

Handbook of Assessment for Language Teachers

Dina Tsagari
Karin Vogt, Veronika Froehlich
Ildikó Csépes, Adrienn Fekete
Anthony Green, Liz Hamp-Lyons
Nicos Sifakis, Stefania Kordia



Preface

There is evidence that language testing and assessment can have a strong impact on the quality of the learning outcomes (Green, 2013; Tsagari & Cheng, 2016). Therefore, it is important to enhance teachers' level of Language Assessment Literacy (LAL) (Inbar-Lourie, 2008), that is, their "ability to design, develop and critically evaluate tests and other assessment procedures, as well as the ability to monitor, evaluate, grade and score assessments on the basis of theoretical knowledge" (Vogt & Tsagari, 2014, p. 377). However, research (Vogt and Tsagari, 2014; Tsagari and Vogt, 2017) has shown that English language teachers often lack sufficient understanding of the nature of LAL and are not familiar with the relevant assessment techniques and methods to further the language learning experiences of their students.

We therefore developed an online, self-study (Moodle-based), training course (<http://taleproject.eu>) that can be used by pre- and in-service teachers of English (and other foreign) languages as well as teacher trainers to raise their awareness and levels of LAL.

The training materials included in this Handbook are based on the online training course and consist of eight chapters plus a glossary, all of which cover a variety of theoretical and practical issues. Each chapter comprises several tasks which acquaint teachers and trainers with a wide repertoire of useful assessment methods, techniques and tools so that they can enhance their language assessment literacy.

References

- Green, T. (2013). Washback in Language Assessment. *IJES*, 13(2), 39-51. DOI: <http://dx.doi.org/10.6018/ijes.13.2.185891>
- Inbar-Lourie, O. (2008). Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses. *Language Testing*, 25(3), 328–402. DOI: 10.1177/0265532208090158
- Tsagari, D. & Cheng, L. (2016). 'Washback, Impact and Consequences'. In E. Shohamy and N. H. Hornberger (eds.), *Encyclopedia of Language and Education*, 3rd Edition, Volume 7: Language Testing and Assessment, pp. 1-13. Educational Linguistics: Springer. DOI:10.1007/978-3-319-02326-7_24-1
- Tsagari, D. & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 18-40.
- Vogt, K. & Tsagari, D. (2014). Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Language Assessment Quarterly* 11(4), 374-402. DOI: 10.1080/15434303.2014.960046.

©TALE course training materials were developed in the TALE project (co-financed by the European Commission within the Erasmus+ Programme, project number: 2015-1-CY01-KA201-011863).

For information and online materials of the TALE project, please visit: <http://taleproject.eu>

Project Coordinator and Partners

TALE Project



TALE Coordinators

University of Cyprus

2015-2017

Department of English Studies

Dina Tsagari

dinatsa@ucy.ac.cy

<http://www.ucy.ac.cy/eng/en/>

2017-2018

Department of Social and Political Sciences

Iasonas Lamprianou

Lamprianou@ucy.ac.cy

<http://www.ucy.ac.cy/sap/en/>

<http://www.ucy.ac.cy/en/>



Erasmus+ Programme, KA2 Strategic Partnership, Project number: 2015-1-CY01-KA201-011863.

This publication and project has been funded with support from the European Commission.

This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

ISBN 978-9925-7399-0-5 (printed)

ISBN 978-9925-7399-1-2 (digital)

TALE Partners

University of Heidelberg

Padagogische Hochschule Heidelberg

Karin Vogt

vogt@ph-heidelberg.de

<https://www.ph-heidelberg.de/englisch/aktuell.html>

<http://www.ph-heidelberg.de>

University of Debrecen

Institute of English and American Studies

Ildikó Csépes

csepes.ildiko@arts.unideb.hu

<http://ieas.unideb.hu/>

<http://www.unideb.hu>

Hellenic Open University

Faculty of Humanities

Nicos Sifakis

sifakis@eap.gr

<https://www.eap.gr/en/>

University of Bedfordshire

CRELLA

Anthony Green

tony.green@beds.ac.uk

<https://www.beds.ac.uk/crella>

OsloMet – Oslo Metropolitan University - formerly known as Høgskolen i Oslo og Akershus (2017-2018)

Department of Primary and Secondary Education

Dina Tsagari

Dina.Tsagari@hioa.no

<http://www.hioa.no/eng/About-HiOA/Faculty-of-Education-and-International-Studies>

<http://www.hioa.no/>

The following were involved in and supported the development of the original ©TALE materials:

Cyprus

● **Department of Social and Political Sciences, University of Cyprus**

Iasonas Lamprianou

● **Department of English Studies, University of Cyprus**

Dina Tsagari, Anna Mavroudi, Claudia Konyalian, Marina Hadjicharalambous, Maria Doukanari, Christina Giannikas

Germany

University of Education Heidelberg

Karin Vogt, Veronika Froehlich

Greece

School of Humanities, Hellenic Open University

Nicos Sifakis, Stefania Kordia

Hungary

Department of English and American Studies, University of Debrecen

Ildikó Csépes, Adrienn Fekete

UK

Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire

Anthony Green, Liz Hamp-Lyons

The following were involved in reviewing the ©TALE Handbook materials:

Germany

University of Education Heidelberg

Karin Vogt, Veronika Froehlich, Ines Sperling

Greece

School of Humanities, Hellenic Open University

Nicos Sifakis, Stefania Kordia

Hungary

Department of English and American Studies, University of Debrecen

Ildikó Csépes, Adrienn Fekete

Norway

Department of Primary and Secondary Education, OsloMet (formerly HiOA)

Dina Tsagari, Therese Marie Tishakov, Siri Mohamad-Roe, Kirstin Reed

UK

Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire

Anthony Green, Liz Hamp-Lyons

Special thanks to all the teachers and teacher trainers who piloted the TALE materials!

TABLE OF CONTENTS:

1.The ABCs of Assessment	1
1.1 Topic 1: Purposes of Assessment	2
Task 1. Purposes of assement	3
Task 2. Approaches to assessment	4
1.2 Topic 2: Types of Assessment	5
Task 1. Assessment tasks	5
Task 2. Assessment task specifications	7
1.3 Topic 3: Recording and Reporting Outcomes	8
Task 1. Systems for recording judgements	8
Task 2. Assessment feedback	9
1.4 Topic 4: Qualities of Assessment	10
Task 1. Thinking about qualities	10
Task 2. Assessment quotes	10
Task 3. Practical steps	11
The ABCs of Assessment - Outcomes	12
Indicative Answers	12
Further Reading	14
References	15
2. Assessing Reading Skills	16
2.1 Topic 1: Introduction to Assessing Reading	17
Task 1. How do we read and what needs to be considered when assessing reading?	17
2.2 Topic 2: Reading Skills and the CEFR	21
Task 1. Familiarization with the CEFR levels	21
2.3 Topic 3: Techniques for Assessing Reading	26
Task 1. Guidelines for designing specific task types to assess reading	26
Task 2. Informal assessment of reading in the language classroom	29
2.4 Topic 4: Evaluating and Designing Reading Tests	30
Task 1. Guidelines for designing specific task types to assess reading	30
Task 2. Designing a reading test task	37
Assessing Reading Skills - Outcomes	39
Indicative Answers	39
Further Reading	45
References	45

3. Assessing Writing Skills	46
3.1 Topic 1: Understanding Writing	47
Task 1. Introduction	47
Task 2. The teacher as a writer	51
3.2 Topic 2: Types and Purposes of Writing Assessment	54
Task 1. Qualities of writing assessments in the classroom	54
Task 2. Writers at the centre of classroom assessments	55
3.3 Topic 3: Four Key Components of Writing Tests	57
Task 1. Designing writing tasks	57
Task 2. Judging what has been written - The scoring procedure	61
3.4 Topic 4: Why the Readers are Important	62
Task 1. Being a teacher-rater	63
Task 2. Designing a rubric	63
Task 3. Feedback on the three texts from different perspectives	69
Task 4. A role for computers?	71
Assessing Writing Skills - Outcomes	74
Indicative Answers	74
Further Reading	78
References	79
4. Assessing Listening Skills	81
4.1 Topic 1: What is Listening	82
Task 1. Listening experiences	82
Task 2. Kinds of listening text	83
Task 3. Students' listening needs	84
4.2 Topic 2: Listening Types	86
Task 1. Types of listening	86
Task 2. Kinds of text and types of listening	86
Task 3. Factors that add difficulty	87
Task 4. The Common European Framework of Reference levels	88
4.3 Topic 3: Finding Materials for Listening	89
Task 1. Finding or making recordings	89
Task 2. Locating recordings on the internet	90
Task 3. Making your own recordings	92
4.4 Topic 4: Using Recordings for Assessment	93
Task 1. Dialogues	93
Task 2. Monologues	93
Assessing Listening Skills - Outcomes	94
Indicative Answers	94
Further Reading	97

5. Assessing Speaking Skills	98
5.1 Topic 1: Introduction to Assessing Speaking	99
Task 1. What characterises and influences speaking performances?	99
Task 2. Speaking skills and the CEFR	103
5.2. Topic 2: Designing and Evaluating Speaking Tests	107
Task 1. Awareness raising of general issues in designing speaking test tasks	107
Task 2. Designing speaking assessment tasks	115
5.3. Topic 3: Giving Feedback to Enhance Learning	115
Task 1. Oral feedback and its effectiveness	115
5.4. Topic 4: Evaluating Speaking Performances	116
Task 1. Scoring benchmarked sample performances	116
Task 2. Teacher assessment vs. peer-assessment and self-assessment	121
Assessing Speaking Skills - Outcomes	123
Indicative Answers	124
Further Reading	127
References	128
6. Providing Feedback	129
Introductory Activity	130
6.1. Topic 1: Feedback for Formative Assessment	130
Task 1. The Role of Feedback in the Learning Process	130
6.2. Topic 2: Feedback Types	133
Task 1. The Role of Feedback in the Learning Process	133
Task 2. Conferencing	135
Task 3. Feedback Sheets/Rubrics	137
6.3. Topic 3: Effective Teacher Feedback	138
Task 1. Characteristics of Effective Teacher Feedback	138
Providing Feedback - Outcomes	143
Indicative Answers	143
Further Reading	146
References	147

7. Alternatives in Assessment	148
7.1. Topic 1: Introduction to Alternatives in Assessment	149
Task 1. Contrasting Traditional and Alternative Teacher Assessment	149
7.2. Topic 2: Teacher Observation	153
Task 1. Teacher Observation as a Form of Assessment	153
7.3. Topic 3: Self- and Peer Assessment	161
Task 1. Using Self- and Peer-Assessment in the Foreign Language Classroom	161
7.4. Topic 3: Portfolio Assessment	166
Task 1. Introduction to Portfolio Assessment in the EFL Classroom	166
Task 2. How to introduce Portfolio Assessment in the EFL Classroom	167
Task 3. The European Language Portfolio	171
Alternatives in Assessment - Outcomes	172
Indicative Answers	173
Further Reading	177
References	178
8. Test Impact	179
8.1. Topic 1: Essential Clarifications	180
Task 1. Using and learning English	180
Task 2. Teaching, assessing and testing English	183
8.2. Topic 2: Defining 'Test Impact'	187
Task 1. What do we mean by 'test impact'?	187
8.3. Topic 3: Understanding 'Washback'	191
Task 1. Key dimensions of the washback effect	191
Task 2. Understanding washback in practice	196
8.4. Topic 4: Broader Consequences of Washback on Teaching and Learning	201
Task 1. The culture of testing	201
Task 2. Where to go from here?	204
Test Impact - Outcomes	205
Indicative Answers	206
Further Reading	208
References	209
Glossary	211

Introduction

Teachers of English Language (ELTs) and teachers of other languages deal with various assessment procedures in their professional lives on a daily basis, for example, they organise and administer classroom assessment activities, use various forms of continuous or formative assessment or are faced with external testing procedures (e.g. school-leaving examinations and international standardised tests). Furthermore, new developments in foreign language teaching as well as European policies on language learning require new competencies of teachers (e.g. the European Language Portfolio, self- and peer-assessment, etc). However, research to date (Fulcher 2012; Jin, 2010, Tsagari & Vogt, 2017; Vogt & Tsagari, 2014) has shown that ELTs have expressed a strong need for training in the area of language testing and assessment (LTA), referred to as 'language assessment literacy', LAL (Vogt & Tsagari, 2014, p. 377). Awareness, therefore, is growing across Europe that there is an urgent need to develop an efficient, relevant, and sustainable LAL training infrastructure for language teachers.

Consisting of a network of experts from six different European countries (Cyprus, Greece, Hungary, Germany, Norway and the UK), the Teachers' Assessment Literacy Enhancement (TALE) project provides a LAL training resource built on an empirical identification of the LAL training needs of pre- and in-service language teachers and trainers across many European countries. The TALE project offers innovative LAL training materials and services (e.g. forum) that language practitioners can access, either through the online mode (<http://taleproject.eu/>), or through the traditional paper version (this Handbook). They can access the online version in either synchronous or asynchronous modes depending on context-specific and accessibility needs.

The materials in this Handbook are the product of two phases: First, a comprehensive literature review on current thinking in LAL and an extensive consultation with language teachers were conducted to ensure that the TALE course would meet their needs. The results of this phase were used as input in the second phase, which involved the development and piloting of online training materials. The pilot feedback was given careful consideration and revisions were made as necessary. Once the project partners had worked with and refined the materials and the evaluation process was complete, the current Handbook was drafted, refined, reviewed and finally published to enable and promote wider implementation of the project beyond the partner organisations. The Handbook contains LAL activities and tasks derived from the implementation of the online pilot training course aiming at developing and enhancing teachers' LAL levels.

The Handbook materials as well as the online course materials are based on distance learning methodology that allows teachers to fit their learning around their work and home lives and to set their own pace of study. They are specifically designed with the teacher as a reflective practitioner in mind (Wallace, 1991). This means that when using the TALE materials, teachers are expected to adapt them to their contexts and then reflect on the outcomes of their assessment processes. Teacher trainers using these materials as part of their teacher training and development programmes are expected to act as facilitators to raise trainees' awareness of language assessment, give possible answers to teachers' questions and provide feedback.

The complete course takes a task-based approach and provides hands-on tasks that contribute to raising teachers' awareness of LAL issues in language classrooms. Every chapter follows a reflective cycle in which teachers draw on their experiences, reflect on their current practices, understanding, attitudes and working contexts. Then they are offered input through various modes, short texts, links to videos, etc. related to reflective tasks. At the end of the chapters, teachers are invited to reflect on what they have learnt.

The TALE course is designed to be flexible and adaptable to match the needs of language teachers in different European and other countries. The complete TALE course consists of 8 chapters which could be covered in a semester/term if used as part of university programmes, e.g. in 8-10 or even 15 weeks in countries/institutions where a term lasts for 10 or 15 weeks respectively. Alternatively, the TALE course materials can be used as part of individual self-study taken at one's own pace and time.

The course includes the following chapters:

1. The ABCs of assessment
2. Assessing reading skills
3. Assessing writing skills
4. Assessing listening skills
5. Assessing speaking skills
6. Providing feedback
7. Alternatives in assessment
8. Test impact

The order of use of the courses is up to the users, but we recommend the above order as courses follow a logical order built on each other. All the chapters have optional tasks which allow the adaptation of the course to form a longer programme of study. TALE chapter materials also offer suggestions to those teachers who also need to assess students with Special Learning Difficulties (SpLDs). Finally, users can refer to the Glossary for the definition of some basic terms.

We hope that you enjoy the TALE materials and find them useful for your assessment purposes!

References

- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. DOI: 10.1080/15434303.2011.642041.
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555-584. DOI: 10.1177/0265532209351431.
- Tsagari, D. & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 18-40.
- Vogt, K. & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402. DOI: 10.1080/15434303.2014.960046.
- Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach*. Cambridge: Cambridge University Press.

1. The ABCs of Assessment

Anthony Green

❖ Aims

This Course introduces some of the concepts you will encounter in the other TALE Courses. It should help you to reflect on the purposes of assessment, the techniques used by language teachers and the essential qualities of effective language assessments.

❖ Expected Outcomes

Upon completion of this Course, you will be able to:

1. Distinguish between tests and other forms of assessment
2. Identify the main purposes of assessment in the classroom
3. Identify the essential qualities of good assessments
4. Identify methods for reporting on performance
5. Summarize characteristics of assessment tasks

❖ Key Concepts

- [Diagnosis](#)
- [Formative assessment](#)
- Test impact
- Placement
- [Practicality](#)
- Progress
- [Reliability](#)
- Reporting results
- Selection
- Specifications
- [Summative assessment](#)
- [Validity](#)

Brief Introduction

This Course introduces **key concepts** that appear in the other TALE Courses. Please look at the list of key concepts in the previous section. If you already have a good understanding of these terms and concepts, you may choose to skip this Course and go directly to other Courses that interest you.

Four Topics are included in this Course. The first is about **different purposes of assessment**. The second Topic presents some of the options available for **creating assessment tasks**. The third covers **recording and reporting** how learners have performed. The fourth Topic is about the **qualities of useful assessments** and how to improve the assessments you make.

You may wish to study the materials included in each Topic, or only those you feel you need to be informed about.

There are **several Activities** throughout each Topic.

Introductory Activity: Your experience

Think about your experience as a teacher, or as a student. Which of these quotes from teachers best describes your feelings about assessment? You may choose more than one answer.

1. Assessment helps to show the teacher who is the strongest and who is the weakest learner in class so that they can give suitable grades or marks.
2. Assessment is one of the most rewarding aspects of teaching: it helps to show how well learners understand what they are studying and how to improve their performance.
3. I think teaching would be more effective if teachers didn't have to give assessments: they use up valuable class time that would be better spent on other things.
4. Assessment is a good way to motivate learners because they will study harder if they know a topic is going to count towards their final grade.
5. Assessment in schools is dominated by external or public examinations: learners care more about the exams than actually learning English.
6. The main point of assessment is to allow school managers or government authorities to judge whether teachers are doing a good job and whether learners have learnt what they are supposed to.
7. None of these quotes describes my feelings.

1.1 Topic 1: Purposes of Assessment



There are **two Tasks** in this Topic. You will read about the **purposes of assessment** in language education and then focus on **assessments** used by teachers in the language classroom.

In education, assessment involves **collecting evidence and making judgements or forming opinions about learners' knowledge skills and abilities**. It often also involves keeping an **informal or formal record** of those judgements. It is a key professional responsibility of all teachers to become effective at assessment. Unless teachers are able to judge what learners know or can do, they neither reasonably decide whether or not their teaching has been successful nor can they choose what to do next to help learners to improve.

Task 1. Purposes of assessment

Why do you **assess** language learners?

Teachers usually assess language learners because they need to make decisions about these learners. The evidence that they collect helps them to make good decisions. For example, teachers may need to decide whether all of the students have successfully learnt the material they just taught them.

Activity 1 - Decisions

Can you think of **five different decisions** that teachers may need to make about language learners for which an assessment might help to inform the decision?

Write them below.

Five assessment decisions
1.
2.
3.
4.
5.

Activity 2 - Assessment terms

How familiar are you with the various **terms** that are used in assessment?

Here is a list of some decisions that people make using language assessments. Can you fill in the gaps in each sentence (1-4) with the correct terms for the function of the assessment (A, B, C or D)?

A. diagnostic	C. progress
B. placement	D. selection

1. When a school head uses an assessment to find out learners' levels of language ability to assign them to the most suitable class group, this is a(n) _____ function.
2. When a college uses results from an assessment to decide which students to admit to the course and which to reject, this is a(n) _____ function.

3. When a teacher gives an assessment to find out which English sounds are difficult for learners and to choose useful practice exercises, this is a(n) _____ function.
4. When a teacher uses an assessment to find out how much the learners in her class have learned (and so where she needs to review material), this is a(n) _____ function.

Suggested answers are provided in the **Indicative Answers** section at the end of this Course.

Learn more about the role of assessment in education systems on the Course focusing on **Test Impact**.

Task 2. Approaches to assessment

Teachers have **two main purposes** for assessing the learners in their classes.

One purpose is to **improve learning** by checking that learners are progressing. They do this so that they can decide whether to give additional help, try a different explanation or use different materials when learners find things difficult, or whether to provide more challenging activities when learners are ready for these.

The other purpose is to **judge how successful** learners have been in mastering the content of a course in order to report this to parents, school management or educational authorities. This usually involves deciding on grades or scores.

The first of these purposes is called **formative assessment** or **assessment for learning**. The second is called **summative assessment** or **assessment of learning**.



Please watch the video accessible via the link below. Then, carry out the Activity that follows.

Video: <https://vimeo.com/187545455>

Activity 3 - Formative or summative?

Look at the following examples of assessment activities. Decide whether you think these are examples of **formative** assessment or **summative** assessment.

Circle your answer.

1. Learners read a text and write a summary of the text in 50 words. Groups of learners compare their summaries and produce a shared group version. Finally, they compare the group summaries with one written by the teacher. **formative or summative?**

2. Learners listen to a recording and answer multiple choice questions about it. The teacher asks the learners to raise their hands when they answered a, b or c to each question. She asks individual learners to explain why they chose their answer. **formative or summative?**

3. Learners choose a topic and prepare a presentation about it. They give the presentation to the other learners in the class. The teacher gives them a score for the quality of their presentation. **formative or summative?**

4. The teacher walks around the class while the learners carry out a grammar exercise on the use of adjectives. She notes the questions to which several of the learners gave incorrect answers. She spends three minutes reviewing uses of adjectives, then asks the learners to look at the questions again. **formative or summative?**

5. The teacher gives a ten question quiz each week on vocabulary that students have studied and notes the scores in her grade book. **formative or summative?**

1.2 Topic 2: Types of Assessment



There are **two Tasks** in this Topic. You will read about different kinds of task that teachers and learners can use in the classroom to find out about learners' abilities. It is important that people are able to describe them in the same ways.

Assessment is not the same thing as testing, grading or scoring. Tests are a particular kind of formal assessment procedure carried out under controlled conditions. There are many other kinds of assessment that teachers can use (ranging from observing learners talking in class to extended projects and portfolios). You can find out more about the choices open to teachers in the Course on **Alternatives in Assessment**.

Task 1. Assessment tasks

In classrooms, when teachers are going to make judgements about learners' language abilities, they first need to collect some evidence. Teachers generate evidence by giving learners **tasks** that involve **language use**. Although, as you will see in these Courses, there is a wide variety of tasks that learners can be asked to carry out, these can all be described in relation to a set of common features.

Tasks have three parts:

Instructions: guidance telling learners what they need to do to complete the task

Input: the material that learners need to look at/read/listen to in order to respond

Expected response: what the learner is expected to do in response to the input to demonstrate his or her language ability.

Activity 1 - Parts of a task

Try this **quiz** about these parts of assessment tasks.

Look at items 1 to 9 below. These are parts of three different assessment tasks. One task assesses **speaking**, the second **reading** and the third, **writing** skills. Can you identify the **instructions**, the **input** and the examples of the **expected response** for each task?

Circle your answer.

- | | | |
|---|---|--|
| 1 | Describe the picture. You have 10 minutes to write your answer. | Instructions
Input
Expected response |
| 2 | Read the text and answer the questions below (5 minutes). | Instructions
Input
Expected response |
| 3 | Why was the man digging in his garden?
A. He was trying to find treasure. <input type="checkbox"/>
B. He wanted to plant some bushes and flowers. <input type="checkbox"/>
C. He was planning to create a pool of water. <input checked="" type="checkbox"/> | Instructions
Input
Expected response |
| 4 | · wake up
· come to school
· have lunch | Instructions
Input
Expected response |
| 5 | A: <i>What time do you have lunch?</i>
B: <i>I have lunch at one o'clock. How about you?</i> | Instructions
Input
Expected response |
| 6 | Work with a partner. Ask and answer what time you usually do these things on weekdays. | Instructions
Input
Expected response |
| 7 | <u>Some children are planting a tree.</u> | Instructions
Input
Expected response |
| 8 | In 2007, a man known only as Andreas K. from Austria was digging in his backyard to make a small pond. While digging, the lucky man found a 650 year-old treasure that included more than 200 rings, belt buckles, gold-plated silver plates and other pieces | Instructions
Input
Expected response |



Instructions
Input
Expected response

Task 2. Assessment task specifications

It is very important to have a shared terminology for describing assessment tasks because this helps teachers to develop **assessment task specifications**.

Specifications are documents that work like recipes for producing sets of assessment tasks. Tasks that are all written based on the same specifications will probably all require learners to use similar skills and will probably be of a similar level of difficulty.

Describing **instructions**, **input** and **expected responses** can help us to prepare specifications.

Compare the three assessment tasks you saw in the previous Activity. Think how this information would help a writer to produce other, similar tasks.

Activity 2 - Compare the tasks from Task 1

Compare the three assessment tasks in Task 1. In what ways are the **inputs** different to each other?

What differences do you notice between the **instructions**?

How are the **expected responses** different?

Which task do you think learners would find most difficult? Why?

Learn more from the Courses on **Reading, Writing, Listening, Speaking** and **Alternatives in Assessment**.

1.3 Topic 3: Recording and Reporting Outcomes



There are **two Tasks** in this Topic. You will read about ways to **report** how learners have performed on language assessments and then look at how **feedback** can be used to help learners improve their performance.

Grades and **scores** are formal ways to document how well learners have performed on an assessment. They are often used to report to people outside the classroom (school managers, parents, governments) about how successful learners have been.

In fact, there are many other ways to record judgements about performance. These can range from noting some of the common errors made by learners to holding a student-teacher or student-parent-teacher **conference** where learners give an account of their own progress. Take a look at this example of a student-parent-teacher meeting:



Please watch the video accessible via the link below.

Video: https://www.youtube.com/watch?v=L_WBSInDc2E

Techniques like the one shown in the video can be more useful than grades for teachers and learners.

Task 1. Systems for recording judgements

Activity 1 - Systems for recording judgements

In the school where you work, or where you study, what **systems** do teachers generally use to **record** their judgements of learner performance? Mark your answers with a tick (✓).

1. Marks or scores (percentages/points)
2. Letter grades (A/B/C)
3. Written comments on learners' work
4. 'Can Do' descriptions and checklists (Can understand simple sentences, e.g. 'Throw the ball')
5. Written progress reports
6. Regular meetings between teachers and individual learners to discuss progress

What do you think are the main advantages and disadvantages of each system?

Which systems do you think can be most useful in helping to improve **language learning**? Why?

Task 2. Assessment feedback

Perhaps the most important reason why we teach people languages is so that they can become better language users. Each time we give them an assessment, we have a chance to help them to improve their performance.

For learners to know how to improve, Grant Wiggins (1998) has suggested that they need to understand three things about their performance:

- what they are doing **well**
- what they are **struggling with**
- what they need to **do differently** in order to be **more successful**

The teacher can help learners to understand these three things by giving them **feedback** that highlights ways to improve. But just telling the learner what she did correctly and incorrectly is not usually very effective. It does not often lead to improvement.

Read the following **examples of feedback** given by a teacher on a learner's essay. Which comments do you think would help the learner to **improve** the essay?

Activity 2 - Effective feedback?

1. This essay is too short. You need to write a lot more to get top marks.
 - a) they are useful in guiding improvement
 - b) they provide minimal guidance for improvement
 - c) there is no guidance for improvement
2. This is a really great essay. Good work!
 - a) they are useful in guiding improvement
 - b) they provide minimal guidance for improvement
 - c) there is no guidance for improvement
3. Your point might be clearer if you first state the view you don't agree with, then explain the weakness in that argument.
 - a) they are useful in guiding improvement
 - b) they provide minimal guidance for improvement
 - c) there is no guidance for improvement
4. You made a mistake with the past tense in line 3: you should have written 'prices fell, but then rose again' not 'prices have fallen, but then rose again'.
 - a) they are useful in guiding improvement
 - b) they provide minimal guidance for improvement
 - c) there is no guidance for improvement

Learn more by taking our Course on **Providing Feedback**.

1.4 Topic 4: Qualities of Assessment



There are **three Tasks** in this Topic. You will watch a short video about the essential qualities of language assessments and connect these to comments and criticisms that people often make about assessments. Finally, you will read a short article discussing ways in which teachers can improve the assessments they use.

Task 1. Thinking about qualities

Think about the following question:

What do you think are four essential qualities of good assessments?

Write down your ideas.

My ideas

Activity 1 - Qualities of assessments

Watch a short video from Cambridge Assessment about the qualities of good assessments.

Did the speaker mention your ideas from Activity 1? Would you add anything to the list?



Please watch the video titled 'Video Shorts #5 Validity, reliability, impact and practicality' accessible via the link below.

Video: <https://www.youtube.com/watch?v=DsgORlf2lDg>

Task 2. Assessment quotes

Read the following **four quotes** from people talking about assessment. Which assessment quality are they talking about in each case? Circle the quality you think they refer to each time. The correct answers are provided in the **Indicative Answers** section.

Activity 2 - Assessment quotes

1. This new assessment procedure looks great, but I can't use it in my classes. It takes two hours to complete and my lessons only last for 45 minutes!
 - a. Validity
 - b. Reliability
 - c. Practicality
 - d. Impact
2. Look at this question: 'Julia Gillard (a) **is**, (b) **was** or (c) **will be** the first woman to be Prime Minister of Australia'. It's supposed to test knowledge of tenses, but I think you need to know about Australia to answer.
 - a. Validity
 - b. Reliability
 - c. Practicality
 - d. Impact
3. My daughter was given an 'F' grade by one teacher for her assignment. She was upset so I showed her work to another teacher of the same year group. He said he thought it should be given a 'B'. I don't know who to believe.
 - a. Validity
 - b. Reliability
 - c. Practicality
 - d. Impact
4. I have an English test tomorrow and I'm really worried and nervous. I cried last night because I'm so scared.
 - a. Validity
 - b. Reliability
 - c. Practicality
 - d. Impact

Task 3. Practical steps

What **practical steps** do you think teachers in your school could take that might make their assessments **better**?

Activity 3 - Improving assessment

Now read a **short article** (2½ pages) by Stephen Bax.



You can find the article, titled 'Integrating assessment into EFL teaching and learning environments', at this website:

<https://tinyurl.com/baxabc>




Activity 4: Reflection

Compare the suggestions in the article with your ideas. Were they the same?

Learn more about the qualities of effective assessments from the other TALE Courses.

The ABCs of Assessment - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
Distinguish between tests and other forms of assessment			
Identify the main purposes of assessment in the classroom			
Identify the essential qualities of good assessments			
Identify methods for reporting on performance			
Summarize characteristics of assessment tasks			

Indicative Answers

Topic 1 Task 1 Activity 2

Here is a list of some decisions that people make using language assessments. Can you fill in the gaps in each sentence (1 - 4) with the correct terms for the function of the assessment (A, B, C or D)?

The correct answer is:

1. When a school head uses an assessment to find out learners' levels of language ability to assign them to the most suitable class group, this is a(n) [B placement] function.
 2. When a college uses results from an assessment to decide which students to admit to the course and which to reject, this is a(n) [D selection] function.
 3. When a teacher gives an assessment to find out which English sounds are difficult for learners and to choose useful practice exercises, this is a(n) [A diagnostic] function.
 4. When a teacher uses an assessment to find out how much the learners in her class have learned (and so where she needs to review material), this is a(n) [C progress] function.
- A. Diagnostic assessments are used to help identify learners' strengths and weaknesses. Teachers can use diagnostic assessments to decide what learners need (and don't need) to work on.
- B. Placement assessments are used by schools to decide how to group learners, usually so that classes are made up of learners with similar levels of ability.
- C. Progress assessments are used to find out how far learners are meeting objectives. Teachers use progress assessments to help them decide what they need to review with learners. Progress assessments are often also diagnostic: teachers want to learn about differences between levels of progress on different objectives.

D. Selection assessments are used to decide which applicants should be given jobs or accepted onto courses.

Topic 4: Activity 2 - Assessment quotes

- | | |
|--|--|
| 1. This new assessment procedure looks great, but I can't use it in my classes. It takes two hours to complete and my lessons only last for 45 minutes! | This is a matter of practicality. If the assessment takes more time than the teacher has available, it is not practical to use it. |
| 2. Look at this question: 'Julia Gillard (a) is , (b) was or (c) will be the first woman to be Prime Minister of Australia'. It's supposed to test knowledge of tenses, but I think you need to know about Australia to answer. | The problem here is validity. You need general knowledge to answer the question. People who have a good knowledge of tenses may get the answer wrong because they don't know about Australian history. |
| 3. My daughter was given an 'F' grade by one teacher for her assignment. She was upset so I showed her work to another teacher of the same year group. He said he thought it should be given a 'B'. I don't know who to believe. | This is a matter of reliability: the learner gets different scores from each teacher for the same performance |
| 4. I have an English test tomorrow and I'm really worried and nervous. I cried last night because I'm so scared. | This is a question of impact. Tests can make some people feel acutely anxious, especially young children |

Topic 3 Activity 2

1. The suggested answer is: **b**. Minimal guidance for improvement

This comment gives minimal guidance for improvement. It does not give the learner enough information on how to improve – it would be more helpful if the teacher suggested the kinds of information the learner could add to make the essay longer: 'In your essay you say that people watch too much TV. Why do you think so? Add two reasons or examples to support your argument.'

2. The suggested answer is: **c**. No guidance for improvement

This comment gives no guidance for improvement. It praises the learner, but it does not show why the essay is good ('The essay was very clear because you gave reasons to support all your arguments') or how the learner can improve his or her work: 'Check your essay to make sure that your use of tenses is correct – then show it to me again.'

3. The suggested answer is: **a**. Useful in guiding improvement

This comment is potentially useful in guiding improvement as it gives a specific suggestion on what the learner can do to make the argument clearer. However, it will only be really helpful if the learner has a chance to act on it by revising the essay: 'Add a paragraph about an idea you disagree with. Then show me your essay again.'

4. The suggested answer is: **b**. Minimal guidance for improvement

This comment gives minimal guidance for improvement as it does not leave any thinking for the learner to do. It might have been more effective to first ask the learner to try to correct the error himself, or to explain his choice of tense. This would encourage the learner to notice and learn from his or her mistake.

Further Reading

Educational Testing Service (2003). Linking Classroom Assessment with Student Learning. Princeton, NJ: ETS (16 pages). You can download it from here:

https://www.ets.org/Media/Tests/TOEFL_Institutional_Testing_Program/ELLM2002.pdf

Green, A. (2014). Exploring Language Assessment and Testing. Abingdon, Oxon: Routledge. Chapters 1 and 2. You can have a look at the first 10 pages of the book here:

<https://www.book2look.com/embed/9781134516629>

Here you can find some very useful examples of **formative assessment**.

Watch this short video from the British Council explaining the concept of **validity**:



Please watch the video accessible via the link below.

Video: <https://www.youtube.com/watch?v=U-HYICnt2Kw>

Here you can find a short lecture on the basic principles of validity, reliability and test fairness and on accommodations for **students with special differences - SpLD** (part of the www.dystefl.eu project)



Please watch the video accessible via the link below.

Video: https://www.youtube.com/watch?v=h1UKjv3UW_8

You may also be interested in how various examination boards support students with SpLD and other difficulties when they take external standardised tests. The following list presents some links with relevant information:



ETS - <https://www.ets.org/disabilities>

Cambridge Assessment English - www.cambridgeenglish.org/help/special-requirements/
& <https://www.youtube.com/watch?v=RxzjENVYOG8> (video)
Pearson Tests of English <https://pearsonpte.com/the-test/faqs/>

You might also like to read:

Kormos, J., & Kontra, E. H. (2008). *Language learners with special needs: An international perspective*. Clevedon: Multilingual Matters.

Kormos, J., & Smith, A. M. (2012). *Teaching languages to students with specific learning differences*. Bristol: Multilingual Matters.

Nijakowska, J. (2010). *Dyslexia in the Foreign Language Classrooms*. Bristol: Multilingual Matters.

Tsagari D. & G. Spanoudis (Eds.) (2013). *Assessing L2 students with learning and other disabilities*. Cambridge: Cambridge Scholars Publishers.

Tsagari, D. & I. Sperling (2017). 'Assessing SLLs with SpLDs: Challenges and Opportunities for Equity in Education'. In Szymańska-Czaplak, E. (ed.), *At the Crossroads: Challenges of Foreign Language Learning*, Series: Second language learning and teaching. Springer. pp 175-188. DOI: 10.1007/978-3-319-55155-5_10.

You might also like to visit various FB sites, e.g.:

Elt Well: https://www.facebook.com/ELTwell/?ref=br_rs

Teaching English to learners with dyslexia and other learning differences:

<https://www.facebook.com/groups/622573647846109/>

Dyslexia: <https://www.facebook.com/DyslexiaWorld/>

Or you might like to take the online course on 'Dyslexia and Foreign Language Teaching' offered by Lancaster University, UK <https://www.futurelearn.com/courses/dyslexia/>

References

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass Publishers.

2. Assessing Reading Skills

Ildikó Csépes & Adrienn Fekete

❖ Aims

In this course, you can learn about what efficient readers do while they are interpreting written texts. The course will show you how to take into account L2 learners' purposes for reading texts as well as their level of proficiency by familiarizing you with some Common European Framework of Reference (CEFR) scale descriptors for reading. Some of the CEFR levels will be illustrated with sample assessment tasks to highlight the requirements of the given levels. To help you enhance your task writing skill, first you will be asked to consider different test formats for assessing reading skills, especially how they can be constructed. Then you will be asked to evaluate some reading tests according to specific criteria, as a result of which you can gain useful insights into test writing so as to be able to start designing your own reading assessment tasks.

❖ Expected Outcomes

Upon completion of this Course, participants will be able to:

1. Distinguish how good readers adapt their reading according to their purpose.
2. Identify features of reading assessment tasks that are in line with good practice.
3. Evaluate reading tests more critically.
4. Design a reading assessment task on their own.

❖ Key Concepts

- Bottom-up processes
- Extracting detailed information
- Extracting specific information
- Getting the general picture
- Productive tasks
- Receptive tasks
- Top-down processes

Brief Introduction

Assessing reading is of central importance to every classroom teacher. Whether you have experience or not in this area, this section has been designed to benefit everyone: it will provide both input for novices and help the more experienced teacher to revise basic principles of and item writing guidelines for assessing reading.

This section focuses on four topics. The first one is intended to give insights into how efficient readers construct meaning from written texts. The second topic will help you familiarize yourself with some CEFR scale descriptors for reading in order to understand the target levels that are typically measured. Then we will focus on different test formats for assessing reading

skills, especially how they can be constructed. Finally, we will deal with item evaluation and design, hoping that you will be equipped with useful insights into test writing so as to be able to design your own tests for assessing reading.

2.1 Topic 1: Introduction to Assessing Reading



There is **one Task** in this Topic containing a total of **five Activities**.

Task 1. How do we read and what needs to be considered when assessing reading?

Grabe & Jiang (2014) list some factors that strongly impact reading abilities and account for individual differences in reading comprehension performance. For example, efficient readers can formulate the main ideas of a text, and they have sufficient motivation to engage in reading and to persist in reading without distraction in order to achieve success in comprehension.

Activity 1

Try to draw up a list of what specific things **efficient readers can do** when they read various texts or **what they have as a resource** to draw on. When you are ready, check how your list compares to the list adapted from Grabe and Jiang (2014). This is provided in the **Indicative Answers** section.

Efficient readers can ...

Efficient readers have ...

The outcomes of text interpretation very much depends on our **purpose** as reading is goal-driven: why we read (and what we read) has an important influence on how we read.

Activity 2

Look at the following **reading skills** (Harmer 1991, pp. 183-184). What do you think each of them involves?

- a) Predictive skills
- b) Extracting specific information
- c) Getting the general picture
- d) Extracting detailed information
- e) Recognizing function and discourse patterns
- f) Deducing meaning from context

Match the definitions (i – vi) to the reading skills (a- f) and write your answers in the box below.

- i. Texts contain many clues that show the meaning of the words that come before or after. For example, “however” introduces an idea that is different or contrary to the one that went before.
- ii. Very often we read in order to find specific bits of information, for example to find out when or where a meeting is to take place. This skill is often called scanning.
- iii. Readers need to work out what unfamiliar words mean, such as specialist words that may be new. Readers also sometimes need to “read between the lines” in order to work out the meaning, for example “It was raining in Manchester, surprisingly” is probably intended as a joke because it was not a surprise at all.
- iv. Before reading we have an idea what is in the text. We can also use clues such as pictures, headings or headlines. We also have expectations about what we are going to read, and while reading we see how the content matches up with those expectations.
- v. A reader often needs to find important information in a text answering questions like how or why. This information is more detailed than individual facts, and can often involve attitudes. For example, when a reader needs to know not just that an article is about a film, but whether the film is good or not.
- vi. This skill involves reading, usually quickly, to find out the main points or ideas in a text, without being concerned about details. It is often called skimming.

i.	ii.	iii.	iv.	v.	vi.

The answer key is provided in the **Indicative Answers** section.

Reading is a receptive skill but readers need to be very active as they typically engage in two processes:

- **Top-down processes** by activating their background knowledge and/or contextual information to predict meaning and/or fill in gaps in their comprehension
- **Bottom-up processes** whereby they process the visual input in the text, such as letters, words, sentences, to check their predictions

Activity 3

What kind of reading skills is each of the above processes associated with?

Match the following reading skills to top-down processes (A), bottom-up processes (B) or both (C).

1. identifying the sound/symbol correspondence in English, and its varieties [_____]
2. inferring context and connections between ideas from writing [_____]
3. recognizing “sight words” - common words that are read quickly and easily [_____]
4. recognizing and interpreting culturally specific references in texts [_____]
5. distinguishing between literal and implied meanings [_____]
6. recognizing cohesive devices and their function in signaling relationships in texts [_____]
7. recognizing rhetorical conventions of writing [_____]
8. recognizing the function of types of writing [_____]
9. recognizing what part of speech a word functions as and the systems of tense, agreement, and pluralization [_____]
10. using reading skills (skimming, scanning, guessing meaning from context) to help in the interpretation of texts [_____]

- A. top-down processes
- B. bottom-up processes
- C. both

The answer key is provided in the **Indicative Answers** section.

When we want to assess L2 learners’ reading ability, we need to take into account several issues.

Activity 4

What kind of issues do you think a teacher needs to keep in mind when assessing the reading skills?

Read the following statements and by drawing on your teaching experience decide whether they are True (T), False (F) or It depends on the context (D). Write T, F or D in the column on the right.

1. The reading material should represent the type of reading material that learners may use outside of the classroom.	
2. When choosing a reading passage, we should avoid topics which are very popular among students.	
3. The reading passage should include no more than 20% unknown words for good level of comprehension.	
4. The difficulty of a reading test is primarily determined by the structural and lexical difficulty of the text.	
5. The outcomes of text interpretation of a long reading passage may be influenced by fatigue in reading.	
6. When we score learners' responses to short answer questions in reading tests, we should assess the content of the response, not the correctness of the writing.	
7. When editing reading passages, the title or subtitles can be changed.	
8. While text authenticity is important, task authenticity is less important.	
9. The questions/test items for a reading passage should be written in a simpler language than the text.	
10. The number of questions for a reading passage (e.g. 300 words) depends on the reading skills to be assessed, but too few questions/items may fail to cover the whole text when we wish to assess detailed understanding.	

The answer key is provided in the **Indicative Answers** section.

Activity 5

Watch a short video (ca. 6 mins.) supported by the British Council that gives an overview of assessing reading. You can revise some of the issues we have dealt with so far as well as hear about some useful guidelines for task design to prepare you for the next task.



Watch a Youtube video by British Council entitled "Assessing reading" (uploaded by English Test - Aptis, 17 June 2016). The video is accessible via the following link:

<https://www.youtube.com/watch?v=fAKOXVpmRgE>

After watching the video, do the following task.

Fill in the missing words (a – j) in the sentences below (1 – 8).

- a. purpose
- b. meaningful
- c. summary writing
- d. scoring
- e. culturally
- f. instructions
- g. advantages
- h. combination
- i. disadvantages
- j. their parents

1. People nowadays seem to read different types of reading as opposed to what _____ did.
2. We can read texts very differently based on our _____.
3. The way we wish to assess reading should be made as relevant and _____ as possible.
4. Appropriate reading assessment includes a suitable and efficient _____ of texts and tasks.
5. Appropriate texts shouldn't be _____ biased to cause offense for some test takers.
6. An increasingly popular reading assessment task is the so-called _____ task, which is an integrated skills activity.
7. The reason why it is important to include a range of tasks in reading assessment is that all tasks have both _____ and _____.
8. Before giving the reading assessment task to test takers, it is important to check that the _____ are clear, enough time is given for completing the task and the _____ of the answers is indicated.

The answer key is provided in the **Indicative Answers** section.

2.2. Topic 2: Reading Skills and the CEFR



There is **one Task** in this Topic containing a total of **2 Activities**.

Task 1. Familiarization with the CEFR levels

When assessing learners' reading skills, the level of achievement we want to measure is typically related to one of the performance levels of the Common European Framework of Reference (CEFR). The CEFR describes what learners CAN DO in terms of reading and typically what text types they are expected to deal with at a given level. A good understanding the CEFR level descriptors thus will help you to locate both your learners and the reading texts you use in your classes in relation to the CEFR.

First, you can familiarize with or revise each of the six main levels of the CEFR for Reading. Then you can see examples for how some of the CEFR levels are represented in actual tests of reading ability.

Activity 1

Match the **level descriptors** containing **CAN DO statements** and the **conditions, limitations** (1 - 6) with the **appropriate CEFR level** (A1, A2, B1, B2, C1, C2). When you finish, compare your answers with the key in the **Indicative Answers** section.

	The Candidate CAN understand	Conditions, limitations	CEFR level?
1.	Straightforward factual texts on subjects related to my field of interest. Everyday material, e.g. letters, brochures and short official documents. Straightforward newspaper articles on familiar subjects and descriptions of events. Clearly written argumentative texts. Personal letters expressing feelings and wishes. Clearly written, straightforward instructions for piece of equipment.	Ability to identify main conclusions and follow argument restricted to straightforward texts.	
2.	Texts on familiar, concrete matters. Short, simple texts e.g. routine personal and business letters and faxes, most everyday signs and notices, Yellow Pages, advertisements.	Restricted mainly to common everyday language and language related to his/her job.	
3.	Wide range of long and complex texts – practically all forms of written language. Abstract structurally complex or highly colloquial literary and non-literary writings.	Few limitations – can understand and interpret practically all forms of written language. Very unusual or archaic vocabulary and phrases may be unknown but will rarely impair understanding.	
4.	Wide range of long, complex texts from social, professional or academic life. Complex instructions on a new, unfamiliar machine or procedure outside my area.	Understanding of details of complex texts usually only if difficult sections are re-read. Occasional use of dictionary.	
5.	Very short, simple texts, typically short, simple descriptions, especially if they contain pictures. Short, simple written instructions e. g. short simple postcards, simple notices.	Single phrase at a time, re-reading part of the text.	
6.	Correspondence relating to my field of interest. Longer texts, including specialized articles outside my field and highly specialized sources within his/her field. Articles and reports on contemporary problems with particular viewpoints.	Range and types of text only a minor limitation – can read different types of text at different speeds and in different ways according to purpose and type. Dictionary required for more specialized or unfamiliar texts.	

Activity 2

Match the following three **sample reading tests** (1-3) to the **appropriate CEFR level** (A1, A2, B1, B2, C1 or C2) that they have been designed for. When you finish, compare your answers with the key in the **Indicative Answers** section.

Sample reading test 1

Read about a government scheme to keep people healthy. Then choose true or false.

Local governments have had to take on a new responsibility; that of ensuring that people stay fit and healthy. Birmingham council is already hot on the case. For the past few years they have been encouraging local residents to enrol on the *Be active* scheme, which allows them to use local leisure services completely free of charge.

Participants must register to obtain a membership card. They can then use the facilities in leisure centres across the city at certain times of the day.

29 leisure centres are signed up to the scheme. Each leisure centre must offer a minimum of one hour of swimming time and an hour of gym time to *Be Active* members. In reality, in some establishments in the more deprived areas, 70% of opening hours are reserved for *Be Active* members. Exercise classes and badminton courts are also made available on top of the standard facilities. *Be Active* classes are also being provided in schools and community centres. There are guided bike rides, and even proposals to roll out activities in public spaces such as rounders games in parks, and buggy pushes for new mothers.

Since it was launched in 2008, a third of the local population, a total of 360,000 people, has signed up for the scheme. 60% of these are from minority groups, and the average age is 49, as opposed to 29, the figure for private gyms. The majority were not previously members of a sports club of any kind, half were overweight or obese, and a fifth considered themselves to be in poor health, indicating that the scheme is reaching the people who need it most. Research also shows that since the scheme was set up, there has been a rise in demand for information about reducing alcohol intake and quitting smoking.

Statistics show that for every £1 spent on the *Be Active* scheme, £23 is saved in the health service. Sadly, the scheme has suffered from budget cuts, and lately it has had to reduce the hours available to members. Nonetheless, other councils are interested in the great strides made by Birmingham council, and officials from Birmingham have been holding workshops with representatives from other councils to spread the word about what can be achieved.

Source: Trigg, N. (2013) *Can free leisure services get people fit?* BBC News <http://www.bbc.co.uk/news/health-22350807>

Choose *true or false*.

1. The *Be Active* scheme started up earlier this year. T/F
2. Everyone in Birmingham is eligible for *Be Active*. T/F
3. Participants must pay a one-off fee to register. T/F
4. Participants can use the leisure facilities at any time of day. T/F
5. Participants can use their *Be Active* membership cards at any sports centre in Birmingham. T/F
6. Participants can only use the leisure facilities for two hours per week. T/F
7. Some sports centres dedicate over half their opening hours to *Be Active* members. T/F

8. Some Be Active activities take place outside leisure centres. T/F
9. The Be Active scheme was not as popular as the council hoped. T/F
10. The Be Active scheme attracts both slim and overweight people. T/F
11. The majority of participants are White British. T/F
12. The scheme is saving the government money in health costs. T/F
13. Less money is available for Be Active now than in the past. T/F
14. The *Be Active* Scheme is currently only available in Birmingham. T/F

Source: http://www.examenglish.com/B2/b2_reading_health.htm

Sample reading test 2

Read the texts below and decide if the information is in text A, B, C or D.

Example:

This café has food which is surprisingly inexpensive.

The correct answer is: C.

Information to find

This café...

1. has a range of drinks from around the world: ____
2. offers world cuisine: ____
3. is a relaxing place to listen to music: ____
4. changes the menu constantly: ____
5. serves dishes which are better than in the original country: ____
6. has irritating music: ____
7. has food from different areas of the same country: ____

A: Pepe's

A Spanish style theme restaurant. Full of plastic decorations and awful photographs of Spain on the walls, the general feeling is of being in a theme park for children not a grown up restaurant. The food though is very good and there is plenty of it. I took a Spanish friend and she was very sceptical but said that the soups and the pastries were better than some of the restaurants she has been in some parts of Spain. They concentrate on a regional menu and all the dishes are from different parts of Spain so you can learn a bit about the country and the origin of the dishes you are going to have. There's more information on the back of the menu.

B: The Red Café

Really nice food and extremely nice staff. The food is international; they have specialities from different countries. Moussaka from Greece and spaghetti with meatballs, Italy of course, were the dishes of the day when we went and these change every day so there's always something new. They really do try to look around for new things and they seem to ask all of their customers where they are from. If it's somewhere interesting then they ask if you have any good recipes. I suppose they really get them all off the internet or recipe books but it's nice that they ask and they do seem genuinely curious.

C: Andre's

This café doesn't have a big range of foods but what there is really good and cheap. What isn't cheap is the beer and wine but the place has beers from just about everywhere, Japan, India, as well as really nice Belgian stuff and wine from everywhere too. An expensive evening if you drink a lot but worth it. There's no real theme for the food, it's just the usual restaurant things and mostly quite simple but nicely done and, unlike the drinks, really quite cheap for what you get. What is great is the atmosphere, the music is really cool and they will play requests if they have the CD. The staff really leave you alone to get on with your conversation and your meal and don't interrupt every two minutes to ask if everything is OK.

D: The Koh I Noor

The only decent Indian Restaurant I could find. Don't be put off by the cost which soon adds up with all the different courses. The only other negative is the constant music which really annoyed me. The food is fantastic and the waiters, all Indians, who seem to speak every European language there is, are really friendly. They sell Indian beer and some not very good wine, they serve the beer really cold – good with spicy food and it's good if a little sweet for my tastes. They have all the traditional things but also a lot of ordinary stuff if you don't like spicy food. You can just have a steak or grilled chicken and they will happily give you chips and a burger if you insist.

Source: <http://www.keronline.hu/>

Sample reading test 3

Susie is writing in her diary about her new job. She works in a television company. Read her diary, then complete the sentences with the names on the left. There is one sentence which you do not need. There are 7 right answers.

DIARY**Monday**

Today I started my new job. It was so exciting! I work in a big room with two men called Dan and Peter. Our boss is someone called Annie. She makes us work very hard. We are all a bit afraid of her.

Tuesday

Today Annie and I had lunch together in the cafe. She is really quite nice. She travels all over the world. I asked her which country she liked best. She said she liked India best.

Wednesday

I did a TV interview with a man and his pet monkey, Popo. The man has taught the monkey to write. Popo wrote his name in front of the cameras. It was amazing!

Thursday

Peter and I went to an Indian restaurant today. It has won a Best Food prize. Peter filmed me while I ate a fabulous meal. Poor thing! He must have felt hungry!


Friday

Annie invited us for a drink to a new bar tonight after work. Peter couldn't come, so we went without him. We had a great time. Dan was so funny. We laughed and laughed all evening.

Peter and Dan	_____	- can make people laugh.
Annie	_____	- travelled to India.
Popo	_____	- wrote a name in a TV show.
Dan and Annie	_____	- won a Best Food prize.
Dan	_____	- went out for a drink.
Susie	_____	- works with cameras.

Source: <http://www.keronline.hu/>

2.3. Topic 3: Techniques for Assessing Reading

 There are **two Tasks** in this Topic containing a total of **four Activities**.

In order to design appropriate reading tests, we need to consider **different test formats** for assessing reading skills, especially **how** they can be constructed.

Task 1. Guidelines for designing specific task types to assess reading

In large-scale testing, it is typical to use **receptive tasks** such as multiple-choice questions, true/false statements, matching tasks, or sequencing tasks. Receptive tasks require L2 learners to choose from a set of options, they do not need to produce language on their own. Some of the learners can simply go guessing and get the answers right, and we can never be absolutely sure that they chose the correct answer for the right reason. However, if the task is well designed and the answer key is unambiguous, we are likely to trust the scores in spite of the doubts mentioned. Although scoring the answers seems to be easy in the case of receptive tasks, some of these tasks are not so easy to design, especially multiple-choice questions. The latter requires talent and considerable experience in item writing, thus we do not recommend this task type in reading assessment designed by teachers for classroom use.

Activity 1

What kind of **guidelines** should one follow when designing **multiple-choice questions, true/false questions, matching tasks** and **sequencing tasks**? If you have limited or no experience in designing such task types, try to guess the answers.

There are some guidelines for item writing (1-20) that you need to match with the task types that they are related to (A – D). In some cases, the guidelines may be matched with more than one task type. Please note that:

- **'items'** are used here to refer to individual questions in one task;
- **'distractors'** refer to wrong answers among the options;
- **'stem'** refers to the main question/statement in multiple-choice questions.

Task types:

- A. multiple-choice questions
- B. true/false statements
- C. matching tasks
- D. sequencing tasks

Item writing guidelines:

1. You need to include minimum one distractor. [_____]
2. You need to include minimum two distractors. [_____]
3. Distractors must relate to the text content to make them plausible. [_____]
4. You can take out words, clauses, sentences or paragraphs from the text. However, the parts taken out must be of the same type. [_____]
5. You must paraphrase the options, you should not use the same wording as in the text. [_____]
6. The wording of the items should be such that it is easy to understand, is below the language level of the text. [_____]
7. Make sure that no item can be answered correctly without reading the text. [_____]
8. Make sure that each question can be answered independently, i.e. the answer to one question does not depend on the answer to another question. [_____]
9. To minimize guessing, include a third option "not stated" / "not mentioned". [_____]
10. Make sure that items do not overlap. [_____]
11. Sequence items in the order they appear in the original text. [_____]
12. If you delete first words or beginnings of sentences, warn test takers in the instruction and do not capitalize the beginning word in the list of alternatives. [_____]
13. Where you take out paragraphs, do not take out the first or the last paragraph, in order to leave enough context for comprehension. [_____]
14. Make sure there is only one correct order for the options. [_____]
15. Do not use absurd options. [_____]
16. Avoid using negative stems. [_____]
17. Try to make the options of the same length, do not let one option stand out among the rest. [_____]
18. Avoid a grammatical mismatch between the alternatives and the stem. [_____]
19. Check the items for punctuation errors and other mistakes in grammar or lexis. [_____]

The answer key is provided in the **Indicative Answers** section.

Productive tasks require learners to produce language of their own and therefore such tasks are considered to be more communicative because they elicit direct performance from test takers. We will look at some guidelines for task design in relation to short answer questions, gap-filling tasks, information-transfer tasks and summary writing.

Learners can produce varied responses to some of these tasks, therefore marking the answers needs more careful checking and sometimes even a task-specific assessment scheme where grammar, punctuation and vocabulary may also be scored, not just the content of the response. While it is possible to work out an unambiguous answer key for most of the selected task types, summary writing is different as it measures both reading and writing skills and so the assessment of the learner's performance will have to follow guidelines for assessing writing. In gap-filling tasks, sometimes the task may be integrated because it is likely to measure reading as well as vocabulary and/or grammar.

Activity 2

What kind of **guidelines** should one follow when designing tasks which require questions with **short answers, gap-filling, information-transfer** and **summary writing**? If you have limited or no experience in designing such task types, try to guess the answers.

Match **guidelines for item writing** (1 – 14) with the **task types** that they are related to (A – D). In some cases, the guidelines may be matched with more than one task type.

Task types:

- A. short answer questions
- B. gap-filling (the text is gapped without alternative answers given)
- C. information-transfer (information needs to be transferred into gapped sentences or charts)
- D. summary writing

Item writing guidelines:

1. Clearly indicate the required length of the answers. [_____]
2. Make sure no item can be answered correctly without reading the text. [_____]
3. Make sure that items do not overlap. [_____]
4. A sentence at the very beginning and at the end must be left intact for lead-in and lead-out. [_____]
5. Make sure you discourage learners in your scoring scheme to lift whole chunks instead of paraphrasing them. [_____]
6. There should be at least five words between gaps to provide enough context. [_____]
7. Award scores for both content and composition features. [_____]
8. Avoid choosing texts that contain a lot of proper names as well as facts and figures. [_____]
9. Avoid questions which demand much imagination. [_____]
10. Avoid Yes/No question. [_____]
11. The wording of the items should be such that it is easy to understand, is below the language level of the text. [_____]
12. If the response is very short (maximised in three words), do not penalise learners for minor punctuation errors. [_____]
13. Make sure all the possible answers are included in the answer key. [_____]
14. Make sure items are related to clearly stated information in the text. [_____]

The answer key is provided in the **Indicative Answers** section.

Activity 3



Go to Glenn Fulcher’s website entitled “Language testing Resources Website”. The link of the website is: <http://languagetesting.info/>

Then select “Videos” from the Resources list on the left. You will see a “Multiplex”, with a list of videos. Scroll horizontally and find “6. Reading”, which is a talk on testing reading by Caroline Clapham (ca. 9 mins).

While listening to the talk, take notes on the following:

- what **reading skills** to test,
- how to choose **reading passages**,
- how to test **text comprehension**.

Write your thoughts here.

My thoughts:

If you need some extra prompts, consult the notes in the **Indicative Answers** section.

Task 2. Informal assessment of reading in the language classroom

When developing L2 learners reading skills in the classroom, it can also be a good opportunity to help them get **feedback** on how well they are doing. Grabe (2009) identifies some **classroom-based informal assessment activities** that can be carried out by teachers. These informal activities include, for example:

- a) keeping a written record of student responses to questions in class after a reading
- b) keeping time of how long students spend on task during free reading or sustained silent reading
- c) asking students to list words they want to know after reading and why
- d) encouraging students to write simple book reports
- e) keeping a record of student reading rate growth
- f) encouraging students to compile a portfolio of reading activities

Activity 4

Do you **use** any of the above suggestions or other techniques not mentioned above to evaluate students' reading ability informally? Is there one activity you particularly **like**? **Why**?
Write your thoughts here.

My thoughts:

2.4. Topic 4: Evaluating and Designing Reading Tests



There are **two Tasks** in this Topic containing a total of **two Activities**.

Task 1. Guidelines for designing specific task types to assess reading

You will evaluate **five reading test tasks**, all related to the same text. The item writers are all English language teachers, who decided by themselves on what CEFR level to aim at with their test and what testing technique to use. In your evaluation, you are advised to take into account the following task design features:

- *CEFR level targeted*: is it appropriate in the light of the task?
- *target domain*: does the task test reading skills and/or something else?
- *task contextualisation in the instructions*: is it suitable, e.g. is the topic of text, purpose of reading given?
- *number of items to cover the text suitably*: are there enough items?
- *working example*: is there one?
- *language level of the items*: is it less or more difficult than the level of the text?
- *wording of the items*: are the items paraphrased, or can be located word by word in the text?
- *distractors (if there are some)*: are they appropriate / plausible?
- *scoring*: is it indicated / appropriate / transparent? Is differentiated scoring justifiable?
- *reliability of marking*: is subjectivity minimised?
- *layout*: is the task layout user-friendly? Is there adequate space given for the answers? Is the text on one page?
- *punctuation/capitalisation*: is it appropriate?

Activity 1

First, read the **original text** entitled “21st century Pet Gadget” (source: Into Europe, Reading and Use of English). Then take **each sample test** (Sample Task 1-5), imagining you are a **test taker** for whom the task was intended. Pay attention to the features listed above and especially consider the **task demands** in the light of the targeted CEFR level.

Write down your **detailed remarks** about each sample test (and preferably about each item/question that you consider problematic) and then compare your notes with the comments made by **expert item reviewers** in the **Indicative Answers** section.

21st Century Pet Gadgets from Japan, the Latest in Telephone Gadgets for the High Tech Pooch

In Tokyo we meet Mint, a black Labrador puppy, and her owner Yoko Sakohata, to try out some of the latest Japanese gadgets designed to make a dog's life that bit easier – (0)

In a country where working late and getting stuck in traffic jams are part of normal daily life, it can be a problem getting home (1) A timer-controlled feeder is one answer, but it doesn't give much personal contact with the lonely pet, (2) Yoko tries out a telephone-controlled feeder which enables you to call home from anywhere, at any time, (3) , then at the press of the button deliver a meal. Mint finds the whole thing a little puzzling at first, but has no qualms about tucking in (4)

Practically everyone has a mobile phone in Japan, and now they've even started making ones for your pet. At the moment they're not

designed to allow you to talk with your pet, (5) Because Japan's PHS network (*Personal Handyphone System*) has thousands of antennae, (6) , it is possible to tell almost exactly where any handset is by measuring the strength of its signal at several adjacent antennae. With one of the mobile handsets fitted to her collar, Mint is taken to a secret location (7) Logged on to the internet, Yoko enters Mint's mobile number onto the website (8) showing where Mint is – in a local playground.

The tracker device can also be used to keep tabs on wandering children and property such as motorcycles. Sadly (9) where the number of antennae make it accurate to within 50 metres – Britain's cellphone networks can't give anything like that degree of accuracy as each cell is much larger.

Sample Task 1 (intended for level B1)

Read the text and answer the following questions using your own words. Please write in complete sentences. There are two points for each question, one point for content and one point for the correct use of grammar and vocabulary.

1. Who are Mint and Yoko?

2. Why can it be a problem to feed your pet in Tokyo?

3. Explain the feeder system that Yoko tries.

4. Why have mobile phones for pets been developed?

5. Who or what could the device also be used for?

6. Could the tracking device be used in Britain? Why or why not?

Answer key: (*content*)

1. Mint is Yoko's puppy/pet or Yoko owns Mint.
2. Working late and long traffic jams are part of everyday life.
3. The feeder is controlled by phone, you can call home from anywhere, talk to your pet, press a button and food is delivered.
4. Mobile phones for pets have been developed so you always know where your pet is.
5. The device could also be used for young children or objects such as motorcycles.
6. No, it couldn't be used in Britain because there aren't enough antennae so it wouldn't be accurate enough.

My comments:

Sample Task 2 (intended for B2)

Answer questions 1-5, using one to six words. Use your OWN words. Spelling counts.
(2 points per question).

1. What do Japanese pet owners find it difficult to do for their pets in a timely fashion?
2. Why is the telephone controlled feeder preferable to the remote-controlled one?
3. Is Mint enthusiastic about the food she is given by the telephone-controlled feeder?
4. Where must owners go to locate the signals given off by their dogs' telephones?
5. Why wouldn't the dog-locating system work in Britain?

Answer key:

1. Feed them. / Give the pets food. /Give them dinner etc.
2. It provides more contact. / It is more personal. / More personal.
3. Yes.
4. The internet. / A website.
5. Antennas are too far apart. /The antennas are too distant/ Too much space between antennas.

My comments:

Sample Task 3 (intended for level A2)

Read each statement below carefully. Place a **T** on the line if you think a statement is TRUE. Place an **F** on the line if you think the statement is FALSE.

1. In Japan it is difficult to take care of pets because of busy daily routine. ____
2. A telephone-controlled feeder helps dogs to feel master's attention. ____
3. Modern cell phones allow people to talk with their pets. ____
4. The tracker device can be used to look-out for children and bicycles. ____
5. The same technology can be used in Great Britain. ____

Answer Key:

1. True
2. True
3. False
4. True
5. False

My comments:

Sample Task 4 (intended for level B2)

Read the text again and decide whether the statements are true (T) or false (F).

(2 points each)

- ___ 1. Available technology lacks in providing you to communicate with your pet.
- ___ 2. The feeder controlled by telephone is comparatively a more efficient gadget than the one controlled by a timer.
- ___ 3. Even Mint got surprised by the feeding gadget and she rejected eating in that way.
- ___ 4. The only concern in the efforts to invent a new gadget is to feed the pets.
- ___ 5. Due to the current technology in Japan, you can easily feed your pet, find its location and even chat with it.
- ___ 6. Britain's network system technically fails to tell the exact place of your properties.
- ___ 7. The tracking system in Japan lets you see the place of your pet by the help of antennae and a few handsets without the need for Internet connection.
- ___ 8. Today, Japan is the only country who has the technology of keeping tracks of things accurately apart from pets as well.

Answer Key:

1. F 2. T 3. F 4. F 5. T 6. T 7. F 8. T

My comments:

Sample Task 5 (intended for level B1)

Read the text on gadgets for dogs in Japan and the questions below. For each question 1- 8 choose the most appropriate answer A, B, C or D, according to information in the text. Mark only ONE answer for each question. Tick the box to mark your answers. Each correct answer is worth 1 point.

1. Why did the author meet Mint and her owner Yoko?

- A because he wanted to help owners have more time with their dogs
- B because he wanted a new black lab puppy
- C because he wanted to test new technology in pet care
- D because he wanted to design a new gadget

2. What can make a dog owner arrive home late to feed their dog?

- A too many cars on the road
- B working in the late morning
- C feeding dogs too often
- D going to work by train

3. Which of the following is a problem with a timer-controlled feeder?

- A It is easily broken
- B It is hard to teach the pet how to use it
- C It is cold and mechanical
- D It is expensive

4. Which of the following is a benefit of a telephone-controlled feeder?

- A It is easy to use because the dog presses a button.
- B It is personal because the dog hears your voice.
- C Dogs find it confusing because they have to eat biscuits.
- D Dogs eat less because they cannot ask for more.

5. What does the mobile phone for your pet allow you to do?

- A call the vet for help
- B talk to your pet on the phone
- C open a door to let your pet out
- D understand where your pet is

6. How did Yoko find Mint's location?

- A In the yellow pages
- B Through a social network
- C With a tracker
- D With a telephone-controlled feeder

7. Which of the following best describes how the writer feels about the availability of this technology in Britain?

- A worried
- B sad

- C surprised
- D disappointed

Answer Key

1. C
2. A
3. C
4. B
5. D
6. C
7. D

My comments:

Task 2. Designing a reading test task

After raising your awareness of good practice in item writing guidelines and various task design issues, you are ready to **design a test task** on your own.

Activity 2

Design a test task that is aimed at measuring students' reading ability at a specific CEFR level of your choice. Before you start working on the task, please take the following into account:

- use only one type of testing technique (e.g. matching, short-answer questions, true/false/not stated, information transfer, gap-filling, etc.)

- the task must include as many items as the text supports (minimum 5 items)
- provide an answer key
- provide clear instructions and indicate the scoring of the items
- design the layout carefully (e.g. do not forget to provide space for the answers)
- do not edit the text for content but feel free to make formal changes to suit the task (e.g. take out words/clauses/sentences when designing a matching task)
- choose your target CEFR level and indicate it clearly

GIANT PANDA FACTS

Giant pandas are chubby mammals that live in a few remote mountainous regions in China. They have thick fur with bright black-and-white markings. A panda's coat acts like a thick winter raincoat. The fur is water-repellent and helps keep a panda warm and dry in cold, wet weather.

Pandas are plant eaters and they feed mainly on a plant called bamboo. Sometimes pandas eat other types of plants and occasionally they eat small mammals. But pandas usually eat only the stems, twigs, leaves, and fresh young shoots of the different types of bamboo. They especially like the tender shoots of young bamboo plants.

Some pandas live as long as 30 years and weigh as much as 117 kg. Full-grown pandas are close to 1.5 m tall when standing up and some grow as tall as 1.7 m. Males and females look alike, but females are a bit smaller than the males.

Pandas usually live alone. Each panda lives in an area that's about one or two miles (1.5 to 3 km) in diameter. Although pandas will share part of their territory with other pandas, they don't usually get too close to each other.

In the spring, pandas search for a mate. They mark their territories with special scent glands to let other pandas know they are ready to mate. Once pandas mate, they separate and the females raise the young alone.

A new-born panda is only about the size of a hamster and weighs about 100 grams. Pandas are born without teeth and with their eyes closed. And they have only a thin covering of hair. It takes a few weeks for the typical black-and-white markings to appear.




Young pandas stay with their mothers for about a year and a half. They learn how to find food, climb trees, and stay away from enemies.

Ranger Rick's Nature Scope

If you need some extra prompts, consult a sample model task in the **Indicative Answers** section.

Assessing Reading Skills - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I am able to distinguish how good readers adapt their reading according to their purpose.			
I am able to identify features of reading assessment tasks that are in line with good practice.			
I am able to evaluate reading tests more critically.			
I am able to design a reading assessment task on their own.			

Indicative Answers

Topic 1: Activity 1

Efficient readers can:

- recognise words efficiently by activating phonological, orthographic, morphological, and semantic knowledge;
- activate their grammar knowledge;
- engage in a variety of strategic processes while reading more challenging texts, such as inferencing, monitoring their interpretation, etc.;
- recognize discourse structuring and genre patterns;
- activate and exploit background knowledge to support their text comprehension;
- interpret text meaning critically to suit their reading purposes;
- use their working memory efficiently.

Efficient readers have:

- a large vocabulary knowledge;
- efficient reading fluency skills;
- a lot of experience with L2 reading.

Topic 1: Activity 2

i.	ii.	iii.	iv.	v.	vi.
e	b	f	a	d	c

Topic 1: Activity 3

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
B	A	B	A	C	B	A	A	B	C

Topic 1: Activity 4

1. The reading material should represent the type of reading material that learners may use outside of the classroom.
True. The purpose for reading a text in an assessment should reflect the purpose a reader is likely to have for reading similar texts in the real world.
2. When choosing a reading passage, we should avoid topics which are very popular among students.
False. We should avoid topic bias and disturbing topics instead, e.g. war, death, politics, religion, culturally sensitive issues.
3. The reading passage should include no more than 20% unknown words for good level of comprehension.
It depends on the context: if the purpose of assessment is gist reading, readers can succeed in the task even if the text contains a good number of unfamiliar words for them. However, detailed understanding of a text would require familiarity with the great majority of words in a given passage, thus unknown words should be less than 20% (5-10%).
4. The difficulty of a reading test is primarily determined by the structural and lexical difficulty of the text.
False. It is important to consider text difficulty in relation to the target proficiency level of the learner. However, the difficulty of a reading assessment task cannot be judged in relation to the text alone. The reading task together with the text will define the difficulty level of the assessment.
5. The outcomes of text interpretation of a long reading passage may be influenced by fatigue in reading.
True. Text length is one of the factors influencing overall test difficulty.
6. When we score learners' responses to short answer questions in reading tests, we should assess the content of the response, not the correctness of the writing.
True. Typically, learners should not be penalised for minor mistakes in spelling and grammar.
7. When editing reading passages, the title or subtitles can be changed.
True. Useful editing guidelines can be found in Green (2014: 112), such as cutting down the length, editing potentially offensive, overly technical or culturally inaccessible content, clarifying ambiguous wording.
8. While text authenticity is important, task authenticity is less important.
It depends on the context. Multiple-choice questions, for example, represent a typical test task that people never do in non-test situations, still large-scale tests prefer to use them as the answers can be scored by machine. In other contexts, multiple-choice questions would be rejected, and other task options would be chosen that resemble more closely what people do when they have read texts (e.g. write a letter or a summary).
9. The questions/test items for a reading passage should be written in a simpler language than the text.
True. We do not want to test the comprehension of the questions primarily but learners' understanding of the reading passage.

10. The number of questions for a reading passage (e.g. 300 words) depends on the reading skills to be assessed, but too few questions/items may fail to cover the whole text when we wish to assess detailed understanding.

True. We should design as many questions/items for one reading passage as the text can support.

Topic 1: Activity 5

1.j 2.a 3.b 4.h 5.e 6.c 7.g & i 8.f & d

Topic 2: Activity 1

1 – B1; 2 – A2; 3 – C2; 4 – C1; 5 – A1; 6 – B2

Source: The Common European Framework of Reference

Topic 2: Activity 2

Sample reading test 1: B2

Sample reading test 2: B1

Sample reading test 3: A2

Topic 3: Activity 1

1	2	3	4	5	6	7	8	9	10
C	A	A, C	C	A, B	A, B	A, B, C, D	A, B, C	B	A, B, C

11	12	13	14	15	16	17	18	19	20
A, B, C	C	C	D	A, C	A	A	A	A, B, C, D	A, C

Topic 3: Activity 2

1	2	3	4	5	6	7	8	9	10	11	12	13	14
A, B, C, D	A, C	A, C	B	D	B	D	D	A, B, C	A	A, C	A, B, C	A, B, C	A, C

Topic 3: Activity 3

- **What reading skills to test:** a variety of skills specified in the test specification such as inferencing, skimming (looking for general meaning) or scanning (looking for specific meaning)

- **How to choose reading passages:** choose texts upon which you can ask good questions; first and foremost consider the purposes of your students; texts should look authentic; use a range of texts (to represent different lengths, genres, topics)
- **How to test text comprehension:** use more test methods, multiple-choice questions are difficult to write; Caroline’s favourite is short answer questions; you can use matching headings to parts of the text, gapped summary for advanced students, information-transfer, or statements that are true/false/not given in the text.

Topic 4: Activity 1

Comments by an expert item reviewer on the sample reading tasks on “21st Century Pet Gadgets”.

Sample Task 1 (intended for B1)

1. There are several other possible answers, the answer key is incomplete.
2. The length of the answers can vary too much, the expected length is not specified.
3. The passive voice is used in 4 out of the 6 sentences, which makes the processing of the questions difficult for the B1 level.
4. It is unclear why only complete sentences are accepted when the main target is text comprehension.
5. The scoring allows for subjectivity in the marking (one point for content and one point for the correct use of grammar and vocabulary).
6. The number of items seems to be insufficient to cover the whole text.

Sample Task 2 (intended for B2)

1. The length of the expected answers is specified, but it allows for too much variation (1 vs. 6 words)
2. The scoring is the same for each answer in spite of the apparent disparity in length – this is unfair.
3. The scoring allows for subjectivity in the marking (there is no guidance given on when to give one point and how to handle spelling mistakes).
4. There are several other possible answers, the answer key is incomplete.
5. In short answer question YES/NO question should be avoided (see Question 3).
6. The number of items seems to be insufficient to cover the whole text.
7. The intended level seems to be OK.

Sample Task 3 (intended for A2)

1. The text seems to be too demanding for the A2 level although the difficulty of the questions is low.
2. The instruction can be worded more efficiently (there is repetition: **Place a T on the line if you think a statement is TRUE. Place an F on the line if you think the statement is FALSE**)
3. Some of the questions are not worded suitable to match the intended level (A2): question 4 and 5 contain passive constructions.
4. There is no indication about how the answers should be scored.
5. There is a 50% chance of getting the items right. A third option could minimise the chance factor.
6. The number of items seems to be insufficient to cover the whole text.

Sample Task 4 (intended for B2)

1. There is no guidance on how to write/mark the answers: write T/F or TRUE/FALSE?
2. The number of items seems to be sufficient to cover the whole text.
3. The wording of some of the statements needs to be changed (e.g. in Question 7 “by the help”)
4. The Answer Key is incorrect for Question 8: there is no explicitly stated information given in the text that supports the statement. The answer is more likely to be “the text does not say”.
5. The scoring of the answer can be subjective as there is no guidance given when to give one score.
6. There is a 50% chance of getting the items right. A third option could minimise the chance factor.
7. The intended level seems to be right.

Sample Task 5 (intended for B1)

- The number of items seems to be sufficient to cover the whole text.
- If the options are full sentences, they need to start with a capital letter and end with a full stop.

Question 1

- The options contain the same phrase, this could be lifted and added to the stem (because he wanted)

Question 2

- Option B has no relation to anything in the text
- Option C is absurd and does not relate to what the text says.

Question 3

- Option A, B and D, all the distractors practically have nothing to relate to in the text.

Question 4

- Option C and D cannot be regarded as benefits even without reading the text as so nobody is likely to choose them.

Question 5

- Option A and C have nothing to relate to in the text. No vet or door is mentioned anywhere in the text.

Question 6

- Yellow Pages, because they are so specific and are not mentioned in the text, are unlikely to distract anybody in Option A.

Question 7

- Option B and D are related and somewhat synonymous therefore both can be accepted as correct. There seem to be two good answers here.

Topic 4: Activity 2

Sample model task (intended for level B1)

Source: Into Europe, Reading and Use of English

You are going to read a magazine article about pandas. Some sentences are missing from the text. Choose the best sentence (A-G) for each gap (1-5) in the article and write its letter in the box. There is one extra sentence that you do not need to use. There is one example (0) at the beginning.

GIANT PANDA FACTS

Giant pandas are chubby mammals that live in a few remote mountainous regions in China. They have thick fur with bright black-and-white markings. **(0)** ____ The fur is water-repellent and helps keep a panda warm and dry in cold, wet weather.

(1) ____ Sometimes pandas eat other types of plants and occasionally they eat small mammals. But pandas usually eat only the stems, twigs, leaves, and fresh young shoots of the different types of bamboo. They especially like the tender shoots of young bamboo plants.

(2) ____ Full-grown pandas are close to 1.5 m tall when standing up and some grow as tall as 1.7 m. Males and females look alike, but females are a bit smaller than the males.

Pandas usually live alone. Each panda lives in an area that's about one or two miles (1.5 to 3 km) in diameter. **(3)** ____

In the spring, pandas search for a mate. They mark their territories with special scent glands to let other pandas know they are ready to mate. Once pandas mate, they separate and the females raise the young alone.

A new-born panda is only about the size of a hamster and weighs about 100 grams. **(4)** ____ And they have only a thin covering of hair. It takes a few weeks for the typical black-and-white markings to appear.

(5) ____ They learn how to find food, climb trees, and stay away from enemies.

Ranger Rick's Nature Scope

- A** Some pandas live as long as 30 years and weigh as much as 117 kg.
- B** Pandas are plant eaters and they feed mainly on a plant called bamboo.
- C** In stormy weather, they sometimes try to find a cave or some other type of shelter.
- D** Pandas are born without teeth and with their eyes closed.
- E** Young pandas stay with their mothers for about a year and a half.
- F** A panda's coat acts like a thick winter raincoat.
- G** Although pandas will share part of their territory with other pandas, they don't usually get too close to each other.

Write your answers here:



0	1	2	3	4	5
F					

Score: 5 points

Answer key:

1. B 2. A 3. G 4. D 5. E

Further Reading

Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press. Chapter 1,2&7.

Green, A. (2014). *Exploring Language Assessment and Testing*. Oxon: Routledge. pp. 95-117.

Weir, C. (1990). *Communicative Language Testing*. New York: Prentice-Hall. pp. 42-58.

Watch the following Youtube webinar by Cambridge English: 'Understanding reading comprehension assessment: what every teacher should know'



<https://www.youtube.com/watch?v= gyYtys805U>

Finally, if you wish to explore ideas on how to **assess SpLD students' reading skills** visit the following website of the DysTEFL project and follow the suggestions in Chapters 9 (Reading and Writing) and 10 (Assessment)



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

or download the DysTEFL training materials and work directly from there:

If you are a trainer: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

If you are a teacher: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

And read the following:

Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning differences. Bristol: Multilingual Matters. Chapter 8. Assessment

References

Alderson, J. C., & Cseresznyés, M. (2003). *Reading and Use of English*. Budapest: Teleki László Foundation and the British Council.

Common European framework of reference for languages: learning, teaching assessment. (2001). Cambridge: Cambridge University Press.

Grabe, W. (2009). *Reading in a Second Language: Moving from Theory to Practice*. Cambridge: Cambridge University Press.

Grabe, W., & Jiang, X. (2014). *Assessing Reading*. In A. J. Kunnan (Ed.) *The Companion to Language Assessment* (pp. 185 - 200). London: John Wiley and Sons.

Green, A. (2014). *Exploring Language Assessment and Testing*. Oxon: Routledge.

Harmer, J. (1991). *The Practice of English Language Teaching*. Essex: Longman.

3. Assessing Writing Skills

Liz Hamp-Lyons

❖ **Aims**

Assessment Literacy refers to the ability and willingness to ask and answer important questions about fundamental issues in assessment. These questions include what and who you are assessing and why you are assessing them; what tool or tools would be suitable for your specific context; how and when those tools should be used; and also what the results will be used for.

In this Course you will think about how you teach and assess writing as part of your teaching of English. Although there are many tests of English, there are more ways to assess students' writing than many teachers are aware of, and some of them are not usually found in formal tests. This Course shows you a variety of writing tests/assessment, and suggests ways that you can look at each of them in order to decide whether they fit the needs of your learner.

❖ **Expected Outcomes**

Upon completion of this Course, you should be able to:

1. Understand the difference between writing as a process and writing as a product;
2. Distinguish between different ways of assessing student writing;
3. Understand the uses of the different types of writing assessments;
4. Understand the key elements of any writing assessment;
5. Be sensitive to the effects of your choices of assessment tools on your learners.

❖ **Key Concepts**

- [Assessment criteria](#)
- [Feedback and feedforward](#)
- [Rating and raters /teacher-raters](#)
- [Rating scales](#)
- [Score reporting](#)
- [Tests and assessments](#)
- [Writing assessment principles](#)
- [Writing tasks](#)

Brief Introduction

The ability to write continues to be a key characteristic of any educated person. Learning to write in one's own language is difficult; learning to write in a second language is even more difficult. Now that so much interaction between people takes place in a 'mixed' language that is partly oral and partly written (email, Twitter, Facebook, etc.), and as more and more schools bring computer-based teaching into the classroom, our definition of 'writing' is getting more complex. Ways of assessing what learners have written are also becoming more complex. This offers us many opportunities, both for more creative ways of formally assessing writing products, and for using assessment formatively to support learners' writing processes.

3.1 Topic 1: Understanding Writing



There are **two Tasks** in this Topic with a total of **five Activities**. These Activities invite you to reflect on issues discussed in this Topic.

Task 1. Introduction

The purpose of this introductory section is to let you check if you understand the difference between the **writing process** and **writing that is a product**. This is an important distinction when it comes to deciding on ways to assess writing.

Activity 1

Below you can see three pieces of writing. All of them have been written by English language learners of about 14-17. First, look at these two questions; then, read each of the texts and answer the two questions:

- a) **Who** was each sample written for?
- b) **Why** was it written?

Sample 1:

Even though globalization affects the world's economy in a very positive way, its negative side should not be forgotten. Discuss.

Everything has two sides and the globalization is not the exception. Our first thoughts about this topic include the process of global "McDonaldisation" and, generally speaking, spreading the one, American, lifestyle across the whole Globe.

Firstly, I would try to concentrate on the positive aspects of globalisation. As far as economy is concerned, institutions like the Global Bank or IMF are always focused on developing the 'Third World' and helping poor people to combat their life obstacles (through loans and donations). Moreover, the world becomes an area of sharing thoughts (e.g. philosophical or economical doctrines), which become popular due to lack of barriers.

However, disadvantages of globalization are also widely known. Some people insist that because of this process, the spirit of countries and nations rapidly disappears. The integrity, established hundreds years ago is on the verge of collapsing. Furthermore, there's a strong lobby of communists who reckon, that the globalization indicates an uncontrolled reign of capitalists and slave work of lower labour-class. We should never forget about the detrimental impact of global investments on the environment – the green house effect or soar rains are triggered by globalization.

To sum up, globalization has both positive and negative influence on our everyday life. I can't agree with the popular statement that we should try to avoid being affected by it. However, we must not forget about our surroundings and local communities. They have a great value which should last forever.

Your answer:

a)

b)

Sample 2:

Date :- 2072-04-02

My Experience of Nepal Earthquake

I was at home reading books. Suddenly, the windows shake and produce noise. I was afraid. I ran towards my mother. We were afraid. The doors were making a lot of noise. The house were shaking. fifty seconds the earthquake come. when the earthquake stop, we came out of the house to open places. The wall were falling, a lot of people were coming out of home. It was too cold. I was listening radio. 7.9 richter scale earthquake was come. many people were dead. we were so sad of their parents. our Dharahara was fell down. Many people were crying of their losing friends, members. It was a bad day. My ^{close} brother Dipu was Dead because of falling wall. We were felling so sad for him and his parents. We were sitting in tent for many days. Small earthquake was coming all days. My heart was running faster. Many small and bigger earthquake make a lot of natural disaster like landslide, flood etc. It is a natural disaster. It can come in any second. I was shocked of earthquake. now, also i am a afraid of earthquake. which kill a lot of people, destroy houses. we were sitting in home now but many people are sitting at tents. children donot go to school because of earthquake. I want to tell that students should go to earthquake. The earthquake broke dream of all people of Nepal

Manish Barman (8)
Roll :- 11

Your answer:

a)

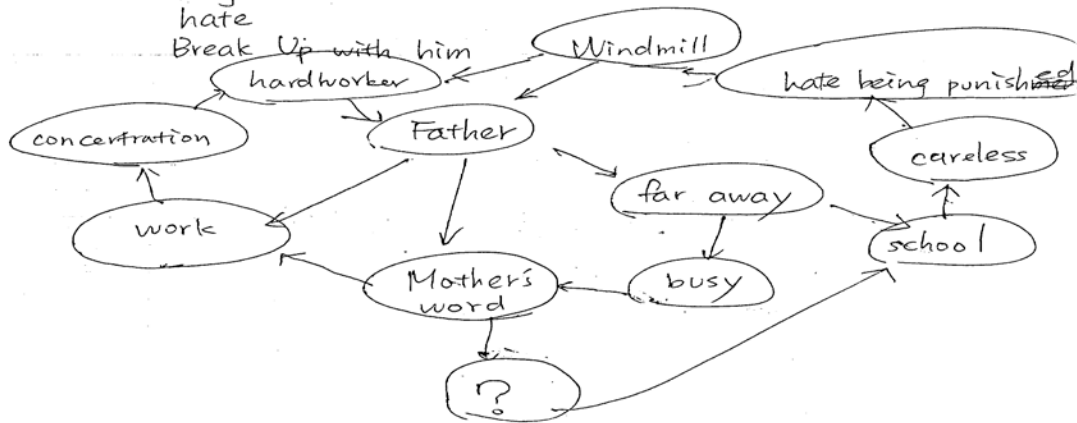
b)

Sample 3:

Uscē 4

Writing Process: Originally "Spiritual Joy land" describing a childhood thing reminded me so many things happened then. I decided to pick something instead of general introduction of my childhood.

- Tree
- Mother
- Andy
- Father
- Grandpa Gouh
- Exchange Year
- host mother
- host sisters
- drinking
- drunken night
- Love
- Disgrace
- hate
- Break Up with him



Your answer:

a)

b)

Activity 2

Even three texts will show you that texts are written for **different people** and for **different reasons**, and that these differences lead to the writer's decisions about what to write about, whether to be formal, informal, technical, newsy, funny, etc. These choices in turn, lead to specific choices of vocabulary, grammar, level of sentence complexity, and many other linguistic aspects.

Look back at the three samples again and try to decide what the **purpose** and **audience** of each of them is. The **purpose** and **audience** for which a person is writing determines:

- content / topic (what to write about)
- genre (e.g., narrative, description, argument, etc.)
- tone/style (formal, informal, humorous, etc..)
- linguistic choices (vocabulary, grammar, etc.)
- rhetorical features (language choices within and between sentences)

When you finish, you can read our answer in the **Indicative Answers** section.

Task 2. The teacher as a writer

In this topic you will think about yourself as a writer: we hope this will make you become more aware of your students as writers.

Activity 3

Think about your own experiences as a writer:

- Do you ever write **voluntarily** (for example, a letter or chatty email to a friend or family member)?
- Have you ever written a letter to a newspaper or local magazine or other writing for publication?
- What kind of writing do you **prefer** to do?



Keep in mind that your students, like you, will have some things they are interested in writing and some they are not; and that these are probably different for different learners. This means that students as writers need to have a **variety of writing opportunities and experiences**. The teacher who enjoys writing for pleasure as well as necessity is a more understanding and creative writing teacher. She or he is probably a better assessor of writing too.

Activity 4

Below you will see short descriptions of four writing teachers and their approaches to the teaching of writing. Try to identify which elements of each teacher's beliefs, attitudes and knowledge have **teaching** as the focus, and which elements have **assessing** as the focus. Some of them are mixed as well.

Descriptions of four teachers:

T1: Teacher 1 teaches children aged 12-13. She believes that writing is a process and that teachers need to help young learners find things that interest them to write about, and that she needs to give them opportunities to get feedback at different stages of their writing.

T2: Teacher 2 teaches children aged 15-16 who will take exams at the end of the year. She knows that the exams for the English subject will emphasize not only ideas but also correctness of language, and she believes that correct grammar and a wide range of vocabulary are essential for the expression of ideas.

T3: Teacher 3 also teaches children aged 15-16 who will take exams at the end of the year. He believes that reading in English is a very good way to get ideas for something to write about in English, so he uses age-appropriate novels as well as non-fiction texts in his writing lessons, and he encourages discussion about the readings as part of the children's planning of their writing.

T4: Teacher 4 teaches learners aged between 16 and 19 who aim to go to university or other advanced education and training. Her teaching centres around 'teaching units' each of which focuses on a single topic, and requires several related writing tasks that are assessed by their peers or themselves with the final piece of writing teacher-assessed.

Your answer:

When you finish, you can compare your answer with ours in the **Indicative Answers** section.

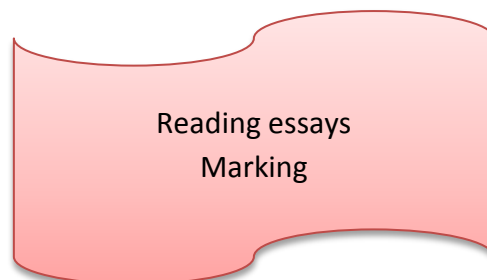
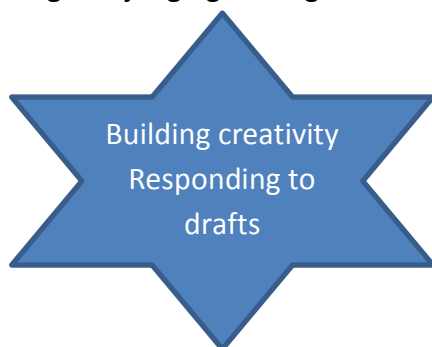
Use Activity 5 below as an opportunity to think about what kind of teacher you are. If you have never thought about yourself as an assessor, this is a good opportunity to think about that too.

Activity 5

Think about your teaching experience thus far and try to provide answers to the following questions:

- a) What do you **like** and what do you **dislike** about **teaching** writing?
- b) What do you **like** and what do you **dislike** about **judging** writing?

See if you can think of at least two things that you like and two things that you dislike about teaching and judging writing.



3.2 Topic 2: Types and Purposes of Writing Assessment



There are **two Tasks** in this Topic with a total of **four Activities**. These Activities invite you to reflect on the issues discussed in this Topic.

Task 1. Qualities of writing assessments in the classroom

Assessment is not the same thing as testing, grading or scoring.

Tests are a particular kind of formal assessment procedure carried out under controlled conditions. There are many other kinds of assessments that teachers can use (ranging from in-class writing, group writing projects, extended projects in which learners draw on speaking and listening, as well as reading in order to write content-rich texts; and portfolios). These are not tests, but ways of **gathering information** to monitor students' progress, and to be able to make sound educational decisions on what those students need next. You can find out more about the choices open to teachers in the course on [Alternatives in assessment](#).

Activity 1

Look back again at the third text from Topic 1 again. Our indicative answer told you that this was a learner's first notes in preparation for writing a personal essay. However, mind-mapping is also a good way of getting people writing (your students; yourself; colleagues).

Think about what is different in this text compared to the other two texts you looked at before. Do you think it would be a good idea to assess this text and texts like it? If not, what use might it be in your writing classroom, and in getting students ready to think about being assessed?

When you finish, you can look at our **Indicative Answers** for some ideas.

Activity 2

Writing is the skill that is most often used for testing because it provides a **permanent record of performance**, and it is generally believed to give an insight into the **thinking** behind the production of the response (see Lawrence, 1972; Deane, 2011).



Watch a video from the British Council Aptis test team (uploaded on 17 June, 2016), which gives a simple overview of some types and purposes of writing assessments, and a summary of things to think about: <https://www.youtube.com/watch?v=5-Tryu8KnIA>

According to the video, what are **four essential things** to think about when planning a good writing assessment?

Your answer:

Task 2. Writers at the centre of classroom assessments

As a teacher, you choose materials and design tasks to interest and motivate the students you teach. But formal tests are designed for a much wider audience, and that makes them more predictable in many ways; it may also make them more difficult. Large-scale test tasks have to be accessible to every person who takes the test, or it wouldn't be fair. The article by Kroll and Reid (1994) will show you of the issues this creates. But in classroom assessment of writing, the teacher, or group of teachers in a school, can develop writing tasks /prompts to suit their learners and make their own writing tasks for their real students.

Activity 3

While you are in your classroom, look at the **students**:

1. How many girls are there? How many boys?
2. What is their age range?
3. How much do you know about what each of them likes to do?
4. How much do you know about their strengths and weaknesses as learners?
5. What else is valuable to know about them as a group before you design a writing task for them that you plan to assess?

Write your thoughts down below if you wish, and then continue reading below.

Your thoughts:

This reflection on your students as people is really valuable to remind you that in your classroom, the way you assess their writing makes a difference to their **learning**, but also to **how they feel about themselves as learners**. When we look at writers on formal tests, you will notice some quite major differences.

Activity 4



Visit Glenn Fulcher's website at <http://languagetesting.info/videos/subs.html> and **watch the video** called "7. Writing", at the bottom of the page. In this video, Liz Hamp-Lyons talks about two key aspects of writing tests, **reliability** and **validity**.

1. Make sure you are clear about the difference between reliability and validity.
 2. What else does Liz Hamp-Lyons say is **important** when assessing writing?
- You can read our answer in the **Indicative Answers** section when you are ready.

3.3 Topic 3: Four Key Components of Writing Tests



There are **two Tasks** in this Topic with a total of **six Activities**. These Activities invite you to reflect on the issues discussed in this Topic.

Task 1. Designing writing tasks

When we plan to create a writing test, or to choose a writing test that is suitable for our own students, we need to keep in mind these four key components: the **task**, the **writer**, the **scoring procedure**, and the **reader(s)**. We thought about writers earlier in this Course, so now let's look at **tasks**. Both direct writing tests and classroom assessments of writing start with the shaping of students' work through well-designed tasks.

Activity 1

Below you can see **three writing tasks**, from different tests and at different levels. Read these tasks and think about:

1. the writing situation
2. the topic
3. the task(s)
4. the wording of the **rubric**
5. the rhetorical specifications
6. the scoring criteria

Based on these six points, would the following writing tasks be suitable for your students? Why or why not?

Writing Task 1

Write a short text about your holidays. Where did you spend your holidays? Where did you stay? What did you do? Who went with you? What didn't you like?

Your answer:

Writing Task 2

Every country in the world has problems with pollution and damage to the environment. Do you think these problems can be solved?

Your answer:

Writing Task 3

You have just joined a new sports club. You have also joined their online forum and received this message from another member. Write your reply in the box below. Remember to keep to 40 words only!

Hi! I've been a member of this club for two years, after my brother gave me a year's subscription as a gift. I really enjoy it! Why did you join?

(On this test, students are also asked to reflect on their own performance of the task.)

Student Reflection: How long did it take you to complete this task? Did you check it carefully? Remember that in the real Aptis test you will need to write three responses.

Think about how we could characterize each of these tasks according to the 6 aspects above.

Although this is a complex list with some unfamiliar terms, discussing it with colleagues, in your school or in an online group, will help you get a feeling for some of the ways that various aspects of a writing task can affect students' written text and, therefore, their score on a writing assessment. Discussing complex issues where there is no single completely 'correct' answer with colleagues is always useful.

Below you can see more about each of these tasks, and there is a link to follow for each of them.

Writing Task 1:

This is Communicative Activities-Writing, Task 11 in the CEFTTrain Project and you can find a sample text, a scoring scale, and some comments on the Project website at



<http://www.helsinki.fi/project/ceftrain/index.php.66.html>

Writing Task 2:

This essay task is on the Cambridge First (FCE) website at



<http://www.cambridgeenglish.org/images/181491-cambridge-english-first-fce-from-2015-writing-part-1.pdf>

If you would like to learn more about how Cambridge: First is scored, you can search online for the Cambridge English: First Handbook, or go to www.cambridgeenglish.org/images/167791-cambridge-english-first-handbook.pdf The section on the assessment of writing, which begins on page 32, includes sample tasks and rubrics, some sample texts written by learners, and examiners' scores with explanatory comments.

Writing Task 3:

This writing task is on an Aptis test prep at website



<https://www.slideshare.net/nursurayakma/assignment-5-aptis-b1-b2>

Note that there are also some suggestions there for people preparing to take the test to reflect on what they have done:

Additional reading if you wish – Follow the link below to read an article by Barbara Kroll and Joy Reid (1994):



<http://documentslide.com/documents/guidelines-for-designing-writing-prompts-clarifications-caveats-and-cautions.html>

This article is open access, and is reproduced from *The Journal of Second Language Writing* (see References). It gives a clear overview of the issues of designing writing tasks (often called 'prompts' among professionals in writing assessment) and will help you think about the questions below.

Activity 2

The three writing tasks

Go back to the three tasks you saw in Activity 1.

Ask yourself:

1. Does this task make **clear** to the writer what she or he has to do?
2. Does the writer know **how** his or her writing will be **assessed**?
3. Would any of these tasks be **suitable** for students you teach? Why? Why not?

If you can, **share your thoughts** with other teachers working on this Course or with other teacher-assessors of writing.

Your thoughts:

Activity 3

Remember that the effects of a writing task should be two-fold: to provide meaningful information about student writing quality; but also, to provide students with meaningful learning opportunities. **Do any of the tasks above provide good learning opportunities?**

Note down your thoughts. You may consider our answer in the **Indicative Answers** section if you wish.

Your thoughts:

Activity 4

Of course, as a teacher you do not need to ‘imagine’ your learners. As we practiced earlier, you can look at them in your classroom and think about what you know about them as people and their needs as writers. But on a test, the readers (usually referred to as ‘raters’) have to assume that all the writing they judge has been written by someone who fits into the mould of the generic hypothetical writer for whom the test was designed.

Think about what **advice** you need to give to your students before they take a large-scale writing test. Where can you find **support** to help you give them sound advice? Of course, first you, the teacher, must understand the writing test your learners are going to take. Try to do this Activity with other teachers, or at least discuss the ideas with some colleagues.

When you feel you know all the important information about that test, you may read our comment in the **Indicative Answers** section.

Task 2. Judging what has been written - The scoring procedure

The scoring of writing often gets the greatest emphasis in a writing test, but you should now realise that many other aspects are important. Because most teachers only experience a very small number of different options for scoring their students’ writing, they may not be aware of the very wide of possibilities that exist.

The first choice any testing organisation will have to make about how to judge/score writing is whether or not to use **criteria**. Criteria are the labels given to aspects of a performance that those responsible for making judgments have decided are essential in reporting the quality of the performance. For example, in Olympic diving judges use two criteria: criterion 1 is **difficulty** of the type of dive chosen; criterion 2 is **execution** of the dive itself. A diver might get high scores on both, or they might get a high score for the difficulty of the dive but a middling score for their skill in executing it, and so on. In writing assessment scoring is more complex, as the criteria used to judge writing will differ from test to test and from school to school or even from classroom to classroom.

Often small groups of teachers will agree to share their views of what scores (or grades, or marks—labels may vary) mean. They might start by agreeing to ‘mark’ all their own students’ writing out of 10. But if three teachers get together to talk about their marks they might find that that they don’t have the same mark on the same paper. Why not? The only way to find out is for them to talk about each piece of writing and see if they can figure out why the other(s) had a different judgment. So—what happens?

Activity 5

If you can, get together with two (or more) other teachers teaching the same age and level of students as you teach. **Read two texts** from each of your classes (so that's 6 texts altogether). Then, **score them** out of whatever you agree on (in general we don't score out of more than 10 in the present day). You will probably want to **keep notes** as you are reading.

Activities 5 and 6 are best done with a **small group of other teachers**. If you are doing the Course alone, and have no group of colleagues to do this Activity with, you may do these Activities using the first 2 texts you saw in Topic 1, Activity 2; or you can look ahead to Topic 5 Activity 5 and use the first 2 texts; or you can visit websites such as the Cambridge: First website. However—there is no substitute to talking to other teachers with similar students and issues.

When everyone has done that, begin with any of the texts and find out what score you each gave. You can use the classroom blackboard or whiteboard to display the scores, or just a big piece of paper. If you all agree on the first text—great! Talk about why you think it deserves that score. Even then you might find out you had different reasons.

At this point, you may read our comment in the **Indicative Answers** section. Then, continue reading.

Activity 6

Continue Activity 5 by discussing one or two pieces of writing where you didn't agree on marks. If possible, make an audio-recording of your discussion. While ideally you will reach agreement, or very close to agreement, it is also useful to listen to and think about very different opinions. Probably you will start to realise that certain points come up again and again, and that some members of your group value particular things about a text (and maybe about texts in general) more than you do. Think about the implications of that.

You may read our comment in the **Indicative Answers** section, as well.

3.4 Topic 4: Why the Readers are Important



This Topic consists of **three Tasks** with a total of **five Activities**.

It's time for us to think about the fourth component of a writing assessment: the **readers**. Although large-scale tests increasingly use partial machine scoring of writing (known as

automated writing evaluation—AWE), most writing in most contexts is scored entirely by human readers, or by a reader (a rater) and a computer-generated score. This means the job of the human readers is still extremely important. In most countries and levels of education, scoring tools have been designed to help readers be reliable raters.

Task 1. Being a teacher-rater

If you have done Activities 5 and 6 in Topic 3, probably you and your colleagues have made notes about what you have judged positively, and what more negatively, in your students' writing: you have been acting as raters as well as teacher-readers. This is what always happens when teachers have to judge students' writing. Increasingly, people feel that the best scoring methods for writing use an analytic (also called multiple trait) rating scale as well as clear instructions specifically chosen to fit the demands of the task and the needs of the writers. In general these days, we usually call those instructions, which on large-scale tests tend to be quite standardised, together with the actual scoring instrument (the rating scale) a **rubric**.

What's really valuable about rubrics and rating scales is the way they build bridges between teachers' **feedback** to formal assessment tools.

Activity 1

Watch a very short video from YouTube on '5 reasons for using a rubric' uploaded by Katie Palacios on 28 October, 2013.



<https://www.youtube.com/watch?v=0zrLUM5CWWE>

How many of the reasons given there are **assessment-related**, and how many are **teaching-related**?

Discuss with other teachers, or reflect personally, on ways you could use rubrics as part of your job as a teacher-assessor. Then you may go to the link below the video to see an interesting answer.

You may read our comment in the **Indicative Answers** section.

Task 2. Designing a rubric

If you want to design a rubric of your own based on what you learned about your students' writing in Topic 4 Activities 5 and 6, you can look at Examples 1 and 2 below, which show rubrics for tests you probably know well.

The combination of a rubric (task instructions) and criteria (rating scale) is an important tool to ensure full access to essential test information for test-takers, and helps ensure **fairness**. Here are two example sets of criteria for you to think about:

Example 1

The Cambridge Assessment ‘Main Suite’ of exams includes Cambridge: First, previously called First Certificate in English or FCE. This exam is aimed around the B1/B2 level of the CEFR, and the criteria for the writing component of the test are used by trained raters for scoring the FCE. You can see the full set of criteria and get good teacher support by downloading the full Teacher Handbook for Cambridge First at a link we have mentioned before:



<http://www.cambridgeenglish.org/images/167791-cambridge-english-first-handbook.pdf>

Example 2

You have met the Common European Framework of Reference for Languages, usually referred to as ‘the CEFR’, before in other Courses. You can download it at



<http://www.coe.int/en/web/common-european-framework-reference-languages/>

The CEFR has illustrative scales for written production, creative writing; and reports and essays. Here is the global (holistic) description for B2 Writing:

I can write clear detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.

You may find more CEFR scales at:



<https://rm.coe.int/168045b15e>

Activity 2

Consider this **different kind of rubric**:

The NorthWest Regional Education Lab/NCTE publish the ‘6+1 traits of writing’ scoring scale for writing: here are the names of the traits and their definitions:

Ideas—the main message

Organization—the internal structure of the piece

Voice—the personal tone and flavor of the author's message
Word Choice—the vocabulary a writer chooses to convey meaning
Sentence Fluency—the rhythm and flow of the language
Conventions—the mechanical correctness
Presentation—how the writing actually looks on the page

You can find more information about these traits in the website at



<http://educationnorthwest.org/traits/trait-definitions>

This generic set of 6 criteria is described at 6 levels, and is frequently used in US school systems for judging the writing of children aged between 8 and 17. These criteria also exist in a much more detailed form, but can be used in very flexible ways. The website has plenty of detail and teacher support.

- Which example do you think would be most helpful to you **as a teacher**?
- Which example do you think would be most helpful to you if you **need formally assess** students' writing in a European country?

A strength of the 6 + 1 approach is that a teacher can choose which of the criteria to use for their own teaching and assessing needs. However, this is a more advanced task.

Because the 6 +1 criteria were developed with children from regular school classes, which included both L1 and L2 learners; and because they are designed to be used across most school levels, deciding which of the criteria at which performance levels are suitable for your students is more difficult. Think about whether they are suitable for the students you teach before using them.

Activity 3

Think about – and talk with colleagues about – what we have discussed so far and **try to build your own rubric** for use with your own students' writing.

What **steps** would you follow? You will find our **Guidelines** at the very end of this Topic.

Your notes:

Activity 4

Now see if you can find a place on the rubric you have made where the **descriptor** we saw in Example 2 would fit:

I can write clear detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.

This Activity doesn't work very well, does it? Go to the **Indicative Answers** section to find out what the problem is.

Activity 5

Now look at the **three student texts** below. From the **rubrics** we have looked at in this course: Cambridge: First; the CEFR; CETrain; Aptis; or any one part of the 6 +1 traits we discussed earlier in this Topic, **choose a relevant rubric or part- rubric** (rating scale within the whole rubric). Use the links we have given you in this course, or other resources you may have. Try to work with at least one other teacher teaching the same level of students as you to **decide whether and how well that rubric would work to usefully judge and give feedback to your own students**.

When you have done this, go to the **Indicative Answers** section to read our answer.

Three texts for scoring

TEXT 1

Gabriela, a student from Romania, wrote this essay as part of preparing for the Cambridge FCE (now called Cambridge: First).

Task: In your English class you have been discussing the pros and cons of different types of transport. For your homework your teacher has asked you to write a composition giving your reaction to the following statement: "Cars should be banned from the centres of all major cities."
Write your composition for your teacher (around 120-180 words).

Gabriela's answer:

Nowadays, most of the people have their own car. In a survey that was made here in Romania, most of those who were questioned said that they couldn't live without a car. It would be too difficult for them to go somewhere because of the madding crowd.

It's true that cars are a must in people's life, but in the same time they are a major factor of pollution. What people haven't thought about is that 1 car pollute less the air than 20 cars. So, it's more convinient to use public transport instead of your own car, if you do think about the nature and the air you breath every second.

Another problem are the frequent traffic jams that happen mostly because of careless drivers or because of the rush hours. No one would like to be in the centre of Bucharest at hours like 7 a.m. or 4 p.m. because it would be like throwing yourself into a labyrinth you could hardly manage to get out of.

People should use more public transport than their own car, first of all because it would be cheaper, second of all because it would reduce pollution rate by aproximately 80% and the city it would be also less noisy. All these improvements would make our life much easier, because an unpolluted air makes our bodies healthier and we can do as much physical effort as usual, but with less energy consumed.

TEXT 2

Torre is a student from Chile who went to Denver Colorado for his final year of high school. In the text he had been asked to write a personal reflection about how he came to be a foreign student in the USA. This was an in-class writing (40 minutes).

Torre's text:

I arrived to Denver, Colorado on the twelfth of march of this year. The first month was terrible, I just wanted it to be a dream, and that when I woke up I would be back in my country, like nothing had occurred.

As days, weeks , and months went by, I began to believe that this was not a dream, and that there was nothing more to do, but resign to it . After seven months here, I am still not used to it . It is true, I have made a lot of friends, and that people here are as friendly as in my country, but it is like taking a fish from the sea and leaving it in the dessert, it is never the same. I cannot forget where I was born. "Miss" is a word very used in my vocabulary, for example: I miss my family, I miss the weather, I miss friends, I miss tastes, etc.. I never thought that something like this could happen to me, and I hope everything goes well for me, I find a second country to love, and someday return to my birth country.

Despite the faith I had when I talked with John, in this moment destiny does not have a purpose for me, but I hope to find it soon.

TEXT 3

Natalya is from Russia and was taking an entrance test for a college which has 'minimum readiness' requirements for English language. These requirements are usually lower than in most European countries.

Task: You will have 50 minutes to plan and write the essay assigned below. You may wish to use your 50 minutes in the following way: 10 minutes planning what you are going to write; 30 minutes writing; 10 minutes re-reading and correcting what you have written.

You should express your thoughts carefully and organize your ideas so that they will make sense to your reader. Correct grammar and sentence structure are important.

The task Natalya chose was: *"It always strikes me as a terrible shame to see young people spending so much of their time staring at television. If we could unplug all TV sets in our homes, children would grow up to be healthier, better educated, and more independent human beings."*

Do you agree or disagree? Explain and illustrate your answer from your own experience, your observations of others, or your reading.

Natalya's text:

I agree that television does have an harmful effect on young people. The young people will watch television then to do homework or prepare themselves for a test. They will make up excuses for not doing the homework or prepare themselves for the test, the excuse might be one of the following: it doesn't matter if I fail the test I will still pass the class, it's only one homework assignment I will be missing, or the teacher shouldn't fail me because I missed one or failed one test. This is how the young people feel, they would watch television before they would read a book. I know they feel that way because all my friends are like that and I use to be like that. I use to watch television then to read a book or do homework until I realized how important an education was. Watching television wouldn't make me do better in school it makes me do worse watching television wouldn't get me a job. Television changes some peoples life styles they try to be like the person on the television. After the program is over they might say I want to be like him or her. They don't realize that is a person playing the role of a character. If the character does something the young person might think that it is done like that in real life. The young person could get hurt doing something that they might seen on television. Then they wouldn't understand why it don't happen the same way for them. Until someone explains to them that certain this on television only happens on television.

Task 3. Feedback on the three texts from different perspectives

1. Feedback to Gabriela from her class teacher

Gabriela, this is a good attempt at the question. To improve this even further you could state whether or not you believe that cars should be banned from major cities. (Your answer implies that you think this but you don't state it explicitly!)

This essay is clearly organised and the points are put into appropriate paragraphs. Be a little careful with your use of contractions (e.g. 'haven't', 'it's') in a discursive essay as this task type needs to be slightly more formal than informal. There are no major language errors here - just some minor ones like word order, the odd spelling mistake or word choice. On the plus side, I like your interesting and ambitious use of language. Do you read a lot of English literature?

Fiona

2. Torre's self-reflection on his writing

Here is what Torre wrote about his own development as a writer during his final high school year:

SELF-REFLECTION

Dear instructor:

Before starting this class, I did not liked to write, for me the meaning of writting was boring. I just did drafts, because I had too. The idea of getting to this course didn't like me, but I had too because it was required for my major.

Now that I am finishing the semester I realize of all the new concepts I discovered from writting classes. My attitude had changed tremendously towards the semester, now I enjoy doing drafts. My imagination takes me where I have never been before, making a flow in my writting that I couldn't find before.

Right now I feel I am a good writer, because I put my feelings in a piece of paper, I had improved a lot. I see me us a person that have found the best way to express feelings through a paper.

My writting has changed a lot, at the begining I did short drafts, I couldn't do detailing, no ideas came to my mind, but now I have to limit my ideas because if not I would never finish writting.

As every writer I have good things, and things you still could make better. I feel that my potential is the detailing, that makes the reader situate deeply, in the context of my draft. I discovered that something I enjoy more than any other thing, is to write personal experiences, and showing what I feel.

On the other hand the weakest thing of my writting, is the limited variety of words I know, the

reason of this is because I am foreign student. I have learned more vocabulary during this course, and it is reflected along the drafts.

To finish I want to say that a good writing is the one that, is well organized, straight to the point in a clear and simple way. I also think that has to attract the readers attention from the beginning.

SINCERELY,

3. Essay exam rater's comments on Natalya's essay

Natalya's essay was scored as an overall 3, the lowest score for minimum readiness. Here are some comments on the writing:

This writer uses the language with relative ease, for example by stating a clear position and offering two arguments to support a point of view. The sentences are mostly clear and the spelling is reasonably accurate. However, the essay does not reach level 4 because the writer does not stay on track throughout the essay. There is only limited skill with sentence structure and punctuation, and only fair command of subject-verb agreement and verb forms. Also, the writer's vocabulary does not seem strong enough to fit the needs of this topic. There is vocabulary repetition and some inexact statement.

Task 4. A role for computers?

Computer technology is increasingly being used for the delivery of writing tests, for raters to score online, and automated writing assessment is being used with some writing assessment. For example, Educational Testing Service (USA) uses the internet-based e-rater® automated writing evaluation engine in combination with human raters to score the writing sections of the TOEFL iBT®. The e-rater engine is also sometimes used to provide the sole score in learning contexts, such as for formative use in a classroom setting where its adaptation is called the Criterion® online essay evaluation system. In the Criterion application, the automated scoring engine is used to generate individualized feedback for students. You can learn more about this on the ETS (Educational Testing Service) Research webpage at



https://www.ets.org/research/topics/as_nlp/writing_quality/

On that webpage there are also links to several other automated assessment and feedback programmes.

One certainty is that as time goes more, more writing assessments will be delivered and scored online. It is important that teachers play a role in planning for computer/internet-based assessment in their local context.

Remember that there is a list of further reading at the end.

Error correction and/or feedback

This course has not looked at feedback, although feedback to writing and writers is a very important part of learning to write well, and good teachers use **feedback** and **feedforward** as key tools to support their students' as they develop their writing skills. Go to the Course on **Feedback** to learn much more about this.

Portfolio assessment of writing

Portfolios are excellent tools for assessing students' writing over the course of whole year or a complete programme of writing instruction. However, they are not simple tools to use, and a full exploration of the use of portfolios is beyond the reach of this introduction to assessment literacy. To find out more information about **portfolio assessment**, go through the **Alternatives in Assessment** Course.

The British Council provides a useful introduction to the use of portfolio with younger learners in the context of the European Language Portfolio:



<https://www.teachingenglish.org.uk/article/portfolios-elt>

Activity 6

Guidelines for building your own rubric

You can build your own **rubric** for use with your own students' writing by following these simple steps:

a. decide which aspects of writing have been important to your discussion about your students' writing (In Topic 4 Activities 5 and 6);

b. look at other rubrics to see if any of the criteria are suitable for your own context;

The criteria you use should be related to the learning outcome(s) that you are assessing. For example, the Cambridge: First scoring instrument looks at 4 areas:

- Content
- Communicative Achievement
- Organisation
- Language

But you might want a clearer criterion/set of descriptions of performance than, for example, 'Communicative Achievement, which is rather vague.

Remember that you should choose the right criteria for the situation you are teaching in. You may want to look back at some actual examples of student work to see if you have omitted any important criteria.

c. Determine your performance levels.

Examples of performance levels may be, for example: Strong; At level; Below level; Weak. Or they may be numbers: 1-2-3-4-5-6. Looking at different rubrics will give you ideas.

d. Make a rubric grid to fill in your rubric. It should look something like this:

	Criterion 1	Criterion 2	Criterion 3	Criterion 4
Good				
At expected level				
Approaching level				
Weak				

e. The hardest part of this task is to create a description for each criterion and each performance level.

You might want to start with the strongest and weakest, and then fill in the middle levels based on your knowledge of common problems. It may be helpful to sort examples of actual student work into three piles: the very best, the poorest and those in between. Try to articulate what makes the good assignments good and the poor assignments poor. BUT remember that creating good rubrics/[rating scales](#) for formal writing assessments is difficult and usually done by specialists with a great deal of experience. This course can only give you a brief introduction, to help you assess your own students' writing as well as possible.

f. After you have used it, look closely at your rubric and the scores students have been awarded, and consider what has happened, and revise it as necessary. Always try to do this with other teachers.

There is a short and excellent explanation of how to design your own rubric by Heidi Goodrich Andrade (2000) called 'Using rubrics to promote thinking and learning' at






<http://www.ascd.org/publications/educational-leadership/feb00/vol57/num05/Using-Rubrics-to-Promote-Thinking-and-Learning.aspx>

Before you leave this Course, here is a reminder to you, as teachers:

What this course can cover is limited and there is much more out there for you, if you are interested, to find out about issues in reading and responding to writing.

Assessing Writing Skills - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I can distinguish between different ways of assessing student writing.			
I can understand the uses of the different types of writing assessments.			
I can understand the key elements of any writing assessment.			
I am sensitive to the effects of your choices of assessment tools on your learners.			

Indicative Answers

Topic 1: Activity 2

The purpose of the **first writing sample** is to display an ability to write a fairly simple text discussing some aspect of a content area (in this case, globalisation). This is a fairly typical kind of text found on a large-scale writing test. The purpose is to generate a large enough amount of writing from each test taker to be scorable on all dimensions of language. This implies it was written to be scored by a writing test rater, or for a class teacher who was preparing students for this writing test. In cases like this the individual identity of the student is not seen as very important.

The purpose of the **second writing sample** is probably to inform a classroom teacher of how well a class of students (of whom this is one) can write a personal reflection or 'story'. Note that this also helps the teacher understand more about each student, because each student's story will be different. This 'teacherly' purpose for the writing is very different from the purpose behind the writing students do for large-scale tests.

The **third writing sample** is very different. Early in a term, 'Elsie', the student, was asked to think about an interesting event in her childhood and plan an essay about it. This was part of a writing class that followed the 'process approach': you can learn more about process approaches to writing at



<https://www.teachingenglish.org.uk/article/approaches-process-writing>

In Elsie's 'text' we can see: a list (and maybe we can see some development of her thinking as the list grows); a mind-map (a kind of spider's web) using some words from the list, with several different ideas spreading out in the web; and at the top right, squashed into the remaining space, a sentence explaining what has been happening as she did this pre-writing. SO - when Elsie did this work, there were several purposes all at the same time. It might be interesting to know that the essay she ended up writing was called 'The windmill he made me' and was a very strong description of her father, the windmill-making, and her relationship with him.

Topic 1: Activity 4

In this key, the elements of each teacher's beliefs, attitudes and knowledge that have teaching as the focus are in **blue**, and those which have assessing as the focus are in **red**, while those which are mixed are in **green**.

T1: Teacher 1 teaches children aged 12-13. She believes that **writing is a process** and that **teachers need to help young learners find things that interest them to write about**, and that she needs to give them **opportunities to get feedback at different stages of their writing**.

T2: Teacher 2 teaches children aged 15-16 who will take exams at the end of the year. She knows that **the exams for the English subject will emphasize** not only **ideas** but also **correctness of language**, and she believes that **correct grammar and a wide range of vocabulary are essential for the expression of ideas**.

T3: Teacher 3 also teaches children aged 15-16 who will take exams at the end of the year. He believes that **reading in English is a very good way to get ideas for something to write** about in English, so he uses **age-appropriate novels as well as non-fiction texts** in his writing lessons, and he **encourages discussion about the readings** as **part of the children's planning of their writing**.

T4: Teacher 4 teaches learners aged between 16 and 19 who aim to go to university or other advanced education and training. Her teaching **centres around 'teaching units' each of which focuses on a single topic**, and requires **several related writing tasks** that are **assessed by their peers or themselves** with the **final piece of writing teacher-assessed**.

Topic 2: Activity 1

Text 3 was a mind-map created in a writing class that followed the 'process' approach. It doesn't have any of the formal textual structures we looked at before. It would be very difficult to assess using any formal scoring system. But mind-mapping and other process techniques can be very useful in the writing classroom. They get learners thinking, and they get learners writing down their ideas, so that they can see them and think more about them. The various techniques in the 'process' approach help students overcome writing anxiety and uncover their

own ideas, opinions and memories. When learners begin to enjoy some kinds of writing, they will feel better even about writing in an exam context. Read more about this at



<http://www3.canisius.edu/~justice/CSTmodule-final/CSTmodule-final17.html>.

Topic 2: Activity 2

Teachers have two great advantages over testing specialists when assessing writing in their classes and with their own students: first, they know the students well enough to be aware of their **strengths and weaknesses**; second, during a term or a whole school year they can assign a wide range of writing tasks in order to see what their students can do in different contexts.

The video ended by suggesting three important issues to remember:

Is the assessment **appropriate**?

Is the way it is scored **accurate**?

Is the process **fair**?

If you are not sure what is meant by any of those terms, go the course titled [The ABCs of assessment](#) and read about them. Note that ' **fairness** ' is one of the most complex concepts in all of [language assessment](#). It is only briefly mentioned here, and to fully understand it you should read several discussions of fairness, as they will cover a range of definitions and viewpoints (see for example: Hamp-Lyons 2000; Song 2014; Suskie 2000, all listed in full in the References).

Topic 2: Activity 4

The key qualities of formal (writing) tests as seen by testing bodies are **reliability, validity, practicality and impact**. In case you need more information on these qualities, check out [Topic 4](#) of [The ABCs of Assessment](#). When you, as a teacher, evaluate tests to see if they are suitable for your students—or perhaps to advise students which tests would be most suitable for them to take, you will also want to look at some more concrete aspects of the test.

Another key thing to think about when selecting or designing a writing test, or preparing students to take an external, formally scored, writing test, is **test impact** . Go to the course on [Test impact](#) to learn more about what impact is, and why it is important.

A fourth area that testing bodies find important is **practicality**. This often means cost, but also refers to practical issues, such as how to get exam papers to test centres around the world on time and securely; how to find enough raters to score writing, etc. Schools have similar if less severe issues, because of their 'home' context, and because the smaller the group being assessed the easier it is to deal with practical issues.

Topic 3: Activity 3

Tasks provide good learning opportunities when (1) they give teachers meaningful information about student writing quality; but also, (2) provide the student with [feedback](#) they can use. There are many kinds of feedback, but this course has suggested that using rubrics within teaching as well as within assessment helps students understand their own strengths and weaknesses.

Topic 3: Activity 4

Of course, as a teacher, it's your job to find out all about the test your students are taking and make sure you understand it. In this course we have offered you basic information about how to assess student writing; but there is much more that you could find out for yourself, and we have added some external links to get you started if you are interested.

Nowadays almost all language tests have websites that provide teaching support as well as test information and samples. It is usually best to use the official websites of the tests you are working with: some of the other sites set up by language schools or private companies may not have all the accurate and up-to-date information. Also, if you are choosing among potential tests you might use, it's a good idea to check how much support they offer to teachers as one of the criteria for making your choice.

Topic 3: Activity 5

If (probably **when**) you find a text where all of you differed on scores, what can you do? Remember that if these marks are only for classroom use (for you as teachers to identify your students' needs) then variation is fine. But what if they are to be reported to the school principal or on the student's official year report? The use of scores is a major consideration in your choices of how to make and record/report judgments.

When groups of teachers run into disagreement about what a student's writing is worth, that's the perfect opportunity to discuss very closely what each of you rewarded and what each of you penalized.

Topic 3: Activity 6

The fact that certain points come up again and again in talking about the qualities of texts, and that teachers value particular things about a text (and maybe about texts in general) more than others, tells us that 'writing' is not a single phenomenon. There have been very many debates in the writing assessment field about what the qualities of 'writing' really are: it seems that there

is no list that everyone can agree on. That is why most researchers agree that writing needs to be judged by multiple criteria. If you would like to read more about holistic and multiple trait/analytic scoring of writing visit Liz Hamp-Lyons' ResearchGate page at



https://www.researchgate.net/profile/Liz_Hamp-Lyons/contributions and look for “Farewell to holistic scoring. Part Two: Why build a house with only one brick?”

Topic 4: Activity 1

In the video, Katie Palacios gives five reasons for using rubrics: increased feedback; students' opportunity to self-assess; allowing students to revise based on clear expectation; aligning assessment with learning objectives; and, to grade faster. One more excellent use of rubrics we should mention is the opportunity they offer to let students work together to discuss their work and to carry out peer assessment.

In the same area of YouTube you will find a large number of videos about rubrics—use them critically, as the content quality is sure to vary.

Topic 4: Activity 4

The descriptor doesn't work very well because **one** piece of writing can never tell you how well someone can write “**on a wide range of subjects**”, and you probably don't know what the individual's “interests” are! Then, the descriptor states **two** things this writer can do: pass on information/give reasons (and you might think that's already two different objectives), and “write a (particular kind of) letter”. We are sure you realise by now that different purposes for writing, different subject matter, different genres, etc., are probably going to need different rubrics. A very broad descriptor like the one here might work well as part of an overall end-of-year assessment, but not for a single piece of writing.

Further Reading

For a theory-driven and detailed treatment of the use of portfolios for teaching writing and for assessing writing, see Liz Hamp-Lyons and William Condon (2000). *Assessing the Portfolio: Principles for Practice, Theory, and Research*. Cresskill, NJ: Hampton Press.

Assessment Reform Group (n.d.). The role of teachers in the assessment of learning. <https://www.aiaa.org.uk/content/uploads/2010/06/The-role-of-teachers-in-the-assessment-of-learning.pdf>

Finally, if you wish to explore ideas on how to **assess SpLD students' writing skills** visit the following website of the DysTEFL project and follow the suggestions in Chapters 9 (Reading and Writing) and 10 (Assessment)



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

or download the DysTEFL training materials and work directly from there:

If you are a trainer: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

If you are a teacher: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

And read the following:

Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning differences. Bristol: Multilingual Matters. Chapter 8. Assessment

References

Andrade, H. G. (2000). Using Rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.

Deane, P. (2011). *Writing assessment and cognition*. ETS Research Report 11-14. Educational Testing Service. <https://www.ets.org/Media/Research/pdf/RR-11-14.pdf>

Hamp-Lyons, L. (2000). Ethics, fairnesses and developments in language testing. In C. Elder, A. Brown, L. Grove, K. Hill, N. Iwashita, T. Lumley and T. McNamara (Eds.), *Experimenting with Uncertainty: Essays in Honour of Alan Davies; Studies in Language Testing 11*: 222-227. Cambridge: UCLES/Cambridge University Press.

https://www.researchgate.net/publication/271214432_Ethics_Fairnesses_and_Developments_in_Language_Testing

Kroll, B. & J. Reid (1994). Guidelines for designing writing prompts: Clarifications, caveats and cautions. *Journal of Second Language Writing*, 3(3), 231-255.

<http://documentslide.com/documents/guidelines-for-designing-writing-prompts-clarifications-caveats-and-cautions.html>

Kunnan, A. J. (2000). Fairness and justice for all. In A. J. Kunnan (Ed.). *Fairness and validation in language assessment*. Cambridge: CUP, pp. 1-13.

Lawrence, M. (1972). *Writing as a Thinking Process*. Ann Arbor: University of Michigan Press.

Song, X-M. (2014). *Test fairness in a large-scale high stakes language test*. PhD. U. of Ottawa.

http://ospace.library.queensu.ca/jspui/bitstream/handle/1974/12229/Song_Xiaomei_201406_PhD.pdf;jsessionid=BD33A00C4EF7DD11ACEE01F917CC8794?sequence=1

Suskie, L. (2000). *Fair Assessment Practices: Giving Students Equitable Opportunities to Demonstrate Learning*. AAHE Bulletin.
http://ctl.ok.ubc.ca/_shared/assets/FairAssessmentPractices45517.pdf

4. Assessing Listening Skills

Anthony Green

❖ Aims

This Course will introduce you to the design and use of tasks for assessing listening skills. It will show you a variety of assessment techniques and suggest how you can decide whether they fit the needs of your learners. The course will help you to choose and adapt material from published books, tasks made by other teachers or to prepare your own listening assessments to use in the classroom.

❖ Expected Outcomes

Upon completion of this Course, you will be able to:

1. Identify some of the problems that learners have with listening
2. Distinguish between different kinds of listening text
3. Distinguish different types of listening
4. Find suitable recordings to use for listening tasks
5. Demonstrate awareness of the advantages and disadvantages of different types of listening task
6. Select or design appropriate tasks to assess your students for different purposes

❖ Key Concepts

- Tests and assessments
- **Assessment for learning**
- Listening
- **Feedback**

Brief Introduction

Before going through the first Topic of this Course, think about the following question.

Introductory Activity: What makes listening difficult?

Why do learners find listening difficult?

Write down five things that make listening more difficult than other skills: speaking, writing and reading.

1. _____
2. . _____
3. . _____
4. . _____
5. . _____



Now follow this link to watch a short presentation about aspects of listening that can make it difficult for learners. <https://tinyurl.com/TALEL1>

4.1 Topic 1: What is Listening?



The main aim of this Topic is to introduce different kinds of listening text and to connect these to students' needs as learners of English. There are three Tasks in this Topic.

They ask you to:

- reflect on your own listening experience
- connect your experiences to a list of kinds of listening text and
- relate your experiences to a list of kinds of listening text and to students' needs as users of English.

Task 1. Listening experiences

Activity 1 - Recent listening experiences

Make a list below of your different experiences of listening to people speaking over the last 24 hours. Who was speaking? Why were you listening to them?

Who was speaking?	Why were you listening to them?

Think about these experiences.

Which involved seeing the person who was speaking?

Which involved other kinds of visual information?

Which involved recordings?

Write your answers in the box below.

Your answers:

Task 2. Kinds of listening text

Activity 2 - A list of kinds of listening texts

Look at this list of kinds of listening texts (this is taken from the Common European Framework of Reference or CEFR. There will be more about the CEFR in Topic 2, Task 4).

public announcements and instructions public speeches, lectures, presentations, sermons rituals (awards ceremonies, religious services) entertainment (drama, shows, readings, songs) sports commentaries (football, cricket, boxing, horse-racing, etc.)	public debates and discussion inter-personal dialogues and conversations telephone/video-conferencing conversations job interviews news broadcasts
---	---

Activity 3 - Matching kinds of text to your experience

Can you match the listening experiences you listed in Task 1 to the kinds of text in the list above?

Were any of your listening experiences from Task 1 not included on this list of kinds of text? What were they?

What kinds of listening text did you listen to most often? Which ones were most important for you? Which were the most interesting?

Share your list with other teachers. Were your experiences similar to other people's?

Write your answers in the box below.

Your answers:

Task 3. Students' listening needs

Now think about your students. Which kinds of listening text are (or will be) most important for them (in English)?

Are there any kinds of text that they do not (or will not) need to listen to in English, or that are less important for them?

Write your answers in the box below.

Your answers:

Activity 4 – Students’ listening

Look again at the list of kinds of listening texts from the CEFR.

Put ticks in the boxes on the left to show which kinds of listening text you think your students need to listen to.

<input type="checkbox"/> public announcements and instructions	<input type="checkbox"/> news broadcasts
<input type="checkbox"/> public speeches, lectures, presentations, sermons	<input type="checkbox"/> public debates and discussion
<input type="checkbox"/> rituals (awards ceremonies, religious services)	<input type="checkbox"/> inter-personal dialogues and conversations
<input type="checkbox"/> entertainment (drama, shows, readings, songs)	<input type="checkbox"/> telephone/video-conferencing conversations
<input type="checkbox"/> sports commentaries (football, cricket, boxing, horse-racing, etc.)	<input type="checkbox"/> job interviews

How much practice do you give your students in listening to these kinds of listening texts?

Activity 5 - Student experiences

If you are teaching a class now, give your students the **list of kinds of listening text** in Activity 3. Ask them to put the kinds of listening into three groups:

1. Very important for me in my future as a user of English
2. Useful for me in my future as a user of English
3. Not useful for me in my future as a user of English

Do your students agree with you about the most important kinds of listening text for them?

4.2 Topic 2: Listening Types



The main aim of this Topic is to introduce different types of listening, to connect these to the kinds of listening text that you looked at in the list presented in Topic 1 and to consider some of the factors that can make listening to different texts easier or more difficult for language learners.

There are four Tasks in this Topic. They ask you to

- Identify different types of listening
- Connect types of listening to kinds of listening text (Topic 1)
- Identify factors that make listening easier or more difficult
- Judge the Common European Framework level of listening tasks

Task 1. Types of listening

When people listen to different kinds of text, they pay attention to different kinds of information. For example, on the train, you may listen to announcements to check the name of the next station, but you may decide to stop paying attention if the announcement is not relevant to you. This is very different from listening to a friend when you are having a conversation. Your friend probably expects you to show your interest by reacting (smiling, nodding, making 'interested' sounds) and asking questions.

Task 2. Kinds of text and types of listening

Activity 1 - Examples of types of listening

We can think about the way we listen in terms of different types of listening. Brown (1994) suggested six types of listening that are important in language classrooms.

Can you connect the six types (1-6) to the examples (a-f)?

- | | |
|--|--|
| 1. Reactive listening, focussing just on the sounds heard | a) A classroom drill in which the students try to copy words and phrases that they hear |
| 2. Intensive listening, focusing on the form of the language | b) Discussing with classmates about what kind of present to buy for your English teacher |
| 3. Responsive listening, showing understanding | c) Following simple instructions: stand up, open the door, pick up your bag, draw a box |
| 4. Selective listening, picking out important information such as dates, times, locations. | d) Listening and identifying how many words you hear or deciding which sounds are questions and which are statements |
| 5. Extensive listening, understanding longer texts and connecting the ideas. | e) Listening to a five-minute lecture on the internet and taking notes |
| 6. Interactive listening combining listening with speaking – and sometimes other skills as well. | f) Listening to announcements in a train station to find the time and platform of your train |

Activity 2 - Most important types of listening

When we listen to different texts, we use different listening skills.

Look again at the list of listening experiences you made in Task 1. Which types of listening were most important in each case?

Please write kind of listening text (e.g. announcements) and listening type (selective listening) below.

Kind of listening text	Listening type

Task 3. Factors that add difficulty

Some kinds of listening text and types of listening skill are easier for language learners than others. It is easy to listen responsively to instructions such as:

‘Please stand behind the yellow line.’

It is more difficult to listen interactively in a discussion on a complex topic like:

‘Would you agree that examinations generally have a beneficial influence on education?’

Several factors can make it easier or more difficult to understand speech. For example, it is more difficult to understand a speaker in a crowded room where other people are also speaking.

Activity 3 - Trouble listening

Can you think of another three factors that make it hard to listen and understand?

1
2
3

Task 4. The Common European Framework of Reference levels

The Common European Framework of Reference for Languages (CEFR) is designed to give people involved in language education around the world a shared way to describe language abilities.

Activity 4: The CEFR



If you are not already familiar with the CEFR, please watch this short video from Canada about its uses.

Video: www.youtube.com/watch?v=JwIVExKgEuQ

Activity 5: CEFRTrain project



Follow this to go to the CEFRTrain project website: www.helsinki.fi/project/ceftrain/

Go to the Training area, read the instructions and click on 'Communicative Activities'. Look at some Listening tasks. Can you judge the CEFR level?

Activity 6: CEFR levels

The CEFR describes easier and more difficult listening activities that learners 'Can Do' at different levels of ability.

Can you match these descriptions to their CEFR level?

Remember that A1 is the lowest and C2 the highest level.

1. Can understand the main points of radio news bulletins and simpler recorded material about familiar subjects delivered relatively slowly and clearly.	C2	C1	B2	B1	A2	A1
2. Can keep up with an animated conversation between native speakers.	C2	C1	B2	B1	A2	A1
3. Has no difficulty understanding any kind of spoken language, whether live or broadcast, delivered at fast native speed.	C2	C1	B2	B1	A2	A1
4. Can understand simple directions relating to how to get from X to Y, on foot or with public transport.	C2	C1	B2	B1	A2	A1
5. Can easily follow complex interactions between third parties in group discussion and debate, even on abstract, complex unfamiliar topics.	C2	C1	B2	B1	A2	A1
6. Can follow speech that is very slow and carefully articulated, with long pauses for him/her to assimilate meaning.	C2	C1	B2	B1	A2	A1

4.3 Topic 3: Finding Materials for Listening



The main aim of this Topic is to introduce some important issues to consider when looking for recordings to use in assessing listening skills. There are three Tasks in this Topic.

They ask you to think about what makes a recording suitable for use in assessments and to compare different kinds of recordings. You will also look at how to find and adapt suitable recordings from the internet and issues to consider when making your own recordings to use for assessment.

Task 1. Finding or making recordings

Teachers can assess some listening skills just by speaking to the class, but it is also important to expose students to different voices and to kinds of listening text (such as conversations) that the teacher alone can't provide.

Many publishers offer additional listening material that can be used for assessment purposes alongside their textbooks, but sometimes this is not enough, or not what the teacher is looking for. Unfortunately, finding suitable recordings for listening can be difficult.

When creating or looking for recordings to assess listening abilities in the classroom, what factors do you think are most important?

Activity 1: Finding or making recordings

Decide whether you agree or disagree with a number of statements about recordings used in assessing listening skills.

1. The recording must include key vocabulary and grammar points I have taught.	I agree	I don't have an opinion	I disagree
2. Only native speakers of English should be on the recording.	I agree	I don't have an opinion	I disagree
3. The people should speak slowly and clearly so that students can understand.	I agree	I don't have an opinion	I disagree
4. The language used has to be completely accurate, with no grammar mistakes.	I agree	I don't have an opinion	I disagree
5. The recording must be authentic: if it is a news broadcast, it ought to come from the BBC or CNN, not from an EFL text book.	I agree	I don't have an opinion	I disagree
6. There shouldn't be any words or expressions that the students are not expected to know.	I agree	I don't have an opinion	I disagree
7. The topic should be something that the students have studied.	I agree	I don't have an opinion	I disagree
8. The recording should be a kind of listening text that the learners have experienced in class before.	I agree	I don't have an opinion	I disagree
9. It's a good idea to use written stories from newspapers or magazines and read them aloud.	I agree	I don't have an opinion	I disagree

Task 2. Locating recordings on the internet

Activity 2: Recordings on the internet

Using the internet, try to find:

- One five-minute talk giving advice on crime prevention
- One personal conversation between two friends about their plans for a holiday

How easy was it to find the recordings?

Was either of the recordings you found made especially for language learners?

What (CEFR) level of listening ability do you think that a learner would need in order to understand these recordings? Is there anything you could do to make the recordings easier for students to understand?

Do you think you would use one or both of these recordings with your students? Why? / Why not?

If you find a written text on the internet, you can adapt it to make it easier for learners. Sound recordings are not easy to find and are more difficult to edit or adapt. There are a lot of examples of quite formal talk on the internet - like lectures and interviews. TV shows and films are also easy to find. But it is much more difficult to find examples of informal conversations that are not specially made for language learners.

Perhaps because of these problems, in a survey, both teachers and students reported to us that listening is the skill that is assessed least often in English language classrooms.

Want to find additional listening material? Check out these **online listening resources**



The British Council Listen and Watch - A range of video, audio and reading materials. Listening materials come with activities and transcripts.

<http://learnenglish.britishcouncil.org/en/listen-and-watch>



BBC Learning English - A range of video and audio material that includes English lessons, dramas and news reports. Mostly for B1 level and above.

<http://www.bbc.co.uk/learningenglish/>



Voice of America: Let's Learn English - A similar mix of teaching material and news to the BBC, but in American English. <http://learningenglish.voanews.com>



Macmillan Education onestopenglish - This publisher's website has a range of free listening material for learners including interviews and dramas. The recordings have transcripts and exercises. <http://www.onestopenglish.com/skills/listening>



ESL student lounge - Listening material designed for learners from elementary to advanced level. Features different varieties of English (mainly American and British). Recordings have transcripts and a variety of exercises. <http://www.esl-lounge.com/student/listening.php>



Randall's ESL Cyber Listening Lab - Another website, with mostly American recordings, that has transcripts and exercises made for learners at different levels.

<http://www.esl-lab.com>



The BACKBONE project - A collection of interviews with native and non-native speakers of English (as well as other European languages) on topics that range from culture and region to education and health, business and technology. Transcripts are provided along with teaching resources.

<http://webapps.ael.uni-tuebingen.de/backbone-search/faces/initialize.jsp>

Task 3. Making your own recordings

Because it can be difficult to find certain kinds of listening text, it may be necessary for teachers to create their own recordings to use for assessment. These days many people have phones that can also be used as video cameras and voice recorders. This makes it reasonably easy to prepare recordings that are good enough to use for classroom assessments (although the quality will of course be much better if you can use professional equipment). Of course, it is always important to test the equipment before making a full recording.

Plan how you would make a recording to assess students' listening abilities. Can you find people to speak on your recording? What kinds of people can you ask?

Activity 3: Advice on making recordings

Your recording will need to be:

- Long enough for you to produce the number of questions that you need.
- At the right level of difficulty for the learners.
- On a topic that learners will not find offensive or distressing (war, death, disease).
- On a topic that is reasonably familiar to the learners. Often a short outline is helpful as a way to introduce a topic– e.g. you will hear a man and a woman talking about food they eat when they are on holiday in other countries.
- Not something the learners have heard before.

Think about the factors that can make listening easier or more difficult listed in Topic 2, Task 3 and the features of useful recordings in Topic 3, Task 1.

You will need to balance the naturalness of the talk on the recording with consideration for the level of the learners. If the learners are A1 or A2 listeners, it is appropriate to ask speakers to use simple language, to speak a little more slowly and more clearly than usual and to leave more pauses between utterances.

It is usually best to make recordings in a very quiet place. Noises that you do not notice while speaking can be picked up by a microphone and make your recording unintelligible.

Avoid using scripts when making recordings. When people read out written texts, it never sounds very natural. If you want people to talk about certain topics, provide them with an outline or guidelines rather than a script. The result should sound more realistic.

Video can add to the interest of the recording. Seeing people speaking can make it easier to understand them and pictures can provide helpful clues, but video can also make editing more difficult.

If you make a lot of recordings, it is worth learning how to edit and adjust them: to remove very long pauses, noises and other unusable parts. Unfortunately, most phones do not offer many audio editing features. Luckily, there are many software packages that do allow you to edit sound files. One that is popular with teachers because it is free, but flexible, is called Audacity. You can download this from audacity.sourceforge.net.

Activity 4: People to speak on your recording

Can you find people to speak on your recording? What kinds of people can you ask?

Choose the people who could help you from the list below.

- Other teachers at your school
- Friends and family members
- Residents of your local area
- Tourists and other visitors to your area
- Teachers in other countries (using video conferencing or telephone)

4.4 Topic 4: Using Recordings for Assessment



The main aim of this Topic is to explore the use of tasks to assess different types of listening. There are two Tasks in this Topic.

They ask you to think about how you use recordings in the classroom to find how well learners understand what the speakers say and to find the causes of difficulties for learners.

Task 1. Dialogues






Watch a presentation about using a dialogue to assess listening:
<https://sway.com/eUK5Osj2as7KfKnm>

Task 2. Monologues



Watch a presentation about using a short lecture to assess listening:
<https://sway.com/s3v0iFoZhvcmmtEq>

Assessing Listening Skills - Outcomes

My Outcomes			
Identify some of the problems that learners have with listening			
Distinguish between different kinds of listening text			
Distinguish different types of listening			
Find suitable recordings to use for listening tasks			
Demonstrate awareness of the advantages and disadvantages of different types of			
Select or design appropriate tasks to assess your students for different purposes			

Indicative Answers

Topic 2 Activity 1 - Examples of types of listening – Answers

- | | |
|---|--|
| 1. Reactive listening, focussing just on the sounds heard | a) A classroom drill in which the students try to copy words and phrases that they hear |
| 2. Intensive listening, focusing on the form of the language | d) Listening and identifying how many words you hear or deciding which sounds are questions and which are statements |
| 3. Responsive listening, showing understanding | c) Following simple instructions: stand up, open the door, pick up your bag, draw a box |
| 4. Selective listening, picking out important information such as dates, times, locations | f) Listening to announcements in a train station to find the time and platform of your train |
| 5. Extensive listening, understanding longer texts and connecting the ideas | e) Listening to a five-minute lecture on the internet and taking notes |
| 6. Interactive listening combining listening with speaking – and sometimes other skills as well | b) Discussing with classmates about what kind of present to buy for your English teacher |

Topic 2 Activity 3 – List of factors that make listening easier or harder

Here is a list of factors that can make listening easier or more difficult for language learners.

Makes listening easier	← →	Makes listening more difficult
One person speaking	← →	Multiple speakers
Clear articulation	← →	Mumbling, unclear speech
Speaking slowly	← →	Speaking quickly
Seeing the speaker	← →	Not seeing the speaker
A familiar topic	← →	A topic the listener knows little about
Restricted vocabulary	← →	Idiomatic expressions, specialist terms
Short, simple utterances	← →	Long, complex utterances
A quiet environment	← →	A noisy environment like a café or airport
An accent that's familiar to the listener	← →	An accent that's unfamiliar

Topic 2 Activity 6 – CEFR levels

1. Can understand the main points of radio news bulletins and simpler recorded material about familiar subjects delivered relatively slowly and clearly.	C2	C1	B2	B1	A2	A1
2. Can keep up with an animated conversation between native speakers.	C2	C1	B2	B1	A2	A1
3. Has no difficulty in understanding any kind of spoken language, whether live or broadcast, delivered at fast native speed.	C2	C1	B2	B1	A2	A1
4. Can understand simple directions relating to how to get from X to Y, by foot or public transport.	C2	C1	B2	B1	A2	A1
5. Can easily follow complex interactions between third parties in group discussion and debate, even on abstract, complex unfamiliar topics.	C2	C1	B2	B1	A2	A1
6. Can follow speech that is very slow and carefully articulated, with long pauses for him/her to assimilate meaning.	C2	C1	B2	B1	A2	A1

Topic 3 Activity 1 – Finding or making recordings: suggested answers

1. The recording must include key vocabulary and grammar points I have taught.

Although it is often best to use recordings that mostly include language that students know, listening is not only a matter of grammar and vocabulary, and assessment should focus on

listening abilities – recognizing and understanding sounds and understanding listening texts, not grammar.

2. Only native speakers of English should be on the recording.

Listening in the classroom should reflect what learners will need to do outside the classroom. All users of English will need to listen to, and understand, non-native speakers as well as native speakers.

3. The people should speak slowly and clearly so that students can understand.

It is reasonable to use slow, simplified language with learners in the early stages of studying English. But it's also important to help learners become familiar with natural-sounding speech as soon as possible so that they become attuned to it.

4. The language used has to be completely accurate, with no grammar mistakes.

It is sometimes argued that learners should only be presented with perfectly accurate speech so that they do not learn bad habits, but natural, fluent speech is characterized by inaccuracy. If learners only hear accurate recordings, they will be unprepared for listening in the real world.

5. The recording must be authentic: if it is a news broadcast, it ought to come from the BBC or CNN, not from an EFL text book.

Hearing the language as it is used outside the classroom is useful, but it can also be overwhelming for learners. Simplified recordings may also be useful. It should be remembered that news broadcasts are only one kind of listening text. They are partly scripted and do not involve much spontaneous conversation.

6. There shouldn't be any words or expressions that the students are not expected to know.

See question 1. Students may benefit from some practice in listening to texts that include a lot of language that they don't know.

7. The topic should be something that the students have studied.

At lower levels learners find it difficult to follow unfamiliar topics. Using new topics in an assessment may prevent learners from showing their listening ability. At higher levels, it may sometimes be important to assess whether listeners are able to listen to learn about new subjects.

8. The recording should be a kind of listening text that the learners have experienced in class before.

Assessments should not generally involve *kinds* of text that learners have never experienced before.

9. It's a good idea to use written stories from newspapers or magazines and read them aloud.

Recordings should reflect spoken language. Spontaneous speech is very different to writing. Reading from a book or newspaper – or from a play or scripted dialogue – is not generally a good way to prepare a listening task. It is usually **better** to improvise based on an outline or list of topics.

Further Reading

Videos



Follow this link to watch a British Council video titled 'Assessing Listening': www.britishcouncil.org/exam/aptis/research/projects/assessment-literacy/listening



Follow this link to watch an interview with Gary Buck, an expert in assessing listening: <http://languagetesting.info/video/main.html>

Brief readings

Green, A. (2013). *Exploring Language Assessment and Testing*. Abingdon: Routledge. Chapter 5.

Flowerdew, J., & Miller, L. (2012). Assessing listening. In Coombe, C., Davidson, P., O'Sullivan, B., & Stoyhoff, S. (Eds.) *The Cambridge guide to second language assessment*. Cambridge University Press. pp. 225-233.

Books

Buck, G. (2003). *Assessing Listening*. Cambridge: University Press.

Field, J. (2008). *Listening in the Language Classroom*. Cambridge: University Press.

Finally, if you wish to explore ideas on how to **assess SpLD students' listening skills** visit the following website of the DysTEFL project and follow the suggestions in Chapters 8 (Listening and Speaking) and 10 (Assessment)



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

or download the DysTEFL training materials and work directly from there:

If you are a trainer: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

If you are a teacher: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

And read the following:

Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning differences. Bristol: Multilingual Matters. Chapter 8. Assessment

5. Assessing Speaking Skills

Ildikó Csépes & Adrienn Fekete

❖ Aims

In this course, you can learn about how L2 learners' speaking performances can be influenced by specific features of the context, for instance the task they need to carry out, or certain individual characteristics of the participants themselves. It is important to bear in mind all this in order to design appropriate speaking assessment tasks to help learners perform to the best of their abilities. Since learners' performance levels are typically assessed in relation to the **Common European Framework of Reference (CEFR)**, you can also familiarize yourself with some scale descriptors for speaking, which will also be illustrated with sample test performances to highlight the requirements of the given levels. Issues in **speaking task design** will be explored through a variety of awareness raising activities that will enable you to design your own assessment tasks in the future. Finally, you can gain insights into how speaking performances can be evaluated based on an **analytic rating scale** as well as in informal ways through oral **feedback** given by the teacher and the learners themselves.

❖ Expected Outcomes

Upon completion of this Course, participants should be able to:

1. Distinguish between specific characteristics of the learner(s), the examiner and the task that can impact on test performance.
2. Identify features of speaking test tasks that are in line with good practice.
3. Evaluate speaking tests more critically.
4. Design speaking test tasks on your own.
5. Evaluate speaking performances based on an analytic rating scale.

❖ Key Concepts

- **Benchmarks**
- **Global questions**
- **Group oral**
- **Individual long turn**
- **Individual mode**
- **Interlocutor**
- **Paired mode**
- **Turns**

Brief Introduction

Assessing learners' speaking ability is of central concern for all language teachers as most people spend more time speaking than reading or writing. When we want to assess L2 learners' speaking skills, we need to focus upon assessing their ability to use the language in **lifelike**

situations, in which they have to perform a variety of **language functions**. In order to make valid assessments, teachers need to have a good understanding of the **nature of speaking**, the basic principles of **task design** as well as **scoring** and **evaluation**. This Course focuses on **four Topics**. The first one is intended to give insights into what **variables** can shape speaking performances in general and in test situations, and what **guidelines** one needs to be aware of in order to follow good practice in speaking assessments. It will also help you familiarize yourself with some CEFR scale descriptors for speaking in order to understand the target levels that can be measured. Then, at the second Topic, we will focus on different **test formats** for assessing speaking skills, especially how they can be constructed. The third Topic is intended to highlight how speaking can be assessed in the language classroom **informally** through feedback given by the teacher and the learners themselves. Finally, we will deal with the **evaluation** of speaking performances by illustrating what assessment criteria can be taken into account for judging the learners' speaking ability.

5.1 Topic 1: Introduction to Assessing Speaking



There are **two Tasks** in this Topic containing a total of **six Activities**.

Task 1. What characterises and influences speaking performances?

Language users typically speak **spontaneously** since in real life we rarely have the chance to prepare for what we want to say. Furthermore, when two people are engaged in a conversation, they take **turns** to initiate and to respond, they may decide to change the topic of the conversation suddenly, or they can interrupt each other in order to take a turn. As a result, conversations may take very different directions depending on the **speakers' intentions**.

Activity 1

What characterises speaking performances? Do a **simple quiz** in which you can consider and/or revise what characterises speaking performances.

What characterises speaking? Read the following questions and choose the most likely answer from A, B or C.

Question 1

L2 speakers often have to face pressure of time and thus they typically

- A. do not speak in sentences.
- B. choose not to listen to their partner(s).
- C. abandon their message(s).

Question 2

What characterises spoken discourse typically?

- A. It contains a great variety of words.
- B. Speakers make long utterances.
- C. It includes a lot of conventional phrases.

Question 3

What is NOT TRUE about the act of speaking?

- A. It is a shared social activity.
- B. It is always reciprocal.
- C. It may contain repetition.

Question 4

Typically when L2 speakers make a mistake/error in normal spoken discourse, it is related to:

- A. pronunciation
- B. grammar
- C. level of formality.

Question 5

When speakers deliver a talk that focuses on factual details, it can be:

- A. a description.
- B. an explanation.
- C. a comparison.

Question 6

In general, when people talk, they search for meaning and meaning is not always clear and explicit.

This is most likely to happen because this is how they try to:

- A. avoid hurting the listener(s).
- B. keep the conversation going.
- C. catch the listener's attention.

Question 7

Which is NOT TRUE? Appropriate topics for chatting:

- A. are very similar in different cultures.
- B. depend on the speaker role and role relationships.
- C. are influenced by the speakers' personality.

The answer key is provided in the **Indicative Answers** section.

A speaking performance of an L2 learner can be **positively** and **negatively** influenced by a number of **factors**, which relate to the **learner**, the **interlocutor** and the **task** at hand. In the classroom, speaking assessment can be carried out in the **individual mode** (a teacher made assessment), in the **paired mode** (two learners interact with each other), in the **group oral mode** (three or four learners interact with one another). Speaking assessment tasks may be varied too as they can relate to a **variety of topics**, may make use of **visual or word prompts**, and can generate **individual long turns** (sustained monologues) or **interaction**, which is based on collaboration between the partners.

Activity 2

Based on your experience as a learner and/or an interlocutor (examiner/teacher), what would you say are **the most important positive and negative aspects** that can be related to specific characteristics of the learners, the interlocutor and the speaking task?

In the following table, there are some examples to start you off. Your task is to write further examples.

	Positive aspects	Negative aspects
The learner	s/he has relevant life experience for carrying out the task	
An examiner/teacher as the interlocutor		usually superior to the learner in terms of L2 proficiency, age and social status
Another learner as the interlocutor		
The task		complex tasks can be cognitively challenging

If you wish, go to the **Indicative Answers** section to see some ideas.

Activity 3

When we want to assess L2 learners' speaking ability, we need to consider a number of issues.

Read some statements below and decide whether they are ' **true** ' or ' **false** '. A third option is also included (i.e. ' **I can't decide** ') for you to choose when you think that whether a statement is true or false depends on the context.

Question 1

When designing speaking assessment tasks, we need to take into account the learner's needs and characteristics.

- a. True
- b. False
- c. I can't decide

Question 2

When assessing L2 learners' speaking ability, we do not need worry about the length of their performance.

- a. True
- b. False
- c. I can't decide

Question 3

When scoring learners' performance we need to take into account what they say and how they say it.

- a. True
- b. False
- c. I can't decide

Question 4

If a learner does not respond to a task well, it is because the task is bad.

- a. True
- b. False
- c. I can't decide

Question 5

The learner can get more support from the examiner/teacher in the one-on-one, individual mode than in the [paired mode](#).

- a. True
- b. False
- c. I can't decide

Question 6

The examiner/teacher can correct the learner's mistakes/errors while they are in interaction.

- a. True
- b. False
- c. I can't decide

Question 7

One type of elicitation technique/task is enough to use in order to sample learners' proficiency.

- a. True
- b. False
- c. I can't decide

Question 8

The scoring of spoken performance is more reliable when an analytic rating scale is used.

- a. True
- b. False
- c. I can't decide

The answer key and our comments are provided in the **Indicative Answers** section.

Activity 4

Watch a short video (ca. 6 mins.) supported by the British Council that gives an overview of **testing speaking** in large-scale testing contexts. You can revise some of the issues we have dealt with so far as well as hear about some **useful guidelines for task design** to prepare you for the next unit.



The video entitled *Assessing Speaking* (uploaded on 17 June, 2016) is available at this address: <https://www.youtube.com/watch?v=fkWibQIID2Y>

After watching the video, carry out a **simple quiz**.

Select the missing words from the the list of words to fill the gaps in the text below.

artwork, audience, imagination, multi-dimensional, person, plan, recorded, simple, tablet, trade-off

Speaking is (1)_____ because you need to take into account the purpose of speaking, the (2)_____ and the conditions under which we speak.

Commonly used tasks include:

- describing a picture, a place, a (3)_____ an action
- telling a story, which is based on a single picture, a series of pictures, (4)_____
- comparing real objects, photographs (5)_____ abstract concepts
- giving personal information

Speaking performances can be live or (6)_____ and they can be delivered face-to-face or via telephone/ (7)_____ /computer.

Complex **rating scales** tend to be better than (8)_____ ones.

The language the test tasks generate should reflect what we (9)_____ to assess.

The development of a speaking assessment task is sometimes a (10)_____: we have to compromise and use alternative solutions.

The answer key is provided in the **Indicative Answers** section.

Task 2. Speaking skills and the CEFR

In order to better understand the target proficiency levels of the **Common European Framework of Reference**, you need to familiarize yourself with (or revise) some CEFR scale descriptors for speaking. Some of the levels will also be illustrated with sample test performances to highlight the specific requirements of the given levels.

Activity 5

Match the descriptors with the levels: A1, A2, B1, B2, C1 or C2.

	The Learner	Which CEFR level?
1.	<p><i>Shows fluent, spontaneous expression in clear, well-structured speech.</i></p> <p>CAN express him/herself fluently and spontaneously, almost effortlessly, with a smooth flow of language. CAN give clear, detailed descriptions of complex subjects. High degree of accuracy; errors are rare.</p>	
2.	<p><i>Relates comprehensibly the main points he/she wants to make.</i></p> <p>CAN keep going comprehensibly, even though pausing for grammatical and lexical planning and repair may be very evident. CAN link discrete, simple elements into a connected sequence to give straightforward descriptions on a variety of familiar subjects within his/her field of interest. Reasonably accurate use of main repertoire associated with more predictable situations.</p>	
3.	<p><i>Relates basic information on, e.g. work, family, free time etc.</i></p> <p>CAN communicate in a simple and direct exchange of information on familiar matters. CAN make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident. CAN describe in simple terms family, living conditions, educational background, present or most recent job. Uses some simple structures correctly, but may systematically make basic mistakes.</p>	
4.	<p><i>Makes simple statements on personal details and very familiar topics.</i></p> <p>CAN make him/herself understood in a simple way, asking and answering questions about personal details, provided the other person talks slowly and clearly and is prepared to help. CAN manage very short, isolated, mainly pre-packaged utterances. Much pausing to search for expressions, to articulate less familiar words.</p>	
5.	<p><i>Conveys finer shades of meaning precisely and naturally.</i></p> <p>CAN express him/herself spontaneously and very fluently, interacting with ease and skill, and differentiating finer shades of meaning precisely. CAN produce clear, smoothly-flowing, well-structured descriptions.</p>	
6.	<p><i>Expresses points of view without noticeable strain.</i></p> <p>CAN interact on a wide range of topics and produce stretches of language with a fairly even tempo. CAN give clear, detailed descriptions on a wide range of subjects related to his/her field of interest. Does not make errors which cause misunderstanding.</p>	

Activity 6

In order to relate speaking performances to the CEFR levels, watch some **sample performances** offered by Cambridge ESOL Exams. Each video is intended to illustrate what typical learners can do in English at the given target CEFR level. The exams are accompanied by detailed notes on each learner's performance that explain how his/her performance fits the target level. While watching a sample performance, take notes on the strengths and weaknesses of the learners and then compare them to the comments made by the Cambridge examiner. Source: Examples of Speaking Performance at CEFR Levels A2-C2. (2009). University of Cambridge ESOL Examinations, Research and Validation Group. Project overview.

Sample test performance 1



<https://www.youtube.com/watch?v=ycc2G4Ryn3Y>

CEFR Level: A2 (Cambridge ESOL Exam: KET)

Learners: Sharissa and Jannis (uploaded on 19 September, 2013)

My comments:

Examiner's comments



<http://www.cambridgeenglish.org/images/149448-ket-for-schools-examiner-s-comments.pdf>

Sample test performance 2



<https://www.youtube.com/watch?v=QwlbZc9MrJA>

CEFR Level: B1 (Cambridge ESOL Exam: PET)

Learners: Victoria and Chiara (uploaded on 19 September, 2013)

My comments:

Examiner's comments



<http://www.cambridgeenglish.org/images/149446-pet-for-schools-examiner-s-comments.pdf>

Sample test performance 3



<https://www.youtube.com/watch?v=EdeZp0n0JHw&index=5&list=PLbnloEgo0noigRg9zP9fmIjcS65m4ldnz>

CEFR Level: B2 (Cambridge ESOL Exam: FCE)

Learners: Victoria and Edward (uploaded on 21 July, 2014)

My comments:

Examiner's comments



<http://www.cambridgeenglish.org/images/173977-cambridge-english-first-speaking-test-examiner-comments.pdf>

Sample test performance 4



<https://www.youtube.com/watch?v=5nGESyDgmdw>

CEFR Level: C1 (Cambridge ESOL Exam: CAE)

Learners: Raphael and Maude (uploaded on 13 May, 2013)

My comments:

Examiner's comments



<http://www.cambridgeenglish.org/images/173976-cambridge-english-advanced-examiners-comments.pdf>

Sample test performance 5



https://www.youtube.com/watch?v=Bx_x-gvLw0

CEFR Level: C2 (Cambridge ESOL Exam: CPE)

Learners: Dirk and Annick (uploaded on 12 September, 2014)

My comments:

Examiner's comments



<http://www.cambridgeenglish.org/images/cambridge-english-proficiency-speaking-test-video.pdf>

5.2. Topic 2: Designing and Evaluating Speaking Tests



There are **two Tasks** in this Topic containing a total of **four Activities**.

Task 1. Awareness raising of general issues in designing speaking test tasks

The **interview** is often considered to be a prototypical exam format although its scope is limited because the examiner is in control of the conversation, initiating and concluding topics and so the flow of information is one-way. Such an imbalance in conversational rights and duties does not reflect the way we use language in everyday communication since language users often initiate, redirect and conclude topics, and they often want to get information, not simply give it. In an interview, the learner is often asked about a number of different topics that s/he can relate to. This test format is probably suitable for eliciting personal information about the learner and getting him/her to express opinions on certain issues.

For reasons of test **validity** and **fairness**, the interlocutor's contributions should be as guided as possible. This means that the questions should be preferably scripted (pre-written) as it may

make a big difference to a learner's performance if the interlocutor paraphrases the questions inappropriately. For example, display questions (to which the answer is known in advance) are not likely to generate extended and meaningful responses.

Activity 1

Watch the following **two interviews** (Sample Video 3.1 and Sample Video 3.2) and take notes on the following:

1. Does the interlocutor use **paraphrasing**?
2. Does the interlocutor use **display questions**?
3. **Who talks more**: the interlocutor or the learner?
4. Do the interlocutor's questions always **help** the learner to perform? If no, why?
5. Does the interlocutor make **unnecessary comments**?

My notes:

Video source: INTO EUROPE: The Speaking Handbook

The videos are **downloadable ONLY AS ZIP FILES** from this website:



http://www.lancaster.ac.uk/fass/projects/examreform/Pages/IE_Speaking_videos_wm.htm

In order to access the videos above, you need to download all the videos in Chapter 3.

When you finish, see our own answers to the above questions in the **Indicative Answers** section.

The following **recommendations** are made for interlocutors when asking questions to a learner (Csépes & Együd, 2004, p. 40):

- Use **global questions** for elicitation.
- Use wh-questions instead of yes/no questions whenever possible.
- Never ask more than one question at a time.
- Do not talk more than necessary: refrain from making unnecessary comments.
- Do not interrupt or finish what the learner wants to say.
- Do not ask questions that require special background knowledge.
- Avoid ambiguous and embarrassing questions.
- Use genuine questions and avoid display questions.
- Maintain eye contact with the learner when talking to him/her.

In contrast to interviews, in **discussions** or **roleplay activities** the interlocutor acts as the learner's partner with whom they have to reach a specific goal based on some kind of opinion or information gap between the two. In such tasks there is a good opportunity for the learner to display his/her oral interactional skills since the tasks allow for a two-way information flow. In collaborative tasks, learners are typically encouraged to initiate, negotiate, argue for and/or against specific ideas, suggestions or propositions. In discussion activities, learners usually express their own opinion, which distinguishes this technique from a role-play activity, where they often take somebody else's role (e.g. a holiday-maker) in order to reach a particular communicative goal (e.g. to discuss where to go and what to do on a trip). The roles featuring in test tasks usually simulate the ones we take in our everyday lives. However, the examiner's superiority in terms of age, language proficiency and authority (due to his/her role, which is difficult to ignore even in an assessment situation) will still prevail in the majority of cases, and potentially limit the learner's language output for psychological reasons.

Activity 2

Watch the following **video** illustrating a **roleplay activity** (Sample Video 6.9).

Video source: INTO EUROPE: The Speaking Handbook

The videos are **downloadable ONLY AS ZIP FILES** from this website:



http://www.lancaster.ac.uk/fass/projects/examreform/Pages/IE_Speaking_videos_wm.htm

In order to access the videos above, you need to download all the videos in Chapter 6.

When you finish, see our own answers to the above questions in the **Indicative Answers** section.

While watching the video, pay attention to the following points and make brief notes:

1. **How much reading** is required of the learner to understand her task?
2. What do you think about the **role** the learner has to take on?
3. In what ways does the roleplay task require the learner to use her **imagination**, if at all?
4. How is the **information-gap/opinion-gap** generated in the role-play?

The **task instructions** for the interlocutor and the learner:

Interlocutor's Card

To be read out for the candidate:
I am a foreign visitor in Hungary. Give me advice on where to spend a weekend in your country. On your sheet there are some ideas but you can suggest other ones too. You must start the conversation when you're ready.

Prompts for the candidate:

- Place
- What to see
- Activities
- Transport
- Your own experience
- Anything else?

For the examiner only:

- Ask for some clarification you consider relevant.
- Challenge the candidate in a polite manner so that s/he should be forced to defend his/her arguments.
- Say you can't drive.
- You don't like walking for hours.
- You must finish the conversation.

Role-card for Candidate

Place

Anything else?

Transport

Activities

What to see

Your own experience

Source: Into Europe: The Speaking Handbook, p. 127.

When we wish to evaluate a learner's overall speaking ability, in addition to interactional skills, his/her ability to produce sustained monologues also needs to be addressed and assessed. For this purpose, we recommend the so-called 'individual long-turn' tasks, which can be based on visual or verbal (i.e. text) prompts. Very often learners are given pictures to describe, compare

and contrast. However, pictures can turn out to be useless if they are not sufficiently challenging, do not contain enough stimuli to elicit language from the learner.

Typical problems with the choice of visuals include the following:

- If the picture is culture-dependent (shows objects and/or activities that have relevance and make sense only for a specific community of speakers), it will require special background knowledge from the learner.
- If the topic of the picture is distressing, sensitive or violent, it may negatively affect the learner's performance.
- Abstract and symbolic pictures can discourage learners, especially at lower levels, from performing to the best of their abilities given the large amount of subjectivity involved in the meaning-making.
- Bizarre, unrealistic situations shown in pictures are unlikely to elicit meaningful, life-like language output from learners.
- If the pictures used for the 'compare and contrast' task-type are too similar, they will fail to elicit a suitable performance both in terms of length and variation in [language use](#) (structures and lexis).

Activity 3

Look at **two individual long-turn tasks** below, which were produced by secondary school language teachers. Give an evaluation of the picture prompts and the guiding questions the interlocutor can choose from. Write down your comments and when ready, compare them to an **expert item reviewer's comments**.

My comments:

Sample individual long turn Task 1

Participants: Candidate and Examiner

These two pictures show passengers who have just got off the bus and a person driving a car along the motorway. Compare and contrast the two pictures. Include the following points:

1. Means of transport
2. Traffic problems
3. Advantages and disadvantages of travelling by car and by bus



Prompts for the interlocutor:

- What means of transport can you list?
- Why do many people use public transport instead of cars?
- Which means of transport do you prefer? Why?
- What traffic problems face especially in the morning?

Sample individual long turn Task 2

Participants: Candidate and Examiner

These two pictures show two turning points in our lives: they show primary school pupils and secondary school students. Compare and contrast the two pictures. Include the following points:

1. The differences between studying in a primary school and in a secondary school
2. Out-of-school activities
3. Timetable and studying

Prompts for the interlocutor:

- What did you prefer: primary or secondary school? (school subjects? classmates?)
- How do children study at different ages?
- How do you learn different subjects?



You can read the expert item reviewer's comments in the **Indicative Answers** section.

Activity 4

Design a collaborative speaking assessment task using pictures or word prompts. Your task must be targeted at a **specific CEFR level** that you need to specify clearly. The task must be fully contextualized to make it life-like. When you design the speaking assessment task, bear in mind the following aspects:

- the appropriacy of the context (relevant for the target age group and level)
- the quality (length and clarity) of the instructions
- the quality of the visual prompts
- the amount of reading
- the nature of the task (realistic or not)
- the expected amount of participant contribution (is there an assumed balance?)

In the following video (Sample Video 5.4), you can watch a **sample speaking performance** on a **paired task**.

Video source: INTO EUROPE: The Speaking Handbook

The videos are **downloadable ONLY AS ZIP FILES** from this website:



http://www.lancaster.ac.uk/fass/projects/examreform/Pages/IE_Speaking_videos_wm.htm

In order to access the videos above, you need to download all the videos in Chapter 5.

The **task instructions** for the interlocutor and the learners:

Imagine that in your school the students would like to go to free afternoon classes. Talk to each other about what courses the school should organize. On your sheets, there are some ideas but you can suggest other things, too. Talk about how useful these courses would be for students at your school and try to agree about the three most useful ones.

Start when you are ready.

Prompts for Candidate A

Driving

Photography

Dress-making

Word-processing

Anything else?

Prompts for Candidate B

Typing

Aerobics

Using the Internet

Anything else?

Self-defence

Source: Into Europe: The Speaking Handbook, p. 94

Task 2. Designing speaking assessment tasks

As a classroom teacher, in most cases you have to schedule the assessment of speaking as part of your daily work. This means that you need to engage all the learners simultaneously while you may select some students whose performance you wish to assess based on some criteria. The classroom context thus calls for task formats that can be managed with a group of learners. The **paired mode** or the **group oral** seem to be a good choice as the teacher can quietly monitor the students while they are doing the task. However, the participants of either format must be presented with instructions and prompts that are capable of eliciting language performance from all. For task design therefore we need to take into account the following:

- potential contributions to the interaction should be balanced, which could be achieved by giving an equal number of visual or word prompts for both/all participants;
- participants should have comparable tasks, i.e. they should be required to do the same thing in order to facilitate a balanced, realistic and smoothly-running exchange between them;
- the interaction should be task-based because it seems to give learners a meaningful purpose to engage in a conversation (e.g. listing, comparing, contrasting, selecting, justifying, modifying etc.);
- the tasks have to be guided but not fully controlled, i.e. learners should have a chance to add something of their own to the exchange;
- learners should be given a chance to voice their own opinion rather than argue for something given that they cannot identify with.

5.3. Topic 3: Giving **Feedback** to Enhance Learning



There is **one Task** in this Topic containing **one Activity**.

Task 1. Oral feedback and its effectiveness

After oral communication practice, the teacher usually gives some kind of **feedback**, which we can regard as assessment instances, also referred to as **formative feedback**. This kind of evaluation is likely to be less structured and systematic because it can vary from learner to learner. It can address the learner as a person (his/her ego) and/or his/her task completion (what s/he did well or badly) in the form of short/extended teacher comments. You can learn more about this topic in the separate course on **Providing Feedback**.

In general education, research has revealed that different types of feedback have variable impact on students' learning. In a video talk, Dylan Wiliam, a Professor of Educational Assessment, highlights which type of feedback is less effective and which seems to work, also explaining why.

Activity 1

Before you watch Prof. Wiliam's talk, consider the following statements whether they are '**true**', '**false**' or whether you '**cannot decide**'. After the video, revisit the statements to see if you were right. The talk is ca. 3 minutes long.

1. The kind of feedback teachers give in schools is very helpful (in the UK).
2. Grades are very useful forms of constructive feedback as students can see where they stand in comparison to other students in their group.
3. "You did very well" is an effective form of student appraisal.
4. Good feedback makes students think.
5. Students' ability is fixed.

Watch Prof. Dylan Wiliam's talk on 'Feedback on Learning' (uploaded on 13 February, 2015) in the following video.



<https://www.youtube.com/watch?v=MzDuiqaGqAY>

After watching the video, you can check your answers in the **Indicative Answers** section.

5.4. Topic 4: Evaluating Speaking Performances



There are **two Tasks** in this Topic containing **one Activity**.

Task 1. Scoring benchmarked sample performances

The most reliable way of assessing speaking performances is by using an **analytic rating scale** because it helps the assessor to judge a learner's performance with respect to separate, clearly identifiable performance features such as pronunciation, interactive communication, grammar and vocabulary. Examination boards usually develop and use their own assessment tools, which are matched to the tasks they employ. In order to help you gain insights into how to score speaking performances, we will use freely accessible sample tests and assessor training materials from the publication "INTO EUROPE: The Speaking Handbook".

Activity 1

First, study and familiarize yourself with the **B2 Level Speaking Assessment Scale** (source: Into Europe: The Speaking Handbook, p. 184.). Please note that in the given assessment scale an empty band (and the corresponding score) must be chosen when the learner performance fails to meet one bullet point in the band descriptor immediately above it. Each criterion is scored once throughout the whole exam.

Appendix 3 B2 Level Speaking Assessment Scale

	COMMUNICATIVE IMPACT	GRAMMAR AND COHERENCE	VOCABULARY	SOUNDS, STRESS, INTONATION
	Candidate...	Candidate...	Candidate's vocabulary...	Candidate...
7	<ul style="list-style-type: none"> • makes entirely natural hesitations when searching for ideas • participates with ease without requiring additional prompting • contributes fully and effectively to the communication 	<ul style="list-style-type: none"> • uses wide range of structures • uses accurate grammar • makes fully coherent contributions 	<ul style="list-style-type: none"> • has wide range • is fully appropriate 	<ul style="list-style-type: none"> • is understood with ease • uses accurate and appropriate sounds and stress • uses a wide range of intonation to convey meaning effectively
6				
5	<ul style="list-style-type: none"> • makes mostly natural hesitations when searching for ideas • requires no additional prompting • in general contributes effectively to the communication 	<ul style="list-style-type: none"> • uses adequate range of structures • makes occasional minor mistakes only • makes adequately coherent contributions 	<ul style="list-style-type: none"> • has appropriate range • is generally appropriate with isolated inappropriacies 	<ul style="list-style-type: none"> • is understood easily with isolated difficulties • uses mostly accurate and appropriate sounds and stress • uses an adequate range of intonation to convey meaning mostly effectively
4				
3	<ul style="list-style-type: none"> • often makes hesitations in order to search for language • in general requires no additional prompting • contributes sufficiently to the communication, occasionally making irrelevant contributions 	<ul style="list-style-type: none"> • uses limited range of structures • makes occasional major and minor mistakes • makes mostly coherent contributions with occasional inconsistencies 	<ul style="list-style-type: none"> • has sufficient but somewhat limited range • is generally appropriate with occasional disturbing inappropriacies 	<ul style="list-style-type: none"> • is understood with some strain • makes mistakes in sounds and stress that occasionally affect comprehensibility • uses a limited range of intonation to convey meaning
2				
1	<ul style="list-style-type: none"> • frequently makes intrusive hesitations when searching for language • requires some additional prompting • frequently makes irrelevant contributions or contributes little to the communication 	<ul style="list-style-type: none"> • uses very limited range of structures • makes occasional major and frequent minor mistakes • makes mainly incoherent contributions 	<ul style="list-style-type: none"> • has limited range • is frequently inappropriate 	<ul style="list-style-type: none"> • is understood with constant strain • makes mistakes in sounds and stress that seriously affect comprehensibility • makes little use of intonation to convey meaning at all
0	<ul style="list-style-type: none"> • no assessable language 	<ul style="list-style-type: none"> • no assessable language 	<ul style="list-style-type: none"> • no assessable language 	<ul style="list-style-type: none"> • no assessable language

Watch the **video** (Video Sample 8.3) below and **give a score** on all the **four assessment criteria** by the end of the exam.

Video source: INTO EUROPE: The Speaking Handbook

The videos are **downloadable ONLY AS ZIP FILES** from this website:



http://www.lancaster.ac.uk/fass/projects/examreform/Pages/IE_Speaking_videos_wm.htm

In order to access the videos above, you need to download all the videos in Chapter 8.

In order to better focus your attention on the specific features of the learner's performance, read **expert judges' comments (justifications for agreed scores)**.

Justifications for Video sample 8.3

COMMUNICATIVE IMPACT:

The Candidate communicates mostly effectively. He makes hesitations only sometimes when searching for a word, but it's not disturbing. However, he makes some irrelevant contributions and needs some additional prompting: *I think the sport is the most healthy in Hungarian, because your conditional will be good.* In Part 2 the comparison could have been more elaborated (the main difference was not mentioned, he just described the pictures). In Part 3 he makes some attempt to give advice: *you can; you should.* Some utterances are difficult to judge because of language problems: *I like the economical or the other thing is to be a doctor; let's see the different of the pictures.*

GRAMMAR AND COHERENCE:

The Candidate's performance is rather extreme with regard to grammar and coherence. He uses the right tense most of the time (e.g. he always uses the present continuous for picture description), whereas some of his utterances lack any kind of coherence: *something more culture; doing something meal; the other one just draw.* He uses a limited range of structures with occasional major and frequent minor mistakes: *That's very big and not so thin so it's very thick; I would talk about the picture and say the different of the pictures; he's don't help her.* He uses wrong word order: *I don't know what's the time and the wrong preposition: in the same time.*

VOCABULARY:

The Candidate is equally extreme with his use of vocabulary. The range the Candidate tries to use is considerable – and so is the number of mistakes he makes. In many cases he searches for the word and uses a description: *fashion style* [instead of trend]; *something more culture; trousers made from jeans.* He frequently confuses verb forms and noun forms of the same lexical item, as well as nouns with adjectives: *economical* [instead of economist]; *the wife will have an argue with his husband; your conditional* [instead of physical condition].

SOUND, STRESS, INTONATION:

The Candidate is understood easily with isolated difficulties. These isolated difficulties mainly come from the unclear, hasty articulation of words at the end of sentences. He makes some mistakes in sounds which affect comprehensibility, the influence of German can be felt. Though he attempts to use intonation patterns, his intonation tends to be rather 'flat'.

Source: Into Europe: The Speaking Handbook, p. 190.

In the light of the expert judges' comments/justification, **review** the scores you gave earlier. If you wish to **modify** any of them, you can do so. **Finalize** your scores.

Finally, look at the **benchmarks** which are the agreed and justified scores given by experts and **compare** them to your own scores.

Benchmarks for Video sample 8.3

Benchmark for communicative impact	Benchmark for grammar and coherence	Benchmark for vocabulary	Benchmark for sound, stress, intonation
4	3	3	4

Source: Into Europe: The Speaking Handbook, p. 191.

Optional Task:

You can do a second round of practice, following the same steps with the following video (Video Sample 8.4).

Video source: INTO EUROPE: The Speaking Handbook

The videos are **downloadable ONLY AS ZIP FILES** from this website:



http://www.lancaster.ac.uk/fass/projects/examreform/Pages/IE_Speaking_videos_wm.htm

In order to access the videos above, you need to download all the videos in Chapter 8.

Here are the justifications (Into Europe: The Speaking Handbook, p. 192.)

Justifications and benchmarks for Video sample 8.3

COMMUNICATIVE IMPACT: 7

The Candidate communicates with ease and **makes relevant contributions** throughout the whole test. Her responses are quick and natural: *the buses are really rare*. **She makes entirely natural hesitations when searching for ideas** using adequate fillers: *well; kind of; so; actually; you see* in the meantime. **She requires no additional prompting.** The Candidate initiates and responds adequately.

GRAMMAR AND COHERENCE: 7

The Candidate **uses a very wide range of structures** appropriately. She is a confident user of the tenses, the passive and conditional sentences. There are examples of subtle use, which are certainly considered top performance at the intermediate level: *the child is being helped by his mother; she is studying by herself; if it was me who lived here; it's worth going there*. **She uses mostly accurate grammar**, although some minor errors occur: *on these pictures; in the autumn; like you see; advise going*. **All her contributions are coherent.**

VOCABULARY: 6

The Candidate uses a **wide range of vocabulary**: *it won't work; get to know; break up with; by herself; kind of like my room; get together; scattered*. However, **there are some isolated inappropriacies**. She uses Hungarian words twice: *nazarénusok; gimnázium* and makes a minor mistake when using: *telephone line* (telephone cord). Apart from these slips her performance demonstrates an attempt to go beyond the vocabulary resource required at intermediate level.

SOUND, STRESS, INTONATION: 7

The Candidate **is understood with ease**. She uses **mostly accurate and appropriate sounds and stress**. Only a few inappropriately pronounced words occur, which are still comprehensible: *bilingual* ['bɪlɪŋwɔl]; *event* ['ɪvənt]. She uses the American accent consistently. **She uses a wide range of intonation to convey intended meaning effectively.**

Source: Into Europe: The Speaking Handbook, p. 192.

Task 2. Teacher assessment vs. peer-assessment and self-assessment

There are several **benefits** of involving learners in the assessment process. When learners are asked to assess themselves and their peers, they are not in the same position as their teacher, therefore the aspects of the performance they can give **feedback** on and/score should be carefully defined. The effectiveness of **self-assessment** and **peer-assessment** of speaking performances is most likely to depend on:

- the transparency or clarity of the **assessment criteria** (explain and illustrate them);
- learners' developing language proficiency, which influences their perceptions and ultimately their observations and ratings;

- whether learners are provided with training methodologically (helping them to use the assessment criteria) and psychologically (explaining why self- and peer assessment can be useful for them);
- allowing the use of L1, especially in low level classes.

Here are some examples for self- and peer assessment tools:

Self-assessment

Name: _____ Class: _____ Date: _____

What I did in this task:

How true are these? Put a tick in the right box.




	3 = true	2 = partly true	1 = not true
I understood his/her partner/group			
I asked questions			
I answered questions			
I knew how to pronounce words			
I knew enough words and phrases			
I knew enough grammar			
I spoke only English			
I spoke fluently			

Things I managed to do well:

Things I still need to work on:

Peer-assessment for discussion activities (adapted from Mok, 2011)

Your name: _____ Name of student: _____

			
His/her voice was loud enough			
He/she could pronounce English sounds and words clearly			
He/she had friendly body language and made eye contact			
He/she could join in the discussion and encourage others to speak			
He/she could use the right words to explain his/her ideas			
He/she could use grammatically correct language			
His/her ideas were relevant to the topic of discussion			
He/she paid attention to others' ideas and responded to them			

The things I liked most about his/her discussion were:

Total smiley faces for him/her:

Peer-assessment for individual long turn tasks

How true are these? Put a tick in the right box.

	3 = true	2 = partly true	1 = not true
S/he understood his/her partner/group			
S/he asked questions			
S/he answered questions			
S/he knew how to pronounce words			
S/he knew enough words and phrases			
S/he knew enough grammar			
S/he spoke only English			
S/he spoke fluently			

Things s/he managed to do well:

Things s/he still needs to work on:




Finally, here is a quotation from Yuh-Mei Chen, a Taiwanese researcher on classroom assessment, who highlights in an effective way why different forms of assessment by learners is likely to enhance learning outcomes:

"Self - assessment is best practiced in tandem with peer and teacher assessment. Interacting with the learning environment, students can judge their performance skills more accurately in relation to those of others. This self-assessment process allows students to seek knowledge from the surroundings to make sense of uncertainty, conflict, and doubt, and to draw a clearer self-profile."

(2008, p. 255)

Assessing Speaking Skills - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I can distinguish between specific characteristics of the learner(s), the examiner and the task that can impact on test performance.			
I can Identify features of speaking test tasks that are in line with good practice.			
I can evaluate speaking tests more critically.			
I can design speaking test tasks on my own.			
I can evaluate speaking performances based on an analytic rating scale.			

Indicative Answers

Topic 1: Activity 1

Question 1

L2 speakers often have to face pressure of time and thus they typically

- A. do not speak in sentences.

Question 2

What characterises spoken discourse typically?

- C. It includes a lot of conventional phrases.

Question 3

What is NOT TRUE about the act of speaking?

- B. It is always reciprocal.

Question 4

Typically when L2 speakers make a mistake/error in normal spoken discourse, it is related to:

- A. pronunciation

Question 5

When speakers deliver a talk that focuses on factual details, it can be:

- B. an explanation.

Question 6

In general, when people talk, they search for meaning and meaning is not always clear and explicit. This is most likely to happen because this is how they try to:

- B. keep the conversation going.

Question 7

Which is NOT TRUE? Appropriate topics for chatting:

- A. are very similar in different cultures.

Topic 1: Activity 2

	Positive aspects	Negative aspects
The learner	*relevant knowledge of the world can support task fulfilment; *extroverted people are more open and flexible, better at communicating with others	*lack of relevant life experience may inhibit the learner in task performance
An examiner/teacher as the interlocutor	can support the learner *by intervening and/or redirecting the conversation; *by maintaining an encouraging, non-judgemental behaviour; *by giving positive backchannel signals	*usually superior to the learner in terms of L2 proficiency, age and social status, which makes the encounter uneven between the examiner and the learner
Another learner as the interlocutor	*beneficial psychological effects (sense of security) due to familiarity with partner and	*increased anxiety due to perceived differences in L2 proficiency levels of the partners

	similarity in terms of age, background, interests *mirrors good language teaching practice	(incomprehensible and/or uncomprehending partner)
The task	*familiarity with task format and topic increases confidence;	*complex tasks can be cognitively challenging, *task prompts (visuals or texts) may be inadequate for lack of details/ideas to talk about, *sensitive/distressing topics may be difficult to talk about

Topic 1: Activity 3

Question 1

Correct answer: True

Feedback: Speaking tasks should cater for the communicative needs of learners and should match their characteristics (e.g. age, background knowledge, level of education)

Question 2:

Correct answer: False

Feedback: If the learner's performance is too short, it will undermine the validity of the assessment score.

Question 3:

Correct answer: I can't decide

Feedback: It depends on the context, i.e. what aspects of the performance we want to assess. However, in general, we should not make a value judgement about the content of what learners say.

Question 4

Correct answer: I can't decide

Feedback: It depends on the context. Sometimes the task may indeed be hopeless, but there are two other main reasons for lack of adequate response from the learner: insufficient language knowledge and lack of ideas/relevant knowledge, i.e. not knowing what to say.

Question 5

Correct answer: I can't decide

Feedback: It depends on the context. Carefully designed, high-stakes speaking tests are standardized, they often follow a pre-specified/scripted scenario in which the potential questions are already given for the examiner. In classroom-based assessments, teachers typically ask their own questions on the spot, but too much help from them may lead to unfair judgements in the end. In the paired mode, the examiner/teacher must withdraw him/herself from the conversation but when peer partners get stuck or one of them tries to dominate the conversation, it is his/her duty to intervene and redirect the conversation. Thus, examiner/teacher support is also available for learners in the paired mode.

Question 6

Correct answer: I can't decide

Feedback: It depends on the context. In large-scale testing, the interlocutor (examiner/teacher) must refrain from pointing out mistakes as this can distract and distress the learner. However, teachers in their own classroom may choose to give a hint at mistakes thus trying to encourage the learner to focus on form.

Question 7

Correct answer: False

Feedback: False. Different task types have their own advantages as well as limitations.

Question 8

Correct answer: True

Feedback: True. The assessment should be based on criteria that are clearly specified, detailing how each criterion is fulfilled at different levels of the rating scale.

Topic 1: Activity 4

1. multi-dimensional
2. audience
3. person
4. imagination
5. artwork
6. recorded
7. tablet
8. simple
9. trade-off
10. plan

Topic 2: Activity 1

The examiner often paraphrases his own questions and sometimes the paraphrased questions do not help the learner at all (e.g. What can you do? What's good in Sándorfalva? What do you do at home in the evenings?). The paraphrasing strategy does not seem to work in **Sample Video 3.1**, where the interlocutor's efforts to help the weak learner (Zsolt) at all costs result in the interlocutor dominating the interaction. In **Sample Video 3.2**, because of the interlocutor's personal interest in the topic of football, the questions become highly specific and therefore the learner (Zoltán) has only limited scope for producing an extended response (e.g. Which league does your football team play in?; How many matches did the team win?; Who is number one?), and some of the questions focus on the knowledge of specific vocabulary items (e.g. Where do you play?; What position do you play?). In order to keep the conversation going, in **Sample Video 3.2** the interlocutor makes unnecessary comments (e.g. We're going to talk a bit; So you have to work hard) and uses display questions as well (e.g. It's a village, isn't it?).

Topic 2: Activity 2

1. There is minimal reading for the candidate (word prompts in bubbles), most of the instructions are given orally by the interlocutor. Long instructions involve text processing skills, which are irrelevant when assessing speaking.
2. The candidate's role: a local inhabitant. It is lifelike and familiar for her. Unfamiliar roles are difficult to act out.
3. She does not need to stretch her imagination as she is talking about things which are familiar for her.
4. The interlocutor has instructions on her card which prompt her to challenge the candidate by asking her to clarify things, negotiate details or make complaints. By doing so the candidate is forced to talk more, filling in information gaps and/or trying to come to an agreement with her partner.

Topic 2: Activity 3

The picture selection for the two tasks show the following problems: there little to say when the action is too simple (Sample 1: a car is going alone on the motorway); there is only one action (Sample 2: children putting up their hands); there is little contrast between the two picture (Sample 1: a standing vehicle vs. a moving one; Sample 2: younger children vs. teenagers, inside the classroom vs. outdoors). Many of the guiding questions do not require the candidate to use the pictures as they are not related (Sample 1: listing means of transport, traffic problems; Sample 2: timetable, studying in a primary school and in a secondary school). Given such picture prompts complemented by unfocussed and unrelated guiding questions, the candidate is likely to end up talking about general themes or him/herself (Sample 1: public transport/what transport s/he uses and why; Sample 2: school life/his/her own school).

Topic 3: Activity 1

1. False
2. False
3. False
4. True
5. False

Further Reading

Csépes, I., & Együd, Gy. (2004). INTO EUROPE: The Speaking Handbook. Series editor: J. C. Alderson. Budapest: Teleki László Foundation & the British Council. Chapter 3, 4, 5 and 6. Available at http://www.lancaster.ac.uk/fass/projects/examreform/into_europe/speaking.pdf

Webinar by Cambridge English: 'Understanding speaking assessment: what every teacher should know'. See the webinar (uploaded on 2 November, 2016) at this address:



<https://www.youtube.com/watch?v=olzUcZ7BBDk>

Finally, if you wish to explore ideas on how to **assess SpLD students' speaking skills** visit the following website of the DysTEFL project and follow the suggestions in Chapters 8 (Listening and Speaking) and 10 (Assessment)



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

or download the DysTEFL training materials and work directly from there:

If you are a trainer: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

If you are a teacher: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

And read the following:

Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning differences. Bristol: Multilingual Matters. Chapter 8. Assessment

References

- Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235–262.
- Csépes, I. & Együd, Gy. (2004). *INTO EUROPE: The Speaking Handbook*. Series editor: J. C. Alderson. Budapest: Teleki László Foundation & the British Council. Chapter 3, 4, 5 and 6.
- Green, A. (2014). *Exploring Language Assessment and Testing*. Oxon: Routledge. pp. 127-160.
- Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press. Chapter 1 & 2.
- Mok, J. (2011): *A case study of students' perceptions of peer assessment in Hong Kong*. *ELT Journal*, 65(3), 230–239.

6. Providing Feedback

Karin Vogt & Veronika Froehlich

❖ Aims

This course focuses on ways in which teachers can provide feedback so that the greatest possible effect is attained from the efforts of everyone concerned. We will touch on multiple ways in which feedback can be given to students and explore how to optimally use these.

❖ Expected Outcomes

Upon completion of this Course, you should be able to:

1. Identify what good feedback is.
2. Successfully communicate with your learners about feedback.
3. Select tools that will help you give more effective feedback to your learners.

❖ Key Concepts

- Feedback
- Feedforward
- Self-reflection
- Feedback sheets and rubrics
- Oral feedback
- Conference

Brief Introduction

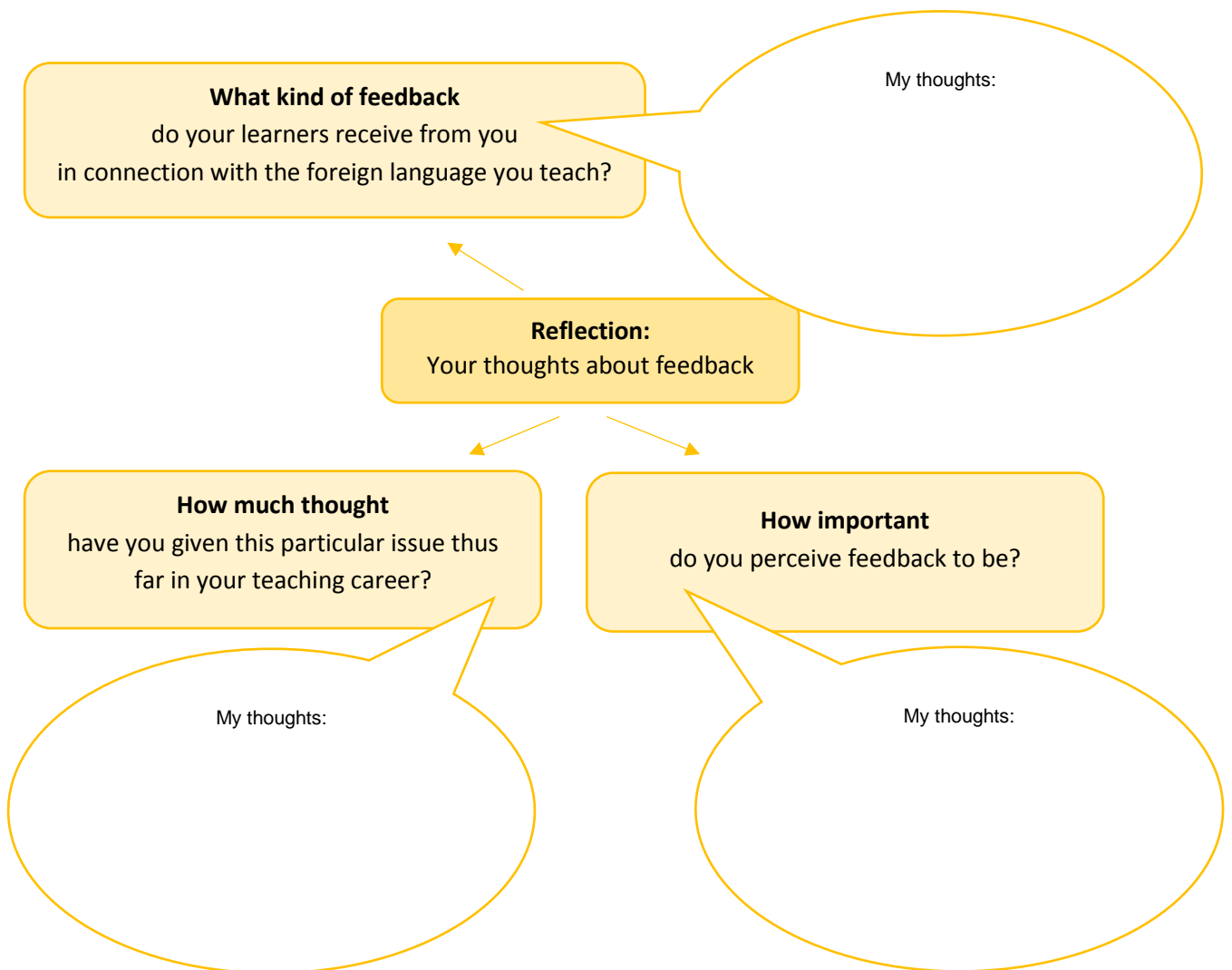
Feedback is any reaction connected to a learner's work. In this course, we will focus on feedback given by the teacher. It can be **spoken, written**, or simply be a **symbol**. Feedback is not limited to the end of a task, activity, or unit. Effective feedback includes useful suggestions that are doable and which learners are able to understand.

Feedback is an important aspect of improving **learner motivation**. In order for our learners to move forward, they require feedback that is both encouraging and helpful – this is often referred to as **feedforward**. The feedback learners are given on their performance generally has a direct impact on them. This is why much thought must be given to making our feedback as significant, comprehensive, and meaningful as possible for our learners.

Both learners and teachers benefit from swift feedback derived from assessment. Interest in learning diminishes when there is a lack of timely and useful feedback. Prompt feedback, including discussion about mistakes, has incredibly positive effects on learners. Walberg (1984) highlights this in the following two quotes: "Instruction that incorporated feedback and correctives was one of [the] most potent." "Knowing what you know and don't know focuses learning" (p. 3).

Introductory Activity

Before beginning the first Topic, consider the following questions:



6.1. Topic 1: Feedback for Formative Assessment

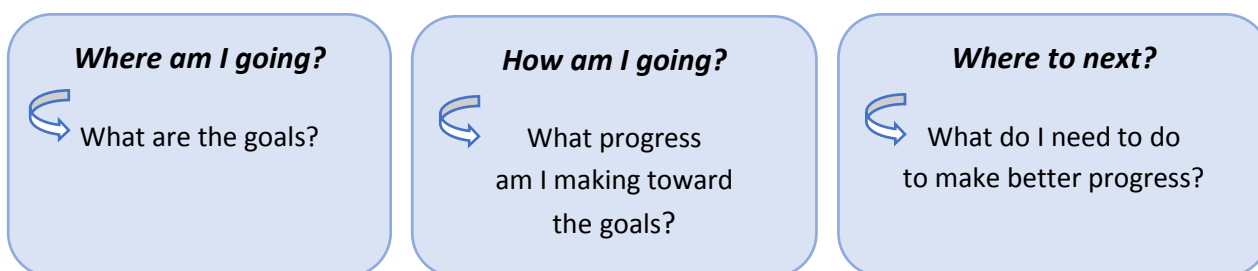


This Topic consists of **1 Task** with a total of **3 Activities**.

Task 1. The Role of Feedback in the Learning Process.

Feedback must be related to the learning aims to ensure that students' learning progresses as much as possible. Hattie and Timperley (2007) maintain that the “main purpose of feedback is to reduce the gap between current understandings and performance and a goal” (p. 86).

For this reason, they suggest that feedback address the questions:



All of the questions above require **learner self-reflection**. It can help learners better understand how to develop their language skills when they have the opportunity to ask questions concerning particular aspects of their learning after every time they are assessed. Learners can set their own learning goals with the help of teacher feedback when learning objectives are longer-term goals.

Be aware, however, that feedback is more effective when there is a **challenge** to learners involved. Thus, teachers need to ensure that they prepare challenging tasks and then give helpful and detailed feedback on those.

Assessments give teachers instantaneous feedback on aspects which learners were taught well (or not so well) by them. This is, in effect, feedback for the teacher on her performance.

Activity 1

Read the following four statements provided by a teacher as feedback on a learner's writing assignment.

	Statement
A	<i>Some of the expressions you used did not sound English, for example...</i>
B	<i>The ideas in your essay were very clearly expressed; you managed to get your point across.</i>
C	<i>Overall, I felt you deserve a B grade.</i>
D	<i>You might want to read some teen-magazines. They are a good way to learn idioms because they use such typical language, and they can be fun, too.</i>

Now, please put the statements above in the correct order (e.g. 1. Strengths = B) below.

	Feedback category	Choose the applicable statement (A-D) from above:
1.	Strengths	
2.	Weaknesses	
3.	Improvements to be made	
4.	Assessment results	

You are welcome to compare your chosen statements to the results in the **Indicative Answers** section.

Activity 2

Can you identify whether the feedback a teacher gives is more likely to be **helpful** or **unhelpful**? Mark the appropriate box in the chart below.

	Statement	Helpful	Unhelpful
1	Do you know the loci method? It might help you remember words better.		
2	vocabulary is your greatest weakness. You need to study harder.		
3	Your use of tenses is a catastrophe.		
4	You did a good job.		
5	vocabulary boxes can be useful for memorizing new words.		
6	Your use of tenses has improved a little because you do not seem to mix them up any more.		
7	I really saw how well you have worked on your connectives.		
8	I think you might want to work on...		
9	You should have prepared better for this test.		

You are welcome to compare your choice of helpful and unhelpful statements to the results in the **Indicative Answers** section.

Activity 3

Let us now focus on **unhelpful/ineffective feedback** - feedback which often fails to make clear to learners how to **improve**. What might some **reasons** for this be?

Reasons for ineffective feedback:



If you would like to learn more about “Using Feedback with Formative Assessments,” read the following article by Albert Oosterhof (2014):

http://www.cala.fsu.edu/modules/assessment_feedback/#page=1

6.2. Topic 2: Feedback Types



This Topic consists of **3 Tasks** with a total of **3 Activities**.

Task 1. The Role of Feedback in the Learning Process

As we have seen in the previous Topic, effective teacher feedback must **support students’ learning** rather than only judge their learning outcomes. Feedback must be **comprehensible** to learners so that they can use it to improve their learning. Feedback should not be a one-way phenomenon.

A foreign-language teacher will naturally select any language or content issues, skills, and comprehension problems that need to be dealt with.

In so doing, she should formulate questions that are **meaningful** and **open-ended**.

The teacher must **communicate clearly and supportively** with her learners, considering the content of her comments and/or written feedback, along with how she communicates this – the volume and tone of her voice, her pace, and her body language.

The teacher should furthermore enable her learners to **reflect** on their learning and support them to **take action** in connection with this.

Teachers can begin by asking open-ended questions; this supports learners' reflective awareness. Some effective beginnings of questions, for example, are:

What did you do well?

Did you do as well as you wanted to?

What might you do differently if you did this again?

What are you still insecure about?

What might help you be more successful in the future?

How can we change this?

Teachers should rephrase learners' remarks to clarify, expand, or summarize these. Teachers should be sure to offer suggestions as to what learners can do to improve their weaknesses and give precise examples for these.

Activity 1

How can a teacher provide feedback that **stimulates self-reflection**?

Read the notes a teacher has made on a learner's assignment, and reformulate the statements so that the feedback the teacher provides is meaningful and helps the learner self-reflect.

Here are the notes the teacher took on Diana's talk:

	Notes: Diana – 3-minute talk on solar energy
A	<i>good eye contact and body language; good introductory question ("Have you ever thought about just how important the sun is for us?")</i>
B	<i>mumbles a little, should enunciate more clearly, particularly at the beginning of sentences: pronounce key aspects and important statements clearly</i>
C	<i>overuses slangy expressions (gonna, wanna): watch examples of formal presentations on YouTube and note relevant language</i>

Now, reformulate the above statements about Diana's work into **feedback that stimulates self-reflection**.

Remember that when providing feedback, we need to:

- communicate clearly by specifying, for example, areas that need improvement
- enable the learners to reflect on their learning, for example, by formulating questions about their strengths, weaknesses, areas for improvement, and ways to improve their performance.

My suggestion for feedback on Diana's work:

You are welcome to consider our thoughts in the **Indicative Answers** section.

Task 2. Conferencing

Conferencing is a method that includes both **self-assessment** and teacher assessment. A **conference** can be used to review pieces of work in **portfolio assessment**, but it can also be used to discuss any other language-learning tasks at school.

The teacher makes an appointment with one (or sometimes a group of) student to talk about a particular topic. This could be about what pieces of recent work might be added to the portfolio or about how a group of learners is progressing with a project. Conferences allow learners to communicate problems, requests for help, or feelings of pride to the teacher. Through conferences, the teacher can gain insight into the **learning process**, **motivation**, or **anxieties** of learners and might be able to help them overcome problems that the teacher may not have realized existed.

Conferences also provide an opportunity to examine learners' **listening and speaking proficiencies**, especially of those who are shy and may not be very active during lessons.

A disadvantage of conferencing is, of course, the **time** necessary to have proper conferences with 20-30 learners. A possible solution is to have conferences with a few different individual learners or groups of learners on different topics; thus covering everyone in the course of the school year.



For examples of effective "**Student-Teacher Conferencing**," watch a video by Micheline Boileau and Carole Parent (uploaded on January 13, 2014) using the following link:

<https://youtu.be/9jQw5dOsW44>

Activity 2



Please watch the video uploaded by kineticstreaming (uploaded on July 19, 2013) accessible via the link below.

In the video, you will observe a primary and a secondary school teacher giving effective feedback to their students. While watching the video, **take notes** on the aspects that you feel **have or have not** made the teachers' feedback **effective**.

<https://youtu.be/mEgVL-nZqFg>

My notes:

Activity 3

Give oral feedback to some of your learners and record yourself doing this. **Analyze and reflect on your “performance.”** Consider the following:

What did you do well?

Where is room for improvement?

Reread the criteria mentioned in Task 1. Analyze your feedback using the following questions:

How did you incorporate important aspects?

Did you keep the order: strengths – weaknesses – improvements – grade?

Did you consider the questions: Where am I going? How am I going? Where to next?

If possible, do this task with a partner.

My thoughts:

Task 3. Feedback Sheets/Rubrics

Using **feedback sheets** is an efficient way to collect from and distribute feedback to learners. Feedback sheets warrant that specific information is methodically collected and given. Feedback sheets help learners reflect on and discuss their learning.

There are generally three parts to feedback sheets:

Descriptions of the
level of performance

Individual
comments

A grade
(optionally)

When designing feedback sheets, teachers must consider that learners might approach the same task differently. Furthermore, the language levels of the learners and how they were prepared both to complete the task and to reflect on their performance must be well thought through.

You are welcome to view an example of a feedback sheet in the **Indicative Answers** section.

6.3. Topic 3: Effective Teacher Feedback



This Topic consists of **1 Task** with a total of **2 Activities**.

Task 1. Characteristics of Effective Teacher Feedback

The type of feedback teachers give to learners will determine whether they can act on it and whether it will thus have an impact on student learning. Teachers often reduce their feedback to short-term rewards in the form of praise and/or smiley faces/stickers or something similar. External rewards like these tend to support competition rather than focus on the learning process. Feedback should reflect the deep understanding that is our ultimate goal.

Activity 1

1. Based on what we have discussed thus far about providing feedback, consider the following question:

Does the feedback I give my learners help them to improve their learning?

My thoughts:

2. After completing the activity, learn more about effective teacher feedback and principles of providing feedback.

Any feedback you give should be personal. Naturally, feedback can also be given to groups or pairs, but this should be in *addition* to and not a *substitute* for personal feedback.

All work done by learners should receive age-appropriate feedback as soon as possible after turning it in (e.g. the day after a test). Feedback should be based on criteria for the task that

both the teacher and the learner clearly understand. The learner must also understand the language used in the feedback.

Feedback should be specific to be helpful to learners. Instead of “too many misspelled words,” exemplify, for example, “You misspelled three words: circus, favorite, and bicycle.” Dignify the incorrect responses of learners by, for example, giving a question for which their response would have been correct: “Brooklyn would have been correct if I had asked for a borough in New York City.” Give the learner a hint: “It’s the largest state far up in the north.” Make the learner accountable for what she doesn’t know: “You didn’t know there were fifty states today, but I’m sure you will when I ask you again tomorrow.” Hold high expectations for all your learners.

Learners should be familiar with a successful (and possibly also an unsuccessful) model of the task. It must be comprehensible to learners what the goal was and why they did or did not reach it. Teachers must have a good understanding of where the learners are and where they need to be in order for this to be effective, and it must be transparent to the learners.

It is important that both strengths and weaknesses are addressed so that learners know what they are already capable of. A combination of positive and negative feedback generally works best: “You wrote very descriptively about the male character, but you need to look up words when you aren’t sure about their spelling.”

Negative feedback should be accompanied by a demonstration of how to perform correctly. When words are used incorrectly, for example, the teacher should give correct word options. When providing learners with direction for improvement, don’t overwhelm them with too many things at once; one goal at a time is generally enough.

Because learners are most interested during and directly following an assessment task, do not allow much time to lapse between the task and the feedback. It is good practice to prepare for ten minutes after the task has been completed to discuss it as a class. This makes it more likely that learners will remember the task, their (incorrect) responses, and the teacher’s correction – a potential ‘teachable moment’. When feedback is given, it is important that the actual task (test, video recording, etc.) is accessible to the learner. When you have more than one class involved in the same assessment, postpone the feedback to the following week, but as soon as possible.

When correcting, make notes for feedback to give the whole class, being sure to maintain a balance between what learners did well and where they need to improve. This helps you be prepared to give both whole class and individual feedback.

Avoid only giving feedback at the end of a unit; the earlier the feedback is given, the more opportunity the learner has to improve. More frequent feedback can be given when multiple shorter assessments are planned rather than few longer ones.

Learners should have time in class to ponder the feedback and readdress the teacher when necessary. Learners should respond to the feedback given by the teacher either orally or in writing; they should remark whether they agree with the feedback and consider what consequences it will have on their learning.

It can be immensely frustrating to provide detailed feedback to learners to find that they don't acknowledge it. This can be prevented when a sheet of colored paper is given to each learner along with the feedback. On this sheet, learners are required to repeat what the teacher's feedback was, what they gained from the feedback in connection with the topic, and what they learned about themselves. Do not finalize the grade until this sheet has been returned.

Principles in Providing Feedback

Feedback needs to:

- Be personal, age-appropriate
- Contain comprehensible criteria
- Be specific
- Honor responses

Provide successful and unsuccessful models:

- What was the goal?
- Why was it reached or not reached?

Combine positive and negative feedback:

- Accompany negative feedback by demonstration of correct performance.
- Do not provide too many directions for improvement – one goal at a time is enough.

Do not allow much time to lapse between task and feedback:

- Discuss as a class after task has been completed (learners remember their responses).
- The task should be accessible to learners.

Give feedback to the whole class, balancing what they did well and where they need to improve.

Avoid only giving feedback at the end of a unit:

- The earlier learners receive feedback, the more opportunity they have to improve.

Learners should be able to address the teacher about the feedback received after having thought about it.

Learners should respond to teacher feedback:

- Do they agree?
- Do they know what consequences it will have on their learning?

It is effective to hold grades back until learners have responded in writing.

Activity 2

Using the text about effective teacher feedback and the summary of Principles in Providing Feedback to reflect on your own teaching, what insights have materialized?

Insights into aspects of “good” feedback:

Insights into aspects of “bad” feedback:

Activity 3

Reflect on your **dos** and **don'ts** about giving feedback.

Do:

Don't:



Watch the following video, which shows a second-grade teacher giving oral feedback on a learner's writing:

"Formative Assessment: A Teacher Provides Oral Feedback on a 2nd Grader's Writing Piece,"
uploaded by Writing City on March 11, 2015:

<https://www.youtube.com/watch?v=tTNKWbhQKYQ>

Consider the following questions:

How effective do you feel her feedback is?




My impression:

Can this serve as a model for you in your classroom?

My thoughts:

Providing Feedback - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I can identify what good feedback is.			
I can successfully communicate with my learners about feedback.			
I can select tools that will help me give more effective feedback to my learners.			

Indicative Answers

Topic 1: Feedback for Formative Assessment

Task 1 - Activity 1

	Feedback category	Choose the applicable statement (A-D) from above:
1.	Strengths	The ideas in your essay were very clearly expressed; you manage to get your point across.
2.	Weaknesses	Some of the expressions you used did not sound English, for example...
3.	Improvements to be made	You might want to read some teen-magazines. They are a good way to learn idioms because they use such typical language, and they can be fun, too.
4.	Assessment results	Overall, I felt you deserve a B grade.

Task 1 - Activity 2

Mark the appropriate box in the chart below.

	Statement	Helpful	Unhelpful	
1	Do you know the loci method? It might help you remember words better.	✓		Where to next?
2	Vocabulary is your greatest weakness. You need to study harder.	✓		
3	Your use of tenses is a catastrophe.		✓	
4	You did a good job.		✓	
5	Vocabulary boxes can be useful for memorizing new words.	✓		Where to next?
6	Your use of tenses has improved a little because you do not seem to mix them up any more.	✓		Where am I going?
7	I really saw how well you have worked on your connectives.	✓		How am I going?
8	I think you might want to work on...	✓		Where to next?
9	You should have prepared better for this test.		✓	

Topic 2: Feedback Types

Task 1 - Activity 1

How might you improve it?

Diana, tell me what you think you did well in your 3-minute talk. (...) Was the introduction OK? (...) Yes, I liked your question best, the one about how important the sun is. That got everyone's attention. Did you keep good eye contact with your audience? (...) I think you did. How about speaking clearly? (...) One could understand you well most of the time, but be sure to pronounce your important sentences really clearly so that everyone gets them. That's also true for the beginnings of some of your sentences. How do you feel about the language you used? (...) Yes, most of the time it was fine, and you used correct sentences with good connectives between your sentences. But you need to pay attention to your style. This is a presentation you give to people you might not know, so you cannot use slang like gonna and wanna. That's too informal. If you like YouTube, you might want to watch examples of formal presentations, really brief ones, and note the typical expressions that many presenters use. Then you can use them yourself to have the perfect presentation.

Task 3 - Activity 1

Please view the example of a feedback sheet below.

Oral Presentation: My Dream School

	Well done	Good	Satisfactory	Needs improvement
1. Organization				
Introduction (Grabber? Is it original?...)				
Main part (Logical? Linear? Structured?)				
Conclusion (Summarizes vital points?)				
2. Presentation techniques				
Clear articulation				
Pace				
Eye contact and body language				
3. Content				
Relevance				
Clarity of ideas				
Ideas				
4. Language				
Fluency (Fillers)				
Vocabulary range/ sentence structure				
Connectives used				
...				
Comments:				
Result:				
Learner:				
The most important aspect for me:				
One thing I would like to improve:				

Further Reading

To read more information on providing feedback as formative assessment, use the following link:



<https://www.edutopia.org/blog/providing-feedback-as-formative-assessment-troy-hicks>

To read more information on praise and feedback in the primary classroom, use the following link:



<https://files.eric.ed.gov/fulltext/EJ906941.pdf>

For a typology of corrective feedback, use the following link:



<http://lrc.cornell.edu/events/09docs/ellis.pdf>

To learn more about helping students take control of their own learning, use the following link:



https://www.learning.ox.ac.uk/media/global/wwwadminoxacuk/localsites/oxfordlearningintitute/documents/supportresources/lecturersteachingstaff/resources/resources/Helping_students_take_control_of_their_own_learning.pdf

Finally, if you wish to explore further ideas on how to **assess SpLD students**, visit the following Web site of the DysTEFL project and follow the suggestions in Chapter 10 (Assessment):



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

You can also download the DysTEFL training materials and work directly from there:

For trainers: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

For teachers: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

You might also like to read:

Kormos, J., & Smith, A. M. (2012). *Teaching languages to students with specific learning differences*. Bristol: Multilingual Matters. (Chapter 8 – Assessment)

Tsagari D. & Spanoudis, G. (Eds.) (2013). *Assessing L2 students with learning and other disabilities*. Cambridge: Scholars Publishers.

Tsagari, D. & Sperling, I. (2017). 'Assessing SLLs with SpLDs: Challenges and Opportunities for Equity in Education'. In Szymańska-Czaplak, E. (Ed.), *At the Crossroads: Challenges of Foreign Language Learning, Series: Second language learning and teaching*. Springer. pp. 175-188. DOI: 10.1007/978-3-319-55155-5_10.

References

- Hattie, J. and Timperley, H. (2007). 'The Power of feedback.' *Review of Educational Research*, 77, 81-112
- Walberg, H. (1984). 'Improving the productivity of America's schools'. *Educational Leadership*, 41(8), 19-27.

7. Alternatives in Assessment

Karin Vogt & Veronika Froehlich

❖ Aims

This Course focuses on alternatives in **assessment of learning**, closely related to **assessment for learning**. The most important and most widely used forms of alternatives in assessment – **teacher observation**, **peer-assessment** and **self-assessment**, and **portfolio assessment** – will be presented and discussed. We will touch on characteristics of each and how they can be implemented in the foreign language classroom.

❖ Expected Outcomes

Upon completion of this Course, you should be able to:

1. discriminate between traditional and alternative teacher assessment
2. critically evaluate and apply different tools as alternatives in assessment
3. successfully communicate with your learners about criteria for self- and peer-assessment
4. design observation sheets and self- and peer-assessment rubrics
5. identify suitable materials and criteria connected with portfolio assessment

❖ Key Concepts

- Assessment of learning
- Assessment for learning
- Teacher observation
- Self-assessment
- Peer-assessment
- Portfolio assessment
- European Language Portfolio

Brief Introduction

During the last two decades, the significance of **alternatives in assessment** as a counterweight to traditional testing has increased in many educational contexts. Alternatives in assessment put a focus on **holistic and learner-oriented assessment**, taking into account more factors than a test in collecting evidence on learner progress and performance. The Course focuses on four Topics: The first Topic gives you a **general introduction** to alternatives in assessment and contrasts them with traditional testing. The second Topic is on **teacher observation**, the most frequent type of alternative assessment, and illustrates ways to systematize this informal form of assessment. The third Topic focuses on **self- and peer-assessment** and how to involve learners in their learning and assessment. The fourth Topic deals with **portfolio assessment**, explaining the different types of portfolios, for example, the **European Language Portfolio**, and showing how to implement them in the foreign language classroom.

7.1. Topic 1: Introduction to Alternatives in Assessment

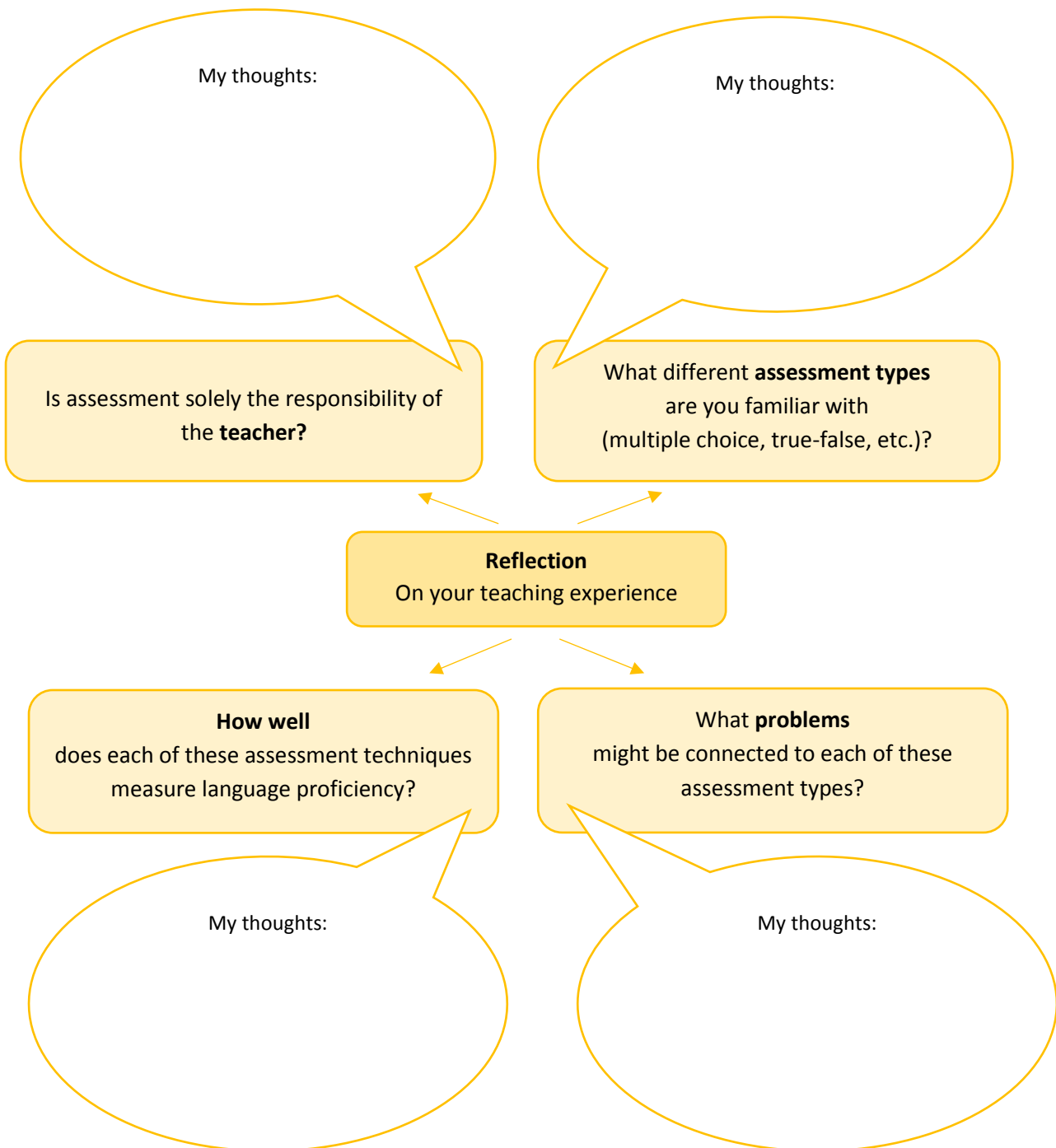


This Topic consists of **one Task** with a total of **four Activities**.

Task 1. Contrasting Traditional and Alternative Teacher Assessment

Activity 1

Before we begin, think about your teaching experience thus far, and try to provide answers to the following questions. Note your thoughts.



You are welcome to consider our thoughts in the **Indicative Answers** section.

When discussing classroom-based assessment, we usually distinguish between **Assessment of learning**, in which the product of learning is assessed (e.g. the extent to which a learner’s text meets the evaluation criteria we have set), and **Assessment for learning**, which aims at bringing assessment and learning closer together. How do these two differ?

Activity 2



Watch the video “*Assessment for learning vs. Assessment of learning*” that clarifies the differences between **Assessment for learning** and **Assessment of learning**, uploaded by CLAS Network on August 31, 2015: <https://youtu.be/wvRJyTExVU>

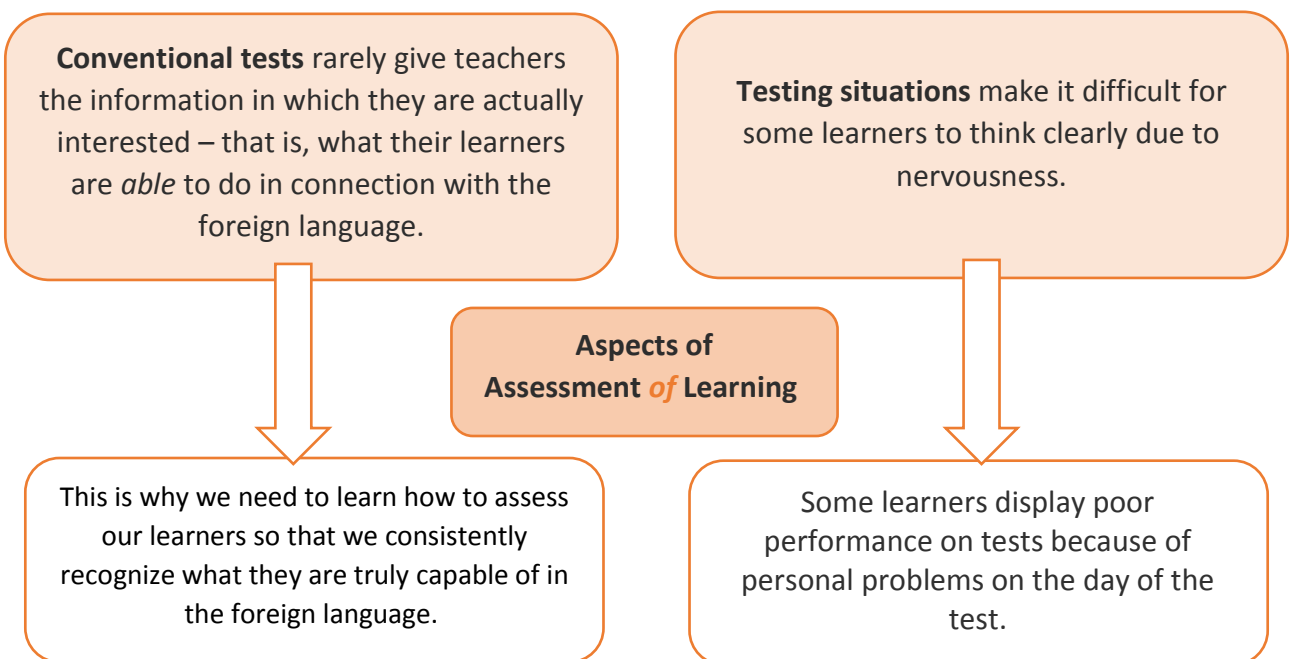
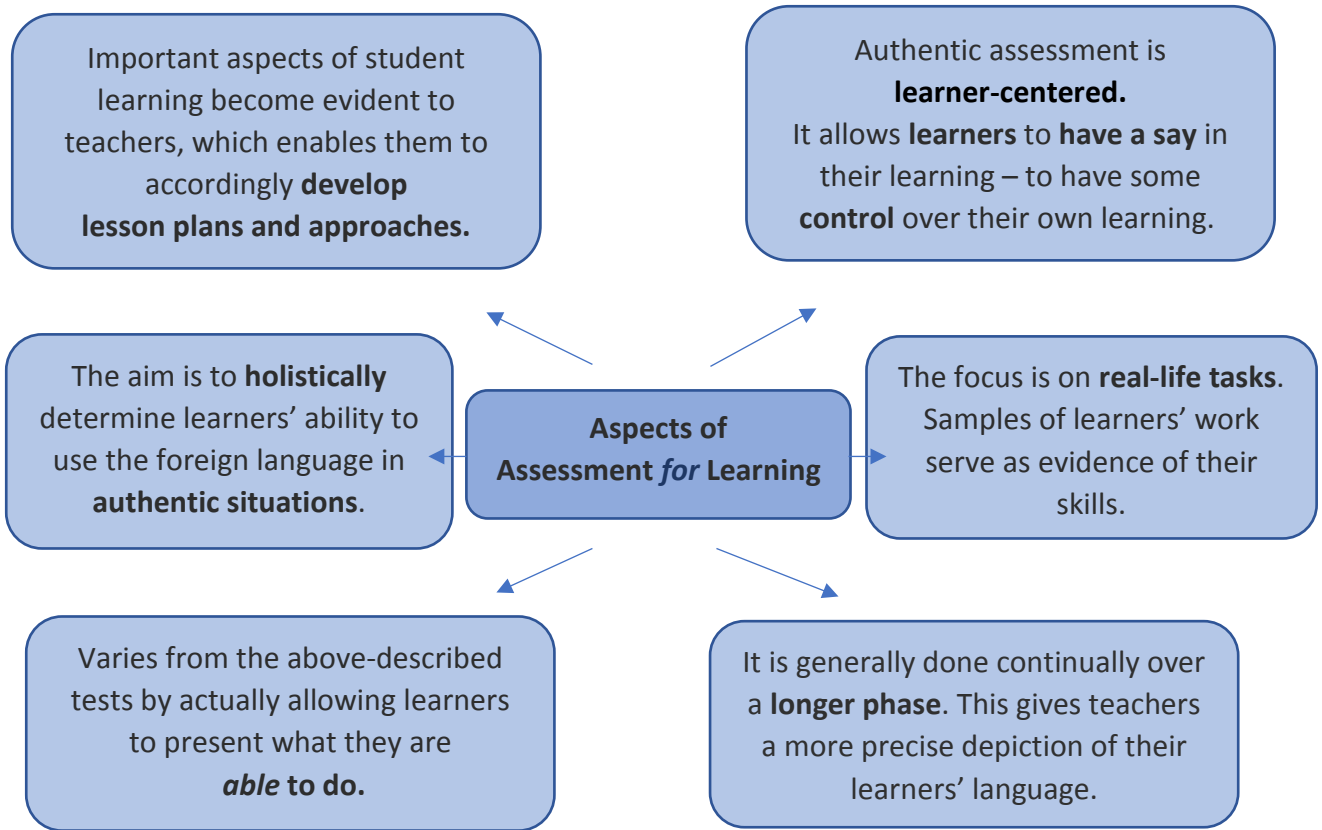
Now you are asked you to **match** particular characteristics of assessment with either assessment *of* learning or assessment *for* learning.

	Characteristics	Assessment of learning	Assessment for learning
1	The questions/tasks are open-ended.		
2	The focus is on the language.		
3	The focus is on assessing isolated skills.		
4	It is learner-centered		
5	The emphasis is on the process of learning.		
6	The goal is to examine the learners.		
7	The focus is on assessing integrated skills.		
8	It is teacher-centered.		
9	The emphasis is on the product of learning.		
10	There is one-way correctness.		
11	The focus is on communication.		
12	The test not only assesses but also teaches the learners.		

You are welcome to check our responses in the **Indicative Answers** section.

While **alternative assessment** (or authentic/informal assessment) tends toward **assessment for learning**, **traditional testing practices** are often associated with **assessment of learning**.

The following figures present typical characteristics of each assessment type.



Activity 3

What do you think are the main advantages of alternative assessment?

In the following activity, you will find a list of the main advantages of alternative assessment. Select those that are most **relevant** in your educational context.

	Advantages of alternative assessment	Most relevant advantage in my educational context
1	It does not interfere with the teacher's lessons.	
2	No additional days for testing need be sacrificed.	
3	Regular classroom activities (role-plays, writing, discussions) are the foundation for alternative assessment.	
4	It reflects exactly what is being done in the classroom.	
5	The teacher receives information on the strengths and weaknesses of each individual learner through multiple sources based on authentic activities.	
6	It can show a learner's development through the work samples and documentation collected.	
7	The teacher can identify both weaknesses and strengths alongside discovering more about learners' personal lives through their work.	

Assessment techniques must be **varied** and **adapted** to the needs of each group of learners and reflect the **goals** aimed at in the particular class. Alternatives to traditional testing can include, for example, video recordings of dialogues, self-assessment, portfolios, and teacher observations, to name a few. These techniques are most effective when learners and teachers **decide together** which technique will be used and when **learners are responsible** for choosing their work to be assessed.

Activity 4

Based on your experience and the information presented so far in this Course, answer the following questions:

What are the **most important aspects** of alternative assessment?

Would alternative assessment be **beneficial** to your learners? Why or why not?

Note your thoughts.

My thoughts:

You are welcome to consider our thoughts about the most important aspects of alternative assessment in the **Indicative Answers** section.

The following Topics focus on some of the most common methods of alternative assessment: **teacher observation, self- and peer-assessment, and portfolio assessment.**

7.2. Topic 2: Teacher Observation



This Topic consists of **one Task** with a total of **four Activities.**

Task 1. Teacher Observation as a Form of Assessment

Teacher observation is one of the most important and most frequent methods of alternative assessment. What exactly is ‘teacher observation’?

Activity 1

Think about teacher observation for a minute and try to provide answers to the following questions:

What **characterizes** this form of assessment?

What **benefits** might it have over traditional testing?

Note your thoughts.

My thoughts:

Now, read the following to learn more about the key aspects of teacher observation:

Teachers ‘informally’ assess their learners in the classroom constantly. This is generally done without any specific criteria in mind and without a focus on particular learners – more or less intuitively. When learners are slow to begin working on a task, teachers realize that they are having difficulties, for example. Because teachers rarely document these occurrences, though, it is impossible for them to measure these observations, and they unfortunately quickly forget them. Systematic, continual recording of such observations to document development of learner ability turns them into assessment. Language use and higher-order thinking skills can be observed when learners discuss and work with one another in task-based activities. This documented development should consist of personal notes with which the teacher can later discuss a learner’s progress with colleagues, parents, or the learner herself. The validity and reliability of teacher observations rise when these are done systematically. Teachers must decide what they will be assessing.

It is impossible to simultaneously assess every aspect of a learner’s performance. Specific criteria must be developed to assess the observed skill; basing assessment on a rough impression of a learner’s ability is unreliable. These teacher observations should be one part of multiple assessment forms, such as self-assessment and traditional classroom tests (Harris & McCann, 1996).

Probably the best thing about using teacher observation as a form of assessment is that it does not create the stressful test situation that traditional testing does. It is also an excellent assessment form of speaking, which is difficult to assess in more traditional testing. The teacher merely moves around among her learners and listens in while they're fulfilling tasks during a regular lesson. It is actually beneficial when learners remain unaware of this so that they produce language without pressure. The teacher can focus on individual learners and document her observations. Her focus might be on a learner's pronunciation, vocabulary range, etc. Another positive aspect of teacher observation is that it allows the teacher to additionally

assess non-linguistic abilities of her learners, such as cooperation, independence, creativity, and more. Harris & McCann (1996) note that "...we also need to think about our students' overall educational development. It is important for learners to develop in terms of language and in terms of attitudes towards learning, towards language, different cultures and other people. We also need to consider students' ability to take responsibility for and organise their own learning" (p. 21). Teachers can choose to give immediate feedback, praise, or to help their learners.

Main characteristics and benefits of teacher observation

- Teacher observation can systematically document the development of a learner's ability.
- Specific criteria must be developed for assessment => systematic teacher observations increase the [reliability](#) and the [validity](#) of the assessment.
- Teacher observation should be one part of multiple assessment forms.
- Teacher observation does not create stressful situations as testing does.
- Teacher observation can focus on individual learners.
- Teacher observation can additionally assess non-linguistic abilities such as co-operation, creativity, etc. => more holistic assessment.

To make their observations more systematic, teachers typically design an **observation sheet** for each learner to keep track of his/her progress. This sheet is usually used for an extended period of time, for example, a semester, and includes separate sections (or columns) in which teachers document the various activities that take place in the classroom at different points in time. A 'comment section' beside each activity is generally extremely useful. Here, teachers can add any thoughts they have while learners are carrying out activities. Later, teachers can return to their comments and identify the extent to which progress has been made. Additional sections for noting dates on which specific [feedback](#) is given to learners or communication with parents (and other stakeholders, if necessary) takes place are also useful.

Activity 2

Consider the various sections that can be included in an observation sheet. How would you design such a sheet for your context?

My thoughts:

Design an **observation sheet** for your learners (without necessarily filling in the various sections).

You are welcome to consider our sample observation sheet in the **Indicative Answers** section.

In addition to the above-described teacher observation, teachers can use so-called '**semi-structured assessment tasks**'. These are also used within regular lessons but particularly for assessment purposes. They are called 'semi-structured' because some of the language the learners will use has been pre-defined. Dialogues, interviews, and role-play situations using language that the learners have previously practiced in class can be used for this. Giving directions, asking about preferences, and other similar classroom discourse in which there are no pre-defined 'correct answers' lend themselves to this type of assessment.

When a teacher listens to a pair or group of learners communicating, for example, she should have pre-conceived criteria on which she will specifically focus, such as pronunciation, the use of complete and correctly formed sentences, and vocabulary.

Using this type of assessment, teachers can check whether they have successfully achieved the learning objectives they set or whether they need to review something. When teachers use the same criteria to assess all of their learners, their evaluations become increasingly reliable.

Activity 3

Consider the scenario below:

You have just done some classroom work on U.S. National Park visitor centers and the activities that are possible in National Parks. Learners are to role-play a dialogue between a tourist and a visitor center assistant in which the tourist asks for advice on activities in this National Park. You have pairs rehearse; then pairs perform their dialogues in small groups. You have focused on the vocabulary connected with various sports activities (hiking, climbing, kayaking, etc.), and you would like to assess learners' pronunciation, how effectively they ask for or give advice on possible activities, and the general accuracy of the language used.

Look at the following sample observation sheet. How might you **adapt** it to document learner performance on this activity?

Task observation form

Class: _____

Date: _____

Activity to be assessed: _____

Name:	Comments on:				Action to be taken

Note your thoughts.

My thoughts:

You are welcome to consider our sample observation sheet in the **Indicative Answers** section.

If you are a pre-service teacher or have little teaching experience, you might want to do the optional activity below:

Optional activity

This is an activity you have designed for your 15-year-old learners (B1 level). The learners are supposed to choose a National Park in the US and give a brief presentation on it. Everybody has different information, so the content of the presentation is new for the audience. The presenter must therefore be appropriately paced, intelligible, and interesting enough for the audience to follow. The presentation is supposed to have an introduction, a middle part, and an ending with time for questions. The learners have had time to rehearse the presentation in pairs, and they have dealt with relevant presentation vocabulary and phrases.

Look at the teacher observation form below and decide which of the criteria chosen for the task are appropriate. Choose a maximum of 4. Note your decision below. If possible, discuss this with a colleague.

My thoughts:

Optional activity

	<i>Excellent (4)</i>	<i>Good (3)</i>	<i>Fair (2)</i>	<i>Poor (1)</i>
Language use	Very good range of vocabulary and structures, used all relevant words and phrases in the presentation	Good range of vocabulary and structures, used most relevant words and phrases in the presentation	Satisfactory range of vocabulary and structures, used some relevant words and phrases in the presentation	Limited range of vocabulary and structures, used no relevant words and phrases in the presentation
Mechanics: spelling, structure, and agreement of nouns and verbs	Excellent mechanics, basically error-free	Good mechanics, only few minor errors	Satisfactory mechanics with some errors, some of which impede understanding	Poor mechanics, frequent and serious errors that impede understanding
Pronunciation	Very clear authentic pronunciation, very comprehensible	Clear, mostly authentic pronunciation, comprehensible	Some pronunciation errors but still comprehensible overall	Many pronunciation issues, hardly comprehensible at times
Presentation style	Very easy to follow, well rehearsed and fluent	Easy to follow, rehearsed presentation with good fluency	At times difficult to follow, not well rehearsed, satisfactory fluency	Very difficult to follow, apparently not rehearsed, not fluent
Structure of presentation	Introduction, main part, and ending with good transitions, time for questions	Introduction, main part, and ending discernible	Some structure apparent, but no clear distinction between introduction, main part, and ending	No distinct parts in presentation, lacks structure overall
Ideas	Creative and original with well developed ideas	Attempt at creativity, some developed ideas	Little expansion of ideas, little creativity	No creativity, minimal ideas
Delivery of presentation	Speaks slowly, very easy to follow, very good enunciation	Speaks slowly, easy to follow, good enunciation	Speed of presentation uneven, not always easy to follow, satisfactory enunciation	Difficult to follow, inappropriate speed, poor enunciation
Total Pts:				

When assessing learners through observation, it is very useful when teachers make notes in a journal/diary. This should be done whenever something 'out of the ordinary' occurs that gives the teacher insight into the learning process of her learners. This is especially beneficial because it allows the teacher to document the details of any such occurrences – outside of any **assessment criteria** otherwise used.

Do you document observations you make in the classroom?

If so, how do you do it? If not, why not? Reflect on your current practice.

My thoughts:

Activity 4

Although teacher observation has plenty of advantages in terms of assessing learner progress, it also has some significant **disadvantages**. Based on your experience as a teacher, what disadvantages might this form of assessment have?

Note your thoughts.

My thoughts:

You are welcome to consider our thoughts in the **Indicative Answers** section.

7.3. Topic 3: Self- and Peer Assessment



This Topic consists of **1 Task** with a total of **3 Activities**.

Task 1. Using Self- and Peer-Assessment in the Foreign Language Classroom

The main aim of **self-assessment** and **peer-assessment** is the involvement of learners in learning. This, in turn, may enhance **learner autonomy**. When **teacher observation** and **self- and peer-assessment** are merged, both information on learner performance and learner ownership regarding assessment processes are enhanced. Self- and peer-assessment can be used both with products of learning (audio samples of a dialogue, a written invitation card, etc.) and for overall learning progress.

With product-oriented assessment, criteria for judging the quality of the performance must be defined in advance, ideally together with the learners.

Activity 1




Consider the following scenario:

Your intermediate (B1 level) teenage learners design a brochure for a Canadian National Park of their choice after having done research. Learners are to evaluate their peers' drafts in pairs. The **criteria** they collaboratively agreed on are:

- Clarity of information / relevance of information
- Comprehensiveness of information
- Characteristics of a brochure (e.g. layout, important elements)
- Comprehensibility of the language used

Using the template provided below, design a **peer-assessment form** that your learners could use to evaluate their classmates' work.

Peer assessment template

Name:			

Keeping the above criteria in mind, what criteria would you write in the column on the left (under the name of the learner whose work is being assessed) to help your learners carry out peer-assessment successfully?

My choice of criteria:

You are welcome to consider our sample peer-evaluation form in the **Indicative Answers** section.



Watch the following video: *“Peer Assessment: Reflections from Students and Teachers,”* uploaded by Jobs for the Future on August 22, 2013

<https://youtu.be/DqWCJZH8ziQ>

What is **your approach** to peer-assessment?

My approach to peer-assessment:



Peer assessment can even be done with **very young learners**.

To see how, watch the following video: *“Peer assessment,”* uploaded by AppalachiaRCC on October 16, 2016

<https://youtu.be/i-HQptt-RtA>

For **process-oriented assessment**, learners can be provided with **can-do statements** as **checklists** with which they rate their abilities or their learning process over time. The learning biography of the **European Language Portfolio** (see Task 5, Portfolio assessment) is a typical example of self-assessment that works this way. Learners rate themselves regularly using statements like: “I can describe my favorite pet,” “I can understand what my teacher tells me to do,” or “I can write some sentences about where I live.”

Activity 2

Consider the scenario below:

Your learners at beginning level (A2) give presentations entitled “My family and I.” In these presentations, each learner presents his/her family members and describes:

his/her own appearance and that of each of his/her family members

his/her favorite activities and those of each of his/her family members

After each presentation, the other learners ask questions about the family members presented. At the end, **every learner assesses his/her presentation using can-do descriptors**. These types of activities can take place at the end of units.

Which of the following **can-do descriptors** would you include in the self-assessment form below?

	Descriptors	To be included
1	I can describe my hobbies.	
2	I can talk about my favorite activity.	
3	I can write about an animal or my pet.	
4	I can talk to my friends about my hobbies.	
5	I can ask and answer questions about where I live.	
6	I can read important information about where I live.	
7	I can describe special celebrations to my teacher and class.	
8	I can answer questions about myself and my family.	
9	I can talk about my family.	
10	I can talk about the color of my hair and eyes and how tall I am.	

Here is an example for a self-assessment form. Include the descriptors from the list above that you find relevant.

I can...				What can you do about it? Please also ask your teacher

When you have finished, take a look at our sample **self-assessment** form for this activity in the **Indicative Answers** section.



Watch this short video on perspectives of students and teachers concerning self-assessment: *“Self-Assessment: Reflections from Students and Teachers,”* uploaded by Jobs for the Future on August 22, 2013.

<https://youtu.be/CkFWbC91PXQ>

Have you ever used this method in your classes? If you are a pre-service teacher, do you think you **might use it**?

Reflection on this method:

Both teachers and learners need opportunities to self- or peer-assess so that they can develop this competence gradually. They can eventually develop their own criteria for evaluation, allowing assessment to be more **learner-oriented**.

To develop **learner autonomy**, self- and peer-assessment must be a continuous, regularly practiced endeavor. This can be time-consuming because time must be spent on the development of **assessment criteria** and on practicing self- and peer-assessment procedures. **Feedback** rounds (peer feedback or teacher-learner conferences) must be considered in lesson planning. The more experienced the learners become, however, the less time will be needed.

Another concern that is often voiced is that learners might rate themselves unrealistically or are biased, for example, in favor of their friends. Hasselgreen (2003) has shown, however, that young learners involved in self-assessment are critical in their ratings and do not tend to rate themselves too positively. **Potential bias** related to peer-assessment can be avoided by **consistently using evaluation criteria**.

Activity 3

How do learners **benefit** from self- and peer-assessment?



Watch two short videos by Dylan Wiliam and Paul Black on self- and peer-assessment.

“Self- and peer Assessment Dylan Wiliam,” uploaded by Mary Land on March 27, 2014:

<https://youtu.be/5P7VQxPqgTQ>

“Paul Black Describes peer and self assessment,” uploaded by Tyson Sprawl on May 5, 2012

<https://youtu.be/NnqHERCFsWM>

While watching, list the **benefits and limitations** of **self-** and **peer-assessment**.

Benefits of self- and peer-assessment	Limitations of self- and peer-assessment

How can these forms of assessment support assessment carried out by the teacher?

Can you identify any limitations related to self- and peer-assessment that might exist in your context?

Note your thoughts.

My thoughts:

You are welcome to view our thoughts on benefits and limitations in the **Indicative Answers** section.

7.4. Topic 3: Portfolio Assessment



This Topic consists of **3 Tasks** with a total of **5 Activities**.

Task 1. Introduction to Portfolio Assessment in the EFL Classroom

The word ‘portfolio’ often brings rise to the image of artists toting their work to display to others. The items they decide to include in these portfolios are generally geared toward those who might potentially view them. Reflections on their own work or reviews (**feedback**) from someone else are sometimes included in these portfolios (Wilson, 2002). What exactly is meant by the term **portfolio assessment** in the foreign language classroom?

A **portfolio** in the foreign language classroom is a collection of a learner's work that displays the progress that learner is making in connection with classroom goals. This assessment method is very learner-centered. A portfolio can contain student work samples, optimally chosen by both the teacher and the learner, that are based on the learning goals. These can, for example, be samples of the learner’s writing, a reading journal, and audio and/or video recordings. Comments on these samples both by the teacher and the learner should be included. The importance of learners selecting the samples to include in their portfolios is great. Through this, they decide which samples of their work document their development over time. This learner autonomy should definitely be encouraged by the teacher.

Activity 1

Below is a sample list of **suitable materials** for a foreign language portfolio. Think about your own classroom(s).

What **materials** could your learners include in their portfolios? Is any of the above **relevant** to your context?

Suitable materials for a Foreign Language Portfolio	To be included	Relevant to my context
Pattern poetry produced by the learner		
Essay on solar energy in draft form		
Artwork or photos related to a topic dealt with in the foreign language classroom		
A dialogue taped during the lesson		
Homework exercises on connectives		
Video of a role-played job interview		
Notes on a teacher presentation		

Task 2. How to introduce Portfolio Assessment in the EFL Classroom

When introducing portfolio assessment in the EFL classroom, the purposes of the portfolio need to be clearly communicated in advance, and both teachers and learners must consider it a valuable assessment tool in order for it to be effective.

The **reflection, (self-) assessment,** and **documenting** functions are usually considered to be the most important functions of the portfolio, but there are also others.

Ownership of the portfolio can increase **learner motivation**. It complements teacher assessment and learner assessment and promotes **interaction** between learners and the teacher. It can also function as a **communication tool** with parents as important stakeholders.

Apart from being learner-centered and allowing ownership of learning, portfolios offer opportunities for self-assessment and therefore **critical thinking processes** (Which work is my best and why?) and for **collaboration processes** with peers, for example, when learners work together on one task for the portfolio. Portfolio assessment allows for a more holistic approach to **language assessment** because it can include multiple dimensions of learning that discrete-point testing cannot encompass.

Activity 2

There are both advantages and disadvantages connected to portfolio assessment. Consider the requirements of your educational context.

What are the **advantages** and **disadvantages** of using portfolios to assess your learners? Fill in the list.

Advantages of portfolio assessment	Disadvantages of portfolio assessment

You are welcome to view our results in the **Indicative Answers** section.

Brown and Abeywickrama (2010) suggest the following **steps for successful portfolio development**:

State objectives clearly: Select one main function of your portfolio in class and connect it to your curriculum. Make the portfolio a compulsory and integral part of the curriculum so that it is taken seriously.

Give guidelines on the materials to include: Give clear selection criteria of materials for the portfolio. Model tasks that are suitable for collection in the portfolio; if possible, show the portfolio of a previous learner as an example of good practice. Be clear about how much choice learners have concerning materials they can select for their portfolios.

Communicate assessment criteria to students: Portfolio assessment incorporates teacher and self-assessment, and the criteria should be the same. These criteria must be communicated (or jointly developed) in advance. Learners can be asked to reflect on and assess their performance, for example, by retelling a story. The teacher and the student discuss the learner's performance during a **conference**, and the teacher gives his or her view and **feedback**. Together, they plan future learning objectives.

Designate time within the curriculum for portfolio development: Allow in-class time for learners to work on the portfolio and for teacher-learner conferences to value this type of **formative assessment**.

Establish periodic schedules for review and conferencing: Reserve time for portfolio work regularly in order to avoid last-minute rushed collections that undermine the idea of process-orientation.

Designate an accessible place to keep portfolios: Make an accessible space in the classroom or elsewhere for the learners to keep their portfolios for practical reasons. It is inconvenient for learners to carry huge and bulky material collections.

Provide positive washback-giving final assessments: On completion of the portfolio and/or a term, a final assessment should be given. The types of grading range from numerical scores to verbal evaluation, depending on the educational context. Numerical scores enable the comparison of performance across students and might be appropriate in otherwise high-stakes assessment cultures due to their face **validity**. In other contexts, a narrative feedback might be more appropriate. To highlight the learner-centeredness of the tool, strengths and areas for improvement should be part of the evaluation.

Activity 3

Consider the above-mentioned steps. How **feasible** are they in your educational context?

Note your thoughts.

My thoughts:

Activity 4

Consider the following scenario:

Your learners have retold a story you have worked on in class as a written text and have included this in their portfolios.

Your job now is to evaluate the written product of the learner.

Design a simple rating scale for a final portfolio assessment. Use a 4-point level scale like the one in the example below to fill it in. Adapt the criteria below to your purposes and context. The first criterion has been done for you, but, of course, you can adapt it as you wish.

Here are possible criteria you could use:

- Learner engagement
- Depth of content
- Creativity
- Learner language: range of structures/vocabulary
- Accuracy (language)

Criterion	Very good	Good	Satisfactory	Needs improvement
Learner engagement: Effort invested in selecting and evaluating work samples				

You are welcome to compare your own rating scale for portfolio assessment to our example in the **Indicative Answers** section.



Watch the video: *“Portfolio sharing,”* uploaded by ESL Literacy on April 17, 2014, to see how a teacher uses this method and how effective she feels it is.

<https://youtu.be/PoHefuvqivw>

Can you imagine using this idea in your classroom?

My thoughts:

Task 3. The European Language Portfolio

The Council of Europe developed the **European Language Portfolio (ELP)** that has the purposes: **evidence, reflection, purpose, and audience**. It is different from regular portfolios because it is designed exclusively for **language learning** and has a more structured content. The ELP incorporates can-do statements from the Common European Framework of Reference (Kohonen, 2002). European Language Portfolios must be accredited by the Council of Europe. They all consist of a language passport that gives evidence of a learner's plurilingual competence, a language biography that is supposed to document a learner's foreign language learning process, and collections of learners' works. These portfolios are closely linked to self-assessment because they feature "I can do" statements in the language biography.

Enter the following link to view an example of a European Language Portfolio for primary school learners (from Ireland).



<http://elp.ecml.at/LinkClick.aspx?fileticket=DNdgQcAwKVs%3d&tabid=2747&language=en-GB>

The next link shows a Web site from the European Center of Modern Languages on European Language Portfolios in general. Teachers can raise learner awareness of learning goals by discussing with learners how they can determine which can-do statement to choose for themselves. Learning strategies can also be a focus of this discussion.



<http://elp.ecml.at/>

Activity 5

Based on your teaching experience, the requirements in your educational context, and on the information from this Course, answer the following questions:

Would you use **portfolios** to assess your learners? **Why** or **why not**?




My thoughts:

What **steps** would you follow to introduce portfolio assessment in your class(es)?

My thoughts:

Alternatives in Assessment - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I can discriminate between traditional testing and alternative teacher assessment.			
I can critically evaluate and apply different tools as alternatives in assessment.			
I can successfully communicate with my learners about criteria for self- and peer-assessment.			
I can design observation sheets and self- and peer-assessment forms.			
I can identify suitable materials and criteria connected with portfolio assessment.			

Indicative Answers

Topic 1: Introduction to Alternative Assessment

Task 1 - Activity 1

Is assessment solely the responsibility of the teacher?

Potential answer: No, it is not the sole responsibility of the teacher (although, in some contexts, it can be seen as the sole responsibility).

What different assessment types are you familiar with (multiple choice, true-false, etc.)?

Examples:

multiple choice

true or false

gap-filling

sentence transformations

essay tests...

What problems might be connected to each of these assessment types?

Potential problems: Discrete point testing types are not holistic enough and may not capture learners' "real" language proficiency. Learners might feel that they are under pressure and might not be able to perform at their best.

Task 1 - Activity 2

	Characteristics	Assessment of learning	Assessment for learning
1	The questions/tasks are open-ended.		✓
2	The focus is on the language.	✓	
3	The focus is on assessing isolated skills.	✓	
4	It is learner-centered.		✓
5	The emphasis is on the process of learning.		✓
6	The goal is to examine the learners.	✓	
7	The focus is on assessing integrated skills.		✓
8	It is teacher-centered.	✓	
9	The emphasis is on the product of learning.	✓	
10	There is one-way correctness.	✓	
11	The focus is on communication.		✓
12	The test not only assesses but also teaches the learners.		✓

Task 1 - Activity 4

The most important aspects of alternative assessment can be summarized as follows:

- Learners communicate actively by performing, creating, or producing something.
- Learners use problem solving to create their responses.
- Tasks are authentic – meaningful, challenging, interesting, replicate real-world contexts.
- Language skills are integrated.
- Both the process and the product are important.
- Assessment criteria are clear and transparent to everyone.

Topic 2: Teacher observation

Task 1 - Activity 3

Task observation form

Class: _____

Date: _____

Activity to be assessed: Role-play National Park visitor center

Name:	Comments on:				Action to be taken
	Pronunciation	Word field activities	Task achievement	Accuracy	
	Feedback given on (date):				
	Spoke with parents on (date):				

Optional activity

The criteria on the scale presented were language use, mechanics (spelling, structure, and agreement of nouns and verbs), pronunciation, presentation style, structure of presentation, ideas, and delivery of presentation. Mechanics is not a suitable assessment criterion because the activity is oral and not written. Ideas is a criterion that does not suit well to a matter-of-fact presentation. Delivery of presentation is feasible as a criterion, but the descriptors represent various categories lumped together. The four criteria we would recommend are language use,

pronunciation, presentation style, and structure of presentation because these embody the requirements of the activity.




Task 1 - Activity 4

Naturally, using **teacher observation** as a form of assessment also has some **disadvantages**. It does take relatively long to assess every learner, which inevitably gives those assessed last an advantage. As with traditional testing, factors that can affect a learner's performance, such as sickness, lack of motivation, etc., can make the assessment unreliable. Furthermore, the order in which teachers evaluate their learners can also negatively impact **reliability**. Sometimes, teachers judge the first learners that perform well to be much better than they actually are – and might judge differently after having assessed several other strong learners. Thus, teachers must be aware that the order in which learners are assessed can create problems.

Topic 3: Self- and Peer Assessment




Task 1 - Activity 1

Peer assessment template

Name:			
The brochure is easy to understand.			
It has a clear structure.			
It has all the relevant/important information (e.g. opening times, prices).			
It is attractive and interesting to read.			

Task 1 - Activity 2

Sample self-assessment form

Name:				What can you do about it? Please also ask your teacher
I can answer questions about myself and my family.				
I can talk about my family.				
I can talk about the color of my hair and eyes and how tall I am.				
I can talk about my favorite activity.				
I can talk to my friends about my hobbies.				
I can ask and answer questions about the place where I live.				

Task 1 - Activity 3

Teacher assessment can be complemented by self- and/or peer-assessment because self- /peer-assessment adds a new perspective to teacher evaluations.

Potential benefits: enhanced learner autonomy, learner ownership of assessment, increased teacher-learner communication (...).

Potential limitations: inhibition to include learners in the assessment process (“teacher knows best”), difficulties to accept changed teacher and learner roles, time constraints, resistance from parents, clashes with prevailing assessment culture, particularly in high-stakes assessment cultures.

Topic 4: Portfolio Assessment

Task 2 - Activity 2

Advantages: [portfolio assessment](#) can be included in regular classwork activities; learners are actively involved in the learning process; enhanced learner-teacher communication; increased insights for teachers into learners’ development processes; more systematic/reliable assessment due to many potential assessment materials

Disadvantages: time needed for learner training; time and effort needed for teacher training; discipline needed to systematically set aside in-class time for conferences and work on portfolios

Task 2 - Activity 4

Sample rating scale for portfolio assessment

Criterion	Very good	Good	Satisfactory	Needs improvement
Learner engagement: Effort invested in selecting and evaluating work samples.				
Inclusion of out-of-class work: Work samples that were produced outside of class (e.g. homework, voluntary work) have been included.				
Depth of content: The work samples show a level of variety and accuracy regarding their content, e.g.				

analysis instead of description.				
Creativity: The work samples show a level of originality and variety regarding their content, text types, etc.				
Language: The vocabulary used is varied and not repetitive, structures are varied, and the language learner takes risks aiming at variation.				
Linguistic accuracy: The language produced is accurate, the communicative message comes across.				

Further Reading



More information on the **reliability** of self-assessment
<http://pareonline.net/pdf/v11n10.pdf>



An article about using journals for learners' reflection and self-assessment
<http://iteslj.org/Techniques/Arciniegas-LearnersJournals.html>



Useful information about portfolios
<http://jfmuller.faculty.noctrl.edu/toolbox/portfolios.htm>



Information about planning and implementing portfolios in your classroom
https://esl-literacy.com/sites/default/files/Portfolio%20Planning%20and%20Implementation_0.pdf



Read about reflective journals and learning logs
http://www.niu.edu/facdev/pdf/guide/assessment/reflective_journals%20and_learning_logs.pdf

Finally, if you wish to explore further ideas on how to **assess SpLD students**, visit the following Web site of the DysTEFL project and follow the suggestions in Chapter 10 (Assessment):



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

You can also download the DysTEFL training materials and work directly from there:

For trainers: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

For teachers: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

You might also like to read:

Kormos, J., & Smith, A. M. (2012). *Teaching languages to students with specific learning differences*. Bristol: Multilingual Matters. (Chapter 8. Assessment)

Tsagari, D. & Spanoudis, G. (Eds.) (2013). *Assessing L2 students with learning and other disabilities*. Cambridge: Scholars Publishers.

Tsagari, D. & Sperling, I. (2017). 'Assessing SLLs with SpLDs: Challenges and Opportunities for Equity in Education'. In Szymańska-Czaplak, E. (Ed.), *At the Crossroads: Challenges of Foreign Language Learning, Series: Second language learning and teaching*. Springer. pp. 175-188. DOI: 10.1007/978-3-319-55155-5_10.

References

Brown, H.D. & Abeywickrama, P. (2010). *Language Assessment. Principles and Classroom Practice* (2nd ed). London et al.: Longman / Pearson Education.

Hamp-Lyons, L. & Gordon, W. (2000). *Assessing the Portfolio. Principles for Practice, Theory and Research*. Creskill, NJ: Hampton Press.

Harris, M. & McCann, P. (1996). *Classroom-based Evaluation in Second-Language Education*. Oxford: Heinemann.

Hasselgreen, A. (2003). *Bergen 'Can Do' project*. Strasbourg: Council of Europe.

Kohonen, V. (2002). The European Language Portfolio: from portfolio assessment to portfolio-oriented language learning. In P. Kaikkonen & V. Kohonen (Eds.), *Quo vadis foreign language learning?* Tampere: University of Tampere, 77-96.

O'Malley, J.M. & Valdez Pierce, L. (1996). *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. White Plains, NY: Addison-Wesley.

Wilson, S. (2002) *Comparing peer, self and tutor assessment in a course for university teaching staff*. Paper presented at the Learning Communities and Assessment Cultures Conference, University of Northumbria 28-30 August 2002.

8. Test Impact

Nicos Sifakis & Stefania Kordia

❖ Aims

This Course invites you to consider the different ways language tests and exams can impact teaching, learning, language use, as well as other domains related to and, at times, dependent on, language testing, such as curriculum designing, coursebook selection and their adaptation in different classroom contexts, language task design and adaptation, and education policy.

❖ Expected Outcomes

Upon completion of this Course, you should be able to:

1. Distinguish between different ways tests impact language learning, teaching and use.
2. Identify positive and negative washback effects of high-stakes tests on teaching and learning.
3. Identify positive and negative washback effects of high-stakes tests on coursebooks.
4. Demonstrate the essential differences (and similarities) between using, learning, teaching, assessing, and testing English as a foreign language.

❖ Key Concepts

- Test impact
- Testing culture
- Washback
- Language use
- Language learning
- Language teaching
- Language assessment
- Language testing
- Standardised tests
- High-stakes tests

Brief Introduction

Without any doubt, we live in an age where taking tests for the purposes of **certification** is all the more important. Certificates are very helpful in boosting up our learners' (and our own!) CVs and professional profiles, but what effect do they have on teaching, learning and using the language?

In this Course, we are going to examine **how tests impact using, learning, teaching, and assessing English as a foreign language**. We will first consider what the terms using, learning, teaching, and assessing mean and what their differences and common characteristics are. Then, we will define **test impact** and establish how it relates to all these characteristics. We will then

focus on the influence of tests on teaching and learning in particular (the term used is 'washback') and link this influence to teachers' and learners' attitudes and, also to the function of coursebooks in the classroom. Depending on the choice of coursebook and the ways we use it to teach the language, washback can be positive or negative. We will look at the various factors that influence washback and reflect on the deeper effect that many tests have on teaching and learning in general.

In this Course you will be invited to consider examples of positive and negative washback with reference to different tests and their corresponding teaching materials. You will also be invited to reflect on the more general impact of tests and testing on teaching and learning. Please note that, throughout the Course, reference will be made to the teaching of English as a foreign language, but everything pointed out in this regard concerns **any language** taught, learnt, or assessed/tested/examined as a foreign language.

8.1. Topic 1: Essential Clarifications



There are **two Tasks** in this Topic containing **five Activities** in total.

Task 1. Using and learning English

The purpose of this introductory section is to introduce you to the ways in which tests can influence different aspects of the whole teaching, learning, and language-use infrastructure.

Admittedly, there are many different types of tests. The ways each one of these types is linked to the classroom can influence, in a positive or negative way, the way teaching is conducted and, ultimately, the way learners are exposed to and engaged with the language.

Activity 1

Without restricting yourself to a particular type of test, reflect for a minute on ***your own previous experience*** with language tests, ***as a teacher*** and ***as a learner***. In order to start thinking about the positive and the negative influence of tests, think about the things that you liked and the things that you disliked (as a learner and as a teacher) about language tests. In the space below, write down your thoughts. Try to bring up at least two things that you liked and two things that you disliked about language tests.

If you want some extra prompt, feel free to consider our own thoughts, in the **Indicative Answers** section. When you finish, continue reading.

	Things I liked about tests	Things I disliked about tests
As a learner		
As a teacher		

In this Course, we will focus on two things. On the one hand, we have the **tests**. On the other hand, we have their influence, their effect or their **impact** on different aspects of teaching and learning. It is important from the beginning to keep those two separate. One way in which tests impact teaching and learning is evident in **our own perceptions, expectations and practices as teachers** and, to a large extent, on the actual **coursebooks and other teaching materials** we use to prepare our learners for a particular test. These coursebooks and teaching materials can be appropriately designed to serve the purposes and aims of the test or they may misinterpret the test's specifications and end up misguiding learners. Another way in which tests can influence teaching and learning is evident in the whole **culture of testing** that tends to dominate foreign language classroom contexts, even when these contexts are not directly linked to a particular test. In what follows we will consider both of the above cases: **the direct impact of a particular test on the teaching materials** and **the broader influence of testing in teaching and learning**.

We should appreciate that tests are developed to serve a particular purpose, and there are ways to evaluate the extent to which they serve this purpose adequately, validly and reliably. At the same time, however, when preparing learners to sit certain tests, what happens is that we sometimes tend to teach to the test rather than teach the language. In this Course we will see how and why this happens with reference to many examples and seek ways to deal with it.

In order to clearly see what happens when certain tests impact teaching, we first need to be able to distinguish between certain processes involved in engaging with a foreign language. Let us start with distinguishing between **using** and **learning** English.

Activity 2

Think about the essential characteristics of each of the following processes surrounding teaching and learning a foreign language. In the spaces below write down your thoughts about what happens when we:

• **Use English?**

• **Learn English?**

How would you define these two processes? **Where** and **when** do they take place? How are they **different** and how are they **connected**? When you finish (and only then!) keep reading.

Using English	Learning English

Let us try to see the different functions in each of the statements presented in Activity 2. Using English involves actual use of the language. This can take place either inside or outside the classroom. By **language use**, we mean the generation of meaningful, spoken or written discourse that is produced for a communicative reason and is comprehensible to our audience. Of course, the reason for producing discourse can change, depending on the context: when we use the language outside the classroom, we are usually guided by our own intentions and respond to a very “real” need to communicate. When we use the language in the classroom, we still produce it, but the “quality” of the language that we will produce is likely to depend on the activities that we are asked to carry out. So, while language use has the same underlying characteristic (i.e., producing language), it can be qualitatively different, depending on where and why we produce it.

As a concept, **language learning** is more complicated than using the language. It essentially involves grasping, or acquiring, the knowledge and skills necessary in order to produce discourse that is meaningful. As a process, learning can take place everywhere and anywhere, at any time and any place. Learning happens inside and outside the classroom and can be a conscious or a subconscious process. For example, carrying out a grammar task in the classroom is intended to lead to conscious learning of that grammatical phenomenon. Alternatively, as the use of the language outside the classroom focuses primarily on communicating meaning, noticing certain grammatical phenomena is likely to be more time-

consuming, but it can also lead to learning (of those grammatical structures) that would be more subconscious than conscious.

This means that the link between learning and using English is a difficult thing to pin down. In its most simplistic form, learning is a prerequisite to using (in other words, in order to produce English we *first* need to learn English). But this is by no means the only way we *learn* a language. We very often **learn by using**, as long as we care to notice what works and what does not work in our communication. In fact, in many ways, learning a language has a more lasting effect when it is intrinsically linked with using that language.

One more thing we should consider when we think about using and learning English is **where** these processes take place. This is important because if we live in a country where English is used all the time (e.g. in a city in the UK, the USA, Australia, etc.), then our use of the language outside a classroom would be much more frequent than if we lived in, say, a city in Greece, Brazil, France or Hungary. In the latter case, we would not be ‘forced’ to use English outside the foreign language classroom so much, which means that we would learn it (and learn to use it) differently. Of course, while this distinction is certainly true, it is jarred by the fact that many learners of English living in contexts where English is not typically used outside the classroom nevertheless do use a lot of English when they use online technologies, e.g., when they play videogames online, and so on.

Task 2. Teaching, assessing and testing English

Activity 3

As with Activity 2, think about the essential characteristics of each of the following processes. In the empty spaces, write down your thoughts about what happens when we:

• **Teach** English

• **Assess** English

• **Test** English

Try to define these three processes. **Where** and **when** do they take place? How are they **different** and how are they **connected**? When you finish (and only then!) keep reading.

Teaching English

Assesing English

Testing English

In its simplest form, **language teaching** refers to the ‘steering’ of the process of learning by the teacher in a classroom setting. This implies that, in the formal sense of the term, teaching is based on an underlying plan (or syllabus) that addresses a particular type of learner or learners, is followed for a specific time and should lead to a specific desirable outcome.

What is the outcome of teaching? In its ideal form, teaching should be tailored to the cognitive and psychological needs of the learners. Probably one of the most central characteristics of teaching is that it usually involves a classroom and centres around one or more coursebooks.

Teaching can have many guises. It can be direct or indirect, e.g., it is direct when the classroom setting involves the teacher explicitly telling learners what to do, and it almost always centres around one or more coursebooks. It can be theory- or practice-driven, depending on whether the teacher follows a top-down or a bottom-up approach to learning. A top-down approach would mean building activities around grammatical theorizing, whereas a bottom-up approach would mean first exposing learners to different uses of language before theorizing, if at all.

Language assessment is the practice of evaluating the extent to which learning and teaching have been successful, focusing on what learners can do with the language, on their strengths rather than their weaknesses. This evaluation usually involves forms other than paper-and-pencil tests, such as portfolios (i.e., compilations of learners’ essays, reports, other assignments that may have been carried out), presentations or demonstrations. Assessment works best if it is continuous (or ‘on-going’), i.e., carried out throughout the entire process of teaching and learning, not just on completion of the learning/teaching process. When assessing is learner-oriented, it helps establish whether learners’ learning needs and styles have been made good use of. Appropriate assessing is also very helpful for teachers because it can facilitate self-evaluation of their own pedagogical decisions.

Language testing refers to the precise measuring of different aspects of learners’ performance. It is usually seen as the end-product of teaching and learning and may either isolate or integrate linguistic, communicative or socio-cultural skills. When we think of tests, we generally

tend to think of tests that address all learners in the same way (they all answer the same questions and carry out the same tasks) and are scored in a ‘standard’, consistent and reliable manner, which makes ranking and comparing learners’ performance easier. These tests are known as “*standardized*” tests.

Because of their nature, standardized tests can equip learners with certification of their various levels of proficiency. A type of standardized test that is central to understanding the impact of tests is the so-called ‘high-stakes test’. **High-stakes tests** are “tests whose results are seen – rightly or wrongly – by students, teachers, administrators, parents or the general public, as being used to make important decisions that immediately and directly affect them” (Madaus 1988, p. 87).

Activity 4

Consider the three statements below. They are simple statements that describe the assumptions we sometimes make about the link between teaching and testing, but they are presented in a jumbled order. Think about the meaning of each statement and, then, put them in their correct order.

A. “ <i>Only what is tested is taught.</i> ”	
B. “ <i>What is not taught is not learnt.</i> ”	
C. “ <i>What is tested should be taught.</i> ”	

When you’re done, compare your answer with ours, below.

Let us consider each in turn. The first assumption is that, **in order to test something, you have to teach it first (statement C)**. It is sometimes the case that syllabus designers who are prompted to raise the status of their courses make the conscious or unconscious decision to link them with a particular high-stakes test. They may then go on to develop EFL courses that centre around the specifications of that individual high-stakes test.

In a similar fashion, **coursebook developers** design coursebooks that encapsulate the essential characteristics of language learning and use specified by those particular high-stakes tests (we will see what those characteristics are below). The extent to which those high-stakes tests have **a positive or negative impact** (or **washback**) on these coursebooks (and, by extension, on teaching and learning) depends on the ways that those coursebooks treat learners’ exposure to language and language learning tasks. For example, if these tasks are reminiscent of the tasks included in high-stakes tests, without any recourse to authentic language communication, then those tests have a **negative washback** on the coursebook.

The second assumption presupposes that teachers (and/or syllabus designers, directors of studies, and so on) endorse a particular coursebook for their course and are dependent

exclusively on that coursebook **without making any adaptations** that consider their learners' preferred ways of learning, their cognitive capabilities etc. (this is represented by statement A). In other words, when a particular high-stakes test designates what is to be taught (statement C) and then only those specifications are taught in the ways that the test is organised (statement A), then that high-stakes test has **a significant (and quite negative) impact** on the teaching, learning, assessing and language use situation. As Hughes (1994) and Spratt (2005) stress, that impact will show in teachers' ways of teaching, its intensity, choice of coursebooks, and use of past-papers. It will also show in learners' practising of the target language and skills, their deciding what to study and what not to study, memorising, worrying, looking for shortcuts, or even cheating.

Finally, the last assumption (statement B) specifies that if something (a language structure or function, a particular genre, etc.) is not taught during the course, then learners will not be exposed to it at all and this will mean that they will not learn it. This is the worse case of **negative washback** certain high-stakes tests can have on teaching and learning.

On the other hand, if a test is based on **authentic communicative use of the language**, then the course that is linked to that test is bound to be **positively** impacted by the test. In other words, the authentic communicative use of the test will be reflected in the coursebooks used in the course, and this will further show in the way language will be taught and used in the foreign language classroom.

Activity 5

Before proceeding to the following Topic, carry out the following activity referring to the key concepts we have discussed so far. When you finish, compare your answer with ours, in the **Indicative Answers** section.

Complete the sentences below using the words in the box. There is one word you do not need.

standardized	assessment	positive	syllabus
negative	context	high- stakes	testing

- Teaching is usually based on a [_____] which specifies what exactly needs to taught.
- To acquire a certificate referring to your level of proficiency, you usually need to pass a [_____] test.
- Language [_____] implies assigning a precise mark or score depending on how well one can carry out a particular task.
- The communicative [_____] may influence the way we use the language.

- e) A [] test is a test whose results are considered to be particularly important and influencing.
- f) Ongoing language [] may offer valuable information about a learner's progress.
- g) A test may have a [] impact on teaching and learning when it focuses on authentic communicative use of the language.

8.2. Topic 2: Defining 'Test Impact'



There is only **one Task** in this Topic containing **four Activities** in total.

Task 1. What do we mean by 'test impact'?

In what ways are using, learning, teaching, assessing and testing English linked? We know from our experience as teachers that the above processes are linked.

Activity 1

Refer to your experience as a teacher and, previously, as a learner. Of all the five processes discussed above, **using, learning, teaching, assessing** and **testing** English, which of those were **more directly related to, dependent on or needed** in a classroom context? Why do you think this was the case?

Think about this for a moment and write your thoughts in the space provided here. When you are done, continue reading.

My thoughts:

In your answer to the above activity, it is possible that you have mentioned that, of all the processes involved in our experience with a foreign language, **testing** has the more **forceful effect** on teaching, learning, even use. Before we go on, let us clarify what we mean when we say 'forceful effect' or impact of tests.

Test impact refers to “any of the effects that a high-stakes test may have on individuals, policies, or practices – within a classroom, a school, an educational system or society as a whole” (Wall 1997: 291).

Activity 2



For this activity, watch a very useful Youtube video entitled “Video Shorts #5 Validity, reliability, impact and practicality” (Cambridge Assessment; Published on 3 Sep 2015).

The video can be accessed via the following link:

<https://www.youtube.com/watch?v=DsgORlf2IDg>

Listen to the speaker briefly define the “four cornerstones of good assessment – **validity**, **reliability**, **test impact** and **practicality**”. In particular, focus on **impact** (from 1:43 onwards). Then consider the following:

This definition of test impact makes it clear that, when we think of test impact, we almost invariably think of **high-stakes tests**. Why do you think this is so? What is so **special** about high-stakes tests that single them out as the ones mainly responsible for impacting, or having a “forceful effect” on language use, teaching and learning? Write your thoughts in the space provided here. When you are done, continue reading.

My thoughts:

Here are some of the thoughts you may have had, while carrying out the previous activity. High-stakes tests are statutory, they are important, they have a special role beyond the classroom, in society at large. They provide certification of competence in the target language, and that certification is often very instrumental in our lives—for example, it can help secure or improve a current job and boost a professional CV, since it proves how well we *know* a foreign language.

Activity 3

In this activity we are going to revisit Wall’s definition of impact. So far we have established it concerns high-stakes tests. Wall mentions that these tests impact ‘individuals, policies, or practices – within a classroom, a school, an educational system or society as a whole’. Read this

comment carefully and try to locate the **TWO** domains involving teaching, learning, assessing and use that are influenced by high-stakes tests.

When you are done, continue reading.

1.

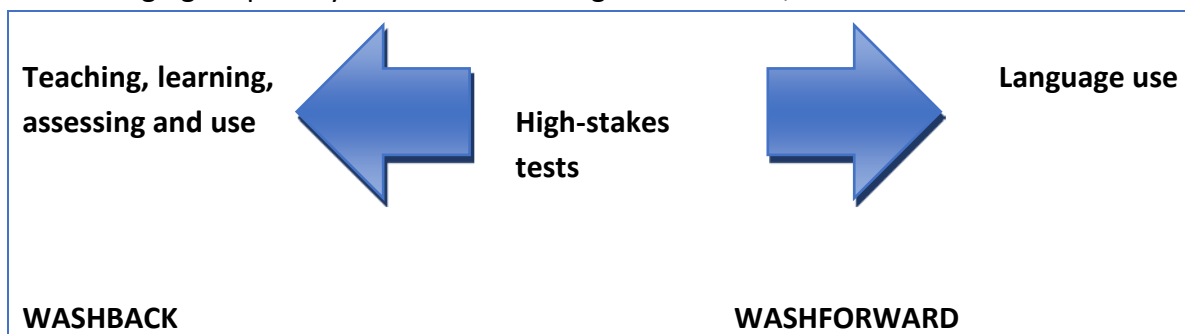
2.

What the above definition of test impact establishes about tests, apart from the fact that they almost always concern high-stakes tests, is that these tests influence the following two aspects:

- **Teaching the language in formal contexts**, i.e. contexts that involve explicit or overt teaching. High-stakes tests can impact the way learners are exposed to the structures and functions of the target language. This is often the result of decision-making at a policy level that involves the design and selection of coursebooks, the way teaching actually unfolds (e.g., an emphasis on testing rather than using and learning the language). Such decision-making is further influenced by established perspectives and attitudes in society in general, which identifies learning a language with having a high-stakes exam certification of proficiency in that language.
- **Using the language once certification has been awarded**. Since the pressure is on preparing learners for a particular high-stakes test, it follows that what these learners acquire is essentially the skills that are necessary to pass that test successfully. What remains to be seen is the extent to which these learners have in fact acquired the skills to use English successfully (i.e., comprehensibly) in different communicative settings some time after the test has been taken.

This means that test impact cuts both ways. First, it involves the effect (positive or negative) that high-stakes tests have on teaching, learning, assessing and use *before* they are taken – in this case, such impact is known as **washback**. And secondly, it involves the (positive or negative) effect that high-stakes tests have on language use some time after they are taken – in this case, such impact is known as **washforward**.

The following figure portrays the two effects high-stakes tests, as described above:

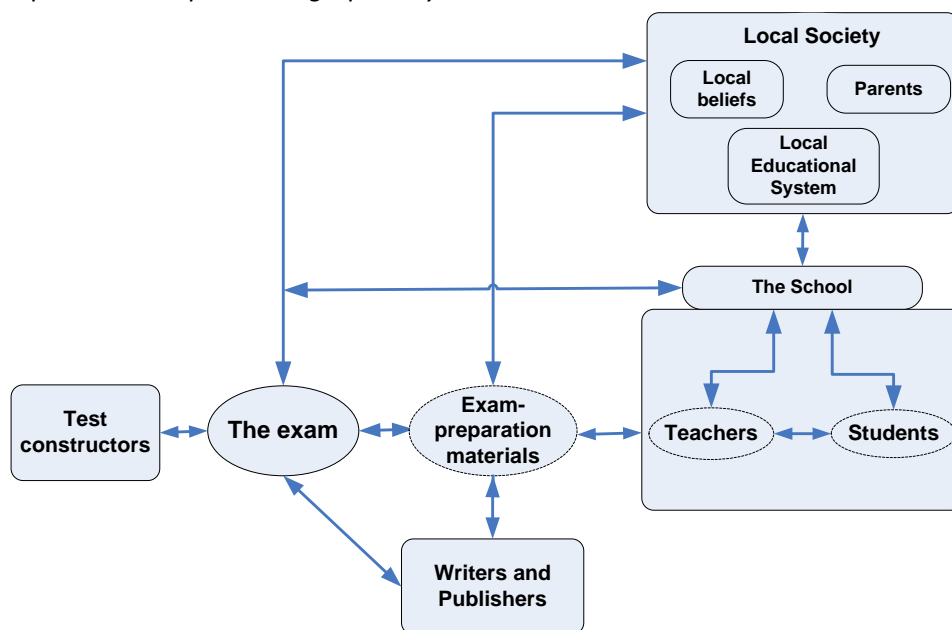


Activity 4

We should stress that the above orientation is a simplification of the real processes going on. The impact of tests is far more complicated. The following extract (from Tsagari, 2009: 221-222) and embedded graph helps further clarify this point, with particular regard to the nature of washback.

As you read, underline the various **parameters** involved. What strikes you as the **most important** and the **most unexpected**? In the space below, note to what extent your own experience agrees with the points made in this extract.

“The following graph illustrates the complex ecology of exam washback as delineated in the present study. In this model, washback is represented as an open loop process identifying the number of participants or stakeholders involved in the process and attempting to portray the relationship between them. However, washback as discussed so far is a complex phenomenon since the relationship among the participants involved is highly dynamic. Therefore, the model presented below is ‘simplified’ to make it possible to represent it graphically.



In this model, the nature of exam washback is circuitous and interactive. Exam washback is indirectly engineered on teaching and learning that takes place in the exam-preparation classroom through the understanding of the exam requirements by various intermediary participants. First and foremost, exam washback is mediated through commercial exam-preparation materials that are shaped by the perceptions of writers and publishers of the exam needs of teachers and students. The exam-preparation materials mediate between the exam intentions and the exam-preparation class. The teacher’s role is also crucial in the process as s/he mediates between material and students. Within this process, washback is also mediated by the school and strengthened by various other stakeholders operating in the wider local community, e.g. parents as well as by the local educational system and local beliefs about the exam and the language tested.”

Tsagari, 2009: 221-222

My thoughts:

What the above excerpt shows is that washback involves **test constructors, examination boards, and coursebook materials** (especially the way they interpret and specify the requirements of any exam). However, washback also involves the **teachers** themselves. In fact, teachers are at the core of every washback situation, since it is the teachers who will interpret, ‘filter,’ and eventually decide whether and to what extent to choose to teach aspects of a given coursebook. This also means that, irrespective of how much a coursebook is affected by washback, the teachers exercise a lot more influence on the content of teaching, and on the methods of teaching. It goes without saying that **learners** are also a major group of stakeholders too – they are at the receiving end (as is portrayed in the model above). Learners are also impacted and are capable of influencing the washback process too since this is reciprocal and essentially interactive.

8.3. Topic 3: Understanding ‘Washback’



There are **two Tasks** in this Topic containing **eight Activities** in total.

Task 1. Key dimensions of the washback effect

Let us consider what is involved in washback. As we have seen, **washback** means that tests can influence teaching and learning. In particular, tests can influence what and how teachers teach and what and how learners learn in formal classroom settings. Washback can be positive or negative.

Positive washback is encountered when test tasks “require the same authentic, interactive language use promoted in the classroom so that there is a match between what is taught and what is tested” (Weigle & Jensen 1997: 205). If a test has positive washback, “there is no difference between teaching the curriculum and teaching to the test” (ibid.). To offer a simple example, if we intend to teach speaking skills, we should test speaking skills and vice versa. Another example of positive washback: if tests demand authentic writing, then coursebooks integrate writing activities of that kind.

On the other hand, **negative washback** occurs when there is a discrepancy between the goals of instruction (as stated in the syllabus or curriculum) and the focus of testing. This often leads to neglecting these goals in favor of preparing for the test (i.e., teaching to the test). An example of negative washback would be a test consisting only of controlled writing (e.g., a dictation exercise or filling in blanks in a given paragraph). That would encourage the learning/teaching of linguistic accuracy rather than actual language.

Where do we see washback? As stated in the above diagram, we can locate washback (positive or negative) in every phase of the teaching and learning process. However, a key role in determining the type of washback is played by the **attitudes, expectations, and decisions** of stakeholders or participants in the teaching and learning process (e.g., teachers, learners, administrators, parents, publishers, materials developers).

Activity 1

The following example presents a typical situation of how the **attitudes** of certain stakeholders (in this case the **learners**), can significantly impact the extent to which a certain exam will have positive or negative washback on their learning. As you read, note your reflections on the following:

- In what way are **learners** (or students) to “blame” for the negative impact of tests on their learning? Can you locate **at which point** of their learning this happens?
- Is the example cited below similar to **your own experience** (as a teacher or learner)?

‘One difficulty related to washback arises because students often have two competing goals. Their long-term goal may be to improve their language proficiency, whether for work, immigration, education or social purposes. Their short-term goal, however, may be to pass a particular high-stakes test with satisfactory marks. These competing goals can lead to questionable behavioural choices. For example, in an intensive English programme I directed in the early 1990, the teachers would sometimes get frustrated because the week prior to the administration of the old paper-and-pencil TOEFL, students would frequently skip class to stay home and study high-level grammar rules and archaic vocabulary items, on the expectation that these elements of English would be tested’. (Bailey & Masuhara 2013: 304)

My thoughts:

As we have shown, **teachers' perceptions and attitudes** also play a key role in determining the impact of tests on teaching and learning.

Activity 2

Tsagari (2011) explored the relationship between the intended influence of the FCE (Cambridge ESOL) exam and teachers' perceptions towards the exam and their classroom practices. In what follows, we present some of the interviewed **teachers' perceptions** about a series of concerns. As you read, answer the questions in the spaces below.

"They expect everything from me. They feel I am the guru and they come to me for psychological help, for practical help for whatever you can imagine. But I think that it's a common attitude and I am not an exception to the rule. They expect everything from me. I don't like this. I try to cultivate a feeling of independence because that's what you are supposed to do and it's right. I don't think people should be dependent on someone else. I don't think this is a right attitude in general." (T1)

"We teach them rules, rules, rules... not the real learning of the language and communication" (T7)

"...grammar, that is Use of English, is the A and Z in FCE If you know Grammar well, you will definitely succeed. There is no way you will fail. You might not be able to speak well, but if you write a perfect grammar paper something will happen. They will pass you... It also gets most of the marks." (T11)

"I give tests following the types found in the exam. For example, when I do Use of English, I give them a short text to insert words, fill in the gaps, or transformations that are adjusted to the demands of the exam." (T10)

"[Students] want to get the Lower [the FCE certificate]. They are not interested in really learning the language They learn the language as if it is a paper that will help them find a better job." (T10)

"As a teacher you are accountable to your employers ... the frontistirio [foreign language centre]... the parents, to your students and to yourself. Maybe you shouldn't but that's how you assess your ability as a teacher, how successful you will be, so, there is... this becomes... your objective but it creates tension, stress and all this." (T13)

"Parents put a lot of pressure on their children to take up English when they are very young ... as a consequence, if there is some sort of inclination, then the child learns... if not then this becomes extremely stressful for the child." (T14)

1. Have you ever had thoughts and feelings similar to those expressed by the teachers above?

2. In what ways do those teachers' expressed understanding of their learners' and other stakeholders' expectations impact their teaching? Is that a positive or negative influence? Why do you think this is the case?

3. In the same way, how does those teachers' understanding and interpretation of the exam's demands and priorities impact the way that they teach?

Apart from teachers' perceptions, **learners' expectations, experiences, and attitudes** towards the importance of high-stakes exams (i.e., preparing for them and getting the certificate at the end) and the **teachers' role** (e.g., way of teaching, providing **feedback**, etc.) are also essential when considering the impact of tests on teaching and learning (washback).

Activity 3

In this activity, we will look at **learners' diaries** from a study carried out by Tzagari (2007: chapter 6). As you read, try to answer the questions below.

"You know sth? I think that I'll be ill at the end of this term. I'll be a computer, trouly! With all these homework and exercises... Please, answer me! I am going to be crazy!" (D5/F- 292:297)

"I'm starting. I feel quite nervous. I've finished the 2nd exercise but I haven't checked it yet. I'll check all of them at the end. I feel very stressed. I have some unknown words. What's "thorough"? I hope I get more than 40% in the test. (I ask too many? Don't I?) I don't know what else to say. I feel like stopping and going straight to home to watch TV. I must say that I find the test more difficult than I expected. My stomach hurts from stress. I've finished. I feel quite disapointed because I believe that I haven't written very well" (D22/F-103:109)

"Tests, compositions, Grammar, reading, listening: all these are useful for FCE, yes, but they want us to practice more and more. They say 'The more you practice, the better for you. You'll pass easier.' Yes, I agree but not with the exaggerating situations. What do you think?" (D5/F-302:305)

"The lesson was very boring and we all waiting the clock to go at 12:00 so we could leave. Today the only thing we done at English was homework and homework and homework. Nothing else. We all sleeping at the lesson and the only thing we wanted was the brake" (D21/F-302:305)

"Also I heard that we will take a test next Thursday so I was ready to said that I hate to write tests and that I will never take and write a test but I didn't do it because then I will understand that I said a very stupid thing because if I want to take the FCE I must to try a lot and very hard and also to do this thing I will pass a lot of happy and bad things and I don't to give up because it has no sense" (D17/M-197:202)

"The only thing I didn't like to the new teacher for us was when she said that if somebody has not the homeworks she will send him to the office [of the school owner]" (D1/M-150:152)

"The lesson was really good. We first did dictation like every Tuesday. It was very easy and I wrote all corrects" (D19/F-104:105)

"At the end Miss Nelly told us the marks from the Thursdays test. I take 85 per cent. I was very happy and this mark gave me the confidence to continue like that" (D21/F-190:192)

"Today we wrote a test at English lesson. It was difficult but I believe that I wrote well. I had many stress about the test and when I started writing I tried to focus to find the correct answers. I was studing for the test all day. When I finished the test I felt like the prisoners when they are released" (D21/F-170:173)

"This day for me was very difficult because we wrote a big test in two units. In all the lesson and during the test I felt embarrassed and sad because I thought that maybe I hadn't write good. Also I said to myself to have patient so to write very good" (D3/M-77:80)

Based on the learners' reflections, in what ways has the high-stakes test they are preparing for influenced:

a) their learning?

b) their motivation for learning?

c) their evaluation of teachers' instructional decisions?

d) the role of the textbook in this whole process?

Let us now consider other areas where **washback** is located. The following activity exemplifies an interesting case.

Activity 4

In a study comparing TOEFL preparation classes with non-TOEFL classes, Alderson & Hamp-Lyons found the following (1996: 288-9):

- Test-taking is much more common in TOEFL classes.
- TOEFL classes are somewhat more routinized.
- There is much more *laughter* in non-TOEFL classes.
- Teachers *talk more* and students have less time to talk in TOEFL classes.
- There is less *turn-taking* and turns are somewhat longer in TOEFL classes.
- Much less time is spent on *pair work* [in TOEFL classes].
- The TOEFL is referred to much more in TOEFL classes.
- *Metalanguage* is used much more in TOEFL classes.

Consider the above observations. What **conclusions** can you draw from them? How is **washback** situated in this study? Write your thoughts in the space below and when you finish, compare your answer with that provided in the **Indicative Answers** section.

My thoughts:

Task 2. Understanding washback in practice

Now, let us focus on what may constitute positive and negative washback.

Activity 5



Visit the website of the Center of Applied Linguistics (CAL) which contains a tutorial entitled ***Understanding Assessment: A Guide for Foreign Language Educators***. Then, go to the **Impact** section and click on **Washback and instruction**. Read the passage and focus on the questions below. The link of the website is the following:

<http://www.cal.org/flad/tutorial/index.html>

1. What is positive feedback?

2. What is negative feedback?

3. How can positive feedback be reinforced?

As a case in point, consider the following **two examples**. In the empty space, state whether these are instances of **positive** or **negative** washback. Why do you think this is so? When you finish, go to the **Indicative Answers** section to read our own comments.

Example 1

The Eurokom exam is an oral exam in one federal state of Germany that was introduced by the local Ministry of Education in 2004. The oral exam is part of the school-leaving exam at age 15 or 16 (equivalent to GCSE exams in the UK) and consists of a production part and an interaction part. The washback was planned (called top-down washback) because the teachers were not using enough oral activities in the EFL classroom. What happened in the years after the exam was introduced was an increase in oral activities, e.g. of presentations that represent one test task, in the classroom. There is still a noticeable amount of oral activities in the regular EFL classroom to date.

Source: Fröhlich, V. (2010). *Washback of an Oral Exam on Teaching and Learning in Germany. The EuroCom in Realschulen in Baden-Württemberg*. Saarbrücken: Verlag Dr. Müller.

Example 2

Oakdale Middle School has a foreign language program offering Spanish and French through level two. Members of the foreign language department were proud of their program but frustrated with students' lack of effort in their classes. The teachers felt many students did not take their foreign language classes seriously; since they weren't studying and practicing much, the teachers knew they were not working up to their potential. During the past year, the teachers began to implement authentic assessments of oral language, reading, and writing. These standards-based assessments, which were given every six weeks, involved the students in real-life tasks.

The department chair sent home parent surveys and found that students were spending more time working on their foreign language skills at home and that they were talking more about their foreign language classes with parents. Parents and students also liked the feedback that they got from the regular assessments. One student commented in an interview that knowing what was going to be on the assessments helped him to prepare effectively and that the results helped him know what to do to improve. At the end of the year, 75% of eighth grade language students placed into level 2 for high school. Overall, the foreign language teachers felt that student performance was improving and they attributed this to the new assessment program. Moreover, the district's foreign language supervisor recognized the school's efforts and was able to find additional funding for curriculum and assessment development the next year.

Source: *Understanding Assessment: A Guide for Foreign Language Educators*. Center of Applied Linguistics (CAL), <http://www.cal.org/flad/tutorial/impact/5exwashback.html>

Thinking about positive washback can be a significant tool in policy making as well. To see how, look at the **case study** presented below.

Crystal is the chair of the combined foreign language departments of three elementary schools in her town. The goal of the language programs in the elementary schools is for students to successfully move into the intermediate levels of the language in middle school, rather than starting at a novice level in 7th grade. Recently, teachers have become concerned about whether 6th graders have high enough levels of language proficiency to move into the intermediate levels the following year. At the most recent meeting of the departments, Crystal and the other teachers discussed instituting a test that the 6th graders will have to pass to place into the intermediate level of their language in 7th grade. Teachers brought up a variety of concerns about the impact of a possible exam.

Source: *Understanding Assessment: A Guide for Foreign Language Educators*. Center of Applied Linguistics (CAL)

Activity 6



Now think about the various stakeholders that would be influenced by positive washback practices. Go to the **Impact** section of the *Understanding Assessment: A Guide for Foreign Language Educators* tutorial (see the instructions mentioned earlier) and click this time on **Stakeholders** to see the “**Stakeholders map**” presented there. Then, go back to the **examples** presented in Activity 5. **Which of these stakeholders are impacted and in what ways? Write your reflections in the space below.**

My thoughts:

Now let’s consider another side of washback, namely, **negative washback**. The problem here begins when teachers and coursebooks use activities that follow the exact testing specifications of the target high-stakes exam and only these. This means that learners are not exposed to other/additional (probably authentic) language use and tasks and are only limited to tasks that follow very closely the format of the final high-stakes exam.

Activity 7

Consider the following case in point. In this case, we will focus on the *Reading and Use of English* paper of the *Cambridge English: First (FCE)* by Cambridge English Language Assessment and consider a coursebook excerpt that attempts to prepare learners for this particular exam. Do the following:



First of all, go to the **Cambridge Assessment English** website. The link is the following:

<http://www.cambridgeenglish.org/>

Put your mouse over **Exams and tests** on the menu on top and click on **First (FCE)**. Then, from the menu on the left, click on **Exam format**. Scroll down the page and read the **specifications** of the “**Reading and Use of English**” paper. Then, to see how the sample paper is structured, click on “**See sample paper**” at the top of the page and, when you download the zip file, open the “cambridge-english-first-2015-sample-paper-1-reading-and-use-of-english v2” pdf file. This page can also be accessed via the following link:

<http://www.cambridgeenglish.org/exams-and-tests/first/exam-format/>



Now let us see some **sample pages of a Greek coursebook** on the same exam. Go to the website of **Grivas Publications**. The link is the following: <https://www.grivas.gr/>

Click on the **Menu** at the top right corner and then click on **Sample Pages**. Select the **Sample Pages for FCE Exams** from the list and open the file entitled “**Reading & Use of English (New Format 2015)**”. In particular, see the first few pages of **Unit 1**. The sample pages of this coursebook can also be accessed via the following link:

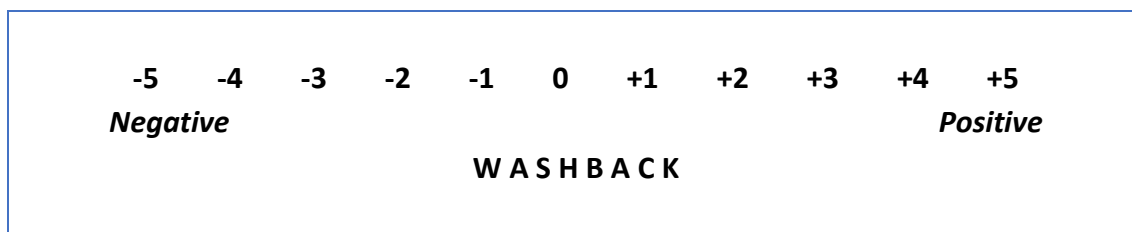
https://www.grivas.gr/assets/files/samples/fce_exams/ReadingUseEnglish_FCE_2015.pdf

What do you notice? **How are the coursebook activities linked to the final FCE test?** To that extent are the coursebook activities interested in exposing learners to **authentic language use**? In what ways are the **language structures and functions** introduced? Why is this coursebook Unit an example of **positive / negative washback**?

Note down your thoughts and when you finish, see our comments in the **Indicative Answers** section.

My thoughts:

As we have seen above, washback can be positive or negative. But this is not a black-and-white, either-or situation. High-stakes tests can have different degrees of influences on coursebook activities. The best way to describe this is to use a **continuum** from completely negative on the one hand to completely positive washback on the other, as follows:



It should be possible to evaluate the extent to which certain coursebook activities portray negative or positive washback in relation to a particular exam. You are now invited to consider the extent to which washback is positive or negative by reflecting on your own teaching context. To help you in this, do the following activity.

Activity 8

If you are teaching a class that prepares learners to take an exam (in particular a high-stakes exam) and would like to find out about the extent to which your teaching is impacted positively or negatively by that exam, follow the steps below:

1. Go to the **testing specifications of a particular paper or section of that exam**. High-stakes exams have websites that offer a lot of information on every conceivable aspect of that exam, e.g., what the exam is supposed to test, how the examination system works, what types of tasks are to be expected, and so on. Read this information but focus on a particular paper (that tests a particular skill or language function), e.g., reading, listening, speaking, writing, use of English or grammar, etc.
2. Go to a **sample paper** of the above exam (as we did in activity 15). Notice how the specifications for the particular paper (skill or function) are reflected in the sample paper.
3. Now go to your **coursebook**. Select a section that focuses on a similar skill or language function.
4. In order to establish whether the coursebook is positively or negatively impacted by the target exam, focus on the following: Does the coursebook focus on teaching language or on testing learners' performance by adopting **the same framework** as the one used by the target test? See the example offered in the previous Activity and use the **continuum** discussed earlier. If you have the time, do the same for other skills, subskills or language functions.

8.4. Topic 4: Broader Consequences of Washback on Teaching and Learning



There are **two Tasks** in this Topic containing **three Activities** in total.

Task 1. The culture of testing

Apart from the ways that specific tests positively or negatively impact teaching and learning, through positive or negative washback, we should also consider the ways in which the whole **culture of testing** (or **testing culture**) can impact, positively or negatively, the attitudes and expectations of different stakeholders.

Activity 1

What factors influence washback?

Spratt (2005) mentions a series of **factors** influencing washback. She groups these factors into different **categories**. In the table that follows, **match** the factors against these categories. When you finish compare your answers with ours, in the **Indicative Answers** section.

Categories		
a) The exam	c) Teachers education and training	e) School
b) Teachers' beliefs & attitudes	d) Resources	
Factors influencing washback		Category
1. Teachers' educational experience		
2. When the exam was introduced		
3. Availability of customised materials and exam support materials such as exam specifications		
4. How much the administrators put pressure on teachers to achieve results		
5. What teachers think about the relationship between the exam and the coursebook		
6. The weighting of individual exam papers (corresponding to different skills and micro-skills)		
7. Understanding of the exam's rationale or philosophy		
8. The amount of time and number of learners in exam classes		
9. What teachers consider effective teaching methods		
10. The "importance" of the exam, its purpose and proximity		
11. The formats employed by the exam		
12. Training in implementing exam-related coursebooks		
13. Reliability and fairness of the exam, its the stakes and usefulness		
14. Cultural factors such as learning traditions		
15. Exam familiarity		

16. The school itself, its atmosphere	
17. Types of coursebooks available	
18. The extent to which the exam contravenes teachers' current teaching practices and their learners' beliefs and expectations	
19. Basic education, ongoing training and more specific training in teaching towards specific exams	

What are the possible **sources of negative washback**? Alderson & Wall (1993) carried out a very interesting study in Sri Lanka. They interviewed teachers and concluded, among other things, that:

- A significant number of teachers do not understand the philosophy or approach of the textbook that they use.
- Many teachers have not received adequate training.
- Many teachers claim to be unable, or feel inadequate, to implement the required methodology.
- Many teachers believe that there are factors in their teaching that prevent them from teaching as they should.
- They claim to be unfamiliar with the nature of the exam.
- They may never have received or, if they have, read the official exam support documents.

So far, we have seen that **coursebooks** that are used to prepare learners for a particular high-stakes exam can be (negatively or positively) influenced by the specifications of that test. This influence is then bound to reflect on the way that teachers will use that coursebook in their classroom.

However, do you think that **any coursebook**, even those that are not overtly linked to specific high-stakes tests, can somehow be **impacted by tests**? And, if they are, to what extent can such impact be **positive or negative**?

Activity 2

Think carefully about these two questions. It would be good, while you do, to consider specific coursebooks that you have been and/or currently using in your own teaching context. Ask yourselves the following questions:

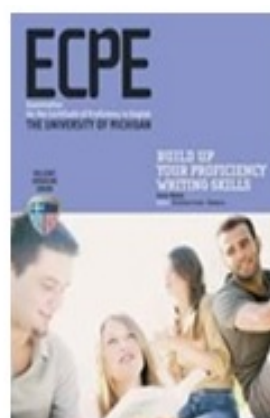
1. How closely linked is this coursebook to a (high-stakes or not) test?
2. How is this link made obvious?
3. If the coursebook is not *overtly* linked to a test, to what extent can it be argued that it is *covertly* linked to some final target test?

4. Assuming there is a covert link between coursebook and test, what impact (or washback) does that test have on the coursebook? How is this impact evident in your view?

Write your thoughts in the space below. Try to give as **specific and detailed information** as you can, this is vital. When you finish, please continue reading.

My thoughts:

Let us try to address the above questions in turn. With regard to the first two questions, it is obvious that not all coursebooks are overtly linked to a particular high-stakes test. Those that are overtly linked use this link as key marketing information, by placing that information on the cover of the coursebook. This is the case, for example, with the following coursebook covers:



The question here, of course, is to try and consider the washback level of the content of these coursebooks. However, it is important to note that, as we have seen, it's not only those coursebooks that *link* to a particular high-stakes test that are influenced by tests. It is perfectly possible for so-called general-purpose coursebooks to also be impacted. In this case, the impact is further prompted by other stakeholders (e.g., teachers) whose broader 'testing culture' determines specific "epistemological suppositions about the dynamic nature of knowledge, as well as assumptions about students, teaching and learning" (Ofra 2008: 387).

Activity 3

We have reached a critical point in our endeavour to establish the washback level of certain coursebooks. **In the list below, identify those types of activities that are a) reminiscent of testing and those types of activities that are b) reminiscent of learning and using language communicatively.**

When you finish compare your answers with ours, in the **Indicative Answers** section. Then, continue reading.

<i>Types of activities</i>	<i>a) or b)?</i>
1. Selecting correct answer	
2. Summarising	
3. Note-making	
4. Checking comprehension	
5. Paraphrasing	
6. Transferring visual information verbally	
7. Guessing	
8. Translating/mediating	
9. Selecting meaningful parts of text	
10. Multiple choice	
11. Multiple matching	
12. Inferring	
13. Cloze (words deleted at regular intervals)	
14. Negotiating	

Task 2. Where to go from here?




This Course offered you a brief introduction to the different ways in which tests can impact teaching and learning. What can you do now to move further along the lines presented above?

- You are prompted to **consider examples from your own coursebooks** and establish:
 - The **washback effect of specific high-stakes tests on your coursebook** (if you're preparing your learners for a particular test)—use the washback continuum offered above.
 - The extent to which your coursebook is impacted by the **broader testing culture** (if your context is not overtly linked to a particular exam)—use the test culture impact continuum offered above.

- Tests can have a negative or positive impact on other things as well, as we have seen. They can impact **policies**, **curricula**, **syllabi**, and **stakeholders’ attitudes**. You are therefore prompted to do the following:
 - Look at the **curriculum** of your teaching situation. To what extent do you see elements that are reminiscent of testing? To what extent is the focus on testing rather than on learning?
 - Consider the more specific details of the **syllabus** (i.e. the day-to-day planner) of your teaching context. Again, to what extent is the focus on testing rather than on learning?
 - Enquire your **learners** and their **parents** or **sponsors** about their expectations from your course—to what extent do they ‘demand’ that the course is more like a test rather than designed for language use?
- Finally, you are invited to consider the articles and papers in the **Further Reading** section. There, you will find more specific information on aspects of washback that might interest you.

Test Impact - Outcomes

Reflect on whether you feel that the expected outcomes of this Course have been achieved.

My Outcomes			
I can distinguish between different ways tests impact language learning, teaching, and use.			
I can identify positive and negative washback effects of high-stakes tests on teaching and learning.			
I can identify positive and negative washback effects of high-stakes tests on coursebooks.			
I can demonstrate the essential differences (and similarities) between using, learning, teaching, assessing, and testing English as a foreign language.			

Indicative Answers

Topic 1: Activity 1

Here are some suggestions of ours. Remember, this activity refers to your own personal experience, and your answers are likely to be different than ours.

	Things I liked about tests	Things I disliked about tests
As a learner	<ul style="list-style-type: none">• They gave me a goal to aim for.• I got very specific feedback on my progress.	<ul style="list-style-type: none">• The more statutory exams were quite stressful at times, especially as the exam day approached.• I was too much involved in ticking boxes and matching answers rather than doing some more “authentic” work with language.
As a teacher	<ul style="list-style-type: none">• They usually act as motivation for my learners to take language learning seriously.• I can use learners’ performance in tests as a gauging mechanism to report their progress and discuss it with parents.• There is a vast selection of coursebooks I can use as preparation material for a particular “important” exam.	<ul style="list-style-type: none">• I find it difficult to design my own tests, which is why I’d often adopt ready-made tests I find online or in other coursebooks or collections of past papers of “important” exams.• My foreign language school director of studies is quite strict about using specific coursebooks and teaching them in specific ways – I don’t feel I am allowed to deviate from them.

Topic 1: Activity 4

The logical order of the above statements is as follows:

- C. “What is tested should be taught”
- A. “Only what is tested is taught”
- B. “What is not taught is not learnt”

Topic 1: Activity 5

a) syllabus, b) standardized, c) testing, d) context, e) high-stakes, f) assessment, g) positive

Topic 3: Activity 4

As the authors claim in the abstract of this paper:

“The study reported in this article set out to investigate common claims that the TOEFL exerts an undesirable influence on language teaching. Our data consist of interviews with teachers and students, and observations made of two sorts of classes: 'normal' language proficiency classes, and parallel classes intended for students preparing to take the TOEFL. We observed both TOEFL preparation classes and non-TOEFL preparation classes by the same teachers in order to be able to separate washback from the TOEFL from any possible effect of individual teacher style. This study leads us to suggest that simple forms of washback hypotheses are too naive: influences on what happens in class are much more complex than unexamined beliefs about washback allow. The results of the study suggest the need for more complex hypotheses about washback.”

Topic 3: Activity 5

Example 1 is an example of positive washback. The actual test reinforces the speaking skill (“production” and “interaction”) and this has an effect on teaching practice. After the test was introduced, teachers were prompted to focus more on teaching various speaking sub-skills (e.g., presentations), and this meant learners becoming more familiar with this particular aspect of language use.

Example 2 is also a case of positive washback. It shows how simply designing a test that assesses an area that learners consider unimportant (such as foreign language classes in this case) can boost learner interest in and appreciation of these classes.

Topic 3: Activity 7

The coursebook begins with an entirely theory-driven presentation of the tenses, on p. 4. This is followed up by a series of drills that follow the right/wrong format and often closely resemble the actual rubrics of the FCE test. To the extent that the coursebook aims at preparing learners for the FCE exam, it presents **positive washback**. However, as regards learning and using the language, the test impacts the coursebook negatively (**negative washback**) in that the latter follows the test specifications without caring for the communicative teaching of the target structures and functions but only for exposing learners to these issues through the format of the final test. Clearly, it is up to the teacher to adapt and use the coursebook materials appropriately.

Topic 4: Activity 1

- a) The exam: 2, 6, 10, 11, 13
- b) Teachers' beliefs & attitudes: 4, 5, 9, 18
- c) Teachers education and training: 1, 7, 12, 15, 19
- d) Resources: 3, 17
- e) School: 4, 8, 14, 16

According to Spratt (2005), one factor that influences washback is the exam itself, how “important” or high-stakes it is, its purpose and proximity, the formats it employs, the weighting of individual papers (corresponding to different skills and micro-skills), when the exam was introduced, and how familiar it is to teachers.

Another factor is teacher beliefs about the **reliability** and **fairness** of the exam, its stakes and usefulness, the relationship between the exam and the coursebook, what they consider effective teaching methods, the extent to which the exam contravenes their current teaching practices, and their learners’ beliefs and expectations.

Yet another factor refers to teachers’ educational experience, access to and familiarity with exam support materials (such as exam specifications) and understanding of the exam’s rationale or philosophy, as well as their basic education, ongoing training and more specific training in teaching towards specific exams (usually organized by exam bodies themselves), and in implementing exam-related coursebooks.

Finally, Spratt refers to what she terms ‘resources’: the availability of customised materials and exam support materials such as exam specifications, the types of coursebooks available, the school itself, its atmosphere, how much the administrators put pressure on teachers to achieve results, the amount of time and number of learners in exam classes, and cultural factors such as learning traditions.

Topic 4: Activity 3

- a) Activity that tests the language: 1, 4, 10, 11, 13,
- b) Activity that helps us learn and use language communicatively: 2, 3, 5, 6, 7, 8, 9, 12, 14

Further Reading

Bedford, J. (2003). Washback – the effect of assessment on ESOL teaching and learning. *Many Voices*, 21, 19-23. Available from: <http://englishonline.tki.org.nz/Media/Files/WXYZ-Files/Washback-the-Effect-of-Assessment-on-ESOL-Teaching-and-Learning-PDF>

British Council India. (2015, April 9). *TEC15: Stephen Carey – Learning to test or testing to learn? Washback and the learner*. [Video file]. Available from: <https://www.youtube.com/watch?v=8f6XfarC-bU>

Elshawa, N. R. M., Heng, C. S., Abdullah, A. N. and Rashid, S. Md. (2016). Teachers' assessment literacy and washback effect of assessment. *International Journal of Applied Linguistics & English Literature*, 5(4), 135-141. Available from: <http://www.journals.aiac.org.au/index.php/IJALEL/article/view/2391/2080>

Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39-51. Available from: <http://revistas.um.es/ijes/article/view/ijes.13.2.185891/153741>

- Pan, Y. (2014). Learner washback variability in standardized exit tests. *The Electronic Journal for English as a Second Language*, 18(2), 1-30. Available from: <http://tesl-ej.org/pdf/ej70/a2.pdf>
- Tsagari, D. (2007). 'Review of washback in language testing: What has been done? What more needs doing? ERIC Document Reproduction Services No. ED 497709. Washington, DC: Center for Applied Linguistics. Available from: <https://eric.ed.gov/?id=ED497709>
- Tsagari, D. (2009). Revisiting the concept of test washback: investigating the FCE in Greek language schools. *Research Notes*, 35, 5-10. Available from: <http://www.cambridgeenglish.org/images/23154-research-notes-35.pdf>
- Tsagari, D. (2009). *The Complexity of Test Washback: An Empirical Study*. Frankfurt am Main: Peter Lang GmbH.
- Tsagari, D. (2012). FCE-exam preparation discourses: insights from an ethnographic study. *Research Notes*, 47, 36-47. Available from: <http://www.cambridgeenglish.org/images/22669-rv-research-notes-47.pdf>
- Tsagari, D. (2014). Investigating the face validity of Cambridge English First in the Cypriot context. *Research Notes*, 57, 23-31. Available from: <http://www.cambridgeenglish.org/images/177881-research-notes-57-document.pdf>
- Tsagari, D. and L. Cheng (2016). 'Washback, Impact and Consequences'. In E. Shohamy and N. H. Hornberger (eds.), *Encyclopedia of Language and Education*, 3rd Edition, Volume 7: Language Testing and Assessment, pp. 1-13. Educational Linguistics: Springer.

Finally, if you wish to explore further ideas on how to **assess SpLD students**, visit the following website of the DysTEFL project and follow the suggestions in Chapter 10 (Assessment)



<http://dystefl.eu/index.php?id=55&L=2%2Fimages%2Fstories%2F3xp.php>

or download the DysTEFL training materials and work directly from there:

If you are a trainer: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainer.pdf

If you are a teacher: http://dystefl.eu/uploads/media/DysTEFL_Booklet_Trainee.pdf

You might also like to read:

- Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning differences. Bristol: Multilingual Matters. Chapter 8. Assessment
- Tsagari D. & G. Spanoudis (Eds.) (2013). Assessing L2 students with learning and other disabilities. Cambridge: Cambridge Scholars Publishers.
- Tsagari, D. & I. Sperling (2017). 'Assessing SLLs with SpLDs: Challenges and Opportunities for Equity in Education'. In Szymańska-Czaplak, E. (ed.), *At the Crossroads: Challenges of Foreign Language Learning*, Series: Second language learning and teaching. Springer. pp 175-188. DOI: 10.1007/978-3-319-55155-5_10.

References

- Alderson, J. C. & Hamp-Lyons, L. (1996). TOEFL preparation courses: a study of washback. *Language Testing*, 13(3), 280-287.
- Alderson, J. C. & Wall, D. (1993). Does washback exist?. *Applied Linguistics*, 14(2), 115-129.

- Bailey, K. M. & Masuhara, H. (2013). Language testing washback: the role of materials. In B. Tomlinson (Ed.), *Applied Linguistics and Materials Development*. London: Bloomsbury, 303-318.
- Hughes, A. (1994). *Backwash and TOEFL 2000*. Commissioned by Educational Testing Service (ETS). University of Reading.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402.
- Madaus, G. F. (1988). The distortion of teaching and testing: high-stakes testing and instruction. *Peabody Journal of Education*, 65(3), 29-46.
- Spratt, M. (2005). Washback and the classroom: the implications for teaching and learning of studies of washback from exams.' *Language Teaching Research*, 9(1), 5-29.
- Tsagari, D. (2007). *Investigating the washback effect of a high-stakes EFL exam in the Greek context: Participants' perceptions, material design and classroom applications*. PhD in Linguistics, Department of Linguistics and English Language, Lancaster University, UK.
- Tsagari, D. (2009). *The complexity of test washback: An empirical study*. Frankfurt am Main: Peter Lang GmbH.
- Tsagari, D. (2011). Washback of a high-stakes English exam on teachers' perceptions and practices. In E. Kiti, N. Lavidas, N. Topintzi & A. Tsangalidis (Eds.), *Selected Papers from the 19th International Symposium on Theoretical and Applied Linguistics*, Thessaloniki: Monochromia, 431-445.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham and D. Corson (Eds.), *Encyclopaedia of language and education. Vol. 7. Language testing and assessment*. Dordrecht, Kluwer Academic, 291-302.
- Weigle, S.C. & Jensen, L. (1997). Assessment issues for content-based instruction. In M. A. Snow & D. Brinton (Eds.), *The content-based classroom: Perspectives on integrating language and content*. White Plains, NY: Addison Wesley Longman, 201-212.

Glossary

A

Assessment criteria

A criterion is a defining characteristic of something, and a standard by which it can be judged or decided. In language assessment we use more than one criterion because no language performance can be properly described with only one criterion. The criteria by which we make judgements of language performances are central to both validity and reliability .

Assessment for Learning

Type of assessment that is process-oriented and learning-oriented. It tends to be formative in nature.

Assessment of Learning

Type of assessment that is product-oriented and focuses on the learning outcome. It tends to be summative in nature.

B

Benchmarks

After detailed individual and group evaluation of a learner's speaking performance, expert judges agree on the learner's scores for all the assessment criteria.

Bottom-up processes

Processing the visual input in the text, such as letters, words, sentences, to check their predictions.

C

Checklist

Type of scoring rubrics that allows for a yes or no decision on certain criteria of a performance.

Conference

Formative assessment method that includes teacher assessment and self-assessment in a talk about learning products or processes

D

Deducing meaning from context

It refers to “reading between the lines” and/or working out the meaning from unfamiliar words

E

European Language Portfolio

A kind of portfolio designed for language learning that includes a language passport, a language biography, and a dossier.

Extensive listening

Understanding longer listening texts and connecting the ideas

Extracting detailed information

Getting information from the text, which is more detailed than individual facts, and can often involve attitudes.

Extracting specific information

It refers to reading for specific information, also referred to as ‘scanning’.

F

Fairness

A ‘fair’ assessment is one which provides everyone who is being assessed with a complete opportunity to demonstrate their knowledge and skills; fair tests/assessments make allowances for differences among individuals; and fair tests have ethical safeguards to prevent them being misused (see for example: Hamp-Lyons 2000; Song 2014; Suskie 2000).

Feedback

The purpose of feedback is always to improve a person’s performance of a skill. Good feedback helps students to understand their own performance and gives them clear guidance on how to improve their learning. Effective feedback includes useful suggestions that are doable and which learners are able to understand. Feedback is not limited to the end of a task, activity, or unit.

Feedforward

Feedback focuses on a current task; feedforward looks ahead toward the next task. Feed forward offers constructive guidance on how to improve.

Formative assessment

Ongoing, process-based type of assessment that feeds into learning.

G

Getting the general picture

By reading the text quickly, readers try to find out the main points or ideas in a text, without being concerned about details. It is often called skimming.

Global questions

These questions are phrased in general terms, have quite a wide scope, thus leaving the learner a lot of room for the answer. As opposed to this question type, we also distinguish specific questions that have a narrow focus. Examples: Why do you think young people use social media? (global question) vs. Which functions of social media are the most popular with young people? (specific question).

Group oral

In order to maximize the efficiency of examining learners orally, 3 (or more) of them are given a discussion activity and their performances are scored simultaneously.

H

High-stakes tests

“Tests whose results are seen – rightly or wrongly – by students, teachers, administrators, parents or the general public, as being used to make important decisions that immediately and directly affect them” (Madaus 1988, p. 87).

I

Individual long turn

The learner has to give an extended performance or sustained monologue based on some kind visual or verbal prompt.

Individual mode

The learner has one partner to interact with, the examiner/teacher.

Intensive listening

Listening that involves focusing on the form of the language

Interactive listening

Combining listening with speaking – and sometimes other skills as well

Interlocutor

The examiner/teacher who is responsible to get a ratable language sample from the learner in an oral exam.

L

Language assessment

The practice of evaluating the extent to which learning and teaching have been successful, focusing on what learners can do with the language, their strengths rather than their weaknesses.

Language learning

The process (and result of) grasping, or acquiring, of the knowledge and skills necessary in order to produce discourse that is meaningful. As a process, learning can take place everywhere and anywhere, at any time and any place. Learning happens inside and outside the classroom, and can be a conscious or a subconscious process.

Language teaching

The 'steering' of the process of learning by the teacher in a classroom setting. In the formal sense of the term, teaching is based on an underlying plan (or syllabus) that addresses a particular type of learner or learners, is followed for a specific time and should lead to a specific desirable outcome.

Language testing

The precise measuring of different aspects of learners' performance.

Language use

The generation of meaningful, spoken or written discourse that is produced for a communicative reason and is comprehensible to a specific audience.

N

Negative washback

Negative washback occurs when there is a discrepancy between the goals of instruction (as stated in the syllabus or curriculum) and the focus of testing. This often leads to neglecting these goals in favor of preparing for the test (i.e., teaching to the test).

P

Paired mode

Two learners are paired up and given one task to carry out. They may be given individual long turn tasks in this format too but then they have to react to the partner's contributions in some ways.

Peer-assessment

The involvement of other learners (peers) in assessment procedures.

Portfolio assessment

Assessment based on a collection of a learner's work that indicates his / her progress in connection with curricular objectives.

Positive washback

Positive washback is encountered when test tasks "require the same authentic, interactive language use promoted in the classroom so that there is a match between what is taught and what is tested" (Weigle & Jensen 1997: 205). If a test has positive washback, "there is no difference between teaching the curriculum and teaching to the test" (ibid.). To offer a simple example, if we intend to teach speaking skills, we should test speaking skills, and vice versa.

Practicality

The extent to which an assessment can continue to be used, given the resources available (such as funding, time, expertise, and people).

Productive tasks

These tasks require learners to produce language of their own, eliciting direct performance from test takers. Task types include, for instance, short answer questions, gap-filling tasks or summary writing.

R

Rating and raters

The processes of scoring or decision-making, and the people who make these decisions

Rating scales

Instruments used to judge writing, comprising a set of domains/criteria and a series of levels of proficiency

Reactive listening

Listening that involves focusing just on the sounds heard

Receptive tasks

These tasks require L2 learners to choose from a set of options, they do not need to produce language on their own, such as multiple-choice questions, true/false statements or matching tasks.

Reliability

The consistency of assessment scores or results

Responsive listening

Listening and showing through words or actions that one has understood

Rubric

A rubric is a set of criteria specifying the desired outcomes of performance of a task, and the levels of achievement that may be seen in each criterion (see also criteria/criterion). A rubric is developed through an empirical process of making judgements and evaluating the qualities of those judgements.

S

Score reporting

Converting scores made on learners' writing into a format that can be delivered to a formal authority

Selective listening

Listening and picking out important information from the input, such as dates, times, locations.

Self-assessment

The involvement of learners in assessment procedures.

Standardised tests

Tests that address all learners in the same way (they all answer the same questions and carry out the same tasks) and are scored in a 'standard', consistent and reliable manner, which makes ranking and comparing learners' performance easier.

Summative assessment

Assessment that looks back and evaluates what has been learnt, usually at the end of a period of teaching and learning

T

Teacher observation

Ongoing, informal assessment that is based on teachers' perception of learner development.

Test impact

The ways in which tests (usually high-stakes tests) can influence different aspects of the whole teaching, learning and language use infrastructure.

Testing culture

Specific "epistemological suppositions about the dynamic nature of knowledge, as well as assumptions about students, teaching and learning" (Ofra 2008: 387) that refer to and are influenced by language testing.

Top-down processes

Activating background knowledge and/or contextual information to predict meaning and/or fill in gaps in comprehension.

Turns

What a speaker says, while holding the floor, counts as one turn, which is followed by the listener's turn. Turn-taking involves processes for constructing contributions, responding to previous comments, and transitioning to a different speaker, using a variety of linguistic and non-linguistic cues.

V

Validity

A judgement, based on evidence, about whether a test is appropriate for a specific group of test takers and purpose

W

Washback

The impact of tests on teaching and learning. In particular, tests can influence what and how teachers teach and what and how learners learn in formal classroom settings. Washback can be positive or negative.

Washforward

The extent to which a test includes and tests language relevant to the post-language learning situation. Proficiency tests, therefore, should have a good washforward effect.