)	17 (1%)	5 (0%)	3 (0%)	1259 (100%)
Pre-intermediate (2014/16)	479 (51%)	245 (26%)	126 (14%)	39 (4%)	49 (5%)	938 (100%)
Pre-intermediate (2014/16) (only assignments)	2059 (65%)	773 (24%)	234 (7%)	72 (2%)	56 (2%)	3194 (100%)

Level		Number of attempts				
Advanced (2014)	378 (54%)	206 (29%)	77 (11%)	21 (3%)	18 (3%)	700 (100%)
Advanced (2016)	1081 (62%)	423 (24%)	140 (8%)	52 (3%)	48 (3%)	1744 (100%)
WS						
Elementary (2015)	4288 (34%)	2874 (23%)	3067 (24%)	1286 (10%)	1175 (9%)	12690 (100%)
Pre-intermediate (2015)	15691 (50%)	7287 (23%)	4887 (16%)	1831 (6%)	1512 (5%)	31208 (100%)
Intermediate (2015)	6627 (50%)	3432 (26%)	2015 (15%)	667 (5%)	516 (4%)	13257 (100%)

Pearson

Student performance on MEL

Student progress on assignments provides an indicator of student learning and achievement over time. Performance on MEL assignments is the only source of student performance data that we have from these institutions. In this section, we describe average student performance by institution and assignment type. We are also able to investigate the relationship between completion of assignments and performance (i.e. scores) on those assignments. Finally, we offer evidence that individual assignment scores are relatively reliable indicators of student learning that can be confidently used by instructors to monitor progress over time.

Overall student performance on MEL assignments/practices

Overall, students had a high performance on practices, as shown by the average scores. In more detail, average performance on practices for AWF was 88% for two of the two year cohorts and 92% for the one year cohort. For Wyższa Szkoła Logistyki, performance ranged from 84% for its one year cohort and from 87% to 94% for the two-year cohorts. For WS, performance was 97% for all three levels.

Table B10: average performance on practices, assignments and tests per institution and class—MEL data

Average performance on tasks			
	Average practice score	**Average assignment score**	**Average test score**
AWF			
Intermediate (2014)	88	–	69
Intermediate (2015)	88	–	70
Intermediate (2016) (only one year covered)	92	–	77
Wyższa Szkoła Logistyki			
Pre-intermediate (2014) (one-year course)	94	–	–
Advanced (2016) (two years course)	84	–	–
Pre-intermediate (2014/16) (one-year course)	87[1]	87[1]	–
Advanced (2014) (one-year course)	93	–	–
WS			
Elementary (2015)	97	–	–
Pre-intermediate (2015)	97	–	–
Intermediate (2015)	97	–	–

Reporting students' performance by score band is important in order to inspect the number of students who failed (a score below 70%) or who achieved the top score (a score above 90%):

- Around 55% of AWF students received an average practice score of 90% or more, and 33% of AWF students received an average practice score of 80–89%.
- 92% of Wyższa Szkoła Logistyki students received an average practice score higher than 70%, and 60% of Wyższa Szkoła Logistyki students received an average score of more than 90%.
- Almost all (93%) WS students received an average practice score higher than 90%.

Table B11: Percentage of students achieving within different score bands in practices per institution — MEL data

Performance band	% of students		
	AWF	Wyższa Szkoła Logistyki	WS
Below 50%	–	2 (2%)	–
50%–59%	2 (1%)	0 (0%)	–
60%–69%	7 (2%)	6 (6%)	–
70%–79%	30 (9%)	8 (8%)	2 (1%)
80%–89%	109 (33%)	25 (24%)	21 (7%)
90%–100%	178 (55%)	62 (60%)	284 (93%)
Total	326 (100%)	103 (100%)	307 (100%)

Note: The results do not change significantly if the zero scores are not removed from the analysis (see Appendix 3 Table 1).

As regards test scores, 61% of AWF students achieved a score above 70%, whilst 39% did not.

Table B12: Percentage of students achieving within different score bands in tests and assignments per institution — MEL data

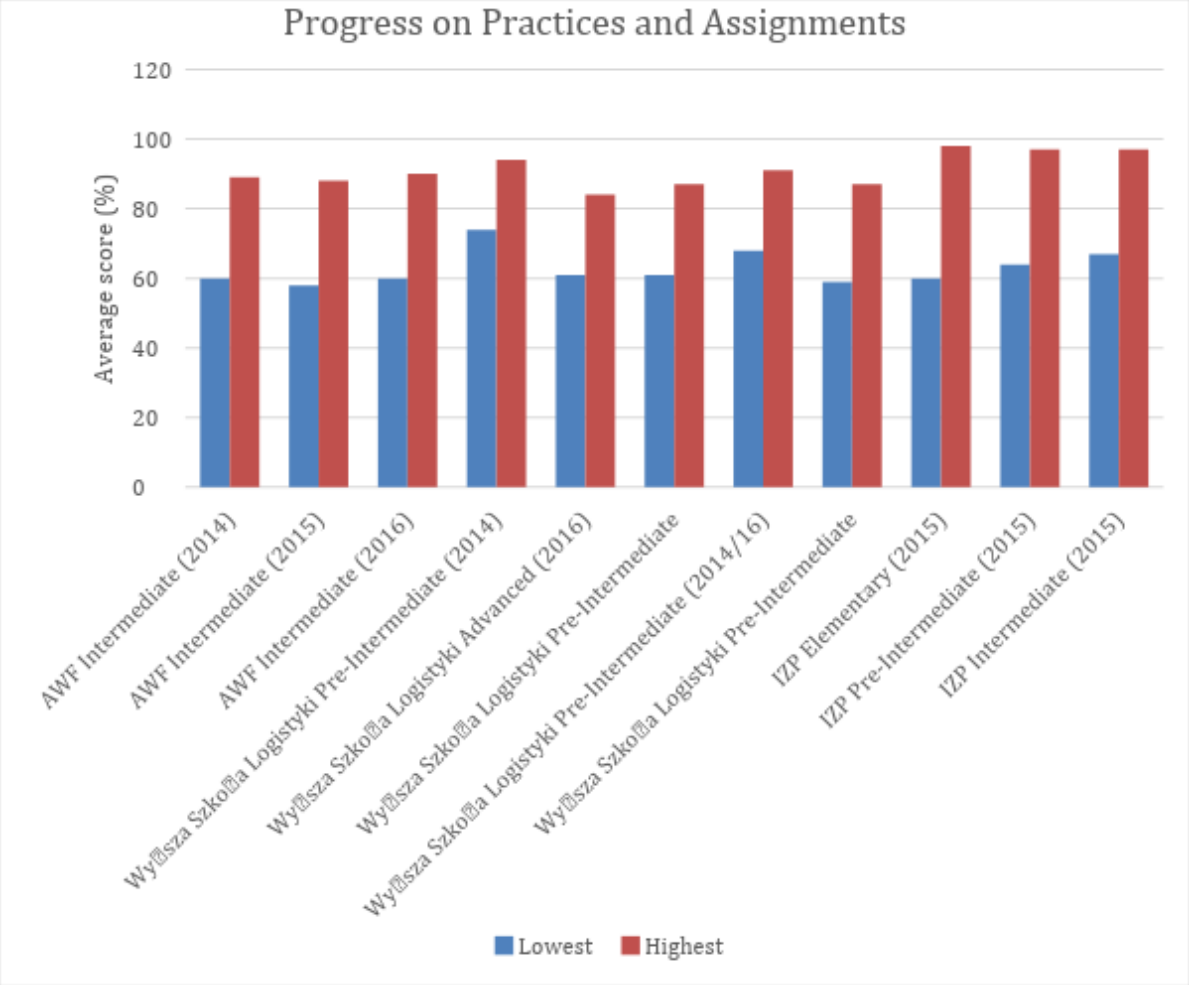
Performance band	% of students	
	AWF Test scores	Wyższa Szkoła Logistyki Assignment scores Pre-intermediate 2015–2016 one-year course
Below 50%	20 (6%)	0 (0%)
50%–59%	35 (11%)	2 (3%)
60%–69%	69 (22%)	0 (0%)
70%–79%	94 (30%)	9 (15%)
80%–89%	69 (22%)	23 (40%)
90%–100%	28 (9%)	25 (42%)
Total	315 (100%)	59 (100%)

Note: The results differ if the zero scores are not removed from the analysis (see Appendix D, Tables D1 and D2).

Student progress on practices and assignments

Progress from students' first to highest attempt was quite astonishing. The average score improvement was around 30% (per level ranged from 21% to 38%). (See Appendix E for details.)

Figure B8: Progress on assignments/practices per class — MEL data



Correlation between students' average practice/assignment scores and percentage of practice/assignments completed[17]

Overall, data indicates that a higher rate of completion of practices and assignments is associated with higher scores. In eight out of 11 cases, there was a high and positive correlation between the average practice score and the percentage of practices completed.

- The correlation for AWF ranged from 0.52 to 0.67, all statistically significant at the $p < 0.001$ level.
- Two of the correlations for Wyższa Szkoła Logistyki were not statistically significant, but one correlation was significant and high ($r = 0.59$, $p = 0.013$).

[17] No correlations were calculated for tests due to the small number tests completed by students.

- The correlations for WS ranged from 0.50 to 0.73, all statistically significant at the $p < 0.001$ level.

An additional correlation was computed for Wyższa Szkoła Logistyki Pre-intermediate level students (the only group of students which completed assignments): there was a statistically significant and positive correlation between the average assignment scores and the percentage of assignments completed. The Pearson correlation was 0.44 ($r [77] = 0.44$, $p < 0.001$), and the Spearman correlation was 0.41 ($rho = 0.41$, $p = .001$).[18]

[18] Both Pearson and Spearman's correlations were computed to accommodate for the skewed distributions of the data. Removing a limited number of outliers does not change the values of the correlations significantly.

Table B13: Correlation between percentage of assignments/practices completed and average score per institution and course — MEL data

	Practices	
	Pearson's r	**Spearman's rho**
AWF (all at the Intermediate level; all are two-year courses)		
Intermediate (2014)	r=0.63, p<0.001	rho=0.59, p<0.001
Intermediate (2015)	r=0.67, p<0.001	rho=0.65, p<0.001
Intermediate (2016)	r=0.52, p<0.001	rho=0.59, p<0.001
Wyższa Szkoła Logistyki		
Advanced (2016) (one year course)	r=0.18, p=0.515	rho=0.11, p=0.691
Pre-intermediate (2014) (two years course)	r=0.59, p=0.013	rho=0.55, p=0.023
Pre-Intermediate (2014/16) (one year course)	r=-0.19, p=0.164	rho=-0.22, p=0.12
Advanced (2014) (one year course)	r=0.83, p<0.001	rho=0.41, p=0.104
WS (all two-year courses, 2015-2017)		
Elementary (2015)	r=0.73, p<0.001	rho=0.69, p<0.001
Pre-Intermediate (2015)	r=0.52, p<0.001	rho=0.61, p<0.001
Intermediate (2015)	r=0.50, p<0.001	rho=0.67, p<0.001

Pearson

Correlation between the average unit practice or assignment

Practices are relatively reliable indicators of students' performance. So, the performance of a student in one practice could be used as an indication of their future performance if they were to take another. For each student, the average practice score was computed for all the exercises of each unit. These average practice scores for each student and for each unit were then correlated between them, in the same way one might correlate the item scores of students in a test. In most cases, the correlations ranged from moderate to high (average correlation of 0.45 for AWF, 0.53 for Wyższa Szkoła Logistyki and 0.41 for WS; most of the correlations were statistically significant, at least at the p = 0.05 level), suggesting that high performance on past practices is a good predictor of high performance on future practices completed for a unit.

The correlations between the unit average practice scores mainly ranged between 0.31 (1st Quartile) and 0.58 (3rd Quartile) for AWF, with a mean of 0.42. One hundred and fifty-nine Pearson correlations were computed; of 131 correlations, around 82% were statistically significant at the p = 0.05 level.

For Wyższa Szkoła Logistyki, the correlations ranged from 0.41 (1st Quartile) to 0.70 (3rd Quartile), with a mean correlation of 0.53. Eighty-six Pearson correlations were computed; of 26 correlations, around 30%, were statistically significant, but the sample size was extremely small — around 15 students for most of the correlations. Once again, it is not surprising that many of the correlations were, therefore, not statistically significant.

For Wyższa Szkoła Logistyki Pre-intermediate students (not shown in the table below), using only the assignments, the correlations mainly ranged between 0.31 (1st Quartile) and 0.47 (3rd Quartile), with a mean of 0.41. Ten Pearson correlations were computed; of six correlations, 60%, were statistically significant, but, again, the sample sizes were small — around 15 students for most of the correlations.

The correlations between the unit average practice scores mainly ranged between 0.21 (1st Quartile) and 0.61 (3rd Quartile) for WS, with a mean of 0.41. Two hundred and ninety-four Pearson correlations were computed; of 177 correlations, around 60% were statistically significant at the p = 0.05 level, but most of the sample sizes were small, with an average sample size of around 15. It is not surprising, therefore, that many of the correlations were not statistically significant.

Table B14: Correlations between the average Unit practice per institution—MEL data

	AWF	Wyższa Szkoła Logistyki	WS
Average correlation	0.45	0.53	0.41
1st Quartile	0.31	0.41	0.21
3rd Quartile	0.58	0.70	0.61

Correlation between average practice score and average test score

We have strong indications of convergence validity when practice and test scores are considered. In other words, the practices and the tests tend to yield consistent information on the ability of the students. This means that practice scores could be an early warning of student performance on MEL tests too. There was a strong association between the average practice score and the average test score. The correlations were moderate to high (0.40 to 0.69) and statistically significant in all cases at the $p < 0.001$ level.

Table B15: Correlations between the average unit practice and test score per institution and course — MEL data

	Pearson's r	Spearman's rho
	AWF	
Intermediate (2014)	r = 0.45, p < 0.001	rho = 0.52, p < 0.001
Intermediate (2015)	r = 0.69, p < 0.001	rho = 0.71, p < 0.001
Intermediate (2016)	r = 0.40, p < 0.001	rho = 0.36, p < 0.001

Correlation between test scores

Tests are relatively reliable indicators of students' performance. The correlations between the test scores across students mainly ranged between 0.38 (1st Quartile) and 0.57 (3rd Quartile), with a mean of 0.46. Seventy-six out of 81 correlations, around 94%, were statistically significant at the p = 0.05 level.

Student and teacher perceptions of Speakout with MEL

This section provides evidence from the student questionnaires, student focus groups and teacher interviews on the perceived impact of Speakout with MEL that maps directly to the learner outcomes of Speakout with MEL — namely access, engagement, positive learning behaviours, achievement, preparation for the next level of learning and achieving one's goal. In turn, this section outlines perceptions of the impact of the Speakout package, the Students' Book and MEL on students, and teacher perceptions of the impact of Speakout with MEL on teachers and teaching.

Before presenting Tables B16 to B21, we summarise some institutional factors noted in the data that have affected implementation. These factors related to infrastructure, training, and teaching and learning.

Infrastructure

- The lack of availability of technology (in one institution in particular), such as lack of equipment to display multimedia resources, project materials, or connect to the internet in classrooms, meant that teachers and students could not take full advantage of all the materials offered through Speakout with MEL.
- The lack of equipment at home, e.g. microphones, prevented some students from completing certain exercises in MEL.
- Registering large numbers of students at once overloaded some institutional systems.

Training

- Teachers welcomed further support on how best to implement blended instruction.

Teaching and learning

- Some students drew conclusions on the usefulness of MEL based on the exercises they were assigned. For those who were not assigned exercises across all English language skills, MEL was not seen as useful.
- Students suggested that if teachers assigned a smaller number of activities in MEL daily that could increase usage. Assigning over 200 assignments at once was overwhelming for some students.
- When speaking was assessed, it seems that it was also taught more often. When it was not assessed, it was not taught as often as other skills, such as writing.

Pearson

Table B16: Speakout with MEL — perceived impact on student outcomes, across data sources

Evidence of impact and enablers	Suggested improvements
Engagement	
• The majority of students agreed or strongly agreed that their confidence had increased since using Speakout with MEL in learning English (74%, 142/192); reading (76%, 145/192); listening (75%, 143/191); writing (59%, 113/191); and speaking (55%, 106/192). • The majority of students suggested that Speakout with MEL significantly or very significantly helped them engage with learning English (77%, 149/194), enjoy learning English (66%, 128/194) and improve their motivation to learn English (70%, 135/193). • One teacher remarked that improvements to a student's speaking skills were always related to teacher involvement, but agreed that Speakout with MEL did support the development of communication skills.	• Although 84% (31/37) of Wyższa Szkoła Logistyki students agreed or strongly agreed that their confidence in speaking had increased, the equivalent for AWF was 49% (71/146). Student responses from AWF therefore skewed the overall finding to suggest that 41% of students (78/191) disagreed or strongly disagreed that their confidence in writing had increased since using Speakout with MEL. • Students were grouped into: detractors (1–6 on a 10 scale), passives (7–8) and promoters (9 and 10), depending on their likelihood to recommend Speakout with MEL. The large majority of students were passives, scoring 7 or 8 on the scale (57%). This large proportion of passive students led to an overall NPS of -11.9. When the data was analysed descriptively by separating students into groups as follows, — 'low' (1–3 on a 10-point scale), 'moderate' (4–7) or 'high' (8–10), results show 6% of students in the low category, 45% of students in the moderate category, and 48% of students in the high category.
Achievement	
• The majority of students in the questionnaire thought that the Speakout with MEL package had either significantly or very significantly helped them improve their English (61%, 115/189); their vocabulary (87%, 169/194), grammar (78%, 151/193), listening (75%, 146/194), writing (70%, 135/194) and speaking skills (65%, 126/194).	• The reasons behind students' responses that Speakout with MEL had helped them to improve their English (39%, 74/189) and their speaking (35%, 68/194) 'to a little extent or not at all' needs further investigation in relation to the materials themselves, teacher training, and implementation. Speakout with MEL could further support pronunciation by including more drills, one interviewee suggested.
Progression	

Evidence of impact and enablers	Suggested improvements
• The majority of students (84%, 161/192) agreed or strongly agreed that Speakout with MEL was preparing them well for the next level of English study, with 80% (154/193) indicating that it did so significantly or very significantly. • 79% (152/192) of students suggested that Speakout with MEL was preparing them well to achieve their goal, with 78% (151/194) indicating that it was doing so significantly or very significantly.	• The progression in challenge of grammar between Intermediate and Upper Intermediate levels could be reviewed, as suggested by two coordinators. This would be particularly useful for students working below the expected level, they suggested.

The Students' Book

Table B17: Student perceptions of the impact of the Student's Book

Evidence of impact and enablers	Suggested improvements
Access and experience	
• Data indicates the accessibility of the Students' Book. Students and teachers in the interviews and the large majority of students in the questionnaire agreed or strongly agreed that the Students' Book was: • up to date (99%, 198/199). Students suggested that updated video material was particularly popular in one institution, e.g., 'The Blind Painter' • interesting (91%, 182/199). Students and teachers noted in interview the appeal of the interesting and engaging stories and dialogue topics, such as 'How to Be a Millionaire' • culturally appropriate (94%, 188/199) • at the appropriate level (90%, 180/199) • relevant to real life (82%, 163/199), with topics such as 'Travelling' or 'How to Find a Job' • easy to navigate, according to several teachers • flexible enough that extramural students, who only meet ten times a semester for 90 minutes at a time, can still cover 50%–70% of the content	• Guidance on how teachers could cover course content using a fast route would be helpful. This would be useful because teachers suggested that contact time is relatively limited, especially with extramural students. • Reviewing the videos so that students can more easily access the humour might be beneficial. Students were more familiar with American humour, making it hard for them to culturally relate to the British comedies in videos, such as Blackadder. Teachers also suggested that the vocabulary used in such videos was challenging for students. • A list of difficult words at the end of each unit, or at the end of the book would be helpful according to one student focus group. It would also help them to practice more easily during exam time, they told us. • The inclusion of academic and technical vocabulary would be useful, teachers suggested. One teacher also suggested that it might be worth reviewing the links between activities in each unit, so that they are closely linked.
Engagement	
• The Students' Book content, mainly the stories, led to high levels of engagement, a finding across student and teacher interviews • Video clips seemed to motivate students in all three institutions through the diversity of topics, which broadened their cultural horizons (according to teachers)	• Reviewing content to further include topics that support students' interests would be beneficial. A student's age and own personal likes/dislikes affected their interest in a topic. For example, and unsurprisingly, mature students were not necessarily attracted to the same topics as young adults. Furthermore, one student noted that the 'Keeping Secrets' topic was too personal and made them feel uncomfortable. Other students in the same focus groups agreed.
Achievement	
• Teachers suggested that students' confidence in speaking had increased	• Answer keys for the exercises in the Students' Book were seen as

Evidence of impact and enablers	Suggested improvements
since using the Students' Book. AWF students confirmed that the dialogue activities in the book helped with their confidence in speaking. • Teachers attributed improvements in students' spoken ability to the book's: • interesting topics, which promote discussion naturally in class • variety of spoken exercises (dialogue, discussion, role play) • grammar and vocabulary exercises, which usually include pair/group speaking options • Both students and teachers thought that the book could be adapted to suit the needs of students operating at different levels. • The activities where students have to plan and produce a topic for discussion themselves were particularly praised by one teacher. • One teacher suggested that pre-intermediate and intermediate books supported struggling students with speaking through repetition of basic structures, engaging topics and questions as discussion prompts. • One teacher noted that the opportunity to watch videos with or without subtitles was useful for differentiated instruction.	useful. Reviewing them for consistency would be helpful, as teachers found errors in them at times. • Consideration could be given to the addition of more pronunciation, writing, vocabulary exercises, and practical English, teachers and students suggested. • Reviewing the Speakout materials for differentiation, and further investigating and training teachers on how they can use them effectively to support the differing needs of students, seems appropriate.
Progression	
• Overall, Speakout was seen as a book that prepares students well for the next level in their studies (see achievement data in this table).	• The level of challenge of grammar could be reviewed, as teachers suggested that higher challenge, particularly from the Intermediate level onwards, would benefit students.

Pearson

MyEnglishLab

Table B18: Student perceptions of MEL's impact on access

Evidence of impact and enablers	Suggested improvements
Access and engagement	
• MEL is accessible according to the large majority of students in the focus groups and in the questionnaire. In the questionnaire: • 95% (190/199) agreed or strongly agreed that they could access it on a computer and 81% (157/193) on a smartphone/tablet. MEL's accessibility on smartphones and tablets allows students to study in their own time and place, which was seen as useful, especially for extramural students • 95% (188/198) agreed or strongly agreed that MEL was easy to navigate • 94% (188/200) that accessing assignments was easy	• Student training materials could be developed that support institutions to raise students' awareness of the most important aspects of Speakout with MEL, as not all students in the focus groups were aware that MEL can be accessed via smartphone or of the existence of listening and video activities. Moreover, further guidance to clarify for students the different purpose of MEL compared to the Students' Book activities and to other materials, as well as how best to use it to improve their learning could be useful. (A few students made direct comparisons between the Students' Book and MEL, even after the interviewer suggested that they served different purposes.) • Students were grouped into: detractors (1–6 on a 10 scale), passives (7–8) and promoters (9 and 10), depending on their likelihood to recommend Speakout with MEL. The large majority of students were passives, scoring 7 or 8 in the scale (57%). This large proportion of passive students led to an overall NPS of -11.9. When the data was analysed descriptively — by separating students into groups as follows; 'low' (1–3 in a 10-point scale), 'moderate' (4–7) or 'high' (8–10), — results show 6% of students in the low category, 45% of students in the moderate category, and 48% of students in the high category.

Pearson

Table B19: Student perceptions of MEL's impact on behaviours

Evidence of impact and enablers	Suggested improvements
Engagement/contributors to positive learning behaviours	
• Students stated that MEL was modern and opened up another window of opportunity to practice English. • 64% (124/195) of students agreed or strongly agreed that they found MEL engaging. • 57% (113/198) of students indicated that they enjoyed learning by completing MEL assignments. • Student responses to the questionnaire also showed that: • 96% (191/200) of students thought that there was a good variety of tests in MEL • 94% (185/196) of students thought that there was a good variety of exercises in MEL • The majority of students also found the following useful or very useful to their learning: • checking their answers on MEL immediately (96%, 190/198) • accessing MEL at any time or place (92%, 183/198) • seeing a summary of their scores and progress (89%, 175/197) • easily seeing their assignment completion dates (78%, 153/196) • communicating with their teacher online (57%, 110/193) • Students' independence skills developed because they were required to take decisions about when and what they learned, according to teachers. • MEL seemed to promote self-monitoring and self-assessment: • several students suggested that viewing their performance on MEL motivated them to constantly improve. Students in one of the focus group said that they checked their score every time they did an assignment.	• MEL exercises could be reviewed to allow for more than one answer at a time and to increase challenge overall. This was a particular request from students who were studying Languages at a Master's level. • The type and variety of activities in Speakout and MEL could be enriched to reduce repetition, according to several interviewees. • Reviewing the setting of the autoscoring system regarding punctuation and spelling would be appreciated. Students suggested that the system penalised them for punctuation errors and incorrect use of auxiliary verbs, when the activities they were completing were assessing a different skill. • Instructions provided for open-ended questions could be reviewed so that they become even more specific, e.g. stating the length of the answer required, students suggested. • Materials to support institutions with student training on MEL could increase usage. Such materials should include and emphasise the usefulness of the gradebook, given that 28% (56/197) of students in the questionnaire reported not using it.

Pearson

Table B20: Student perceptions of MEL's impact on achievement

Evidence of impact and enablers	Suggested improvements
Achievement	
• MEL freed time in class for speaking, since grammar and vocabulary were practised and checked through MEL. • Students and teachers found the correspondence between the book's practices and those in MEL helpful. • According to the majority of students in the questionnaire, MEL exercises were helpful in the development of different language skills: • 91% (179/197) found the grammar exercises useful or very useful • 90% (177/196) found the vocabulary exercises useful or very useful • 71% (140/198) found the writing exercises useful or very useful • 89% of students (175/197) found it useful or very useful to see a summary of their scores and track their progress on MEL • 84% (167/198) agreed or strongly agreed that MEL feedback helped them understand how to improve their English • 96% (190/198) found the ability to immediately check answers to be useful or very useful • 91% (179/197) found the ability to repeat activities useful or very useful • 78% (155/198) of students agreed or strongly agreed that MEL helped them to understand the content covered in class. • 64% (124/195) of students found the pronunciation-coach videos helpful. • Students in one focus group found the vocabulary exercises particularly helpful in revising difficult words learnt in class.	• To further support differentiation and engagement, MEL activities could be enriched so that they include extra tasks at different levels of challenge and a wider type of activities, such as quizzes and games, which one teacher was using to further support their teaching. • Teacher materials and training could further emphasise the value of the common errors and diagnostic reports to teaching and learning, given that teachers report rarely using them. • The level of challenge of grammar exercises in MEL should be reviewed to reduce the need for teachers to supplement these activities. • The pronunciation coach videos/ exercises could be reviewed. Students suggested that the accents in the pronunciation exercises could sometimes be difficult to understand and 36% (71/195) did not find the pronunciation coach videos useful.

Teacher perceptions of the impact of Speakout with MEL on teachers and teaching

The Students' Book was perceived to be accessible by teachers and engaging to them and students. Teachers seemed to like the subject matter of the units, their structure and that activities covered all the skills including speaking, reading, and writing.

Teachers also suggested that the Students' Book supported monitoring and assessment, particularly of speaking. One teacher noted that the amount of speaking activities meant that teachers were monitoring students' speaking ability more than they were able to do with the book they used before Speakout with MEL. Teachers in two institutions also thought that the language bank supported revision activities in the run-up to assessments.

However, there seemed to be a view that Speakout might not be as technology oriented and challenging as other English courses. As mentioned elsewhere in this report, Speakout's grammar content was seen as problematic. Language Leader was favoured as a more technology oriented and academic book.

The Teacher's Book was used by all teachers, and many of its features were considered to be useful. Several teachers suggested that the background information or 'notes' and suggestions for further activities, such as games, supported and diversified their teaching. The test bank was used to create internal, customised tests. Teachers did not use the tests in the guide in their entirety because students could obtain the answer keys. Some teachers felt, however, that even the test bank tests were relatively easy for some groups of students. Table B21 outlines the perceived impact of MEL on teachers and teaching.

Table B21: Teacher perceptions of MEL's impact on teachers and teaching

Evidence of impact and enablers	Suggested improvements
• Interviewees suggested that MEL reduced teacher workload and supported monitoring and assessment. They thought that MEL freed up their time from marking homework and tests; they did not have to print exercises or tests anymore; and had to provide less feedback to students, since MEL provided it for them. • Interviewees suggested that MEL also supported improvements in teaching because it freed time to support the development of English skills, such as speaking and technical vocabulary, which might have 'suffered' otherwise. It also allowed for other teaching and learning approaches, such as games during contact time, which made learning more fun. • Interviewees suggested that MEL Version 2 allowed all students to automatically be assigned to a task, reducing the time needed to assign practices. • Interviewees noted how the gradebook served several purposes. It allowed teachers to note who was accessing MEL and keep track of scores, for example. This was seen as particularly useful in understand students' weaknesses and providing extra tasks on MEL, or in regular or revision classes.	• Teachers who perceived the autoscoring system as being 'strict', and felt that students' scores needed altering, spent additional time in reviewing scores and providing feedback. • Teachers suggested that a quicker way to identify issues, in particular assignments at a student level, would be welcomed, as currently they perceived it to be time-consuming.

Pearson

Appendix C: Rasch fit statistics for the five questions of the motivation scale

	parameter	Outfit	Outfit_t	Outfit_p	Outfit_pholm	Infit	Infit_t	Infit_p	Infit_pholm
1	Learning	0.78	-1.76	0.08	0.47	0.78	-2.22	0.03	0.11
2	Using	1.01	0.07	0.95	1.00	1.02	0.22	0.82	1.00
3	Succeeding	0.90	-0.85	0.40	1.00	0.93	-0.72	0.47	1.00
4	Completing_level	1.29	2.48	0.01	0.09	1.30	2.75	0.01	0.03
5	Achieving_goal	0.93	-0.57	0.57	1.00	0.97	-0.25	0.81	1.00
6	Cat1	1.54	1.29	0.20	0.79	1.41	7.98	0.00	0.00
7	Cat2	1.39	1.68	0.09	0.47	1.41	13.76	0.00	0.00

Pearson

Appendix D: Percentage of students achieving within different score bands in practices per institution when 0% scores are included in the analysis.

Table D1: Percentage of students achieving within different score bands in MEL practices (zeros not removed)

Performance band	% of students		
	WS	AWF	Wyższa Szkoła Logistyki
Below 50%	–	3 (0.9%)	–
50%–59%	–	7 (2.1%)	3 (3.3%)
60%–69%	2 (1.0%)	11 (3.4%)	4 (4.3%)
70%–79%	7 (2.3%)	40 (12.3%)	10 (10.9%)
80%–89%	40 (13.1%)	108 (33.1%)	22 (23.9%)
90%–100%	257 (83.7%)	157 (48.2%)	53 (57.6%)

Table D2: Percentage of students achieving within different score bands in tests and assignments—MEL data (zeros not removed)

Performance band	% of students		
	AWF (test scores)	Wyższa Szkoła Logistyki (test scores)	Wyższa Szkoła Logistyki (assignment scores)
Below 50%	143 (43.6%)	17 (32.7%)	23 (39%)
50%–59%	33 (10.1%)	5 (9.6%)	8 (13.6%)
60%–69%	49 (14.9%)	8 (15.4%)	7 (11.9%)
70%–79%	51 (15.5%)	7 (13.5%)	6 (10.2%)
80%–89%	37 (11.3%)	6 (11.6%)	8 (13.6%)
90%–100%	15 (4.6%)	9 (17.3%)	7 (11.9%)

Pearson

Appendix E: Student progress on practices and assignments per institution and course, MEL data

Table E1: Student progress on practices and assignments per institution and course —MEL data

	Progress on practices	Progress on assignments
WS		
Elementary (2015)	38 (60→98)	–
Pre-intermediate (2015)	33 (64→97)	–
Intermediate (2015)	30 (67→97)	–
AWF		
Intermediate (2014)	29 (60→89)	–
Intermediate (2015)	30 (58→88)	–
Intermediate (2016)	30 (60→90)	–
Wyższa Szkoła Logistyki		
Pre-intermediate (2014)	20 (74→94)	–
Advanced (2016)	23 (61→84)	–
Pre-intermediate	26 (61→87)*	28 (59/87)*
Pre-intermediate (2014/16)	23 (68→91)	–