

Centrum voor
Taal ^{nt} 2 en
Onderwijs

KU LEUVEN



taal:
unie

CNaVT
Certificaat Nederlands als Vreemde Taal

50th ALTE Meeting and Conference Day
50e ALTE-bijeenkomst en -conferentie

The impact of language tests on
education, migration, and society

*De impact van taaltoetsen op onderwijs,
migratie en maatschappij*

20th–22nd September 2017
20 – 22 september 2017

Hosted by the Certificate Dutch as a Foreign Language, Centre for Language and Education
(KU Leuven)

*Organisatie: Certificaat Nederlands als Vreemde Taal, Centrum voor Taal en Onderwijs
(KU Leuven)*

Programme | Programma

Welkom op de 50e ALTE-bijeenkomst en -conferentie in Leuven

ALTE, het Certificaat Nederlands als Vreemde Taal (CNaVT) en het Centrum voor Taal en Onderwijs (CTO) verwelkomen u op de 50e ALTE-bijeenkomst en -conferentie in Leuven.

Het CNaVT is een project van de Taalunie en toetst en certificeert wereldwijd leerders van het Nederlands als Vreemde Taal. Het CNaVT biedt taakgerichte, domeinspecifieke taaltoetsen aan van niveau A2 tot en met niveau C1. Het Centrum voor Taal en Onderwijs, waar de CNaVT-toetsen worden ontwikkeld, zet in op gelijke onderwijskansen en leerwinst voor alle leerlingen. Deze aanpak staat al 25 jaar centraal in de missie van het CTO en daarom is het wel toepasselijk om het thema impact te kiezen voor deze bijeenkomst en conferentie.

We kijken ernaar uit om dit belangrijke onderwerp de komende week met u te bespreken. We zijn blij en trots dat we u een gevarieerd programma kunnen aanbieden met een aantal befaamde experts uit het vakgebied. Zij zullen vanuit hun eigen perspectief bijdragen tot een discussie over de impact van taaltoetsen op onderwijs, migratie en maatschappij. Wij zijn ervan overtuigd dat de presentaties en workshops waardevolle ideeën zullen opleveren en er een debat op gang zal gebracht worden dat positieve effecten zal hebben op ons toekomstig werk.

Als het cliché waar is dat de beste idealistische debatten gehouden worden op café of in een restaurant, is Leuven de perfecte omgeving voor dit ALTE-evenement. Wij raden u dan ook aan om naar buiten te gaan en de culinaire en culturele geneugten van deze middeleeuwse stad te verkennen.

Als er iets is dat we kunnen doen om uw verblijf aangenamer te maken, aarzel dan niet om ons aan te spreken.

Hartelijk welkom!



Steven Verheyen
Certificaat Nederlands als
Vreemde Taal
Certificate Dutch as a
Foreign Language



Mariet Schiepers
Centrum voor Taal en
Onderwijs, KU Leuven
Centre for Language and
Education, KU Leuven

Welcome to the 50th ALTE Meeting and Conference day in Leuven

ALTE, the Certificate Dutch as a Foreign Language (CNaVT) and the Centre for Language and Education (CLE) welcome you to the ALTE's 50th Meeting & Conference day in Leuven.

The CNaVT is commissioned by the Dutch Language Union to provide A2 – C1 certification of Dutch language proficiency worldwide, using task-based and domain-specific exams. The Centre for Language and Education, where the CNaVT tests are developed, seeks to ensure that language education offers equal opportunities for all learners and benefits all learners equally. This approach has been central to CLE's mission for the past 25 years and so it seemed appropriate to choose the concept of impact as the theme for the meeting and conference day this time.

We are looking forward to discussing this important topic in the course of the week and we are happy and proud to present a varied programme that includes a number of renowned experts from our field. They will contribute to a discussion of the impact of language tests on education, migration, and society from their own perspectives. We are confident that the presentations and workshops will generate thought-provoking ideas and will stimulate debate that will have positive outcomes for our future work.

If the cliché is true that idealistic debates work better in little bars and restaurants, Leuven is the perfect setting for this ALTE event and we encourage you to go out and explore the culinary and cultural delights that this medieval city has to offer.

If there is anything we can do to make your stay more comfortable, do not hesitate to ask us.

A very warm welcome!



Dr Nick Saville
ALTE Secretary-General

ALTE 50th Meeting & Conference, Leuven, September 2017

Wednesday 20th September, 2017

Registration, breaks & lunch will take place at the entrance of the MSI (Mgr. Sencie-instituut, Erasmusplein 2)

All sessions will take place at the MSI (Mgr. Sencie-instituut, Erasmusplein 2)

Time	Session	Room
8.45–9.00	Registration	MSI Entrance
9.00–11.00	Standing Committee Meeting (elected members of committee only)	MSI 02.23
11.00–11.15	Coffee	MSI Entrance
11.15–13.15	Standing Committee Meeting continues	MSI 02.23
13.15–14.30	Lunch	MSI Entrance
14.30–16.00	Executive Committee Meeting (elected members of committee only)	MSI 02.23
16.00–16.15	Coffee	MSI Entrance
16.15–18.00	Executive Committee Meeting continues	MSI 02.23
18.15–19.15	City walk with professional guide	Meeting point = MSI Entrance
19.15–20.30	Beer Benchmarking at Bar Domus (Tiensestraat 8)	Meeting point = stairs of the City Hall (Grote Markt) at 19.15

ALTE 50th Meeting & Conference, Leuven, September 2017

Thursday 21th September, 2017

Registration, breaks & lunch will take place at the entrance of the Erasmushuis building (Blijde-Inkomststraat 21)
All sessions will take place at the MSI (Mgr. Sencie-instituut, Erasmusplein 2)

Time	Session	Room
08.30–9.00	Registration	Erasmushuis Entrance
9.00–9.15	Official Opening	MSI 00.08
9.20–11.00	Parallel SIGs	
	SIG on Technology in Language Assessment (any ALTE delegate)	MSI 02.18
	LSP SIG (any ALTE delegate)	MSI 02.15
*This meeting will last until 11.20	*LAMI, Teacher Training and Young Learners (joint SIGs) (any ALTE delegate)	MSI 02.23
11.00–11.30	Coffee	Erasmushuis Entrance
11.30–13.00	Parallel SIGs	
	QMS working group (any ALTE delegate)	MSI 02.18
	Special Requirements and Circumstances SIG (any ALTE delegate)	MSI 02.15
	CEFR SIG (any ALTE delegate)	MSI 02.23
13.00–14.00	Lunch	Erasmushuis Entrance
14.00–15.30	Parallel Workshops	
	Professor Lourdes Ortega Measuring Language Learning for Social Justice: What do We Need from Testers?	MSI 02.18
	Professor Jan Hulstijn How to deal with problematic features of the CEFR?	MSI 02.23
15.30–16.00	Coffee	Erasmushuis Entrance
16.00–16.15	Reports from SIGs	MSI 00.08
16.15–18.00	Annual General Meeting – ALTE Members only	MSI 00.08
19.30–23.00	Conference dinner at De Hoorn (Sluisstraat 79)	Meeting point = stairs of the City Hall (Grote Markt) at 19.10

ALTE 50th Meeting & Conference, Leuven, September 2017

Friday 22nd September, 2017

The impact of language tests on education, migration, and society

Registration will take place at the entrance of the Erasmushuis building (Blijde-Inkomststraat 21).

All presentations & sessions will take place at the MSI (Mgr. Sencie-instituut, Erasmusplein 2).

Break and lunch will take place at the entrance of the Erasmushuis building and at the entrance of the MSI.

Time	Session	Room
8.30–9.00	Registration	Erasmushuis Entrance
9.00–9.20	Opening Address Mariet Schiepers , Centre for Language and Education, KU Leuven Dr Nick Saville , ALTE Secretary-General	MSI 03.18
9.20–10.00	Professor Lourdes Ortega A Transdisciplinary Agenda for Language Testing? Nudges from a Multilingual SLA Perspective	MSI 03.18
10.00–10.40	Professor Jan Hulstijn Construct and measurement of language proficiency: perspectives from BLC Theory and corpus linguistics	MSI 03.18
10.40–11.00	Coffee	Erasmushuis Entrance & MSI Entrance
11.00–11.40	Professor Constant Leung Language assessment for social affiliation?	MSI 03.18
11.40–12.20	Dr Bart Deygers Just Testing. Applying theories of justice to high-stakes language tests	MSI 03.18
12.20–13.30	Lunch	Erasmushuis Entrance & MSI Entrance
13.30–14.10	Professor James Simpson Language assessment for adult migrants: Issues and implications	MSI 03.18
14.10–14.50	Professor Kris Van den Branden Energy for learning? The impact of assessment on learning in compulsory education	MSI 03.18
14.50–15.00	Closing remarks and introduction to the workshops Dr Nick Saville , ALTE Secretary-General	MSI 03.18
15.15–16.45	Parallel Workshops	
<i>Low-educated learners & fairness</i>	Professor Cecilie Hamnes Carlsen & Professor Jeanne Kurvers Giving low-educated learners a fair chance	MSI 02.23
<i>Educational policy</i>	Professor Constant Leung Assessing additional language performance	MSI 02.18
<i>Low-educated learners & migration</i>	Professor James Simpson Literacy & speaking tests for adult migrants	MSI 01.23
<i>International language Policy & Classroom practice</i>	Dr Koen Van Gorp (Michigan State University) & Steven Vanhooren (Dutch Language Union) Toward sustainable language assessment: Linking an international language policy with a language assessment policy in schools	MSI 03.18

Parallel workshops on Thursday

Professor Lourdes Ortega (Georgetown University)

Measuring Language Learning for Social Justice: What do We Need from Testers?

In this workshop, the workshop leader, a member of the field of second language acquisition, and the attendees, members of the field of language testing, will interact in a structured way around four case study scenarios, with the goal to end the session with a provisional list of most-urgent pending questions for language testers whose expert answers would help extend knowledge about the proficiency of elite and circumstantial multilinguals and put pressure on governments and institutions for more socially just policies that support multilingual minorities. The provisional list will pair questions, as much as possible, with proposed justice-affirming language assessment practices.

The scenarios will speak to the following issues. Many of our societies treat monolingual-like proficiency as the linguistically normal state of human beings and as the marker of national loyalties. So do many of our language tests. This status quo turns the back on many minoritized groups, including immigrants, refugees, Indigenous Peoples, and Deaf communities, who share the experience of learning, unlearning, relearning, and nonlearning multiple languages – on and off over the lifespan. They also share the experience of injustice arising in part – albeit not exclusively – from living with multiple languages. This is because, having evolved multilingual rather than monolingual-like proficiencies, their complex multilingual abilities are ill-measured. Perceived as unmeasurable, and potentially uneducable, the language abilities of multilinguals also readily become a target of discrimination and easily turn into socially allowable excuses for other types of oppression. In the reverse, and paradoxically, foreign language learners are supported through traditional pedagogies to develop relatively monolingual-like proficiencies in the new language. Their emerging competencies fit existing tests better. This also gives way to unjust valorizations and paradoxes. For example, foreign language learners may be judged more competent speakers of the (standard) new language than minoritized members of the originary speech communities. Research has amassed from fields as varied as second language acquisition, bilingualism, educational linguistics, and sociolinguistics regarding, on the one hand, what is needed in order to learn a majority language (often in the face of no support and under extreme pressure and unrealistic expectations) and, on the other, what is needed to support the human right to the mother tongue(s) and to bilingualism. Research has also helped us understand how social dynamics such as linguisticism, language ideologies, and intersectionality lead to the violation of language-related human rights for minoritized, multilingual groups and enable the paradox of multilingualism: lived as a curse by some and as a boon by others, even within the same society or for the same languages. So, what are the responsibilities for these habitual social practices and what are the consequences of this accumulated research knowledge for the language testing field?

The goal of the workshop is to foster self-reflection and to generate action-oriented knowledge regarding opportunities that language testers have to support useful measurement of the learning, unlearning, relearning, and nonlearning of multiple languages by people from all walks of life, but particularly by members of minoritized communities, whose multilingual proficiencies are both a source of strength and a vulnerability in our present world.

Professor Jan Hulstijn (University of Amsterdam)

How to deal with problematic features of the CEFR

The Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001) is best known for its six levels of language proficiency but it offers much more than that. It renders a plan for foreign language education with truly humanistic goals, in an effort “to promote mutual understanding and tolerance, respect for identities and cultural diversity through more effective

international communication” (Council of Europe, 2001, p. 3). The two scholars who initially conceptualized and constructed the CEFR during a period of over 30 years preceding the CEFR’s official publication, John Trim and Jan Van Ek, wrestled with various educational, political and linguistic needs and demands, suffering from a lack of linguistic theory or empirical evidence from first and second language acquisition research (Trim, 2013). Their pioneering work, for which they deserve our lasting respect, was based on a number of practical compromises. The authors could not avoid therefore, that the CEFR comprises several notions that are difficult to reconcile with one another or poorly supported by empirical evidence. In this workshop, in which I will adopt a constructive rather than critical attitude towards the CEFR, the audience will choose any of the topics listed below for examination and discussion. They are concerned with the tensions, inherent in the CEFR (Hulstijn, 2015, Chapter 10), including the recent pilot version (2016) of the “CEFR Illustrative Descriptors”:

1. The relationship between “language activities” (CEFR, Chapter 4) and “general and linguistic competences” (Chapter 5).
2. The absence of a relationship between the six CEFR proficiency levels and language acquisition as a developmental process.
3. The relationship between cognitive/intellectual skills and the six CEFR proficiency levels.
4. The tension between the humanistic idea underlying the CEFR and the realities of educational testing and educational policy.
5. The irony of history, exhibited in the original strive towards convergence (resulting in the language-neutral CEFR document in 2001) and the subsequent strive towards divergence, evidenced by the publication of ‘Reference Level Descriptions’ (RLDs) for individual languages, all authorized by the Language Policy Unit of the Council of Europe.
6. The virtues and vices of static language-proficiency framework in a dynamic world.

Plenary presentations on Friday

Professor Lourdes Ortega (Georgetown University)

A Transdisciplinary Agenda for Language Testing? Nudges from a Multilingual SLA Perspective.

Fifteen scholars from the field of second language acquisition (SLA) united around a collective position paper published last year (Douglas Fir Group, 2016). We acknowledged that compelling changes have taken place in the nature of language learning and teaching, fueled by sweeping world patterns in technology, globalization, and mobility. We sought to articulate a framework for how the field of SLA might respond to such profound changes, explicitly reframing the study of language learning and teaching as inseparable from people’s emergent multilingual lives. We concluded that SLA must make commensurate changes in its disciplinary research agendas.

In this talk, I will nudge the ALTE audience to reflect on the corollaries of this proposal for the field of language testing. First I will examine that transportability from SLA to language testing of two key features of the vision: transdisciplinarity and multilingualism. Transdisciplinarity calls for risk-tolerant traveling in and out of the boundaries of fields in order to generate socially useful knowledge that is more than just the sum of disciplines. How can the field of language testing engage in transdisciplinary exchanges, given that it is often seen – at least by outsiders – as expertise driven and highly technical? Multilingualism requires a reimagining of both “language” and “competence.” What roles can language testers play in inspiring new imagined yet measurable contents for these two constructs, which have always been central to the daily activities of the language testing field? I will then take up possible

contributions of the new transdisciplinary ethos for language testing across the three layers proposed by the Douglas Fir Group: the micro level of social activity, the meso level of sociocultural institutions and communities, and the macro level of emotionally imbued ideological structures. What might it take for the various communities of language testers to embrace the challenge to measure – and to debunk mismeasuring – complex multilingual abilities for use that are impacted simultaneously at all three levels? What might language testing look like in the future, if it is informed by a semiotic understanding of communicative abilities and an emotional-cognitive-social understanding of multiple-language learning?

I hope to show that the field of language testing is uniquely positioned and equipped to help illuminate the emergent competencies of multilingual people's across their languages. The ability of the field to make a positive impact in today's world depends crucially on the development of a transdisciplinary agenda that carries into the long-term future. Language testing and SLA can be effective allies in this enterprise.

Professor Jan Hulstijn (University of Amsterdam)

Construct and measurement of language proficiency: perspectives from BLC Theory and corpus linguistics.

In the first part of the presentation, I will take the audience on a theoretical, philosophical flight, high above the daily worries of educational policy and language assessment. As Karl Popper (1959) said, scientific inquiry is a matter decreasing ignorance. Language-proficiency theories and accompanying empirical research of the past thirty years have somewhat decreased our ignorance. BLC Theory (Hulstijn, 2015) defines the notion of language proficiency in terms of two independent dimensions: (i) basic language cognition (BLC) versus higher language cognition (HLC), and (ii) core versus periphery. By doing so, BLC Theory breaks down the notion of "native speaker" in extralinguistic terms (e.g., age, level of education) and in linguistic terms. Perhaps "the" native speaker only exist in what all native speakers have in common (BLC), not in the many domains and levels where they differ (HLC). BLC Theory thus offers a window on the role of language in literate societies. In my current empirical work, I am discovering that the size of BLC is fairly large, perhaps roughly as large as B1 in the CEFR.

After this theoretical flight at high altitudes of abstractness, I will try to safely land on the solid soil of educational practices. I will propose that the linguistic flesh to the bones of the six levels of the CEFR should consist, in its core, of an intimate integration of vocabulary and grammar, rather than separate lists of words and grammatical constructions. Corpus linguists should help us define, probabilistically, the CEFR levels in terms of this lexico-grammar. I will end the presentation with several propositions, which may form the input for a discussion during the afternoon workshop.

Professor Constant Leung (King's College London)

Language assessment for social affiliation?

Educational assessment in many parts of the world has taken a pro-learning stance in the past 15 years or so. Jurisdictions as far apart as Hong Kong, New Zealand and Scotland have adopted formative assessment policies that are designed to promote student learning. High(er) quality learning is generally regarded as a desirable outcome in education. Seen in this light, there has been a palpable move towards putting assessment at the service of wider educational and social goals.

In this talk I will examine how far the prevailing fundamental concepts in additional/second language assessment with particular reference to linguistic minority students (e.g. assessing English as an additional language in the UK or Flemish/French as an additional language in Belgium) are part of this development. It will be argued that many of the assumptions embedded in additional

language assessment are largely concerned with abstracted functional-transactional language use. In the contemporary demographic and socio-cultural conditions in Western Europe (and elsewhere), language assessment should also embrace situated language practices that can take account of students' personal investments in language use. Taking account of recent work in language education, multilingualism and translanguaging (e.g. the Douglas Fir Group, 2016; García and Li Wei, 2014), I will provide a sketch of a language assessment approach that can contribute to the fostering of local identities and social affiliation.

Dr Bart Deygers (KU Leuven)

Just Testing. Applying theories of justice to high-stakes language tests.

Is it just for a university to demand that international L2 students meet language requirements that are not met by all L1 students, who are exempt from taking the test? Is it just for a country to raise the language requirements for citizenship to a literacy level that de facto excludes people who have not had access to organized education or schooling?

These are but a few moral and ethical considerations that language testers and policy makers grapple with today. To date, however there have been relatively few attempts at formulating principles of justice that could apply specifically to language testing. Yet, in a socio-political climate where high-stakes language requirements abound, it is important to have a clear concept of what just testing means and how it can be operationalized.

During the presentation I propose six principles of justice for test developers and score users. These principles are based on theories of distributive justice that focus on human rights, fairness, equal opportunity, and dignity. The overarching aim of this presentation is to advance the debate on justice, and to provide a consistent way of considering ethical and moral dilemmas that language testers and policy makers face.

Professor James Simpson (University of Leeds)

Language assessment for adult migrants: Issues and implications.

This presentation is about language assessment for bilingual adults in migration contexts, particularly those with little or no formal educational experience. I begin with an overview of key issues and a summary of current research on high-stakes L2 testing for adult migrants, including discussion of the use of language tests (actual or de facto) for citizenship and naturalisation purposes.

This is pertinent at a time of large-scale high-stakes testing of adult learners who are migrants to the global north and west. What are the implications of such testing, in the lives of adult migrants? I then turn specifically to the assessment of speaking skills, in a study of adult learners of English for speakers of other languages (ESOL) in the UK. I employ the notions of knowledge schema and frame in discourse to highlight two areas of interest in testing the speaking skills of these students: divergent interpretations of the test event by learners; and variation in interlocutor behaviour. What are the implications for testing speaking, of the findings from this study?

Professor Kris Van den Branden (KU Leuven)

Energy for learning? The impact of assessment on learning in compulsory education.

Assessment should be considered an integral part of the range of pedagogical activities that teachers develop in the classroom and which tend to have a strong impact on students' learning motivation and growth. Both from a cognitive and socio-emotional perspective, assessment strongly influences

student learning. In this presentation I will explore the many ways in which assessment can be shown to have a positive or negative influence on the effort and mental energy students invest in learning at school, and on the learning that results.

Parallel workshops on Friday

Professor Cecilie Hamnes Carlsen (Høgskulen på Vestlandet, Kompetanse Norge) Professor Jeanne Kurvers (Tilburg University)

Giving low educated learners a fair chance: reducing the negative impact of high stake testing on low educated L2-learners.

Adult learners with limited schooling and low levels of literacy have long formed part of the immigrant population. This is only natural given the fact that for many, war and conflict in the home country is both the direct cause of limited schooling and low levels of literacy, and the reason why people apply for shelter in safer countries.

Due to a growing tendency towards more and stricter language requirements in Europe, low-literate learners and refugees alike, now need to pass language tests to gain democratic rights, like citizenship, and even human rights, like permanent residency, family reunification and housing. It is therefore urgent that the professional community of language test developers and -researchers take this easily marginalized group into account in the construction and validation of standardized language tests (Carlsen, 2015, 2016).

Low-literate learners are an understudied population in SLA-research, and what we know about second language acquisition is almost exclusively based on research on highly literate, often highly educated learners (Young-Scholten 2007; Tarone et al 2009, 2010; Allemano 2013). Research in cognitive psychology however, has shown that literacy affects the way the mind works, and in particular, the way we process language (Vinogradov 2011; Kurvers 2002; Kurvers, van de Craats & van Hout 2014). Low literate learners perform lower on cognitive tests in general, and on verbal tests in particular (Ostrosky-Solis et al. 1998; Ardila et al. 2010). This may be due to a lack of print literacy on the one hand and on a lack of test literacy on the other. Several scholars have pointed to the fact that low-literate learners' lack of familiarity with the testing situation is bound to affect their scores on tests (Allemano 2013; Mishra, Singh & Pandey 2012)

The aim of this workshop is twofold: First, we want to present some research into how low-educated learners learn a new language, how they differ from educated learners in how they perform on cognitive tests and language tests and in the way they process language. Second, we will use these research results as a starting point for a discussion of language test tasks that would be more or less appropriate to include in a language test where part of the population is low literate.

Professor Constant Leung (King's College London)

Assessing additional language performance: What can language assessment descriptors tell us?

In many education systems teachers are expected to work with prescribed or statutory assessment frameworks and rating scales that apply to all students. This approach can raise questions of usefulness and validity in educational settings with linguistically diverse students.

In this workshop I will provide participants with an opportunity to compare two sets of assessment descriptors: mainstream English (school subject) and English as an additional language, and to trace

their underlying language models. The main aims are (a) to identify the similarities and differences between these sets of descriptors in terms of knowledge and skills, (b) to reflect on the curricular and language concerns embedded in the descriptors, and (c) to explore aspects (if any) of language knowledge and use that may be missing from the point of view of additional language learning and use. There will be opportunities for group discussion and hands-on activities.

Professor James Simpson (University of Leeds)

Literacy and speaking tests for adult migrants.

In this workshop participants will explore issues relating to the testing of literacy and oral communication skills for adult migrants. Using video and hands-on work with data, we will examine alternatives to standardised tests for literacy, and consider in more depth the implications for the testing of oral interaction raised in the plenary session.

The session will be framed around McNamara and Ryan's (2011) questions about fairness and justice which should be asked of any language test for migrants: Does it test what it should? Should it test what it does?

Dr Koen Van Gorp (Michigan State University)

Steven Vanhooren (Dutch Language Union)

Toward sustainable language assessment: Linking an international language policy with a language assessment policy in schools

Dutch is the official language in Flanders (Belgium) and the Netherlands, and is learned as a foreign language in more than 40 countries worldwide. Since 1980, the Dutch and Flemish governments have set a joint policy with respect to the Dutch language. The Dutch Language Union is the governmental organization shaping and enacting this policy. It facilitates experts in developing a shared vision on teaching and assessing language competencies. Furthermore, it is responsible for the testing (and certifying) of language proficiency in Dutch as a foreign language.

This session will explore the opportunities and hurdles of developing a shared vision on sustainable language education and assessment for the 21st century. Having a clear vision and policy about language competencies is one thing, implementing this policy in the teaching practices across countries and even worldwide is another.

We will focus on language assessment policy as a key concept in implementing this policy. Building on Kris Van den Branden's idea of renewing students' energy for learning, assessment in sustainable education should build connections between where the learner is going, how he or she is going there, and where to next (Hattie, 2009; Van den Branden, 2015). To inform this dialogue between teachers and learners, a broad repertoire of assessment procedures and methods is needed (e.g., peer-, co- and self-assessment, observations, portfolio and tests). We will explore the role of broad and alternative forms of classroom-based assessment, as well as standardized testing in the development of a language assessment policy in schools.

We will investigate the impact of a language assessment policy on both learners and teachers or school teams, look for ways to improve teacher and school practices, and formulate tips for policy makers.

About the plenary speakers

Professor Lourdes Ortega

Lourdes Ortega is a Professor in the Department of Linguistics at Georgetown University, where she mentors teachers and researchers and investigates key questions about second language acquisition and foreign language education. She is originally from Spain and taught her native Spanish in Greece for nearly a decade and, since 1993, also her nonnative English in diverse geographies in the United States. Her main research interests are in understanding linguistic and socioeducational influences that impact on youth and adults' cognitive, social, and educational well-being when they learn new languages later in life. Key interests include second language writing and systematic research synthesis. In the last few years she has been applying knowledge from bilingualism and from usage-based linguistics to the investigation of second language development. Two central questions in this new agenda are: How does experience shape language learning? What counts as success in bi/multilingual acquisition, and who is to tell? Lourdes was co-recipient of the Pimsleur and the TESOL Research awards (2001) and has been a doctoral Mellon fellow (1999), a postdoctoral Spencer/National Academy of Education fellow (2003), and a senior research fellow at the Freiburg Institute of Advanced Studies (2010). She is Past Journal Editor of *Language Learning* (2010-2015). Recent journal articles have appeared in *Foreign Language Annals*, *Journal of Second Language Writing*, *Language & Cognition*, *System*, and *World Englishes*. Her most recent books are *The Usage-based Study of Language Learning* and *Multilingualism* (Georgetown University Press, 2016). A revised edition of her *Understanding Second Language Acquisition* (1st edition with Hodder, 2009) is underway with Routledge. She is also currently working on *The Cambridge Handbook of Bilingualism*, with co-editor Annick De Houwer. Lourdes has been a plenary speaker at the American Association for Applied Linguistics conference in 2010, the World Congress of Applied Linguistics (AILA) conference in 2014, and the International TESOL convention in 2015.

Professor Jan Hulstijn

Jan Hulstijn is professor emeritus of second language acquisition at the University of Amsterdam. His research was mainly concerned with explicit versus implicit accounts of secondlanguage learning and with vocabulary learning. His recent and current work is concerned with the construction and testing of his theory of language proficiency (BLC Theory). For information, please visit his webpage at: <http://www.uva.nl/profile/j.h.hulstijn>

Professor Constant Leung

Constant Leung is Professor of Educational Linguistics in the Dept. of Education and Professional Studies, King's College London. He is a Fellow of the Academy of Social Sciences (UK). Before taking up teaching positions in higher education he taught in schools and worked as advisory teacher and manager in local government. He was the founding chair of the National Association for Language Development in the Curriculum. His research interests include education in ethnically and linguistically diverse societies, additional/second language curriculum and assessment, language policy and teacher professional development. He serves as Editor of Research Issues of *TESOL Quarterly*, Senior Associate Editor of *Language Assessment Quarterly*, and as a member of the Editorial Boards of *Assessment in Education*, *Australian Review of Applied Linguistics* and the *Modern Language Journal*.

Dr Bart Deygers

Bart Deygers is a language testing researcher at the University of Leuven, and the chair of the CEFR special interest group in ALTE. His primary research interests include justice, validity, the CEFR, and university entrance language testing. Recent articles have appeared in *Language Testing*, *Language Assessment Quarterly*, *Assessing Writing*, and *The Modern Language Journal*.

Professor James Simpson

James Simpson is a Senior Lecturer in Language Education at the School of Education, University of Leeds, where he has worked since 2004 and where he leads the Language Education academic group. His main research interests are the teaching and learning of English for multilingual students in migration contexts and the sociolinguistic study of urban multilingualism. He is the co-author of *ESOL: A Critical Guide* (OUP, 2008, with Melanie Cooke), the editor of *The Routledge Handbook of Applied Linguistics* (2011), and the co-editor (with Anne Whiteside) of *Adult Language Education and Migration: Challenging Agendas in Policy and Practice* (Routledge, 2015).

Professor Kris Van den Branden

Kris Van den Branden is a professor of linguistics and a teacher educator at the Faculty of Arts of the Katholieke Universiteit Leuven. At the same university, he is one of the academic promoters of the Centre for Language and Education. He has published a wide range of articles on task-based language teaching, and is one of the volume series editors (together with Martin Bygate and John Norries) of *Taskbased language teaching: Issues, research and practice* (published by John Benjamins). He is the editor (together with Elke Peters) of the journal *ITL International Journal of Applied Linguistics*.

Impact: perspectives

There are many ways to consider the impact of language tests, and people may vary quite a lot in their approach to this topic. That is why we asked researchers, policy makers, test takers and test developers to give us their take on what test impact is, what it does, and what it should be. Collected, these texts offer a stimulating range of empirical observations, thought-provoking ideas, fundamental critiques, and idealistic aspirations.

Thank your inspiring contributions, Nick, Kamran, Kimberly, Cecilie, Ricky, Henna, Emyr, Ina, Petra, Steven, Laurie, Gemma, Sabrina, and Gerriet.

Nick Saville, Cambridge English, UK

Impact research investigates and seeks to understand the effects and consequences which result from the use of assessment in educational contexts and throughout society. As a field of enquiry, it appeared in the language testing literature as an extension of washback in the 1990s.

Milanovic and Saville (1996) proposed an early model of test impact designed to meet the needs of examination providers. By conceptualising impact within the process of validation, there was an attempt to integrate impact research into routine procedures for accumulating validity evidence. This model recognised that a proactive approach is needed to achieve intended effects and consequences. Saville (2009) proposed the concept of “positive impact by design” as a key feature of an expanded impact model, that starts from the premise that assessment systems must be designed from the outset to achieve positive impacts and takes an ex ante approach to anticipating the possible consequences of using the test in particular contexts. This builds on and expands Messick’s idea of achieving “validity by design as a basis for washback”.

Adequate specification and communication of the focal constructs is crucial for ensuring that the test is appropriate for its purpose and contexts of use and to counter threats to validity. However, although appropriate construct representation is a necessary condition for achieving the anticipated outcomes, it is not sufficient and impact by design highlights the importance of implementing assessment systems that explicitly incorporate considerations related to the social and educational contexts of learning/teaching and test use. This entails effective communication and collaboration between all stakeholders.

Understanding the nature of context within educational systems and the roles of stakeholders in those contexts are clearly important considerations. It is now widely recognised that educational processes take place within complex systems with dynamical interplay between many sub-systems and “cultures” and so an understanding of the roles of stakeholders as participants is a critical factor in bringing about intended changes. In conducting impact research the aim is to understand better the interplay between the macro and micro contexts within the society where the tests are being used and to determine which elements facilitate or hinder the desired outcomes.

Finally, the ability to change in order to improve educational outcomes or mitigate negative consequences associated with examinations is perhaps the most important dimension of the model. Managing change is therefore critical to this “theory of action” and it requires examination providers to work closely with stakeholders in their own contexts.

Kamran Khan, University of Leicester, UK

For the last seven years, I have researched the impact that citizenship tests have on the lives of migrants in the UK. In that time, I have been involved in two studies. The first was a yearlong ethnography following a migrant through the citizenship process. This study charted his reflections on the citizenship test right through to the end of the citizenship ceremony. In the second study, I

interviewed 158 people as part of an ESRC project at the University of Leicester (www2.le.ac.uk/departments/sociology/research/uk-citizenship-process). During the data collection for these studies, I have sat down with over 150 people, all of whom were involved in the citizenship process whether hoping to take the test some day or having already passed, and looked them in the eye as they have recounted their experiences, expectations, hopes and fears. In relation to the test impact of citizenship testing, I will outline three reflections on my work below.

The first point worth mentioning is that citizenship testing is highly exclusionary even prior to taking the test. We have found some participants, who were participative citizens in other ways, unable to deal with the test whether due to fear about their lack of computer literacy or the immense pressure of the financial burden and time required which is multiplied in the case of a test fail. Rather than providing a facilitative pathway, there are some for whom the testing procedure exacerbates their marginalisation.

The second point is that while tests are one form of assessment, a testing regime opens up of other forms of assessment within a broader architecture of belonging. Tests themselves provide a technical assessment of language, and ideological goal of privileging a dominant language while asking the individual to symbolically demonstrate their capacity to assimilate (Khan & Blackledge, 2015, 2018). We have found too that testing conditions and elongates the naturalisation process. That is to say, assessment permeates other aspects of the process including during citizenship ceremonies and even after the ceremony, through passport interviews. In the administration and paperwork around testing and naturalisation processes, there is the certification of language proficiency, references of good character and of course, the application itself. While individuals are engaged in the citizenship process there are aspects of their life that may be suspended such as travelling, finances and access to certain jobs this is all in addition to the uncertainty and anxiety around how long the process will take. This 'psychic life of citizenship policy' (Fortier, 2017) is embodied by the test and experienced by the individual as one encounter with the State within a broader process, which can at times appear to be never ending.

The third point is that the responsibility for test preparation has shifted towards the individual in the UK. Whereas there had previously been 'ESOL (English for Speakers of Other Languages) with Citizenship' classes for lower level learners, this has been abolished and now all applicants must take the citizenship test. For those with work long hours due to their socioeconomic circumstance and/or caring responsibilities, test preparation becomes a highly onerous task. We have seen how women of colour, often mothers, from non-European, non-English speaking countries have especially been disproportionately affected in lacking time and resources to learn and having had a potentially facilitative pathway taken away. Thus, while citizenship tests are often talked about as instruments of integration, they can also act as a double-edged sword which consolidates marginalisation.

Kimberly, test taker

The first time I took the test I didn't speak as fluently as I do now, but my grammatical knowledge was just as good and I paid more attention to syntax. But I could not say a word because I was so nervous. After the first part of the oral test, I did not remember what I had said. I was so disappointed and thought: OK, this is not going to work. I had worked so hard, and if I did not pass, everything would just end.

Cecilie Hamnes Carlsen, Kompetanse Norge, Norway

I want our test to be fair, just and to promote learning. I want it to open doors to education and jobs. I want it to give all test-takers, regardless of educational background or aims, a chance to show their language abilities. I want learners with little prior schooling and low levels of literacy to have a chance to get a good score in listening and speaking. I want teachers to trust that when they prepare their

students for the test, they will learn the language. I want learners to have enough material to practice so that they won't need to fear the unexpected on the day of the test.

I do *not* want our test to be a gatekeeper to democratic rights and safety for those who need shelter - our test measures language, not integration or the will to integrate. I want our test certificate to make it impossible for test users to hide behind subjective and undefined language demands. I want society, employers and policy makers to know what the test measures, and what it does not measure, what good use and what misuse of test results is.

To achieve this, we work to meet the professional standards of good test development, as described for instance in the 17 Minimum Standards of ALTE. We listen to teachers, call upon their advice, and make them know how valuable it is. We strive towards a good dialogue with policy makers and employers, so that, at least, we can prevent misuse of test results based on ignorance. We give courses in assessment and take part in the public debate in an attempt to raise test literacy and consciousness about possible misuse of tests so that more people will speak up against injustice.

Ricky van Oers, Radboud University, the Netherlands

Language and knowledge of society tests applied as a condition for naturalisation in Germany, the Netherlands and the United Kingdom lead to the exclusion from full-fledged citizenship, and the accompanying rights, of certain categories of immigrant. Part of the immigrant population is barred from becoming citizens for reasons of age, intellectual capacity, war trauma, money or time. The exclusion produced by citizenship tests forms a barrier to equality and the execution of democratic rights, and therefore creates problems from a liberal-democratic perspective. Furthermore, it makes one wonder whether the tests are effective means to achieve the main goal the tests are meant to achieve: increasing immigrant integration. After all, barring immigrants from naturalisation will not increase their language skills, nor their overall integration in the host society. It is moreover questionable to what extent the knowledge tested will contribute to integration. For instance, the level of language skills strived for is often too low for actual participation in the employment market and employers attach no or only little value to the certificates obtained.

At the same time, however, the Dutch case shows that the duty to integrate under the Integration Act and the obligation to pass a test as a condition for naturalisation have led to the participation of large numbers of immigrants in funded language courses. The Dutch government stopped funding language and integration courses per 1 January 2013. Statistics have shown a considerable drop in the numbers of people following courses and passing the exam since. Furthermore, as not passing the exam within the time limit of three years set by the Integration Act is fined and may have consequences for the right of residence, few immigrants aim to achieve a higher level of language skills than A2, the minimum level required. Funding of language and integration courses hence appears to be crucial for integration.

Henna Tossavainen, University of Jyväskylä, Finland

An issue relating to test impact is the unintended use of test scores, which we have tried to overcome in the North Sami language test in the Finnish National Certificates LS/foreign language test system.

North Sami is currently tested as other foreign language tests (e.g. French, Spanish, Italian) via the two official languages, Finnish and Swedish. Consequently, candidates will have to know either of these languages to be able to participate. The Sami language has official status in some municipalities in Finnish Lapland, but has no official recognition elsewhere in the country. The fact that the Sami language is tested via the two official languages underlines the non-official status of the language and does not really support the rights the Sami have to their language. This is particularly so, because the test has been taken by L1 Sami speakers who do not have any other official means of showing their language skills.

There is a rather straightforward solution to the problem, which is to redesign it to a one-language L2 test, similar to the existing L2 Finnish and Swedish tests. It would have the implication that the test could also be taken by further test taker groups with no Finnish or Swedish skills. This process has already started in the test system and it will take some time. However, when it is launched it will hopefully the different test taker groups' participation and also the promotion and maintenance of this minority language.

Emyr Davies, CBAC-WJEC, Wales

Welsh is a minority language, and candidates of the Welsh for Adults test are obviously fewer in number than those taking widely used languages. Approximately 1,500 candidates a year take the tests across 4 levels (A1-B2) in centres across Wales. We would *like* our tests to have an impact a micro- and macro- level.

In terms of impact on individual candidates, the tests should strengthen motivation and confidence, improve learning, give learners achievable goals, and raise *ambition* of learners and their tutors. The tests should also have an effect on classroom practice, e.g. spoken interaction has a higher weighting in these tests (around 50%), in order to ensure that tutors give adequate focus on this skill. If tutors will 'teach to the test', then that will still have a positive washback and raise standards.

On a societal level, we would *like* our tests to encourage candidates to *use* the target language, rather than learning Welsh as a hobby or academic interest. The very existence of tests raises awareness and status of Welsh learners in the wider society and adds a *seriousness* to the whole endeavour of learning a minority language, in a context where everyone can revert to English. However, the widest and most long-term 'ripple' (in a concentric model of impact), is the effect on the Welsh language itself. It is a stated aim of Welsh Government to increase the the number of Welsh speakers, and raising the number and attainment of adult Welsh learners is an important strand in their strategy. The tests are not just a measure of this, but a means to an end.

Ina Ferbežar and Petra Likar Stanovik, Univerza v Ljubljani, Slovenia

Being high-stakes, Slovene language tests are automatically seen as gatekeepers. As such, the main aim of test takers is to pass the exam for the doors it opens, but not to learn the Slovene language. At the University of Ljubljana, the Centre for Slovene as a Second and Foreign Language would like candidates to realize that the tasks on our tests might be useful for real-life communication. That is why we regularly ask test takers about their Slovene language needs and practices, and their answers are taken into consideration when developing tests and designing courses.

Laurie, test taker

It's normal that students need a certain language level before they can attend university. It's important for the students themselves, because if they can't understand Dutch, what can they do here? Nothing! If they don't know the language they will definitely not succeed!

Steven Verheyen, CNaVT, Belgium

Because of their summative nature, the CNaVT tests are often used for gatekeeping purposes and as a result likely to elicit teaching to the test. From these realizations follows the task-based and domain-specific character of the CNaVT tests. By employing functional tasks that stakeholders recognize to be meaningful for the domain to which entrance is sought, the CNaVT aims to foster motivation for learning, raise the standards for language teaching, and open doors by providing test takers the opportunity to demonstrate their mastery of the language skills required to successfully operate in the target domain. For some test takers doors will be closed, however. Although because of its

functional, domain-specific nature the CNaVT tests can serve as safeguards for test takers with little chances of success in the target domain, the use of language tests for gatekeeping purposes is never foolproof. This is why the CNaVT entertains a cyclical renewal and validation procedure which involves yearly interaction with relevant stakeholders, invests in research into the fairness of its tests, and also engages in communication with stakeholders about what its tests *cannot* do.

Gemma Macho Aguillo, Department of Education, Basque Country

The certificate of proficiency in Basque EGA has the aim to assess the language ability at the C1 level according to the CEFR of people who need to study, work or live where Basque is used as the language of communication. Candidates need Basque to study or work or they live where it is used as the language of communication. Due to the circumstance that Basque is a minoritized language, candidates can have it as a second or other language or can also be native speakers.

Research about the impact of the test is, for many reasons, interesting for the examination board, policy makers, and all other stakeholders (candidates, employers, etc.). There are more than 100,000 people who have already got the EGA certificate during more than 30 years and it has become the most popular certification of Basque. A very frequent question made to speakers of Basque is: You can speak Basque, yes, but do you have EGA? This can be an indicator of the desired positive impact of the test as regards to social acceptance, but all other forms of impact and washback must also be taken into account.

Sabrina Machetti, Università per Stranieri di Siena, Italy

The CILS Centre (University for foreigners of Siena) is responsible for producing proficiency exams of Italian as a second and foreign language, in Italy and around the world. These exams are given to approximately 24,000 candidates annually. All the CILS exams test all areas of language ability, from A1 to C2. CILS exams and their results have a significant impact on the educational or professional opportunities of individual test takers. Since 2010, the CILS A2 is required for the issuance of a permit for long-term residents. The awareness that the CILS exams can have consequences on educational systems, and on society more widely has growing year after year. At the same time, we are aware that assertions about the impact in language testing need empirical evidence continuously. This is especially true for the educational context, where our tests directly influence learning and teaching processes. As such, the CILS Centre is involved in research projects concerning the effect and consequences a test can have beyond the classroom and immediate learning context.

Gerriet Janssen, Universidad de los Andes, Colombia

Bachman and Palmer (1996) describe three ways in which tests impact different stakeholders: washback, feedback, and the decisions –uses– that are made about test-takers based on their exam scores. I want to argue that this conceptualization describes at its heart the process of assessment–and–learning. I also want to emphasize that the implementation of this process in localized contexts needs to truly acknowledge the different structures that build and define these different contexts—and that in terms of assessment literacy, we all still have a lot to learn.

To begin, I value Bachman and Palmer’s understanding of the construct of impact as being comprised of cycles of washback, feedback, and test use. Starting with (positive) washback and feedback, stakeholders can work towards learning gains in well-reasoned and clearly stated target domains, reflected across many assessment moments (i.e., feedback). Possibly, one of these moments could be a national, standardized exam. Today, I see most publishers taking great strides to state clear learning objectives in their coursebooks and then create ties between these objectives, their tests’ constructs, and the specific feedback they give concerning a test-taker’s performance. In this vein in post-conflict Colombia, though not related to language instruction, is the recent case of citizenship competencies

that have been defined as a target construct, to be developed in classrooms and assessed on nationally administered exams. These sorts of examples highlight an ideal type of unified assessment–and–learning culture, which could or should be the foundation of our schools’ curriculum.

Much more problematic, however, is the element of test-uses and decision-making. It is easy to provide many examples of test misuse, such as: “the Dean chose the 110 TOEFL score and C2 IELTS band score for entrance into our graduate program in [Colombian university] because that’s what they do at Harvard.” Uffffff. In decisions such as these, I see the heartbreaking duality of desiring to replicate what is done elsewhere so as to become this other thing, all the while not considering how local and foreign contexts and structures are fundamentally different. And when I say different I mean really, breathtakingly different: a regional Colombian university is simply structurally a really different creation, embedded within a really different environment, with different opportunities and possibilities, with actors who arrive with very different life experiences than that other university over there in that enclave of Cambridge. As a result of these wide differences, obstacles emerge, especially in projects related to assessment literacy and specific questions of appropriate test use / decision making. I think these local characteristics are oftentimes, or nearly always, left to one side. For me, this situation suggests that qualities such as co-constructed, localized, and similar, become incredibly important when thinking about the impacts of test-use, decision-making, or assessment literacy.

From here, I keep finding myself jumping to several final thoughts: Do we really need tests that create these sorts of test-use and decision making situations? What happens in terms of these same test-use and decision-making situations, when we don’t exactly know what “C1” or “C2” or “advanced language proficiency” mean? When “C1” or “C2” may well have different localized social implications? When this entire design has been imported from abroad, with no regard for the way things work here? With these questions in mind, I think that we may want to question our own assessment literacy, as part of the ongoing assessment literacy agenda.

See also

Khan, K., & Blackledge, A. (2015). ‘They Look into our lips:’ Negotiation of the citizenship ceremony as authoritative discourse. *Journal of Language and Politics*, 14(3), 382-405.

Saville, N. (2009). Developing a model for investigating the impact of language assessment within educational contexts by a public examination provider. Bedfordshire: University of Bedfordshire dissertation.

van Oers, R. (2014). *Deserving Citizenship. Citizenship Tests in Germany, the Netherlands and the United Kingdom*. Leiden/Boston: Brill.

Notes

Notes

