

The MATSDA/Centro Studi Mugello Conference

Materials for Testing

Wednesday June 21st - Thursday June 22nd, 2023

Rufina near Florence

Day One - Wednesday June 21st

Registration (08.30-09.00)

Welcome (09.00-09.30)

Brian Tomlinson (Anaheim University, University of Liverpool) in Michelangelo

Plenary Presentation 1 (09.30-10.15)

Brian Tomlinson (Anaheim University, University of Liverpool) in Michelangelo

Testing for Acquisition

In my presentation I'll be using the term 'testing' to represent any form of measuring of what learners can do and I'll be using the term 'acquisition' to refer to the development of the ability to communicate in an L2. I'll be using these terms to help me focus both on ways of finding out about the effectiveness of learners' ability to communicate and on ways of ensuring that the process of preparing for a test, being tested and gaining feedback on the test facilitates the learners' development of communicative competence.

I'll be concentrating mainly on informal testing on programmes of language learning and arguing that such tests should not imitate the testing procedures of high stakes examinations which are constrained by their need to justify reliability and validity. Instead they should aim to gain an indication of what the learners can do with the target language whilst focusing primarily on contributing positively to the learners' progress.

Plenary Presentation 2 (10.15-11.00)

Hitomi Masuhara (University of Liverpool) in Michelangelo

'Real L2 speakers don't speak like the ones in the listening test!?' - The Backwash Effect of a 'Native Speaker' Norm in High-Stake Tests on Materials, Teachers and Learners

This session focuses on the increasing tensions between listening pedagogy and exams that seem to be based on a 'native speaker' norm in inner circle English speaking countries and the real-life necessity to listen and comprehend ever growing varieties of global Englishes.

A similar problem exists not only in relation to global varieties but even within an L2 inner circle country like the UK. L2 visitors and students in the U.K. often acknowledge the existence of, what I call, 'the gap phenomenon' - i.e. a huge gulf between listening experience in classrooms/tests and in real life.

The common assumption is that time and exposure will eventually overcome this gap problem. According to our recent research in the University of Liverpool and in the available literature, however, this is often not the case. For example, even after 6 months or longer, the students and visitors in our study report difficulty in listening to varieties of English. What is more concerning is the fact that their negative experience often stimulates unwillingness to communicate and deters L2 immersion.

My session will:

1. briefly report some research evidence of the ‘gap phenomenon’ and its negative effect on L2 real life listening experience;
2. discuss how the gap seems to be expanding and the problems deepening;
3. explore ways of improving pedagogy, materials and the testing of listening.

Coffee Break (11.00-11.30)

Parallel Presentations 1 (11.30-12.00)

1 Meera Srinivas (The English and Foreign Languages University, Hyderabad) in Michelangelo

Designing Tasks for Second language (L2) Writing Assessment

The presentation focuses on two item types used for assessing L2 writing proficiency in an intermediate level proficiency test. The two item types- the direct writing tasks (writing a message, letter and essay) that rely on actual samples of writing to judge different facets of writing ability, and the indirect writing item (paragraph reconstruction) that tests the ability to understand the organization of a text - were used to test discourse knowledge and competence, and functional ability to communicate meaningfully in writing.

The first part of the presentation describes the assessment context and purpose, the rationale for the selection of the item types and tasks, and the process of creating the tasks. This is followed by an in-depth analysis of sample tasks, task components and the performance descriptors developed for evaluating the tasks. The presentation concludes with a discussion of the implications for designing L2 writing assessment tasks.

2 Mukhtar Adam (University of York) in Botticelli

Employing Discourse Completion Tasks (DCTs) to Test Pragmatic Competence: A Case of Sudanese Migrants in the UK.

In research, discourse completion tasks are employed to test language learners’ proficiency when it is not possible to conduct an ethnographic study to observe language learners’ interactions in naturally occurring situations. This presentation will provide a report about employing DCTs to test the pragmatic competence of Sudanese migrant learners during their mutual interactions with English native speakers in the UK.

The presentation will provide a rationale for employing DCTs in research, a definition of DCTs, statements of their advantages and drawbacks and reports of the findings of the study that employed DCTs to examine the pragmatic competence of Sudanese learners.

3 Stephanie Ashford (Baden-Württemberg Cooperative State University (DHBW)) in Dante

Teaching and Testing L2 Writing Skills in the Age of Neural Machine Translation

With the growing sophistication of machine translation technology and other AI-powered writing tools, translations can easily be generated that may be more than just ‘good enough’. It is hardly surprising, then, that the use of such services is spreading in education, raising questions for the curriculum, academic integrity, and fairness in assessment. This presentation focuses on the implications for teaching and testing of professional writing skills where English is a second language (L2). With ‘fitness-for-purpose’ in mind, Stephanie Ashford identifies areas of a typical higher education Business English curriculum and assessment policy that would need tweaking, and presents examples of task types suitable for teaching and testing L2 writing skills.

Parallel Presentations 2 (12.05-12.35)

1 Freda Mishan (University of Limerick) in Michelangelo

Intercultural Competence: To Test or Not to Test?

Intercultural competence is a skill that is intrinsically individual and personal. While often ‘taught’ under the broad umbrella of language learning, it has less to do with language than with critical attitudes to culture/s – one’s own and that of others – and the embracing and understanding of ideologies and values underpinning them. It is not a skill that equates directly with language proficiency – some so-called ‘native speakers’ can have less intercultural sensitivity than a person who is less linguistically proficient. The area of language pedagogy that intercultural competence most closely intersects with is pragmatics – and the two are equally tricky to teach, to learn – and to test. In this paper then, I will debate whether intercultural competence can be assessed and if so, what sorts of instruments might be devised and used for this. I will illustrate my talk with some samples of such assessments, drawing on current research and practice in this area – and may even involve the audience in self-testing their own intercultural competence!

2 Geetha Durairajan (University of English and Foreign Languages, Hyderabad) and Shree Deepa (University of Hyderabad) in Botticelli

Learning to Assess Language Potentiality Through Student Feedback: A Case Study

Traditionally language education has limited itself to evaluating only the skill of proficiency. It is necessary to raise the awareness of language users and also evaluate the potential of language to be used in a constructive, neutral and destructive manner. This is the difference between being merely human and growing to becoming humane.

This paper is an attempt to present our struggle to create, trial and fine-tune a test item type, in both formative and summative contexts to test language potentiality. Using the cross cultural pragmatics discourse completion questionnaire (Blum-Kulka, House and Kasper, 1989) as a base, we modified it to create context specific situations with grammatically accurate yet humanely varying responses, termed Discourse completion tasks (DCQ) (Deepa, 2022). The data for this paper will include student feedback on the modified DCT, and end with the

further modifications attempted, informed by insights gleaned from scenario based testing (Banerjee, 2018).

Parallel Presentations 3 (12.40-13.10)

1 Asma Aftab (Kinnaird College for Women, Lahore) in Michelangelo

Experiencing the Wild, Natural Countryside: Assessment of Communicative Power

Linguists have indicated that language use is purposeful and reflects engagement, and learners should be trained and assessed keeping these aims in mind. This assumption renders assessment tasks complex since communication is multifaceted and a skill which cannot be directly or linearly measured. Thus, in order to make the testing of ESL more appropriate, I analyzed communicative acts which highlighted that effective communication involves using relevant, significant, mentally stimulating or emotionally charged and original content and tools; and, being strategically manipulative. Based on this analysis, the proposed paper equates language assessment with 'wild, natural countryside' implying evaluation of language skills involving flexible, unpredictable, and effective responses, and affective and cognitive engagement. The paper additionally discusses the application of the presented assessment procedures for testing first year undergraduate ESL students who were studying in an exclusive examination oriented educational context where grades are valued more than the acquired skills.

2 Sharon Hartle and Emanuela Tenca, University of Verona in Botticelli

Developing Items to Assess Reading Skills

The aim of this presentation is to share the journey involved in the development of a series of B2 items, created for use in the 2022 Italian national selection procedure of secondary school teachers. These specific items were designed to assess reading skills and what is often referred to as 'Use of English' (Cambridge): the ability to recognise lexico-grammatical usage. The process began with the identification of the stakeholders, their levels and perceived needs, and continued with the decision as to item type. Only then were the items developed and the tests created and used. Key questions involved in this process were:

What type of assessment is required?

What is the most reliable way to develop the items?

Our aim in this presentation is to describe the steps involved in our process but also to help participants plan assessment materials for their own learners.

Lunch and Poster Presentations (13.15-14.15)

Parallel Presentations 4 (14.15-14.45)

1 Anna Costantino (University of Greenwich, London) in Michelangelo

Harnessing Testing Materials to Enhance Students' Learning Experiences and Agency

This presentation shares the insights gained from a teacher research enquiry focused on developing language testing materials using Moodle Quiz and an online Italian portfolio in a

British higher education setting. Initially, the redesign of testing materials was prompted by the need to implement alternative assessments during the pandemic. The testing arrangements have remained in place, prompting an exploration of the benefits of blurring the boundaries between summative tools, like Moodle Quiz, and formative ones, such as an online portfolio, as well as the boundaries between classroom activities and assessment tasks. The presentation details samples of testing activities designed as open-ended tasks to engage learners cognitively and affectively (Tomlinson, 2011) and foster intercultural awareness, creativity, and autonomy (Abrar-ul-Hassan, 2021; Chong, 2018). Finally, the presentation discusses whether it is possible to harness assessment tasks as learning per se (Assessment as Learning) (Earl, 2013) to enhance language students' learning experiences and agency.

2 Umida Abdurakhimova (Westminster International, University in Tashkent) in Botticelli

Issues of Assessment of Children with Disabilities

One of the most important and unalienable requirements for disabled people and for successful socialization, full inclusion in society, and efficient self-realization in a range of professional and social activities is education. Changes in the curriculum and modifications for the evaluation process are essential for gaining access to a high-quality education. The current research examines challenges encountered while evaluating students with disabilities. To address the difficulties in adopting assessment accommodations, the study analyzes limitations and offers further ideas to make education inclusive. One of the recommendations given in the scope of this research is that several nations use various accommodations. Based on the yearly progress of impaired students, consistency in the usage of accommodations and the consistency in creating evaluation standards should be guaranteed.

Parallel Presentations 5 (14.50-15.20)

1 Claudia Saraceni (University of Bedfordshire) in Michelangelo

Developing Assessment Materials to Empower Voice and Promote Citizenship of the Language Classroom

In many ways, the concept of *voice* is intrinsically linked with that of *citizenship*, as it involves shared ownership of language(s), cultures and diverse points of view in an equal partnership amongst the members of a community. This also, therefore, entails empowering its members to make choices that can shape and determine the principles and dynamics of such community.

This paper considers the above concepts in the micro-context of the language classroom, and it focuses on the role of assessment as a tool to promote the development of learner voice drawing a parallel between the assessment of the performance of language learners and that of language teachers in training. This is discussed in particular in relation to teacher development and the role of materials in the assessment of teaching practice as a tool to enable learner-teachers to make choices, hence take responsibility for their learning in an equal partnership with their learners as *citizens of the classroom community*.

2 Shree Deepa (University of Hyderabad) and Geetha Durairajan (English and Foreign Languages University, Hyderabad) in Botticelli

Enabling and Evaluating Language Potentiality: Learnings from an Anthrologic Course

Task based language teaching has been used in the context of language education to enable only proficiency, usually restricted to skills/capabilities. Adults, who are language users and not learners, need to build their ‘verbal disposition’ through working with ‘language potentiality’ in anthropogenic contexts in mainstream higher education (mHE). The paper will begin with a critique of pedagogic practices in a parallel course and go on to examine the attempts made in an mHE-anthropogenic course, that is a course in which the emphasis is on ‘living a life of a cohabitant alongside other beings peacefully on planet earth’ (Deepa, 2022) and which in this case enabled and tested language potentiality with an objective to build the verbal disposition of students. The tasks used in this anthropogenic course, samples of materials used for both teaching and testing along with students’ reflective feedback, will be presented to provide a replicable micro-sample of materials development and performance evaluation which enabled us teachers to get into the teaching-learning-testing loop where our learners, at one point, also became our teachers.

Tea Break (15.20-16.00)

Plenary Presentation 3 (16.00-16.45)

Neus Figueras Casanovas (Former Departament d Educació, Generalitat de Catalunya and University of Barcelona) in Michelangelo

Aligning Language Education to the CEFR: What Does That Mean for Materials?

The publication of the [CEFR Companion Volume](#) in 2020 (CEFR CV), which revised, updated and expanded the CEFR 2001, triggered new interest in the CEFR twenty years after its publication and raised numerous issues and challenges for all professionals in the field of language education, including materials developers. Whilst it is early to assess the impact of the CEFR CV, many in the profession have voiced their concerns as to how it could or should be used and implemented in different contexts, for different purposes and by different stakeholders.

This session will present the contents of a document published with the main aim to help address some of the concerns above. The [Handbook](#) for aligning language education was released in 2022 and the possible implications of the document for materials developers will be discussed. The handbook outlines the processes to follow when aligning curricula, materials and assessments to the CEFR and provides guidelines to complete them. Most importantly, the handbook argues for a comprehensive learning system (CLS, [O’Sullivan](#) 2020), and states that success in language learning depends on the close alignment of three key elements in language education, often regarded as independent of one another: curriculum, assessment and delivery, the latter including teaching approaches, teacher training and materials. The role of materials in this scenario, with a possible key role to facilitate the interaction between these three key elements needs to be explored.

Plenary Presentation 4 (16.50-17.35)

Dave Allan (Norwich Institute for Language Education (NILE)) in Michelangelo

LOLA (Learning Oriented Language Assessment) Approaches to Test Design and Assessment Procedures for Listening Skills

Throughout my career as a teacher and trainer in ELT I have been fascinated by the nature of listening and have continued to explore how we can best teach, test and assess listening skills.

In the last 15 years I have been closely involved in research and in major projects to develop LOLA (Learning Oriented Language Assessment) approaches to test design and assessment procedures in both national and institutional contexts. In this talk I would like to share with those attending the 2023 MATSDA conference a brief overview of LOLA principles, some examples of where LOLA approaches combined with new technology have been demonstrably effective and, as the main focus of this talk, the ways in which LOLA can be put into practice in the testing, assessing and learning of listening skills.

I will be giving examples of ways of testing receptive skills which are in line with LOLA principles, but with my major focus on ways of testing listening that offer possibilities for more detailed feedback for both learners and teachers.

Drinks and Poetry Reading (17.45 – 19.15) at the Garden Bar

Day 2 - Thursday June 22nd

Registration (09.00-09.30)

Plenary Presentation 5 (09.30-10.15)

Carolyn Westbrook (British Council) in Michelangelo

Doing Old Things in New Ways or Doing New Things in Old Ways? - Using the Interaction and Mediation Scales in the CEFR Companion Volume to Design Integrated Assessment Tasks.

Many teachers engage students in active learning by requiring them to do interactive, integrated skills tasks such as role plays or information exchange tasks as well as presentations requiring research on a topic. However, although the use of integrated tasks in assessment has increased in recent decades (Gebriel and Plakans, 2009), such tasks may be less frequently used in assessments than discrete skills which are designed to test one skill at a time. There may be good reasons for this, not least because discrete skills may be easier to test but, possibly, also because the constructs involved in integrated skills were ill-defined and criteria for assessing these satisfactorily were lacking. However, with the publication of the Common European Framework of Reference for Languages: Learning, Teaching, Assessment Companion Volume (CEFR CV) (Council of Europe, 2020) scales describing performance in both interaction and mediation have been expanded considerably, thus making it easier for teachers to measure performance on tasks involving these two modes of communication.

In this presentation, I will look at how we can use the scales in the CEFR CV (Council of Europe, 2020) to assist with identifying the constructs involved in integrated tasks and how these constructs can feed into the design of tasks while the scales can assist with designing rating scales for assessing performance. We will then consider whether, by using these scales, we are doing new things in old ways or whether we are, in fact, doing old things in new ways.

Plenary Presentation 6 (10.20-11.05)

Barry O'Sullivan (British Council) in Michelangelo

Building a Locally Appropriate Comprehensive Learning System

For learning systems to function successfully they should be seen as forming a single entity called a Comprehensive Learning System (CLS). Such a system will contain three core elements (curriculum, delivery, assessment), which are based on a clearly articulated philosophy of the centrality of the individual and supported by fully defined models of language ability/progression and measurement. Learning systems are situated within specific educational and social contexts, which are defined by key stakeholder groups within the context. Of these stakeholder groups, the most important, as suggested above, must always be the individual learners. Ensuring that the system meets the need of each learner adds to the complexity of the design and implementation process.

In this presentation, I will briefly outline the thinking behind the CLS, while focusing on the CEFR as a reference document to be used when defining locally appropriate standards. In addition, I will demonstrate the way in which the British Council works to ensure that these elements are combined with our awareness of the needs of individual learners through our learner-focused and test/materials-focused development and delivery policies. The simplicity of the CLS concept is easily grasped but less easily made to work. This presentation will offer a clear series of steps designed to make building truly integrated learning systems possible.

Coffee Break (11.05 – 11.35)

Parallel Presentations 6 (11.35-12.05)

1 Sakae Onoda (Juntendo University, Japan) in Michelangelo

Assessment Tasks in Japanese Language Education

Based on extensive teaching experience at senior high school and university levels, the presenter will describe a variety of tasks adopted in English classrooms and examination settings in Japan to evaluate learners' performance and achievement and discuss some of the advantages and disadvantages these tasks entail. Tasks may include dictation, listening comprehension questions, cloze tests, partial dictation, inference questions, interviews, Show and Tell, transfer activities, issue logs, collocation knowledge tests, and group oral exams. In addition to describing and discussing these tasks, a hands on-experience session will be provided for participants to try out some of the tasks, and video-taped learner engagement in the tasks will be shown to enable the audience to have a good feel for these activities. The presenter will invite critiques from the audience of some of the activities presented and elicit suggestions for compensating for some of their disadvantages.

2 Paulina Waclawik (University of Warsaw) in Botticelli

Implications of the Representation of Space in Language Textbooks

In the paper, we seek for a spatial definition of the target culture in language materials and to do so we analyse the places depicted in the book along with the way they are presented, context they appear in and the level of detail used to talk about them. We propose a method of representation of space in language textbooks by mapping traditional approaches to researching their content (exemplified here by the study by Risager, 2018) onto a Geographic Information System (GIS) traditionally used for spatial analyses. It is our assumption that space and place shape human experience (Tuan, 2001). To account for that, in the study, space is included in two ways: as geographical space mentioned and referred to in the materials and as space of the textbooks themselves, i.e., their arrangements, the page's layout. To put our discussion on a more concrete footing, we present the elaborated method of spatial analysis and show its

pedagogical application on the example of a language textbook called „The World We Live in” by Ogasawara and Hiroe (2013).

3 Dale Brown (Kanazawa University) and Tim Stoeckel (University of Niigata) in Dante

Lexical Profiling TOEIC Tests: A Flemma-based Variable Bandwidth Approach

Lexical profiling studies examine the frequency level of words in a text in order to estimate the number of words a learner may need to know so that lexis is unlikely to hinder comprehension. Most such studies have utilized word-family based wordlists and analysed the vocabulary of texts in terms of 1,000-word bands. For authentic texts used by advanced learners, this methodology may be appropriate. However, for passages in tests typically taken by less proficient learners, use of the word family as the lexical unit may underestimate lexical difficulty and 1,000-word bands may be too crude a yardstick. This paper will therefore present an analysis of 12 TOEIC Listening and Reading Test official practice papers using the flemma as the lexical unit and word bands that vary in size. The presentation will discuss the lexical demands of the TOEIC test and demonstrate how different methodological choices provide contrasting findings.

Parallel Presentations 7 (12.10-12.40)

1 Natasha Engelbrecht (Rhodes University, South Africa) in Michelangelo

When Learning-Cultures of Performativity do not Align - The Case of German Foreign Language Learning in South African Higher Education

CEFR-aligned commercial textbooks occupy a central position in almost all German foreign language (GFL) courses at South African universities. While these textbooks espouse communicative outcomes, their assessment predominantly focus on grammar, syntax and lexis on the premise that memorisation and language awareness scaffold meaning making in the foreign language. Target learners are both driven to succeed in tests and to apply their declarative knowledge in communicative situations. Conversely, most South African students are accustomed to a passive learning culture acquired in school and are generally motivated by assessment. Following a doctoral case-study on textbook praxis in the first-year GFL course at Rhodes University, I examine how students’ personal interests, life-load, lack of self-efficacy and scant opportunities to use German outside the classroom inform their approach to learning. Tensions are uncovered between students’ lived realities, their aspirations to gain language proficiency and attitudes toward assessment. Practical examples of blended learning and collaboration with students are presented as ways to mitigate these misalignments.

2 Kamilia Mebarki (University of Limerick) in Botticelli

Investigating Intercultural Competence in the Algerian Secondary School EFL Textbook ‘New Prospects’

The promotion of intercultural learning plays an important role in foreign language materials mainly in teaching English as a foreign language (EFL) as intercultural competence (IC)

enables EFL learners to communicate effectively with people coming from different linguistic and cultural backgrounds.

The overall aim of this research is on one hand to investigate IC in the Algerian secondary school EFL textbook 'New Prospects', on the other hand to explore teachers' attitudes and beliefs towards teaching IC with respect to the textbook 'New Prospects' alongside two other EFL textbooks used in the Algerian secondary school.

To estimate the textbook's potential of developing IC, the study relies on three models developed by the researcher which are 'Cultural Contexts in EFL Textbooks', 'Cultural Themes in EFL Textbooks', and 'Intercultural Competence in EFL Textbooks'. The preliminary result of the research reveals that the textbook under investigation is satisfactory in terms of enhancing learners' intercultural competence.

Parallel Presentations 8 (12.45-13.15)

1 Claudia Fernandez (University of Illinois, Chicago) in Michelangelo

Micro-Evaluation of Tasks as Tools for Assessment

The purpose of this presentation is to report the results of a micro-evaluation of a Final Task as performed by students of Spanish in an intermediate course in higher education. Based on Ellis (2011), a micro-evaluation of a task investigates the extent to which the task achieves the objectives set out for it. In the case of the context where this evaluation took place, the task investigated was the Final Task of the teaching module which was used as a summative assessment tool ("Design an object that solves an everyday problem and give a presentation to the class"). Data was collected on students' opinions about the task, how the task was implemented by the instructor, and the extent to which the task was performed to completion. The results shed light on the effectiveness of these task-based materials to evaluate what students can do in the target language

2 Houcine Benlaib (University of Limerick) in Botticelli

Teachers' Beliefs about Gender Representation in the Algerian EFL Textbook

"My Book of English". Case study: First Year Middle School

For decades, EFL textbooks have been the subject of investigation as they are a vital instrument in education that provides knowledge in the classroom and are also regarded as the core of the learning-teaching process. Textbooks are dual-purpose tool; they can be used as an educational tool or as a tool for socialization as they mirror different issues in the vast world of research.

Gender representation in EFL textbooks has been the focus of many studies internationally (Sunderland et al., 2015). However, very little seems to be done in the Algerian context. In this regard, the current study aims to investigate and compare the representation of male and female characters in the Algerian EFL textbook "My Book of English".

Lunch and Poster Presentations (13.15-14.15)

Parallel Presentations 9 (14.15-14.45)

1 Leila Anjomshoa (Islamic Azad University, Kerman, Iran) in Michelangelo

The Relationship between Test Anxiety and Mental Well-Being among Iranian EFL Learners

In the context of language learners' mental well-being, research has recently focused on a number of different learning variables. The direction of the relationship between learners' mental well-being and test anxiety is, however, still open. To address this need, the current study examined the relationship between test anxiety and mental well-being in the EFL context of Iran. Study participants included 250 Iranian students enrolled in B.A. English language courses in English language teaching, translation, and literature at Islamic Azad University, Kerman Branch. The data were derived from Valero's test anxiety questionnaire (1999) and Tennant et al.'s (2006; 2007) Warwick-Edinburgh mental well-being questionnaire. After analyzing the data, the results demonstrated a significant negative correlation between test anxiety and mental well-being. An increase in test anxiety led to a decrease in participants' mental well-being. Moreover, participants' sex and major were not meaningful mediators affecting the relationship between test anxiety and mental well-being. Therefore, minimizing test anxiety and optimizing learners' mental well-being in educational contexts requires teachers and curriculum designers to consider students' readiness and acceptance of academic tests.

2 Maria Stec (University of Silesia, Poland) in Botticelli

Online Assessment of Ph.D. Students on ESP/EAP Courses

Language assessment of Ph.D. students can vary due to a series of factors. The development of the appropriate materials is one of them, following DBR (design-based research). The idea for this paper originated from the English online courses implemented at the University of Silesia during the pandemic period and nowadays.

The aim of the paper is to depict testing materials implemented during the ESP (English for Specific Purposes) and EAP (English for Academic Purposes) online courses on the Moodle platform developed for post-graduate students. The research is diagnostic with a descriptive and an evaluative function. The paper seeks to address the following question: *What are the challenges in materials development for testing Ph.D. students? What is the best form of testing their progress?* The project will use a mixed type of research, including an author's self-reflection analysis. It is hoped that the results will provide valuable insights and highlight the need to further develop digital language learning materials at the university level.

Parallel Presentations 10 (14.50-15.20)

1 Fabio Loi (Oxford University Press) in Michelangelo

Introducing the Oxford Test of English

Abstract to follow.

2 Kasia Wozniak (ALTE - Association of Language Testers in Europe) in Botticelli

Developing Language Testing Materials

For over 10 years, ALTE members and other test developers have been using ALTE's 'Manual for language test development and examining for use with the CEFR', which was produced by ALTE on behalf of the Language Policy Division, Council of Europe in 2011.

This invaluable resource has helped many institutions and individuals develop and use materials for language testing. This manual explains the process of test development and how the link to the CEFR can be built into it. As part of our mission, ALTE is committed to ensuring quality in language testing materials, and this presentation will highlight some of the key features of the manual. ALTE always welcomes feedback on this document for a revision to consider the latest developments in test production. The presentation will also introduce ALTE's new Services Unit, which can provide additional help for organisations who are interested in developing language test materials.

Question and Answer Session (15.30-16.15) in Michelangelo

An opportunity to ask further questions to the plenary speakers.

Plenary Presentation 7 (16.15-17.00)

Alan Maley (Educational Consultant) in Michelangelo

Testing and Education: A Contradiction in Terms?

I shall explain why and in what ways I believe that education and testing are fundamentally at odds, referring to the work of educationists from Jespersen and Dewey to Robinson and Holt. Testing is essentially rooted in a control/predictability model, whereas education rests on an exploratory/unpredictability approach.

The current hegemony of high-stakes, high-cost, high-stress testing seems unfortunate and unhelpful therefore, if our aim is education rather than certification. I shall discuss alternative assessment ideas and practices aimed at reducing negative effects on learners, and restoring some educational value to both teachers and learners. These will include reference to the work of James Dean Brown, Antony Kunnan and other lesser-known workers.

Round Up (16.15.16.30)

Brian Tomlinson (Anaheim University, University of Liverpool) in Michelangelo

Poster Presentations (13.15-14.15 each day)

1 Catarina Castro (Polytechnic of Leiria) and Fausto Caels (Polytechnic of Leiria)

The Development of a Placement Test to Identify Chinese Students' Proficiency Level in Portuguese

In recent years, learning Portuguese has raised more and more interest on the part of the Chinese community, not only in Macau (where it will be an official language until 2049), but also throughout Mainland China. Consequently, the international mobility of Chinese students to Portugal has increased significantly. According to official data, 758 Chinese students attended higher education programs in Portugal in 2015/2016, increasing to more than 1000 in 2020/2021.

Currently, several Portuguese higher education institutions have university cooperation protocols with China and Macau, as is the case of the Higher School of Education and Social Sciences of the Polytechnic of Leiria (ESECS-IPL). Every year, this institution welcomes a significant number of Chinese students in its courses, after they have acquired an initial level

in Portuguese language (level A2) in China or Macau. Nevertheless, their communication skills in Portuguese tend to be quite variable upon arrival.

Within this context, this poster aims to describe a placement test, created by ESECS-IPL's teachers – drawing on official assessments of Portuguese as a Foreign Language – whose main purpose is to identify students' proficiency level in Portuguese. The results of the tests are crucial to signal students in need of additional support, subsequently provided by the school. It also helps teachers to fine-tune the course syllabuses, according to students' prior knowledge.

2 Sakae Onoda (Juntendo University, Japan)

Issues Relating to Questions on Entrance Examinations in Japan

The presenter will first show and describe typical questions used in entrance examinations to Japanese senior high schools and universities and encourage the audience to critique them. The questions will include those measuring vocabulary, pronunciation, grammar, reading, listening, essay writing, and translation skills. The presenter will explain why these questions are used in Japan and review potential problems with their reliability, validity, and practicality and discuss related pedagogical issues, including washback effects on classroom teaching and achievement tests taken by students. The presenter will invite an exchange of views with the audience regarding these issues and elicit possible solutions. The session will be an opportunity to expand the horizons of both the presenter and the audience regarding testing.

3 Giorgia Andreolli (University of Verona)

What Do We Talk About When We Talk About Materials? Emerging Teacher Perspectives on Assessment

References to assessment in teacher discourse may be made either overtly or covertly. Over recent years the focus on assessment literacy has expanded (Tsagari and Vogt 2017; Fulcher 2012; Berry, Sheehan, and Munro 2019), underlining the need for integrating teachers' perspectives. This proposal is for a poster, which presents and discusses the findings of an ongoing mixed-methods research on digital materials in ELT. On the whole, the project investigates the teacher-technology relationship by focusing on the beliefs and reported practices of teachers of English (Halverson et al. 2014) in Italian secondary education. The poster draws on the qualitative strand of the research, which consisted of semi-structured interviews. A thematic analysis of the transcripts enabled an exploration of the intersections between pre-defined issues (e.g., theoretical underpinnings of language teaching and learning, aspects of planning with technology, and materials adaptation) with emerging themes. Among these, references to assessment (both formative and summative, see Hartle 2020) were a key element which will be explored in the attempt to map its occurrence in relation to other relevant themes. Doing so, the study seeks to answer the question of how digital materials use and assessment may be (implicitly or explicitly) connected, as well as to reflect on implications for future research and practice.