ALTE Bibliography of Members' Work

JUNE 2022 VERSION

ALTE has collated references of papers, journal articles, books, etc., written by representatives of member organisations. They have been categorised into the following areas:

ALTE welcomes further additions from members to this list, or requests for works to move into another, or a new, category.

Key ALTE Documents

ALTE (1994) *The ALTE Code of Practice* [BG] [CA] [CY] [DE] [DK] [EE] [ES] [EU] [FI] [FR] [GA] [GL] [IT] [LT] [LV] [NL] [NO] [PL] [RU] [SL] [SV], available online: https://www.alte.org/Materials

ALTE (2001) *Development and Descriptive Checklist for Tasks and Examinations: General*, available online: https://www.alte.org/Materials

ALTE (2001) *ALTE Quality Assurance Checklists, Unit 1: Test Construction*, available online: https://www.alte.org/Materials

ALTE (2001) *ALTE Quality Assurance Checklists, Unit 2: Administration and Logistics*, available online: https://www.alte.org/Materials

ALTE (2005) *ALTE Materials for the Guidance of Test Item Writers*, available online: https://www.alte.org/Materials

ALTE (2007) *Minimum Standards for Establishing Quality Profiles in ALTE Examinations* [BG] [CA] [CY] [CZ] [DE] [EE] [ES] [EU] [FI] [FR] [GA] [GL] [HU] [IT] [LT] [NL] [NO] [PL] [RU] [SL] [SV], available online: https://www.alte.org/Materials

ALTE (2011) *Manual for Language Test Development and Examining* [EU] [FR] [DE] [NL], available online: https://www.alte.org/Materials

ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, available online: http://events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

ALTE (2018) *Procedures for Auditing*, available online: https://www.alte.org/Materials

ALTE (2020), *The History of ALTE – the Association of Language Testers in Europe: the first 30 years*, available online: https://www.alte.org/Our-history

ALTE (2020) *Principles of Good Practice* [EU], available online: https://www.alte.org/resources/Documents/ALTE%20Principles%20of%20Good%20Practice%20Online%20(Final).pdf

ALTE (2021) *Collated Papers for the ALTE 7th International Conference, Madrid*, available online: https://www.alte.org/IntConfProceedings

ALTE Members (1999) *Multilingual Glossary of Language Testing Terms*, Studies in Language Testing volume 6, Cambridge: UCLES/Cambridge University Press.

Beresford-Knox, N (2015) The role of Quality Management in ensuring accurate and meaningful test scores, *Research Notes* 59, 40–44.

Docherty, C and Barker, F (Eds) (2016) *Language Assessment for Multilingualism: Proceedings of the ALTE Paris Conference, April 2014*, Studies in Language Testing volume 44, Cambridge: UCLES/Cambridge University Press.

Galaczi, E D and Weir, C J (Eds) *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press.

Milanovic, M and Weir, C J (Eds) (2004) *European Language Testing in a Global Context: Proceedings of the ALTE Barcelona Conference, July 2001*, Studies in Language Testing volume 18, Cambridge: UCLES/Cambridge University Press.

Parizzi, F (2010) Training auditors: The ALTE system, *Research Notes* 39, 28–31.

Taylor, L and Weir, C J (Eds) (2008) *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity – Proceedings of the ALTE Berlin Conference, May 2005*, Studies in Language Testing volume 27, Cambridge: UCLES/Cambridge University Press.

Taylor, L and Weir, C J (Eds) (2009) *Language Testing Matters: Investigating the Wider Social and Educational Impact of Assessment – Proceedings of the ALTE Cambridge Conference, April 2008*, Studies in Language Testing volume 31, Cambridge: UCLES/Cambridge University Press.

CEFR

ALTE (2002) *The ALTE Can Do Project*, available online: https://www.alte.org/Materials

ALTE (2002) *ALTE Can Do Statements* [CA] [DE] [DK] [ES] [FI] [FR] [IT] [NL] [NO] [PT] [SV], available online: https://www.alte.org/Materials

Althaus, H-J (2018) Warum C1 keine Lösung ist: Der Nachweis von Deutschkenntnissen für den Hochschulzugang, der GER und warum sie nicht zusammenpassen, in Brandt, A, Buschmann-Göbels, A and Harsch, C (Eds) *Der Gemeinsame Europäische Referenzrahmen für Sprachen und seine Adaption im Hochschulkontext*, Bochum: AKS-Verlag, 78–98.

Beacco, J-C and Porquier, R (Eds) (2007) *Niveau A1 pour le français: (utilisateur/apprenant élémentaire); un référentiel*, Paris: Didier.

Beacco, J-C, Blin, B, Houles, E, Lepage, S and Riba, P (2011) *Niveau B1 pour le français: (utilisateur/apprenant indépendant); niveau seuil; un référentiel*, Paris: Didier.

Beacco, J-C, Bouquet, S and Porquier, R (2007) *Niveau B2 pour le français: (utilisateur, apprenant indépendant); un référentiel*, Paris: Didier.

Beacco, J-C, De Ferrari, M, Lhote, G and Tagliante, C (2006) *Niveau A1.1 pour le français: publics adultes peu francophones, scolarisés, peu ou non scolarisés; référentiel et certification (DELF) pour les premiers acquis en français*, Paris: Didier.

Beacco, J-C, Lepage, S, Porquier, R and Riba, P (2008) *Niveau A2 pour le français: (utilisateur/apprenant élémentaire); niveau intermédiaire; un référentiel*, Paris: Didier.

Breton, G, Lepage, S and North, B (2008) *Cross-language benchmarking seminar to calibrate examples of spoken production in English, French, German, Italian and Spanish with regard to the six levels of the Common European Framework of Reference for Languages (CEFR)*, available online: https://rm.coe.int/168045a0cd

British Council, UKALTA, EALTA and ALTE (2022) *Aligning Language Education with the CEFR: A Handbook,* available online: https://www.alte.org/Materials

Carlsen, C H (2010) Discourse connectives across CEFR-levels: A corpus-based study, in Bartning, I, Martin, M and Vedder, I (Eds) *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*, EUROSLA Monograph Series 1, available online: http://eurosla.org/monographs/EM01/191-210Carlsen.pdf

Carlsen, C H (2010) *Learner corpora, contrastive rhetoric and the CEFR – what's the link?*, paper presented at IATEFL TEA SIG-EALTA conference, Barcelona, October 2010.

Carlsen, C H (Ed) (2013) *Norsk profil. Det felles europeiske rammeverket spesifisert for norsk. Et første steg*, Oslo: Novus Forlag.

Carlsen, C H, Deygers, B, Vîlcu, D and Zeidler, B (forthcoming) The application of the CEFR in the European labour market, in Huhta, A, Erickson, G and Figueras, N (Eds) *Sauli Takala Commemorative*.

Carlsen, C H, Deygers, B, Zeidler, B and Vîlcu, D (forthcoming) CEFR-based language requirements in the European labour market, in Huhta, A, Erickson, G and Figueras, N (Eds) *Sauli Takala Commemorative*.

Casanova, D, Crendal, A, Demeuse, M, Desroches, F and Holle, A (2010) Validation empirique d'un test de français langue étrangère en regard du Cadre européen commun de référence pour les langues, in *Actes du 22e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 14 au 16 janvier 2010, Braga, Portugal*, Neuchâtel, Switzerland: ADMEE-Europe, 881–902.

Crendal, A (2005) Vers la multiréférentialisation du TEF, in *Points Communs n°24*, Paris: Chambre de commerce et d'industrie de Paris.

Demeuse, M, Artus, F, Desroches, F, Crendal, A, Renaud, F and Casanova, D (2006) Évaluation et autoévaluation en français langue étrangère: approche empirique du Cadre européen commun de référence pour les langues du Conseil de l'Europe à travers le TEF et l'échelle d'autoévaluation DIALANG, in *Actes du 19e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 11 au 13 septembre 2006, Luxembourg, Luxembourg*, 1–7, available online: https://www.researchgate.net/publication/272488221_Evaluation_et_autoevaluation_en_fran cais_langue_etrangere_approche_empirique_du_Cadre_europeen_commun_de_reference_ pour_les_langues_du_Conseil_de_l%27Europe_a_travers_le_TEF_et_l%27echelle_d%27a utoevaluatio

Demeuse, M, Desroches, F, Crendal, A, Renaud, F, Oster, P and Leroux X (2004) L'évaluation des compétences linguistiques des adultes en français langue étrangère dans une perspective de multiréférentialisation, in *Actes du 17e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE Europe), 18 au 20 novembre 2004, Lisbonne, Portugal*, Neuchâtel, Switzerland: ADMEE-Europe.

Deygers, B and Van Gorp, K (2013) The influence of the CEFR on rating scale design, in Colpaert, J, Simons, M, Aerts, A and Oberhofer, M (Eds) *Language Testing in Europe: Time for a new Framework?*, Antwerp: University of Antwerp, 97-103.

Deygers, B and Van Gorp, K (2015) Determining the scoring validity of a co-Constructed CEFR-based rating scale, *Language Testing* 32 (4), 521–541.

Deygers, B, Carlsen, C H, Saville, N and Van Gorp, K (2018) The use of the CEFR in Higher Education: A brief introduction to this special issue, *Language Assessment Quarterly* 15 (1), 1–2.

Deygers, B, Carlsen, C H, Van Gorp, K and Saville, N (Eds) (2018*)* Language tests for academic enrolment and the CEFR, *Language Assessment Quarterly* 15 (1), Special Issue.

Deygers, B, Van Gorp, K and Demeester, T (2018) The B2 level and the dream of a common standard, *Language Assessment Quarterly* 15 (1), 44–58, available online: https://doi.org/10.1080/15434303.2017.1421955

Deygers, B, Zeidler, B, Vîlcu, D and Carlsen, C H (2017) The CEFR and its impact on university entrance policies: a state of affairs in Europe, *Language Assessment Quarterly* 1, 1-13.

Deygers, B, Zeidler, B, Vilcu, D and Carlsen, C H (2018) One Framework to Unite Them All? Use of the CEFR in European University Entrance Policies, *Language Assessment Quarterly* 15 (1), 3–15, available online: http://dx.doi.org/10.1080/15434303.2016.1261350

Folny, V (2012) Procesos institucionales de diseño, elaboración administración y calificación de pruebas de certificado de niveles C, in Ortega Calvo, A (Ed) *Niveles C: currículos, programación, enseñanza y certificación*, Madrid, Spain: Ministerio de Educación, Cultura y Deporte de España, 293–313.

Folny, V (2015) *Adossement des épreuves d'expression orale et écrite du Test de connaissance du français (TCF) sur les Niveaux de compétences linguistiques canadiens (NCLC) et correspondance avec les niveaux du Cadre européen commun de référence pour les langues (CECRL)*, rapport interne, Sèvres: Centre International d'études pédagogiques, 61.

Folny, V, Jaegher, D D, Butasova, A and Buta, J (2011) Utilisation du CECRL pour développer des tests dans six langues pour les professeurs des écoles: avantages et limites, *Synergies Europe* 6, 51–61.

Glaboniat, M (2005) Sprachbegegnungen auf Schreibtischen von Entscheidungsträgern. Ein Plädoyer für international vergleichbare Sprachbeschreibungsstandards, *ide* 2, 32–43.

Glaboniat, M (2006) Sprachkenntnisse beschreibbar, vergleichbar und messbar gemacht!? Über den Nutzen und die Grenzen der Europaratsskalen im Bereich des Prüfens und Zertifizierens, in Krumm, H-J and Portmann-Tselikas, P (Eds) *Begegnungssprache Deutsch: Motivation, Herausforderung, Perspektiven*, Innsbruck, Vienna, Bolzano: StudienVerlag, 81–94.

Glaboniat, M (2013) Testen, Prüfen und Beurteilen von Deutschkenntnissen — unter dem besonderen Aspekt des GER, in Ahrenholz, B and Oomen-Welke, I (Eds) *Deutsch als Fremdsprache*, Baltmannsweiler: Schneider Verlag Hohengehren, 287–297.

Glaboniat, M (2016) Der Gemeinsame europäische Referenzrahmen für Sprachen (GER) und Profile deutsch – 12 Jahre später, in Drumbl, P (Ed) *Heterogenität in Lernsituationen. Konferenzbeiträge IDT 2013*, Bolzano: Bozen-Bolzano University Press, 7–22.

Glaboniat, M and Müller, M (2006) Note „Sehr gut!" – Aber in Bezug worauf? Referenzrahmen und Profile Deutsch in ihren Auswirkungen auf Prüfungen und Tests, *Fremdsprache Deutsch* 2006/34, 14–21.

Glaboniat, M, Müller, M and Wertenschlag, L (2003) "Profile deutsch" - Lernzielbeschreibungen und sprachliche Mittel für Deutsch als Fremdsprache auf vier Niveaustufen des europäischen Referenzrahmens, in Schneider, G and Clalüna, M (Eds) *Mehr Sprache - mehrsprachig - mit Deutsch. Didaktische und politische Perspektiven*, Munich: Iudicium Verlag, 247–255.

Grego Bolli, G (2004) Il Consiglio d'Europa, il Quadro Comune Europeo e la Certificazione linguistica. L'esperienza dell'Università per Stranieri di Perugia e dell'ALTE, in Jafrancesco, E (Ed) *Le tendenze innovative del Quadro Comune Europeo di Riferimento per le Lingue e del Portfolio*, ATENE: Edilingua, 45–63.

Grego Bolli, G (2006) Il "Quadro" e la valutazione, in Mezzadri, M (Ed) *Integrazione Linguistica in Europa*, Torino: Utet, 221–246.

Grego Bolli, G (2007) La certificazione linguistica in relazione al Quadro Comune Europeo di Riferimento, *Babylonia* 1, 38–39.

Grego Bolli, G (Ed) (2008) *Esempi di produzioni orali. A illustrazione per l' italiano dei livelli del Quadro comune europeo di riferimento per le lingue del Consiglio d'Europa*, Perugia: Guerra Edizioni.

Grego Bolli, G (2008) Progetti Europei: nuove prospettive sulla scia del Quadro Comune Europeo di Riferimento per le lingue, in Ciliberti, A (Ed) *Un mondo di italiano*, Perugia: Guerra Edizioni, 25–48.

Grego Bolli, G (2009) Productions orales illustrant les 6 niveaux du Cadre européen commun de référence pour les langues (Commentaires sur les niveaux attribués en italien), in *Productions orales illustrant les 6 niveaux du Cadre européen commun de référence pour les langues*, SÉVRÉS: Centre International d'Etudes Pédagogiques (CIEP), 1–22.

Grego Bolli, G (2010) I progetti del Consiglio d'Europa in ambito di educazione linguistica e il Profilo della lingua italiana, in Spinelli, B and Parizzi, F (Eds) *Profilo della lingua italiana. I livelli di riferimento del QCER A1, A2, B1, B2*, Firenze: La Nuova Italia, 1–10.

Grego Bolli, G (2010) Il progetto del Profilo della lingua italiana, Il Centro per la Valutazione e la Certificazione Linguistica (CVCL) e il Consiglio d'Europa. Breve storia e ragioni di una collaborazione di lunga data, in Spinelli, B and Parizzi, F (Eds) *Profilo della lingua italiana. I livelli del QCER A1, A2, B1, B2*, Firenze: La Nuova Italia, 113–123.

Huhta, A (2012) Common European Framework of Reference, in Chapelle, C (Ed) *The Encyclopedia of Applied Linguistics*, Chichester, UK: Wiley-Blackwell.

Huhta, A (2016) Using the Common European Framework of Reference in the evaluation of educational achievement in foreign and second languages, in *Proceedings of the Second*

International Conference for Assessment & Evaluation: 'Learning Outcomes Assessment', Riyadh, Saudi Arabia, 1-3 December 2015, Riyadh: Saudi Arabia Ministry of Education, 324–342.

Huhta, A and Figueras, N (2004) Using the CEF to promote language learning through diagnostic testing, in Morrow, K (Ed) *Insights from the Common European Framework*, Oxford: Oxford University Press, 65–76.

Kecker, G (2011) *Validierung von Sprachprüfungen: Die Zuordnung des TestDaF zum Gemeinsamen europäischen Referenzrahmen für Sprachen*, Frankfurt am Main: Peter Lang.

Kecker, G (2016) Der GeR als Referenzsystem für kompetenzorientiertes Testen: Was bedeutet der Bezug zum GeR für eine Sprachprüfung?, *Zeitschrift für Fremdsprachenforschung* 27 (1), 13–37.

Kecker, G and Eckes, T (2010) Putting the Manual to the test: The TestDaF–CEFR linking project, in Martyniuk, W (Ed) *Aligning Tests with the CEFR: Reflections on Using the Council of Europe's Draft Manual*, Studies in Language Testing volume 33, Cambridge: UCLES/Cambridge University Press, 50–79.

Khalifa, H, ffrench, A and Salamoura, A (2010) Maintaining alignment to the CEFR: the FCE case study, in Martyniuk, W (Ed) *Aligning Tests with the CEFR: Reflections on Using the Council of Europe's Draft Manual*, Studies in Language Testing volume 33, Cambridge: UCLES/Cambridge University Press, 80–101.

Lim, G, Geranpayeh, A, Khalifa, H and Buckendhal, C (2013) Standard setting to an international reference framework: Implications for theory and practice, *International Journal of Testing* 13, 32–49.

Mavel, M and Riba, P (2008) L'harmonisation du DELF et du DALF sur les niveaux du Cadre européen commun de référence pour les langues, in Taylor, L and Weir, C J (Eds) (2008) *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity – Proceedings of the ALTE Berlin Conference, May 2005*, Studies in Language Testing volume 27, Cambridge: UCLES/Cambridge University Press, 79–93.

North, B and Docherty, C (2016) Validating a set of CEFR ilustrative descriptors for mediation, *Research Notes* 63, 24–30.

North, B and Panthier, J (2016) Updating the CEFR descriptors: The context, *Research Notes* 63, 16–23.

Platon, E (2008) Validarea competenţelor lingvistice în conformitate cu Cadrul European Comun de Referinţă şi cu Paşaportul lingvistic European, in *România Orientale XXI, Atti del Convegno La lingua rumena: proposte culturali per la nuova Europa, Roma, 15-17 novembre 2007*, Roma: Bagatto Libri, 306–313.

Puigdomènech, L (2014) El Marc europeu comú de referència i el nivell C2: l'experiència de la nova prova de nivell superior, *Llengua i Ús* 54.

Riba, P (2016) *Niveaux C1/C2 pour le Français (utilisateur expérimenté) niveau autonome/maîtrise: éléments pour un référentiel*, Paris: Didier.

Saville, N (2016) *Developing and validating test materials within a Common Framework of Reference (based on the ALTE manual)*, paper presented at the NEAA, Hangzhou, China, November 2016.

Saville, N (2017) *Appropriate uses of the CEFR in planning and implementing language education reforms*, paper presented at the 52nd Regional Language Centre (RELC) International Conference, Singapore, March 2017.

Shackleton, C (2018) Linking the University of Granada CertAcles listening test to the CEFR, *Revista de Educación* 381.

Tagliante, C (2007) *L' évaluation et le Cadre européen commun*, New Edition, Paris: CLE International.

Tagliante, C and Debyser, F (2001) *Portfolio européen des langues: pour jeunes et adultes (Conseil de l'Europe)*, Paris: Didier: Centre Régional de Documentation Pédagogique de Basse-Normandie.

Tagliante, C and Deleuse, J (2002) Les portfolios européens des langues, *Revue internationale d'éducation de Sèvres* 31, 138–142, available online: https://doi.org/10.4000/ries.1979

Vîlcu, D (2014) Relaționarea examenelor de limba română ca limbă străină la CECR, in Platon, E and Arieșan, A (Eds) *40 de ani de limba română ca limbă străină la UBB. 1974-2014, Cluj-Napoca, 3-4 octombrie 2014*, Cluj-Napoca: Casa Cărții de Ştiinţă, 20–29.

Vodičková, K (2011) Modelling language devices for the Czech reference level description at level C1, in *Профессионально ориентированное обучение иностранному языку и переводы в ВУЗе*, Moscow: Российский университет дружбы народов, 469–471.

Vodičková, K (2015) Zkoušky z češtiny a Společný evropský referenční rámec pro jazyky: Na co se spolehnout? [Exams in Czech and the CEFR: What to rely on?], *CASALC Review* 2, 176-188, available online: https://drive.google.com/file/d/10roHF5VILIzf4p2tp9GNPyJAOP97aJSJ/view

Vodičková, K, Pečený, P and Nováková, J (2012) Specifikace Certifikované zkoušky z češtiny pro cizince a Společný evropský referenční rámec [The Czech Language Certificate Exam specifications and the CEFR], in *Výuka a testování cizích jazyků vkontextu Společného evropského referenčního rámce (SERR)*, Praha: ÚJOP UK, 97–100.

Zimmermann, S and Marks, D (2018) Integrierte Prüfungsleistungen mit dem GER beurteilen: Was die Kann-Beschreibungen nicht können, in Brandt, A, Buschmann-Göbels, A and Harsch, C (Eds) *Der Gemeinsame Europäische Referenzrahmen für Sprachen und seine Adaption im Hochschulkontext*, Bochum: AKS-Verlag, 158–166.

Diagnostic Tests

Huhta, A, Luoma, S, Oscarson, M, Sajavaara, K, Takala, S and Teasdale, A (2002) DIALANG - A Diagnostic Language Assessment System for Learners, in Alderson, J C (Ed) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies*, Strasbourg: Council of Europe, 130–145.

Vodičková, K (2014) Designing diagnostic tests in Czech for young migrant learners, in *Efficiency and Responsibillity in Education*, Praha: Česká zemědělská univerzita, 891–897.

Fairness, Justice and Ethics

Arras, U and Müller, A (2012) Fairness bei der Leistungsmessung: Der TestDaF für Menschen mit Behinderung, *Babylonia* 12 (3), 70–74.

Arras, U, Müller-Karabil, A and Zimmermann, S (2013) On equal footing? Accommodations for disabled candidates in the TestDaF, in Tsagari, D and Spanoudis, G (Eds) *Assessing L2 students with learning and other disabilities*, Newcastle upon Tyne: Cambridge Scholars Publishing, 271–286.

Breton, G, Grego Bolli, G and Perlmann Balme, M (2010) All different - all equal?, *Forum Sprache*, 5–19.

Carlsen, C (2007) Language testing – a matter of ethics, in Carlsen, C and Moe, E (Eds) *A Human Touch to Language Testing*, Oslo: Novus Forlag, 97–108.

Carlsen, C (2008) The role of testing in an egalitarian society, *Research Notes* 34, 2–6.

Carlsen, C (2018) *Giving LESLLA-learners a fair chance in testing*, paper presented at 12th Annual LESLLA Symposium, Granada, 8–10 September, 2016.

Carlsen, C and Moe, E (Eds) (2007) *A Human Touch to Language Testing*, Oslo: Novus Forlag.

Carlsen, C H and Rocca, L (2021) Language test misuse, *Language Assessment Quarterly,* 18:5, 477-491. DOI: 10.1080/15434303.2021.1947288

Carlsen, C H and Rocca, L (2022) Language test activism. *Language Policy*. DOI: 10.1007/s10993-022-09614-7

Casanova, D, Aw, A, Bourras, M and Demeuse, M (2018) Nos items à l'accent québécois difR-ils?, in *Actes du 30e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 10 au 12 janvier 2018, Esch-sur-Alzette, Luxembourg*, Neuchâtel, Switzerland: ADMEE-Europe.

Desroches, F, Crendal, A, Renaud, F, Casanova, D, Demeuse, M and Artus, F (2006) Les biais culturels dans les tests internationaux de français langue étrangère: diagnostic et analyse, in *Actes du 19e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 11 au 13 septembre 2006, Luxembourg, Luxembourg*, Neuchâtel, Switzerland: ADMEE-Europe.

Deygers, B (2017) Just testing: Applying theories of justice to high-stakes language tests, *ITL: International Journal of Applied Linguistics* 168 (2), 143–163.

Deygers, B (2019) Fairness and social justice in English language Assessment, in Gao, X, Davidson, C and Leung, C (Eds) *Second International Handbook of English Language Teaching*, Heidelberg and New York: Springer.

Deygers, B, Van den Branden, K and Van Gorp, K (2017) University entrance language tests: a matter of justice, *Language Testing* 35 (4), 449–476, available online: https://journals.sagepub.com/doi/10.1177/0265532217706196

Eckes, T (2010) Die Beurteilung sprachlicher Kompetenz auf dem Prüfstand: Fairness in der beurteilergestützten Leistungsmessung, in Aguado, K, Schramm, K and Vollmer, H J (Eds) *Fremdsprachliches Handeln beobachten, messen, evaluieren: Neue methodische Ansätze der Kompetenzforschung und der Videographie*, Frankfurt am Main: Peter Lang, 65–97.

Eckes, T (2011) Facetten der Genauigkeit: Zur Reliabilität der Beurteilung fremdsprachlicher Leistungen, *Deutsch als Fremdsprache* 48, 195–204.

Georges, S (2013) Fairness and validity from the viewpoints of different actors in assessment: A French reaction to Kane's validation approach, in Chatterji, M (Ed) *Validity and Test Use: An International Dialogue on Educational Assessment, Accountability and Equity*, Bingley: Emerald, 17–53.

Glaboniat, M (2008) Müssen wir für die Deutschprüfung jetzt auch noch Österreichisch lernen? Der plurizentrische Ansatz im Bereich des Sprachentestens und Zertifizierens, in *Le Bulletin de l'ADEAF 2008/100. L'Association pour le Dévelopement de l'Enseignement de l'Allemand en France*.

Glaboniat, M (2019) Sprachprüfungen „made in Austria" als Beitrag zur Förderung der deutschen Sprache. Zur Bedeutung von ÖSD-Prüfungen unter besonderer Berücksichtigung des plurizentrischen Ansatzes, in Ammon, U, Sambe, S and Schmidt, G (Eds) *Förderung der deutschen Sprache weltweit. Vorschläge, Ansätze und Konzepte*, Berlin: De Gruyter.

Glaboniat, M and Mitteregger, B (2014) „Österreichisches Deutsch" in internationalen DaF-Prüfungen. Die Rolle des ÖSD in der Plurizentrik-Diskussion, in *informationen zur deutschdidaktik (ide). Sprachliche Bildung im Kontext von Mehrsprachigkeit. Österreichisches Deutsch und Plurizentrik* 2014/3, Innsbruck, Vienna, Bolzano: StudienVerlag, 88–98.

Kiefer, P (2008) Sprachtests müssen fair sein. Am Beispiel der TELC-Prüfungen im Bereich Migration und Einbürgerung, *Hessische Blätter für Volksbildung* 1, 67–73.

Klein, G (2013) Do gender, age and first language predict the results in the Deutsch-Test für Zuwanderer (DTZ)?, in Galaczi, E D and Weir, C J (Eds) *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press, 389–404.

Vodičková, K (2016) A je ta zkouška spravedlivá? aneb Spravedlivost vjazykovém testování [Is the exam fair? On fairness in language testing], *CASALC Review* 1, 46–50, available online: https://drive.google.com/file/d/0B1pJ9zo3zZe-MFNNUlVNeFRzc2s/view

General

Althaus, H-J (2004) Der TestDaF, in DAAD (Ed) *Die internationale Hochschule: Ein Handbuch für Politik und Praxis* Vol. 8, Bielefeld: Bertelsmann, 80–87.

Arieșan, A (2017) The Process of Setting the Time for A1-B2 RFL Exams, *Studia Universitatis Babeș-Bolyai, Philologia* 2/2017, 139-152.

Arras, U (2010) Subjektive Theorien als Faktor bei der Beurteilung fremdsprachlicher Kompetenzen, in Berndt, A and Kleppin, K (Eds) *Sprachlehrforschung: Theorie und Empirie – Festschrift für Rüdiger Grotjahn*, Frankfurt am Main: Peter Lang, 169–179.

Arras, U (2013) Introspektive Verfahren in der Sprachtestforschung, in Aguado, K, Heine, L and Schramm, K (Eds) *Introspektive Verfahren und Qualitative Inhaltsanalyse in der Fremdsprachenforschung*, Frankfurt am Main: Peter Lang, 74–91.

Arras, U and Grotjahn, R (2004) TestDaF: Jüngste Entwicklungen, in Casper-Hehne, H and Koreik, U (Eds) *DSH und TestDaF als hochschulbezogene Prüfungssysteme für Deutsch als Fremdsprache: Standortbestimmungen und Entwicklungslinien*, Hohengehren: Schneider, 205–225.

Arras, U and Kecker, G (2016) Leistungsmessung, Bewertung, Selbstevaluation, in Burwitz-Melzer, E, Mehlhorn, G, Riemer, C, Bausch, K-R and Krumm, H-J (Eds) *Handbuch Fremdsprachunterricht* 6, Tübingen: Narr Francke Attempto, 393–398.

Barker, F, Salamoura, A and Saville, N (2015) Learner corpora and language testing, in Granger, S, Gilquin, G and Meunier, F (Eds) *The Cambridge Handbook of Learner Corpus Research*, Cambridge: Cambridge University Press, 511–534.

Carlsen, C (2007) The place of reliability in a unified (Messickian) view of construct validity, in Carlsen, C and Moe, E (Eds) *A Human Touch to Language Testing*, Oslo: Novus Forlag, 75-86.

Carlsen, C (2012) Proficiency level – a fuzzy variable in computer learner corpora, *Applied Linguistics* 33 (2), 161–183.

Carlsen, C and Moe, E (2014) Assessing Norwegian, in Kunnan, A J (Ed) *The Companion to Language Assessment*, Chichester: Wiley Blackwell.

Carlsen, C H and Moe, E (2019) *Vurdering av språkferdigheter,* Bergen: Fagbokforlaget.

Cope, L and Somers, A (2011) Effective pretesting: An online solution, *Research Notes* 43, 32–35.

Corrigan, M and Crump, P (2015) Item analysis, *Research Notes* 59, 4–9.

Davidson, F, Turner, C and Huhta, A (1997) Language Testing Standards, in Clapham, C and Corson, D (Eds) *Encyclopedia of Language and Education, Volume 7: Language Testing and Assessment*, Dordrecht/Boston/ London: Kluwer Academic Publishers, 303–312.

Demeuse, M, Desroches, F, Crendal, A, Renaud, F and Casanova, D (2005) La fiabilité de l'évaluation des compétences linguistiques pour des adultes non francophones: présentation d'un protocole d'évaluation, in *Actes du 18e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 24 au 26 octobre 2005, Reims, France*, available online: https://www.lefrancaisdesaffaires.fr/wp-content/uploads/2016/05/3.TEF-Admee2005-Fiabilit%C3%A9.pdf

De Wachter, L and Heeren, J (2013) Een taaltest als signaal: De ontwikkeling en implementatie van een strategische taalvaardigheidstoets aan de KU Leuven, *Levende Talen Tijdschrift* 14 (1), 19–27.

De Wachter, L, Heeren, J Marx, S and Huyghe, S (2013) Taal: een noodzakelijke, maar niet de enige voorwaarde tot studiesucces: De correlatie tussen de resultaten van een taalvaardigheidstoets en de slaagcijfers bij eerstejaarsstudenten aan de KU Leuven, *Levende Talen Tijdschrift* 14 (4), 28–36.

Deygers, B (2012) Taaltoetsen ontwikkelen: tips en valkuilen, in Mottart, A (Ed) *Zesentwintigste conferentie Het Schoolvak Nederlands*, Ghent: Academia Press, 316–322.

Deygers, B and Kanobana, S (2010) Taaltoetsen. Waarom, wat en hoe?, in Peters, E and Van Houtven, T (Eds) *Taalbeleid in het hoger onderwijs: de hype voorbij?* Leuven: Acco, 19–41.

Docherty, C and Corkill, D (2015) Test construction: The Cambridge English approach, *Research Notes* 59, 10–14.

Eckes, T (2003) Qualitätssicherung beim TestDaF: Konzepte, Methoden, Ergebnisse, *Fremdsprachen und Hochschule* 69, 43–68.

Eckes, T (2014) Die onDaF–TestDaF-Vergleichsstudie: Wie gut sagen Ergebnisse im onDaF Erfolg oder Misserfolg beim TestDaF vorher?, in Grotjahn, R (Ed) *Der C-Test: Aktuelle Tendenzen/The C-test: Current trends*, Frankfurt am Main: Peter Lang, 137–162.

Eckes, T (2015) Validität: Flexionen eines polymorphen Konzepts, in Böcker, J and Stauch, A (Eds) *Konzepte aus der Sprachlehrforschung – Impulse für die Praxis: Festschrift für Karin Kleppin*, Frankfurt am Main: Peter Lang, 449–468.

Eckes, T, Ellis, M, Kalnberzina, V, Pižorn, K, Springer, C, Szollás, K and Tsagari, C (2005) Progress and problems in reforming public language examinations in Europe: Cameos from the Baltic States, Greece, Hungary, Poland, Slovenia, France, and Germany, *Language Testing* 22, 355–377.

Eckes, T and Grotjahn, R (2006) A closer look at the construct validity of C-tests, *Language Testing* 23, 290–325.

Elosua, P and Peñalba, A (2018) Language competence assessment in minoritised language revitalisation contexts: The case of Basque, *Journal of Multilingual and Multicultural Development* 39 (7), 629–640, available online: https://www.tandfonline.com/doi/full/10.1080/01434632.2017.1417415

Engelen, C and Van Splunder, F (2018) Ruimte voor creatieve oplossingen. Taal in het Vlaamse hoger onderwijs: beleid en praktijk, in *THeMa Tijdschrift voor Hoger onderwijs & Management* 25 (3), 19–23.

Folny, V (2016) Conception d'une enquête pour évaluer le niveau de français des instituteurs, in Maurer, B (Ed) *Les approches bi-plurilingues d'enseignement-apprentissage: autour du programme école et langues nationales en Afrique (ELAN-Afrique)*, Université Paul-Valéry, Montpellier, France: Éditions des archives contemporaines, 393–410.

Fortiana, B, Jornet, N, Pi, B, Puigdomènech, L and Sánchez, A (2009) L'avaluació de llengües als segle XXI: noves perspectives, nous reptes, *Llengua i Ús* 45, 12–22.

Georges, S (2012) Comment fabrique-t-on un test?, *Le français dans le monde* 383, 52–53.

Glaboniat, M (1994) Das Österreichische Sprachdiplom und ein Blick auf die internationale Testlandschaft, *ÖDAF-Mitteilungen* 1994/2, 29–37.

Glaboniat, M (1995) Das Österreichische Sprachdiplom Deutsch als Fremdsprache, *Stimulus, Mitteilungen der Österreichischen Gesellschaft für Germanistik* 1995/1, 29–37.

Glaboniat, M (1998) *Kommunikatives Testen im Bereich Deutsch als Fremdsprache. Eine Untersuchung am Beispiel des Österreichischen Sprachdiploms Deutsch*, Innsbruck: StudienVerlag.

Glaboniat, M (1998) Kommunikatives Testen: Ein Widerspruch? Merkmale und Grenzen kommunikativer Tests, *ÖDAF-Mitteilungen* 1998/2, 27–38.

Glaboniat, M (2001) Neue Ansätze in fremdsprachlicher Leistungsmessung. Testtheoretische Grundlagen und Beispiele, in Narrog, H and Fuchs, B (Eds) *Freiräume nutzen, neue Wege suchen - Methodik und Leistungsmessung von DaF in Japan*, Sapporo: Hokkaido University, 123–137.

Glaboniat, M (2006) Das Papier nicht wert. Zum Problem schulischer Leistungsbeurteilung und neue Chancen durch Qualitäts- und Leistungsstandards, *ide* 2006/4, 32–51.

Glaboniat, M (2007) Landeskunde im Bereich des Sprachentestens und Zertifizierens, *ÖDaF-Mitteilungen* 2007/2, 46–54.

Glaboniat, M (2009) „Wie viel Standardisierung braucht ein Test?" Vom Nutzen der Standardisierungstendenzen im Bereich des Sprachentestens, *ÖDaF-Mitteilungen. Visionen.*

Gegenwart und Zukunft von DaF/DaZ in Österreich. Sonderheft zur XIV. IDT, August 2009, Vienna: ÖDaF, 41–51.

Glaboniat, M, Kleinberger, A and Ziehr, H (2016) ÖSD-Prüfungen – Prüfungsqualität und Prüfungsteilnehmende im Fokus, in Schweiger, H (Ed) *In die Welt hinaus. Festschrift für Renate Faistauer zum 65. Geburtstag*, Vienna: Praesens Verlag, 45–54, available online: https://campus.aau.at/cris/publication/0f4de0c759e005ab0159e10ea78d0438

Glaboniat, M, Müller, M, Rusch, P, Schmitz, H and Wertenschlag, L (2002) *Profile deutsch. Niveaustufen A1–B2. Version 1.0*, Berlin/ Munich/Vienna/Zurich/New York: Langenscheidt.

Glaboniat, M and Peresich, C (2018) German as a Foreign Language: Das Österreichische Sprachdiplom Deutsch, in Sigott, G (Ed) *Language Testing in Austria. Taking Stock. / Sprachtesten in Österreich: Eine Bestandsaufnahme*, Frankfurt am Main: Peter Lang, 349–369.

Glaboniat, M, Perlmann-Balme, M and Studer, T (2013b) Zertifikat B1. Deutschprüfung für Jugendliche und Erwachsene: Standard Setting. Ein Arbeitsbericht, *Zeitschrift für Interkulturellen Fremdsprachenunterricht. Didaktik und Methodik im Bereich Deutsch als Fremdsprache* 2013/1, 72–75, available online: http://tujournals.ulb.tu-darmstadt.de/index.php/zif/article/view/69/65

Glaboniat, M and Sigott, G (2012) Reifeprüfung Deutsch. Das Konzept der kriterienorientierten Bewertung. Dimension und Niveaus, in *Die Zeitschrift für den Deutschunterricht in Wissenschaft und Schule*, Innsbruck, Vienna, Bolzano: StudienVerlag, 130–140.

Grego Bolli, G (1994) Presupposti teorici e metodologici al modello di certificazione della conoscenza dell'italiano proposto dall'Università per Stranieri di Perugia, *Educazione Permanente*, 46–54.

Grego Bolli, G (1994) La struttura, i livelli, la composizione delle prove della Certificazione dell'Università per Stranieri di Perugia, *Educazione Permamente*, 67–80.

Grego Bolli, G (2004) Measuring and evaluating the competence in Italian as a foreign language, in Milanovic, M and Weir, C J (Eds) *European Language Testing in a Global Context: Proceedings of the ALTE Barcelona Conference, July 2001*, Studies in Language Testing volume 18, Cambridge: UCLES/Cambridge University Press, 271–283.

Grego Bolli, G (2014) Assessing Italian, in Kunnan, A J (Ed) *The Companion to Language Assessment*, Chichester: Wiley Blackwell, 2,018–2,028.

Grego Bolli, G (2014) Using Quality Management Systems to improve test development and standards and to promote good practice: A case study of testing Italian as a foreign language, *Language Learning in Higher Education* 4 (1), 179–206.

Grego Bolli, G (2017) La lettura e la verifica della comprensione nella prospettiva del Language Testing, con un riferimento alle prove INVALSI nella scuola secondaria di primo

grado, in Covino, S (Ed) *Per una didattica mirata della lingua. Acquisizione, misurazione e valutazione delle competenze linguistiche. Seminario triregionale dell'ASLI Scuola (Firenze, Perugia, Catania, 20-23 aprile 2015)*, Firenze: Franco Cesati Editore, 63–83.

Grego Bolli, G, Machetti, S and Rini, D (2017) Researching Italian language assessment in China and Italy: focusing on CELI and CILS examinations, in Bagna, C, Chapedi, N, Savati, L, Scibetta, A and Visigalli, M (Eds) *La lingua italiana per Studenti internazionali: aspetti linguistici e didattici*, Perugia: OL3 Edizioni, 28–36.

Grego Bolli, G, Machetti, S, Rini, D and Bandini, A (2017) The evaluation of Chinese students of Italian L2 in China and in Italy. Practices at the Universities for Foreigners of Perugia and Siena, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 118–123, available online: http://events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Grego Bolli, G and Milanovic, M (1996) Il testing linguistico in un contesto internazionale: lo schema europeo dell'ALTE e la certificazione dell'Università per Stranieri di Perugia, *IL VELTRO* 3-4, 408–412.

Grego Bolli, G and Pelliccia, F (Eds) (2005) *CELI. Breve guida ai certificati di italiano L2*, Perugia: Guerra Edizioni.

Grego Bolli G, and Spiti, M G (2004) *La verifica delle competenze linguistiche. Misurare e valutare nella certificazione CELI*, Perugia: Guerra Edizioni.

Gutierrez Eugenio, E and Saville, N (2017) *Policy review: The role of assessment in European language policy: a historical overview*, available online: http://www.meits.org/policy-papers/paper/the-role-of-assessment-in-european-language-policy-a-historical-overview

Hasselgreen, A, Carlsen, C and Helness, H (2004) *European Survey of Language Testing and Assessment Needs. Report: Part One – General findings*, available online: http://www.ealta.eu/documents/resources/survey-report-pt1.pdf

Huhta, A (2008) Diagnostic and formative assessment, in Spolsky, B and Hult, F M (Eds) *The Handbook of Educational Linguistics*, Chichester: Wiley Blackwell, 469–482.

Huhta, A (2014) Administration, scoring, and reporting, in Kunnan, A J (Ed) *The Companion to Language Assessment*, Chichester: Wiley Blackwell.

Huhta, A (2015) Assessment culture in Finland - a mixture of traditions and innovations, in Huizhong, Y and Zhixiang, Z (Eds) *Sociology of Language Testing*, Shanghai: Shanghai Foreign Language Education Press, 266–301.

Huhta, A and Hildén, R (2016) Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa [Language examinations and other large-scale language assessments in Finland], in Huhta, A and Hildén, R (Eds) *Kielitaidon arviointitutkimus 2000-luvun Suomessa.*

AFinLA-e. Soveltavan kielitieteen tutkimuksia 9, 3–26, available online: https://journal.fi/afinla/article/view/60844

Huhta, A, Kalaja, P and Pitkänen-Huhta, A (2006) Discursive construction of a high-stakes test: The many faces of a test-taker, *Language Testing* 23 (3), 326–350.

Huhta, A, Kohonen, V, Kurki-Suonio, L and Luoma, S (Eds) (1997) *Current Developments and Alternatives in Language Assessment. Proceedings of LTRC 96*, Jyväskylä: University of Jyväskylä.

Huhta, A and Suontausta, T (1996) Students' Proficiency in L2, in Marsh, D, Oksman-Rinkinen, P and Takala, S (Eds) *Mainstream Bilingual Education in the Finnish Vocational Sector*, National Board of Education, 85–117.

Huhta, A and Tarnanen, M (2009) Assessment practices in the Finnish comprehensive school – what is the students' role in them?, in May, S (Ed) *LED2007: Refereed conference proceedings of the 2nd International Conference on Language, Education and Diversity*, Hamilton, New Zealand: Wilf Malcolm Institute of Educational Research (WMIER), University of Waikato.

Jones, N and Saville, N (2016) *Learning Oriented Assessment: A Systemic Approach*, Studies in Language Testing volume 45, Cambridge: UCLES/Cambridge University Press.

Jornet, N (2000) "L'examen de nivell avançat del Certificat internacional de català: procés de desenvolupament i descripció de la prova" *Llengua i Ús* 18.

Kecker, G (2010) Validität und Validierung von Sprachprüfungen, in Berndt, A and Kleppin, K (Eds) *Sprachlehrforschung: Theorie und Empirie – Festschrift für Rüdiger Grotjahn*, Frankfurt am Main: Peter Lang, 129–146.

Kecker, G (2016) Was macht eine gute Sprachprüfung aus? Qualitätssicherung beim TestDaF, in Drumbl, H, Kletschko, D, Sorrentino, D and Zanin, R (Eds) *IDT 2013 (Bd. 7: Lerngruppenspezifik in DaF, DaZ, DaM)*, Bozen, Italien: Bozen-Bolzano University Press, 145–164.

Kecker, G, Depner, G, Marks, D, Schwarz, L and Zimmermann, S (forthcoming) Die deutsche Sprache weltweit fördern: Was können Sprachprüfungen dazu beitragen?, in Ammon, U, Sambe, S and Schmidt, G (Eds) *Förderung der deutschen Sprache weltweit. Vorschläge, Ansätze und Konzepte*, Berlin: De Gruyter.

Lim, G (2015) Determining standards in assessing language ability, *Research Notes* 59, 20-23.

Medrea, A, Platon, E, Sonea, I, Vîlcu, D and Vesa, V (2008) *Teste de limba română ca limbă străină - A1, A2, B1, B2*, Cluj-Napoca: Editura Risoprint.

Mègre, B (2009) The TCF (Test de Connaissance du Français) Making, Analyzing, and Implementing, *Assessment and Evaluation in Language Education* 1, 51–61.

Mègre, B and Georges, S (2011) Les certifications en français langue étrangère du ministère de l'Education nationale: des outils qui se sont adaptés aux besoins des publics, *Les cahiers de l'ASDIFLE* 23, 127–139.

Mègre, B and Portelli, S (2016) Les enjeux du développement de tests de langues non certifiants: l'exemple du test français Ev@lang, *Revue internationale d'éducation de Sèvres* 71, 15–18, available online: https://doi.org/10.4000/ries.4543

Mègre, B and Riba, P (2015) *Démarche qualité et évaluation en langues*, Paris: Hachette.

Pečený, P, Nováková, J and Vodičková, K (2013) Oblasti zvyšování kvality jazykové zkoušky na příkladu Certifikované zkoušky z češtiny pro cizince (CCE) [Increasing the quality of language examinations demonstrated on the Czech Language Certificate Exam (CCE)], in *Zvyšování kvality výuky a testování cizích jazyků (včetně češtiny pro cizince)*, Prague: ÚJOP UK, 87–92.

Plassmann, S (2007) telc – Sprachprüfungen mit Niveau, *Hessische Blätter für Volksbildung* 4, 357–364.

Plassmann, S (2011) Aktuelle Methoden der Testmethodik und Qualitätssicherung am Beispiel des Deutsch-Test für Zuwanderer, *Deutsch als Fremdsprache* 48 (1), 23–29.

Platon, E, Arieșan, A and Fodor, L (2015) *Limba română - evaluarea competenței lingvistice la ciclul primar (EVRO-P5)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E, Chircu, A and Popa, M (2015) *Limba română – evaluarea competenței lingvistice în învățământul secundar (P6)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E and Vîlcu, D (Eds) (2011) *Procesul de evaluare a limbii române ca limbă nematernă (RLNM) la ciclul primar. RLNM: P4 - ciclul primar*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E and Vîlcu, D (Eds) (2011) *Procesul de evaluare a limbii române ca limbă nematernă (RLNM) la ciclul gimnazial. RLNM: P5 - ciclul gimnazial*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E and Vîlcu, D (Eds) (2011) *Procesul de evaluare a limbii române ca limbă nematernă (RLNM) la ciclul liceal. RLNM: P6 - ciclul liceal*, Cluj-Napoca: Editura Casa Cărții de Știință.

Puigdomènech, L (2014) L'anàlisi de corpus com a mètode de recerca qualitativa per al desenvolupament d'exàmens: l'exemple de la prova de nivell C2, *Llengua i Ús* 55.

Puigdomènech, L (2015) L'avaluació de la competència lingüística comunicativa: principis bàsics i propietats dels exàmens, *Llengua i Ús* 56.

Riba, P, Josselin, C, Georges, S and Pecheur, J (2012) Méthodes d'évaluation et enjeux de certification: dossier, *Le Français dans le monde* 383, 46–55.

Robinson, R (2013) Innovation in language test development, *Research Notes* 52, 7–13.

Rocca, L (2013) La certificazione linguistica CELI, *Aggiornamenti* 4 (13), Sprachenzentrum der Universität Bamberg, Bamberg, 26-39.

Saville, N (2010) Auditing the quality profile: From code of practice to standards, *Research Notes* 39, 24–28.

Saville, N (2016) *Mixed methods and language assessment: A new frontier*, paper presented at the 2nd International Mixed Methods International Research Association (MMIRA) Conference, Durham University, Durham, UK, August 2016.

Saville, N (2017) *Setting and monitoring standards within a professional community of practice*, paper presented at Academic Forum on English Language Testing in Asia (AFELTA), Taipei, Taiwan, June 2017.

Sundberg, G (2007) Assessing and testing second language proficiency in the Nordic countries, in Carlsen, C and Moe, E (Eds) *A Human Touch to Language Testing*, Oslo: Novus Forlag, 87–96.

Sundberg, G and Corell, M (1999) *Handbok för produktion av språktest*, Stockholm: Department of Scandinavian Languages, Stockholm University.

Sundberg, G and Corell, M (1999) *Testordlista. Terminologi inom språkfärdighetstestning*, Stockholm: Department of Scandinavian Languages, Stockholm University.

Tagliante, C, Mègre, B, Breton, G, Dupleix, D and Houssa, C (2004) Les tests de langues, *Revue internationale d'éducation de Sèvres* 37, 115–121, available online: https://doi.org/10.4000/ries.1474

Van Gorp, K and Deygers, B (2013) Task-Based Language Assessment, in Kunnan, A J (Ed) *The Companion to Language Assessment*, Chichester: Wiley Blackwell, 578–593.

Vîlcu, D (2015) Romanian as a foreign language in the context of language assessment in Europe, *Studia Universitatis Babeş-Bolyai, Philologia* 2/2015, 65–76.

Vodičková, K (2013) Как обеспечить качество языковых екзаменов? [How to ensure the quality of language examinations?], in *Иновации в преподвании иностранных языков студентам-юристам*, Moscow: РПА Минюстра России, 41–43.

Vodičková, K (2013) Vývoj jazykových zkoušek – nikdy nekončící příběh? [Language examination development – a never ending story?], in *Sborník Asociace učitelů češtiny jako cizího jazyka (AUČCJ)*, Praha: Akropolis, 205–211.

Weir, C J (2013) Measured constructs: A history of Cambridge English language examinations 1913–2012, *Research Notes* 51, 2–11.

Zeidler, B (2016) Getting to know the minimally competent person, in Docherty, C and Barker, F (Eds) *Language Assessment for Multilingualism: Proceedings of the ALTE Paris Conference, April 2014*, Studies in Language Testing volume 44, Cambridge: UCLES/Cambridge University Press, 251–269.

Grammar

ALTE (2001) *Individual Component Checklist for use with ONE task: Structural Competence*, available online: https://www.alte.org/Materials

ALTE (2001) *Individual Component Checklist: Structural Competence*, available online: https://www.alte.org/Materials

Vodičková, K (2011) Can grammar competence tap into discourse competence, especially coherence?, in *Инностранные языки в современном мире*, Moskva: MAKS Press, 101–106.

Impact

Ashton, K (2013) Reflections on the European Survey on Language Competences: Looking back, looking forwards, *Research Notes* 52, 20–23.

Carlsen, C (2009) Crossing the bridge from the other side: the impact of society on testing, in Taylor, L and Weir, C J (Eds) *Language Testing Matters: Investigating the Wider Social and Educational Impact of Assessment – Proceedings of the ALTE Cambridge Conference, April 2008*, Studies in Language Testing volume 31, Cambridge: UCLES/Cambridge University Press.

Culej, J B (2013) The European Survey on Language Competences in Croatia: Results and implications, *Research Notes* 52, 16–20.

Ellis, S, Hawkey, R and Docherty, D (2014) Iterative impact studies of Progetto Lingue 2000, *Research Notes* 58, 15–19.

Hildén, R, Rautopuro, J and Huhta, A (2017) Arvosanan ansaitsemme – asteikolla vai ilman? [We deserve a grade – based on a scale or not?], in Britschgi, V and Rautopuro, J (Eds) *Kriteerit puntarissa*, Joensuu :Finnish Educational Research Association, 63–80.

Huhta, A and Ahonen, S (forthcoming) Arviointi kielipoliittisena ja kielikoulutuspoliittisena toimintana [Assessment as language policy and language education policy activity], in Saarinen, T, Pöyhönen, S, Kangasvieri, T and Nuolijärvi, P (Eds) *Kielen ja kielikoulutuksen politiikkaa [Policies of language and language education]*, Tampere: Vastapaino.

Huhta, A, Kalaja, P and Pitkänen-Huhta, A (2000) "Tuntuu tosi tosi tosi hyvältä … jipii" – ylioppilaskokeiden ja niihin osallistumisen rakentuminen abiturientin päiväkirjamerkinnöissä. ["It feels really really really good ... yippee" – the construction of the Matriculation examination and the test-preparation process in a candidate's diary entries], in Kalaja, P and Nieminen, L (Eds) *Kielikoulutus – kieli koulussa. [Language education – language in school]* AFinLA Yearbook 2000, Publications of the Finnish Association of Applied Linguistics 58, Jyväskylä, 221–242.

Jones, N (2013) The European Survey on Language Competences and its significance for Cambridge English Language Assessment, *Research Notes* 52, 2–7.

Khalifa, H and Docherty, C (2016) Investigating the impact of international assessment: A convergent parallel mixed methods approach, in Moeller, A J, Creswell, J W and Saville, N (Eds) *Second Language Assessment and Mixed Methods Research*, Studies in Language Testing volume 43, Cambridge: UCLES/Cambridge University Press, 269–295.

McKenna, S (2013) The European Survey on Language Competences and the media, *Research Notes* 52, 23.

Radu, A (2017) The Adequacy of the Test of Romanian as a Foreign Language to the Test Takers, *Studia Universitatis Babeş-Bolyai, Philologia* 2/2017, 129–138.

Saville, N (2010) Developing a model for investigating the impact of language assessment, *Research Notes* 42, 2–8.

Saville, N (2012) Applying a model for investigating the impact of language assessment within educational contexts: The Cambridge ESOL approach, *Research Notes* 50, 4–8.

Saville, N (2016) *Imagine a future where testing makes our student journeys the best in class*, paper presented at Navitas, Hanoi, Vietnam, November 2016.

Saville, N and Gutierrez Eugenio, E (2016) The European Commission's 'Study on comparability of language testing in Europe' (2015), *Research Notes* 63, 3–12.

Saville, N and Khalifa, H (2016) The impact of language assessment, in Tsagari, D and Banerjee, J (Eds) *Handbook of Second Language Assessment*, Boston/Berlin: De Gruyter Mouton, 77–94.

Szpotowicz, M (2013) The European Survey on Language Competences – the Polish experience, *Research Notes* 52, 13–16.

Tarnanen, M and Huhta, A (2008) Interaction of Language Policy and Assessment in Finland, *Current Issues in Language Planning* 9 (3), 262–281.

Van Splunder, F and Verguts, C (2017) Testing the Test: How Politics Influenced the Reception of an English Test for Lecturers, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 51–56, available online: events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Language for Specific Purposes (LSP)

ALTE (2018) *Guidelines for the Development of Language for Specific Purposes Tests*, available online: www.alte.org/Materials

ALTE (2022) *Illustration for the Guidelines on Language for Specific Purposes Tests,* available online: www.alte.org/Materials

Arras, U (2012) Im Rahmen eines Hochschulstudiums in Deutschland erforderliche sprachliche Kompetenzen - Ergebnisse einer empirischen Bedarfsanalyse, in Tinnefeld, T (Ed) *Hochschulischer Fremdsprachenunterricht: Anforderungen – Ausrichtung – Spezifik*, available online:
hochschulfremdsprachenunterricht.blogspot.com/search/label/41%20Arras

Artus, F, Demeuse, M, Maréchal, M, Casanova, D, Crendal, A, Desroches, F, Holle, A (2011) Évaluer la compétence écrite en français des étudiants non-francophones en situation académique, in *Actes du 23e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 12 au 14 janvier 2011, Paris, France*, available online:
www.researchgate.net/publication/272488358_Evaluer_la_competence_ecrite_en_francais_des_etudiants_non-francophones_en_situation_academique

Carlsen, C H (2018) The adequacy of the B2-level as university entrance requirement, Special issue of *Language Assessment Quarterly* 15 (1).

Casanova, D (2011) Les certifications en français professionnel de la Chambre de commerce et d'industrie de Paris. Du programme à l'action, *Les cahiers de l'asdifle* 23, 163–175.

Casanova, D, Artus, F, Demeuse, M et Maréchal, M (2011) Comment évaluer les compétences en français académique d'étudiants non francophones souhaitant poursuivre leurs études en France?, *Les cahiers de l'asdifle* 23, 37–46.

Casanova, D, Aw, A and Demeuse, M (2018) Quand le numérique défie la mesure. Comment veiller à la qualité de certifications en langue professionnelle au format numérique ?, in *Actes du 30e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 10 au 12 janvier 2018, Esch-sur-Alzette, Luxembourg*, available online:
www.researchgate.net/publication/325570395_Quand_le_numerique_defie_la_mesure_Comment_veiller_a_la_qualite_de_certifications_en_langue_professionnelle_au_format_numerique

Cheung, K Y F (2017) Conclusions and recommendations, in Cheung, K Y F, McElwee, S and Emery, J (Eds) *Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press, 216–232.

Cheung, K Y F and Elliott, M (2017) *Academic literacy and language proficiency in testing: Overlapping and diverging constructs*, paper presented at the Association of Language Testers in Europe (ALTE) International Conference, Bologna, Italy, May 2017.

Cheung, K Y F and McElwee, S (2017) What skills are we assessing? Cognitive validity in BMAT, in Cheung, K Y F, McElwee, S and Emery, J (Eds*) Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press, 35–80.

Cheung, K Y F, McElwee, S and Emery, J (Eds) (2017) *Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press.

Devine, A, Taylor, L and Cross, B (2017) The biomedical school applicant: Considering the test taker in test development and research, in Cheung, K Y F, McElwee, S and Emery, J (Eds) *Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press, 17–34.

Deygers, B (2017) Hoe effectief zijn universitaire toelatingsvoorwaarden voor internationale studenten? *Over Taal* 56 (3), 19–21.

Deygers, B (2017) Talige toelatingsvoorwaarden voor internationale universiteitsstudenten. Praktische en didactische implicaties, available online: www.tijdschriftles.nl/inhoud/tijdschrift_artikel/LE-35-203-11/Niet-steeds-even-effici-nt

Deygers, B (2017) University entrance language tests: examining assumed equivalence, in Davis, J, Norris, J, Malone, M, McKay, T and Son, Y (Eds) *Useful Assessment and Evaluation in Language Education*, Washington, DC: Georgetown University Press, 169–184.

Deygers, B (2017) Validating university entrance policy assumptions. Some inconvenient facts, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 46-50, available online: events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Deygers, B (2018) How institutional and interpersonal variables impact international L2 students' language gains at university, *System* 76, 91–102.

Deygers, B and Malone, M (forthcoming 2018) Language assessment literacy in university admission policies, or the dialogue that isn't, *Language Testing*.

Deygers, B, Van den Branden, K and Peters, E (2017) Checking assumed proficiency: comparing L1 and L2 performance on a university entrance test, *Assessing Writing* 32, 43–56.

Fyfe, M, Devine, A and Emery, J (2017) The relationship between test scores and other measures of performance, in Cheung, K Y F, McElwee, S and Emery, J (Eds) *Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press, 143–180.

Huhta, A (1996) Validating an EFL C-test for students of English philology, in Grotjahn, R (Ed) *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* volume 3, Bochum: Universitätsverlag Dr. N. Brockmeyer, 197–234.

Huhta, A (2009) Analysis of the quality of testing English for aviation purposes in Finland, *Australian Review of Applied Linguistics* 32 (3), 1–26.

Jacobs G, Deygers B, Rymenams S, Strobl C and Vande Casteele A (2006) Van spelling tot stijl: (vreemde-)taalhulp voor studenten economie, in Van Ootegem, L (Ed) *Wegwijzer wetenschappelijk werk : bronnen voor papers en verhandelingen in de bedrijfseconomie*, Tielt: Lannoo, 309–316.

Marks, D (2015) Prüfen sprachlicher Kompetenzen internationaler Studienanfänger an deutschen Hochschulen – Was leistet der TestDaF? *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 20 (1), 21–39.

McElwee, S, Fyfe, M and Grant, K (2017) The consequences of biomedical admissions testing on individuals, institutions and society, in Cheung, K Y F, McElwee, S and Emery, J (Eds) *Applying the Socio-cognitive Framework to the BioMedical Admissions Test: Insights from Language Assessment*, Studies in Language Testing volume 49, Cambridge: UCLES/Cambridge University Press, 181–215.

Pilkinton-Pihko, D, Virkkunen-Fullenwider, A and Huhta, A (2008) Linking an English for science and technology placement test to the CEFR, in Vogel, T, Satchell, R and Moravciková, A (Eds) *Proceedings from 9th CercleS conference, Frankfurt/Oder, Germany, 2006. Language Centres at Universities: Crossing Bridges, Integrating Cultures*, Göttingen: CercleS, 283–300.

Plassmann, S and Zeidler, B (2014) Taking decisions: Assessment for university entry, *Language Learning in Higher Education* 4 (1), 237–255.

Saville, N (2017) *English in the international(ised) workplace*, paper presented at the Multilingualism in University and the Workplace event, Salamanca, Spain, May 2017.

Saville, N (2017) *Key research-based insights into BULATS*, paper presented at the BULATS Agents Annual Meeting, Paris, France, June 2017.

Schuth, E, Heppt, B, Köhne, J, Weinert, S and Stanat, P (2015) Die Erfassung schulisch relevanter Sprachkompetenzen bei Grundschulkindern – Entwicklung eines Testinstruments,

in Redder, A, Naumann, J and Tracy, R (Eds) *Forschungsinitiative Sprachförderung und Sprachdiagnostik – Ergebnisse*, Münster: Waxmann, 93–112.

Séguis, B (2018) *Language assessment of health professionals: The issue of communicative competence*, paper presented at the ALTE 51st Meeting and Conference, Babes-Bolyai University, Cluj-Napoca, Romania, 13 April 2018.

Ursa, A (2017) Being on Guard vs Being in Blinkers: Perspectives on the Intercultural Competence with Foreign Medical Students in Romania. Aspects of evaluation, *Studia Universitatis Babeş-Bolyai, Philologia* 2/2017, 184–196.

Van Splunder, F (2013) De taaltest Engels voor docenten [English Language Testing for Lecturers], *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid. Bijzonder nummer: Verantwoordelijk bestuur in het Hoger Onderwijs* 6, 111–115.

Vodičková, K (2014) Basic Principles in Assessing Language for Specific Purposes, in *Jazyk i parvo*, Moskva: Moskovskij gosudarstvennyj universitet imeni M. V. Lomonosova, 12–13.

Listening

ALTE (2001) *Individual Component Checklist: Listening*, available online:
www.alte.org/Materials

ALTE (2001) *Individual Component Checklist for use with ONE task: Listening*, available online: www.alte.org/Materials

Arras, U (2015) Mitschreiben als hochschulrelevante Sprachhandlung – kann das eine Sprachprüfung testen?, in Knapp, A and Aguado, K (Eds) *Fremdsprachen in Studium und Lehre: Chancen und Herausforderungen für den Wissenserwerb*, Frankfurt am Main: Peter Lang, 209–245.

Eckes, T (2014) Examining testlet effects in the TestDaF listening section: A testlet response theory modeling approach, *Language Testing* 31, 39–61.

Geist, B (2009) Schwierigkeiten und typische Fehler chinesischer TestDaF-Teilnehmer – am Beispiel des Prüfungsteils Hörverstehen, in Yu, X (Ed) *TestDaF-Training und Studienvorbereitung: Beiträge zur chinesisch-deutschen Fachkonferenz: „TestDaF-Training und Studienvorbereitung" vom 11. bis 12. Oktober 2008 am Deutschkolleg der Tongji-Universität Shanghai*, München: Iudicium, 47–62.

Geranpayeh, A and Taylor, L (Eds) (2013) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press.

Glaboniat, M (2008) Wer nicht hören kann, muss üben … Zur Schulung der HV-Kompetenz aus der Sicht des Deutsch als Fremdsprache-Unterrichts, *ide* 2008/1, 51–65.

Kecker, G (2015) "He listens well who takes notes" – Mit welchen Aufgabenformaten kann Hörverstehen in Vorlesungen an der Hochschule valide getestet werden?, in Böcker, J and Stauch, A (Eds) *Konzepte aus der Sprachlehrforschung – Impulse für die Praxis: Festschrift für Karin Kleppin*, Frankfurt am Main: Peter Lang, 511–526.

Lim, G S and Khalifa, H (2013) Criterion-related validity, in Geranpayeh, A and Taylor, L (Eds) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press, 303–321.

Marking and Rating

ALTE (2001) *ALTE Quality Assurance Checklists, Unit 3: Marking, Grading & Results*, available online: www.alte.org/Materials

ALTE (2001) *ALTE Quality Assurance Checklists, Unit 4: Test analysis and Post-examination Review*, available online: www.alte.org/Materials

Arras, U (2009) What's on a rater's mind? Die Erforschung von Beurteilungsstrategien und ihre Bewusstmachung durch Schulungsmaßnahmen als Voraussetzungen für die Testvalidität, *Zeitschrift für Angewandte Linguistik* 50, 33–45.

Arras, U, Marks, D and Zimmermann, S (2009) BeurteilerInnen in den Kopf geschaut: Wie das Verfahren des Lauten Denkens im Rahmen von Beurteilungsschulungen eingesetzt werden kann, *DaF-Brücke: Zeitschrift für Deutschlehrerinnen und Deutschlehrer Lateinamerikas* 11, 5–9.

Bärenfänger, O and Kecker, G (2004) Beurteilerkonsistenz und Beurteilerübereinstimmung bei der Bewertung von Kurzantwortaufgaben: Eine empirische Studie zum Bewertungsverhalten von Beurteilern im Subtest Hörverstehen des Tests Deutsch als Fremdsprache (TestDaF), *Zeitschrift für Fremdsprachenforschung* 15, 237–268.

Carlsen, C (2004) *Guarding the Guardians. Rating scale and rater training effects on reliability and validity of scores of an oral test of Norwegian as a second language*, unpublished PhD thesis, University of Bergen, available online: bora.uib.no/handle/1956/2085

Casanova, D and Demeuse, M (2011) Analyse des différentes composantes influant sur la fidélité de l'épreuve d'expression écrite d'un test standardisé de français langue étrangère, *Mesure et évaluation en éducation* 34 (1), 25–53.

Casanova, D and Demeuse, M (2016) Évaluateurs évalués. Évaluation diagnostique des compétences en évaluation des correcteurs d'une épreuve d'expression écrite à forts enjeux, *Mesure et évaluation en education* 39 (1), 57–94.

Casanova, D, Demeuse, M and Mommée, A (2017) Évaluer en continu les correcteurs d'une épreuve à forts enjeux pour mieux les accompagner, in *Actes du 29e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 25 au 27 janvier 2017, Dijon, France*, Neuchâtel: ADMEE-Europe, 59–96.

Casanova, D, Brown, G, Boumaaz, A and Demeuse, M (2016) Comment l'analyse des données de production permet d'améliorer la formation continue des évaluateurs d'une épreuve d'expression orale, in *Actes du 28e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 13 au 15 janvier 2016, Lisbonne, Portugal*, available online: www.researchgate.net/publication/309735310_Comment_l'analyse_des_donnees_de_produ

ction_permet_d'ameliorer_la_formation_continue_des_evaluateurs_d'une_epreuve_d'expression_orale/download

Demeuse, M, Artus, F, Crendal, A, Casanova, D and Renaud, F (2008) Évaluer les productions orales en français langue étrangère (FLE) en situation de test. Étude de la fidélité inter-juges de l'épreuve d'expression orale du Test d'Évaluation de Français (TEF) de la Chambre de commerce et d'industrie de Paris (CCIP), in *Actes du 20e colloque international de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE-Europe), 9 au 11 janvier 2008, Genève, Suisse*, available online: www.researchgate.net/publication/255709697_Evaluer_les_productions_orales_en_francais _langue_etrangere_FLE_en_situation_de_test_Etude_de_la_fidelite_inter- juges_de_l'epreuve_d'expression_orale_du_Test_d'Evaluation_du_Francais_TEF_de_la_C hambre/download

Deygers, B, Van Gorp, K, Joos, S and Luyten, L (2013) Rating scale design: a comparative study of two analytic rating scales in a task-based test, in Galaczi, E D and Weir, C J (Eds) *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press, 271–287.

Eckes, T (2004) Beurteilerübereinstimmung und Beurteilerstrenge: Eine Multifacetten-Rasch-Analyse von Leistungsbeurteilungen im „Test Deutsch als Fremdsprache" (TestDaF), *Diagnostica* 50, 65–77.

Eckes, T (2004) Facetten des Sprachtestens: Strenge und Konsistenz in der Beurteilung sprachlicher Leistungen, in Wolff, A, Ostermann, T and Chlosta, C (Eds) *Integration durch Sprache (Materialien Deutsch als Fremdsprache)* 73, Regensburg: Fachverband Deutsch als Fremdsprache, 485–518.

Eckes, T (2005) Examining rater effects in TestDaF writing and speaking performance assessments: A many-facet Rasch analysis, *Language Assessment Quarterly* 2, 197–221.

Eckes, T (2008) Rater types in writing performance assessments: A classification approach to rater variability, *Language Testing* 25, 155–185.

Eckes, T (2009) On common ground? How raters perceive scoring criteria in oral proficiency testing, in Brown, A and Hill, K (Eds) *Tasks and criteria in performance assessment: Proceedings of the 28th Language Testing Research Colloquium*, Frankfurt am Main: Peter Lang, 43–73.

Eckes, T (2017) Rater effects: Advances in item response modeling of human ratings – Part I (Guest Editorial), *Psychological Test and Assessment Modeling* 59, 443–452.

Eckes, T (2018) Rater effects: Advances in item response modeling of human ratings – Part II (Guest Editorial), *Psychological Test and Assessment Modeling* 60, 29–32.

Elliott, M and Stevenson, L (2015) Grading and test equating, *Research Notes* 59, 14–20.

Galaczi, E and Khabbazbashi, N (2016) Rating scale development: A multistage exploratory sequential design, in Moeller, A J, Cresswell, J W and Saville, N (Eds) *Second Language Assessment and Mixed Methods Research*, Studies in Language Testing volume 43, Cambridge: UCLES/Cambridge University Press, 208–232.

Geranpayeh, A (2011) The impact of online marking on examiners' behaviour, *Research Notes* 43, 15–21.

Jones, N (2016) 'No More Marking': An online tool for comparative judgement, *Research Notes* 63, 12–15.

Lepage, S and North, B (2005b) Un outil concret pour les formateurs, *Le français dans le monde* 339, 35–36.

Lim, G S (2012) Developing and validating a mark scheme for writing, *Research Notes* 49, 6–10.

Zeidler, B (2016) Bewerterkalibrierung in kleinen Prüfungen, in Drumbl, H and Hornung, A (Eds) *IDT 2013. Deutsch von innen, Deutsch von außen. Beiträge der XV. Internationalen Tagung der Deutschlehrerinnen und Deutschlehrer Bozen, 29 Juli−3 August 2013*, Bolzano: Bozen-Bolzano University Press, 203–226.

Migration, Immigration and Integration

ALTE (2016) *Language tests for access, integration and citizenship: An outline for policy makers*, available online: www.alte.org/Materials

Council of Europe and ALTE (2018) *Linguistic integration of adult migrants: Requirements and learning opportunities*, available online: https://rm.coe.int/linguistic-integration-of-adult-migrants-requirements-and-learning-opp/16809b93cb

Glaboniat, M and Wintersteiner, W (2012) Empfehlung Migration: KandidatInnen mit einer anderen Erstsprache als Deutsch. Ein Vorschlag zur Vorgangsweise und Orientierung, in. Saxalber, A und Wintersteiner, W in Kooperation mit dem BIFIE Wien (Eds) *Reifeprüfung Deutsch*. informationen zur deutschdidaktik (ide) 2012/1, Innsbruck/Vienna/Bolzano: StudienVerlag, 66–70.

Grego Bolli, G (2013) Migration policies in Europe in relation to language requirements. The project Italiano, lingua nostra: impact and limitations, in Galaczi, E D and Weir, C J (Eds) *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press, 45–61.

Huhta, A, Tammelin-Laine, T, Hirvelä, T, Neittaanmäki, R and Stordell, E (2017) Aikuisten maahanmuuttajien kotoutumiskoulutuksen tulosten ennustaminen lähtötason arviointien perusteella [Predicting the outcomes of adult immigrants' integration training from placement assessments], *Aikuiskasvatus* [*Adult Education*] 37 (3), 190–203.

Machetti, S and Rocca L (2016) *Integration of migrant citizens, from language proficiency to the knowledge of society: the Italian case, procedures from Council of Europe Language Policy Unit*, paper presented at The Linguistic Integration of Adult Migrants: Some Lessons from Research, Strasbourg, 30 March–1 April 2016.

Rocca L (2017) From illiteracy to A1 within the migration context, procedures from the LESSLA Symposium, in Sosinski (Ed) *Alfabetizacion y aprendizaje de idiomas por adultos*, Granada: Editorial Univesitad de Granada, 207–218.

Rocca L (2017) *Guida alla certificazione CELI in contesto migratorio*, Competenza linguistica e competenza civica Livella A2, Torino: Loescher.

Rocca L and Van Avermaet P (2013) Language testing and access, in Galaczi, E D and Weir, C J (Eds) *Language Testing Frameworks:Proceedings of the ALTE Kraków Conference*, *July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press, 11–36.

Tammelin-Laine, T, Huhta, A, Neittaanmäki, R, Hirvelä, T, Ohranen, S and Stordell, E (2018) Predicting placement accuracy and language outcomes in immigrants' L2 Finnish education, in Davis, J, Norris, J, Malone, M, McKay, T and Son, Y (Eds) *Useful Assessment and*

Evaluation in Language Education, Washington, DC: Georgetown University Press, 149–165.

Ullakonoja, R, Nieminen, L and Huhta, A (2013) Suomen kielen oppimismotivaatio, luetunymmärtäminen ja kirjoitustaito venäjää kotikielenään puhuvilla peruskoululaisilla [Motivation to learn Finnish, reading comprehension and writing ability among comprehensive school students with Russian as the home language], *Kasvatus* [*Education*] 44 (5), 508–521.

Vodičková, K (2014) Designing diagnostic tests in Czech for young migrant learners, in *Efficiency and Responsibility in Education*, Prague: Česká zemědělská univerzita, 891–897.

Vodičková, K and Kostelecká, Y (2014) Diagnostic Tests in Czech for the Children of Immigrants Attending Primary Schools, *E-Pedagogium* 4, 38–51.

Multilingualism and Plurilingualism

De Backer, F, Slembrouck, S and Van Avermaet, P (2020) Functional use of multilingualism in assessment: Opportunities and challenges. *Research Notes 78,* Cambridge Assessment English.

Fiorenza, E and Diego-Hernandez, E (2020) The challenge of assessing plurilingual repertoires: The EVAL-IC project. *Research Notes 78,* Cambridge Assessment English.

Gallagher, A. (2021) *Twenty-Five Years of Language Policies and Initiatives in Ireland. 1995-2020.* Paper presented to Irish Association for Applied Linguistics AGM, 27 November 2020.

Gandini, EAM (2020) Assessing learners' bilingual skills across different levels of the CEFR. *Research Notes 78,* Cambridge Assessment English.

Glaboniat, M and Lorenz-Andreasch, H (2007) DaM-DaF-DaZ. Das „multilinguale Klassenzimmer" – eine Herausforderung, in Glaboniat, M, Rastner, E M and Winstersteiner, W (Eds) *Wir sind, was wir tun. Deutschdidaktik und Deutschunterricht vor neuen Herausforderungen*, Innsbruck, Vienna, Bolzano: StudienVerlag, 123–142.

Roche, J, Jessen, M, Weidinger, N, Behrens, H, Haberzettl, S, Hasselhorn, M, Ifenthaler, D, Kapica, N, Kecker, G, Wolfgang, K, Madlener, K, Pagonis, G, Schug, M, Skoruppa, K, Terrasi-Haufe, E and Thissen, F (2016) Zur Entwicklung eines interaktiven Verfahrens der Sprachstandsermittlung bei mehrsprachig aufwachsenden Kindern: Von der Idee zu ersten Umsetzungsschritten, *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 21 (2), 127–142.

Saville, N (2017) *Linking learning and assessment to support plurilingual education*, paper presented at the 11th International Symposium on Bilingualism, University of Limerick, Limerick, Ireland, June 2017.

Saville, N (2017) *How can multilingualism be supported through language education in Europe?*, paper presented at the 18th Association Internationale de Linguistique Appliquée (AILA) World Congress, Rio de Janeiro, Brazil, July 2017.

Saville, N (2017) *The construct of multilingualism in applied linguistics and language testing: os assuntos do multilinguismo importam*, paper presented at the 39th Language Testing Research Colloquium (LTRC), Bogotá, Colombia, July 2017.

Saville, N (2018) *Global approaches to multilingualism and standardisation*, paper presented at the Global Approaches to Multilingualism and Standardisation conference, Cambridge, UK, April 2018.

Saville, N and Seed, G (2021) Language Assessment in the Context of Plurilingualism. In Piccardo, E, Germain-Rutherford, A and Lawrence, G (Eds) *The Routledge Handbook of Plurilingual Language Education.* New York and London: Routledge.

Seed, G (2020) What is plurilingualism and what does it mean for language assessment? *Research Notes 78*, Cambridge Assessment English.

Seed, G and Holland, M (2020) Taking account of plurilingualism in Cambridge Assessment English products and services. *Research Notes 78,* Cambridge Assessment English.

Reading

Alderson, J C, Haapakangas, E-L, Huhta, A, Nieminen, L and Ullakonoja, R (2015) *The Diagnosis of Reading in a Second or Foreign Language*, New York and London: Routledge.

Alderson, J C and Huhta, A (2011) Diagnosing strengths and weaknesses in second and foreign language reading: What do second language acquisition and language testing have to offer?, in Roberts, L, Pallotti, G and Bettoni, C (Eds) *EuroSLA Yearbook volume 11*, Amsterdam: John Benjamins, 30–52.

Alderson, J C, Huhta, A and Nieminen, L (2016) Characteristics of weak and strong readers in a foreign language, *The Modern Language Journal* 100 (4), 853–879.

ALTE (2001) *Individual Component Checklist: Reading*, available online: www.alte.org/Materials

ALTE (2001) *Individual Component Checklist for use with ONE task: Reading*, available online: www.alte.org/Materials

Arras, U (2014) Qualitätssicherung beim Testen fremdsprachlicher Leseverstehenskompetenz, in Merzig, B, Felfe, M and López Barrios, M (Eds) *Cruce de fronteras – Grenzgänge – Cruzando fronteiras*, Córdoba, Argentinien: Comunic-arte, 879–886.

Bax, S and Weir, C J (2012) Investigating learners' cognitive processes during a computer-based CAE Reading test, *Research Notes* 47, 3–14.

Eckes, T (2015) Lokale Abhängigkeit von Items im TestDaF-Leseverstehen: Eine Testlet-Response-Analyse, *Diagnostica* 61, 93–106.

Eckes, T (2015) Eine Replik auf den Kommentar zum Beitrag „Lokale Abhängigkeit von Items im TestDaF-Leseverstehen", *Diagnostica* 61, 110–111.

Green, A, Khalifa, H and Weir, C J (2013) Examining textual features of reading texts – a practical approach, *Research Notes* 52, 24–39.

Grego Bolli, G, Rini, D and Spina, S (2017) Predicting readability of texts for Italian L2 students: A preliminary study, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 272–278, available online: http://events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Huhta, A, Alderson, J Charles, Nieminen, L and Ullakonoja, R (2016) The role of background factors in the diagnosis of SFL reading, in Banerjee, J and Tsagari, D (Eds) *Contemporary Second Language Assessment*, London: Continuum, 125–146.

Huhta, A and Randell, E (1996) Multiple-choice summary: A measure of text comprehension, in Cumming, A and Berwick, R (Eds) *Validation in Language Testing*, Clevedon: Multilingual Matters, 94–110.

Khalifa, H and Weir, C J (2009) *Examining Reading: Research and Practice in Assessing Second Language Reading*, Studies in Language Testing volume 29, Cambridge: UCLES/Cambridge University Press.

Todea, A (2017) Developing an analysis grid for B1/ B2 reading tasks and items, *Studia Universitatis Babeș-Bolyai, Philologia* 52 (2), 167–182.

Speaking

ALTE (2001) *Individual Component Checklist: Speaking*, available online: www.alte.org/Materials

ALTE (2001) *Individual Component Checklist for use with ONE task: Speaking*, available online: www.alte.org/Materials

ALTE (2014) *The CEFR Grid for Speaking*, available online: www.alte.org/Materials

Arras, U (2006) Was macht eine Aufgabe schwierig? Schwierigkeitsdeterminanten in Tests zur Überprüfung mündlicher Kompetenzen am Beispiel von TestDaF-Aufgaben des Subtests Mündlicher Ausdruck, in Oesterreicher, M and Zahn, R (Eds) *Lingua Franca – Lingua Academica: Mehrsprachigkeit im europäischen Hochschulraum*, Dokumentation der 24. Arbeitstagung 2006 des AKS, Bochum: AKS-Verlag, 211–225.

Arras, U (2007) Zur Revision des Subtests „Mündlicher Ausdruck" der Prüfung „Test Deutsch als Fremdsprache" (TestDaF), *Studienkolleg: Zeitschrift zur Pädagogik und Didaktik studienvorbereitender Kurse für ausländische Studierende* 11, 5–29.

Arras, U (2011) Mündliche Kompetenzen in der Fremdsprache fair messen: Überlegungen und Vorschläge zur Qualitätssicherung, *Babylonia* 11 (2), 80–85.

Bolton, S, Glaboniat, M, Lorenz, H, Perlmann-Balme, M and Steiner, S (2008) *MÜNDLICH. Mündliche Produktion und Interaktion Deutsch. Illustration der Niveaustufen des Gemeinsamen Europäischen Referenzrahmens*, Berlin/Munich: Langenscheidt.

Chambers, L and Ingham, K (2011) The BULATS Online Speaking Test, *Research Notes* 43, 21–25.

Eckes, T (2005) Analyse und Evaluation sprachproduktiver Prüfungen beim TestDaF, in Kühn, I, Lehker, M and Timmermann, W (Eds) *Sprachtests in der Diskussion*, Wittenberger Beiträge zur deutschen Sprache und Kultur, Bd. 4, Frankfurt am Main: Peter Lang, 60–93.

Eckes, T (2010) The TestDaF implementation of the SOPI: Design, analysis, and evaluation of a semi-direct speaking test, in Araújo, L (Ed) *Computer-based assessment (CBA) of foreign language speaking skills*, Luxembourg: Publications Office of the European Union, 63–83.

Galaczi, E and ffrench, A (2011) Context validity, in Taylor, L (Ed) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 112–170.

Georges, S (2013) Evaluer la production orale au travers d'une démarche scientifique, *Revue française de linguistique appliquée* XVIII (1), 47–58.

Glasson, N and Galaczi, E D (2015) 'Seeing the images with different eyeballs': Using text-based vs picture-based tasks in the revised CAE Speaking test, *Research Notes* 62, 23–31.

Grego Bolli, G, Lloyd J and Rini, D (2016) Meeting standards in scoring speaking: improving the quality of examinations in Italian, in Docherty, C and Barker, F (Eds) *Language Assessment for Multilingualism: Proceedings of the ALTE Paris Conference, April 2014*, Studies in Language Testing volume 44, Cambridge: UCLES/Cambridge University Press, 467–489.

Kalaja, P, Huhta, A and Pitkänen-Huhta, A (2003) Kokeeseen osallistujan monet kasvot: ylioppilaskokeiden ja niihin osallistumisen rakentuminen abiturienttien puheessa [The construction of the Matriculation Exam and taking the test in the candidates' spoken discourse], in Koskela, M and Pilke, N (Eds) *Kieli ja asiantuntijuus* [*Language and expertise*], AFinLA Yearbook 2003 [Publications of the Finnish Association of Applied Linguistics 61], Jyväskylä, Finland: Suomen soveltavan kielitieteen yhdistys AFinLA, 93–115.

Kallio, H, Šimko, J, Huhta, A, Karhila, R, Vainio, M, Lindroos, E, Hilden, R and Kurimo, M (2017) Towards the phonetic basis of spoken second language assessment: temporal features as indicators of perceived proficiency level, *AFinLA-E: Soveltavan Kielitieteen Tutkimuksia* 10, 193–213.

Khalifa, H and Salamoura, A (2011) Criterion-related validity, in Taylor, L (Ed) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 259–292.

Lepage, S and North, B (2005) *Séminaire pour le calibrage des productions orales par rapport aux échelles du Cadre européen commun de référence pour les langues*, available online: rm.coe.int/168045a0ce

Platon, E (2016) Rolul atelierelor de benchmarking în standardizarea procesului de evaluare a producţiilor orale, in Facultăţii de Filosofie, Universitatea din Novi Sad & Editura Vasiliana '98, Iaşi (Eds) *Limbă şi identitate românească în continua reconfigurare culturală a Europei. 40 de ani de limba română ca limbă străină la Universitatea „Alexandru Ioan Cuza" din Iași*, Novi Sad: Universitatea din Novi Sad, 252–275.

Platon, E, Vasiu, L and Păcurar, E (2015) *Evaluarea competențelor de comunicare orală în limba română - învățământul primar (EVRO- P1)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E, Manole, V, Varga, C and Lazăr, A (2015) *Evaluarea competențelor de comunicare orală în limba română - învățământul secundar (EVRO-P2)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Taylor, L and Galaczi, E (2011) Scoring validity, in Taylor, L (Ed) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 171–233.

Vasiu, L (2017) Considerations on conversational competence: The case of Romanian as a foreign language (RFL), Level A1, *Studia Universitatis Babeș-Bolyai, Philologia* 2, 153–166.

Vidaković, I and Galaczi, E D (2013) The measurement of speaking ability 1913-2012, in Weir, C J, Vidaković, I and Galaczi, E D, *Measured Constructs: A History of Cambridge English Language Examinations 1913-2012*, Studies in Language Testing volume 37, Cambridge: UCLES/Cambridge University Press, 257–346.

Vodičková, K (2012) Kriteriální rysy mluvených projevů na úrovni C1 včeštině [Criterial features in spoken performances at the C1 level in Czech], in *Čeština jako cizí jazyk VI*, Praha: FF UK, 209–215.

Zimmermann, S (2009) Falsche Vorbereitung? – Erkenntnisse aus Teilnehmerleistungen der Prüfungsteile Mündlicher und Schriftlicher Ausdruck im Test Deutsch als Fremdsprache (TestDaF), in Yu, X (Ed) *TestDaF-Training und Studienvorbereitung: Beiträge zur chinesisch-deutschen Fachkonferenz: „TestDaF-Training und Studienvorbereitung" vom 11 bis 12. Oktober 2008 am Deutschkolleg der Tongji-Universität Shanghai*, München: Iudicium, 63–83.

Statistics and Psychometrics

Arras, U, Eckes, T and Grotjahn, R (2002) C-Tests im Rahmen des „Test Deutsch als Fremdsprache" (TestDaF): Erste Forschungsergebnisse, in Grotjahn, R (Ed) *Der C-Test: Theoretische Grundlagen und praktische Anwendungen*, Bochum: AKS-Verlag, 175–209.

Bell, C (2015) A modern perspective on statistical malpractice detection, *Research Notes* 59, 31–35.

Eckes, T (2005) Evaluation von Beurteilungen: Psychometrische Qualitätssicherung mit dem Multifacetten-Rasch-Modell, *Zeitschrift für Psychologie* 213, 77–96.

Eckes, T (2006) Rasch-Modelle zur C-Test-Skalierung, in Grotjahn, R (Ed) *Der C-Test: Theorie, Empirie, Anwendungen/The C-test: Theory, empirical research, applications*, Frankfurt am Main: Peter Lang, 1–44.

Eckes, T (2007) Konstruktion und Analyse von C-Tests mit Ratingskalen-Rasch-Modellen, *Diagnostica* 53, 68–82.

Eckes, T (2008) Assuring the quality of TestDaF examinations: A psychometric modeling approach, in Taylor, L and Weir, C J (Eds) (2008) *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity – Proceedings of the ALTE Berlin Conference, May 2005*, Studies in Language Testing volume 27, Cambridge: UCLES/Cambridge University Press, 157–178.

Eckes, T (2009) Many-facet Rasch measurement, in Takala, S (Ed) *Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (Section H)*, Strasbourg: Council of Europe/Language Policy Division, Section H 1–52.

Eckes, T (2010) Rasch models for C-tests: Closing the gap on modern psychometric theory, in Berndt, A and Kleppin, K (Eds) *Sprachlehrforschung: Theorie und Empirie – Festschrift für Rüdiger Grotjahn*, Frankfurt am Main: Peter Lang, 39–49.

Eckes, T (2010) Standard-Setting bei C-Tests: Bestimmung von Kompetenzniveaus mit der Prototypgruppenmethode, *Diagnostica* 56, 19–32.

Eckes, T (2011) Item banking for C-tests: A polytomous Rasch modeling approach, *Psychological Test and Assessment Modeling* 53, 414–439.

Eckes, T (2012) Examinee-centered standard setting for large-scale assessments: The prototype group method, *Psychological Test and Assessment Modeling* 54, 257–283.

Eckes, T (2013) A study of differential item functioning in the TestDaF reading and listening sections, in Galaczi, E D and Weir, C J (Eds) *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, Studies in Language Testing volume 36, Cambridge: UCLES/Cambridge University Press, 362–388.

Eckes, T (2015) *Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments* (Second edition), Frankfurt am Main: Peter Lang.

Eckes, T (2016) Exploratorische und konfirmatorische Faktorenanalysen, in Caspari, D, Klippel, F, Legutke, M K and Schramm, K (Eds) *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch*, Tübingen: Narr Francke Attempto, 341–355.

Eckes, T (2017) Setting cut scores on an EFL placement test using the prototype group method: A receiver operating characteristic (ROC) analysis, *Language Testing* 34, 383–411.

Eckes, T (forthcoming) Many-facet Rasch measurement: Implications for rater-mediated language assessment, in Aryadoust, V and Raquel, M (Eds) *Quantitative Data Analysis for Language Assessment* volume 2, New York and London: Routledge.

Eckes, T and Baghaei, P (2015) Using testlet response theory to examine local dependence in C-tests, *Applied Measurement in Education* 28, 85–98.

Eckes, T and Grotjahn, R (2006) C-Tests als Anker für TestDaF: Rasch-Analysen mit dem kontinuierlichen Ratingskalen-Modell, in Grotjahn, R (Ed) *Der C-Test: Theorie, Empirie, Anwendungen/The C-test: Theory, empirical research, applications*, Frankfurt am Main: Peter Lang, 167–193.

Grego Bolli, G, Pelliccia, F and Hardcastle, P (2007) Using Classical Test Theory to Establish Test Version Equivalence and Stability in a European language test suite - the Case of the CELI, in Carlsen, C and Moe, E (Eds) *A Human Touch to Language Testing*, Oslo: Novus Forlag, 58-75.

Kecker, G (2014) Aktuelle Entwicklungen in der Messung von Sprachkompetenzen: Einführung, *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 19 (2), 1–4.

Mavel, M, Bessonneau, P and Lalanne, C (2009) Approches psychométriques de l'évaluation à l'aide de test, *Les après-midis de LAIRDIL No. 14, Le suivi des apprenant/es par les systèmes numériques*, 19–44.

Stanca, L and Vîlcu, D (2017) Statistical Analysis of Romanian as a Foreign Language Pretest and Exam Data, *Studia Universitatis Babeş-Bolyai, Philologia* 2/2017, 93–112.

Vîlcu, D (2017) The Calculation of Cut Off Score in Romanian as a Foreign Language Pretest and Examinations, *Studia Universitatis Babeş-Bolyai, Philologia* 2/2017, 113–128.

Zeidler, B (2017) Cut scores for combined constructs, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 219–223, available online: events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Teaching and Learning

Alanen, R, Huhta, A and Tarnanen, M (2010) Designing and assessing L2 writing tasks across CEFR proficiency levels, in Bartning, I, Martin, M and Vedder, I (Eds) *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*, EUROSLA Monograph Series 1, 21-56, available online: http://eurosla.org/monographs/EM01/EM01home.html

Alanen, R, Huhta, A, Jarvis, S, Martin, M and Tarnanen, M (2012) Issues and challenges in combining SLA research and language testing, in Tsagari, D and Csepes, I (Eds) *Collaboration in Language Testing and Assessment, Language Testing and Evaluation* volume 26, Frankfurt: Peter Lang, 15–30.

Deygers, B (2010) Stage lopen in NT2-workshops aan de UGent: een onderzoek naar de effecten van een coachingstag, in Van Hoyweghen, D, Bogaert, N, Cajot, G, Daems, F and De Moer, A (Eds) *Naar taalkrachtige lerarenopleidingen. Bouwstenen voor een taalbeleid*, Mechelen: Plantyn, 1–15.

Galaczi, E D and Swabey, M T (2015) Capturing the ephemeral: Standardising the assessment of teaching practice, in Wilson, R and Poulter, M (Eds) *Assessing Language Teachers' Professional Skills and Knowledge*, Studies in Language Testing volume 42, Cambridge: UCLES/Cambridge University Press, 113–141.

Glaboniat, M (2002) Ein Sprachkenntnisnachweis zur Befreiung von den geplanten Deutschkursen, *Teaching Human Rights. Informationen zur Menschenrechtsbildung* 2002/13, 14–15.

Glaboniat, M (2004) Der Konjunktiv im Mutter- und Fremdsprachenunterricht. Didaktische Betrachtungen aus verschiedenen Perspektiven, *ide* 2004/4, 41–61.

Glaboniat, M, Rastner, E M and Wintersteiner, W (Eds) (2007) *Wir sind, was wir tun. Deutschdidaktik und Deutschunterricht vor neuen Herausforderungen*, Innsbruck, Vienna, Bolzano: StudienVerlag.

Huhta, A (2007) Itsearviointi osana kielitestiä – tutkimus käyttäjien suhtautumisesta tietokoneen antamaan palautteeseen kielitaidon itsearvioinnista [Self-assessment as part of a language test – a study of user reactions to computerised feedback on self-assessment], in Salo, O-P, Nikula, T and Kalaja, P (Eds) *Kieli oppimisessa – Language in learning. AFinLA yearbook 2007*, Publications of the Finnish Association of Applied Linguistics 65, 371–388.

Khalifa, H (2008) Testing teaching knowledge: developing a quality instrument to support professional development, in Taylor, L and Weir, C J (Eds) *Multilingualism and Assessment: Achieving Transparency, Assuring Quality, Sustaining Diversity – Proceedings of the ALTE Berlin Conference May 2005*, Studies in Language Testing volume 27, Cambridge: UCLES/Cambridge University Press, 191–202.

Martyniuk, W (2016) 'Learning through languages' Conference of the European Centre for Modern Languages (ECML), *Research Notes* 63, 30–33.

Pižorn, K and Huhta, A (2016) Assessment in educational settings, in Tsagari, D and Banerjee, J (Eds) *Handbook of Second Language Assessment*, Boston/Berlin: De Gruyter Mouton, 239–254.

Plassmann, S (2017) Welche Sprache lehren wir? Zur Rolle der Sprachkurse und -prüfungen im Prozess der Integration, in Middeke, A, Eichstaedt, A, Jung, M and Kniffka, G (Eds) *Materialien Deutsch als Fremd- und Zweitsprache 97*, Göttingen: Universitätsdrucke Göttingen, 325–338.

Saville, N (2016) *Breaking out of the box: understanding relationships between learning and assessment*, paper presented at the IATEFL Testing, Evaluation and Assessment Special Interest Group (TEASIG), Aigle, Switzerland, October 2016.

Saville, N (2017) *A systemic model for connecting assessment with teaching and learning*, paper presented at the 4th Annual International Conference of the Asian Association for Language Assessment (AALA), Taipei, Taiwan, June 2017.

Saville, N (2017) *Teaching knowledge in the API Economy*, paper presented at the 18th Association Internationale de Linguistique Appliquée (AILA) World Congress, Rio de Janeiro, Brazil, July 2017.

Saville, N (2017) *Multilingualism post Brexit: language as part of lifelong learning in Europe*, paper presented at the European Day of Languages event, Brussels, Belgium, September 2017.

Tarnanen, M and Huhta, A (2011) Foreign language assessment practices in the comprehensive school in Finland, in Tsagari, D and Csepes, I (Eds) *Classroom-Based Language Assessment*, Frankfurt: Peter Lang, 129–146.

Vîlcu, D (2011) Evaluarea elevilor din clasa a VIII-a. Între examenul public și cel „social”, in *Romậna ca limbă nematernă (RLNM), Actele Conferinței de lansare a proiectului „Perfecționarea cadrelor didactice din învățământul preuniversitar care predau limba română minorităților naționale”*, Cluj-Napoca: Editura Casa Cărții de Știință, 284–300.

Zimmermann, S (2009) Vom Lehrwerk zum Unterrichtskonzept: Überlegungen zur Unterrichtsgestaltung mit dem Lehrwerk TestDaF und Studienvorbereitung, in Yu, X (Ed) *TestDaF-Training und Studienvorbereitung: Beiträge zur chinesisch-deutschen Fachkonferenz: „TestDaF-Training und Studienvorbereitung" vom 11. bis 12. Oktober 2008 am Deutschkolleg der Tongji-Universität Shanghai*, München: Iudicium, 110–124.

Technology

Alderson, J C and Huhta, A (2005) The development of a suite of computer-based diagnostic tests based on the Common European Framework, *Language Testing* 22, 301–320.

Casanova, D, Crendal, A, Holle, A and Demeuse, M (2011) *Élaboration d'une version électronique équivalente à la version papier-crayon d'un test de français langue étrangère à forts enjeux*, available online: www.lefrancaisdesaffaires.fr/wp-content/uploads/2016/12/%C3%89laboration-d%E2%80%99une-version-%C3%A9lectronique-%C3%A9quivalente-%C3%A0-la-version-papier-crayon-d%E2%80%99un-test-de-fran%C3%A7ais-langue-%C3%A9trang%C3%A8re-%C3%A0-for.pdf

Desroches, F, Casanova, D, Crendal, A, Renaud, F and Demeuse, M (2005) *Mise en équivalence des versions papier-crayon et électronique d'un test de français langue étrangère et mise en perspective des résultats par une procédure d'auto-évaluation*, available online: www.researchgate.net/profile/Dominique_Casanova/publication/228452132_Mise_en_equivalence_des_versions_papier-crayon_et_electronique_d%27un_test_de_francais_langue_etrangere_et_mise_en_perspective_des_resultats_par_une_procedure/links/546ddf070cf2a7492c560ce6/Mise-en-equivalence-des-versions-papier-crayon-et-electronique-dun-test-de-francais-langue-etrangere-et-mise-en-perspective-des-resultats-par-une-procedure.pdf?origin=publication_detail

Eckes, T (2010) Der Online-Einstufungstest Deutsch als Fremdsprache (onDaF): Theoretische Grundlagen, Konstruktion und Validierung, in Grotjahn, R (Eds) *C-Test: Beiträge aus der aktuellen Forschung/The C-test: Contributions from current research*, Frankfurt am Main: Peter Lang, 125–192.

Grego Bolli, G (2002) CommuniCAT: a Computer Adaptive Testing for Italian as a foreign language, in *L'Italia nella lingua e nel pensiero. LIBRI E RIVISTE D'ITALIA* 2, Roma-Torino: Ministero per i beni culturali e ambientali, 607–614.

Jordan, S, Hughes, G and Betts, C (2011) Technology in assessment, *Research Notes* 43, 2–6.

Saville, N (2017) *Data and devices: The 4th Industrial Revolution*, paper presented at the Japan Association for Language Teaching (JALT) International Conference, Tsukuba, Japan, November 2017.

Saville, N (2017) Digital Assessment, in Carrier, M, Damerow, R M and Bailey, K M (Eds) *Digital Language Learning and Teaching: Research, Theory, and Practice*, New York and London: Routledge, 198–207.

Xu, J and Seed, G (2017) Integrating technology with language assessment: Automated speaking assessment, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 286–291,

available online: http://events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Walczak, A (2015) Computer-adpaptive testing, *Research Notes* 59, 35–39.

Vocabulary

Huhta, A (2007) The Vocabulary Size Placement Test in DIALANG: why do users love and hate it?, in Carlsen, C and Moe, E (Eds) *A Human Touch to Language Testing*, Oslo: Novus Forlag, 44–57.

Huhta, A and Mäntylä, K (2009) Miten tutkia englannin sananmuodostusta: tutkimusvälineiden vertailua [How to study English word-formation: comparing research instruments], *AFinLA-e* 1, 19–32.

Mäntylä, K and Huhta, A (2014) Knowledge of word parts, in Milton, J and Fitzpatrick, T (Eds) *Dimensions of Vocabulary Knowledge*, New York: Palgrave McMillan, 45–59.

Papp, S and Nicholson, G (2011) Vocabulary acquisition in children and Cambridge ESOL's wordlists for tests for young learners aged 7–14, *Research Notes* 46, 13–22.

Writing

ALTE (2001) *Individual Component Checklist: Writing*, available online: https://www.alte.org/Materials

ALTE (2001) *Individual Component Checklist for use with ONE task: Writing*, available online: https://www.alte.org/Materials

ALTE (2014) *The CEFR Grid for Writing Tasks v3.1 (analysis)*, available online: https://www.alte.org/Materials

ALTE (2014) *The CEFR Grid for Writing Tasks v3.1 (presentation)*, available online: https://www.alte.org/Materials

Arras, U (2007) *Wie beurteilen wir Leistung in der Fremdsprache? Strategien und Prozesse bei der Beurteilung schriftlicher Leistungen in der Fremdsprache am Beispiel der Prüfung Test Deutsch als Fremdsprache (TestDaF)*, Tübingen: Narr.

Arras, U (2009) Wie es zu einer Beurteilung kommt: Ein Forschungsbericht zu Strategien bei der Beurteilung schriftlicher Leistungen im Kontext der Prüfung TestDaF, in Hunstiger, A and Koreik, U (Eds) *Chance Deutsch: Schule – Studium – Arbeitswelt (Materialien Deutsch als Fremdsprache 78)*, Göttingen: Universitätsverlag, 179–196.

Arras, U (2015) Exzerpieren als Schreibkompetenz im Hochschulkontext aus Sicht eines Sprachtests, in Böcker, J and Stauch, A (Eds) *Konzepte aus der Sprachlehrforschung – Impulse für die Praxis: Festschrift für Karin Kleppin*, Frankfurt am Main: Peter Lang, 409–434.

Chambers, L (2011) Composition and revision in computer-based written assessment, *Research Notes* 43, 25–32.

Cooze, M (2011) Assessing Writing tests on scoris®: The introduction of online marking, *Research Notes* 43, 12–15.

Eckes, T (2012) Operational rater types in writing assessment: Linking rater cognition to rater behavior, *Language Assessment Quarterly* 9, 270–292.

Eckes, T, Müller-Karabil, A and Zimmermann, S (2016) Assessing writing, in Tsagari, D and Banerjee, J (Eds) *Handbook of Second Language Assessment*, Boston/Berlin: De Gruyter Mouton, 147–164.

Eckes, T, Weiss-Motz, F and Whelan-Mostofizadeh, S (2009) Ermittlung fairer Ergebnisse im Prüfungsteil Schriftliche Kommunikation des Deutschen Sprachdiploms, *Deutsche Lehrer im Ausland* 56, 15–22.

Endres, H (2012) A comparability study of computer-based and paper-based writing tests, *Research Notes* 49, 26–33.

ffrench, A, Bridges, G and Beresford-Knox, J (2012) Quality assurance: Writing examiners, *Research Notes* 49, 11–17.

Glaboniat, M (1999) Auf der Suche nach Objektivität – Zur Benotung und Bewertung schriftlicher Textproduktionen, *Fremdsprache Deutsch* 1999/19, 20–23.

Holzknecht, F, Huhta, A and Lamprianou, I (2018) Comparing the outcomes of two different approaches to CEFR-based rating of students' writing performances across two European countries, *Assessing Writing* 37, 57–67.

Huhta, A, Alanen, R, Tarnanen, M, Martin, M and Hirvelä, T (2014) Assessing learners' writing skills in a SLA study: Validating the rating process across tasks, scales and languages, *Language Testing* 31 (3), 307–328.

Huhta, A, Harsch, C, Leontjev, D and Nieminen, L (forthcoming) *The Diagnosis of Writing in a Second or Foreign Language*, New York and London: Routledge.

Leontjev, D, Huhta, A and Mäntylä, K (2016) Word derivational knowledge and writing proficiency: How do they link?, *System* 59, 73-89.

Lepage, S, Marty, R, Perlmann-Balme, M and Stelter, C (2011) Évaluation de la production écrite dans le cadre de l'Enquête européenne sur les compétences linguistiques (ESLC), *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 19 (2), 58–77.

Lim, G S (2015) Aspects of the revision process for tests of Writing, *Research Notes* 62, 32–35.

Platon, E, Sonea, I, Radu, A and Tărău Ș (2015) *Evaluarea competențelor de comunicare scrisă în limba română - ciclul primar (EVRO-P3)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Platon, E, Sonea, I, Tărău, Ș (2015) *Evaluarea competențelor de comunicare scrisă în limba română - învățământul secundar (EVRO-P4)*, Cluj-Napoca: Editura Casa Cărții de Știință.

Seed, G (2012) Perceptions of authenticity in academic writing test tasks, *Research Notes* 49, 17–25.

Vîlcu, D, Arieșan, A and Vasiu, L (2017) The use of test taker productions in redesigning writing assessment grids. A corpus based study, in ALTE (2017) *Learning and Assessment: Making the Connections – ALTE 6th International Conference, 3-5 May 2017: Conference Proceedings*, 112-117, available online: events.cambridgeenglish.org/alte2017-test/perch/resources/alte-2017-proceedings-final.pdf

Zimmermann, S (2010) Schreiben im Test – Schreiben im Studium: Wie können Anforderungen an das akademische Schreiben in Tests abgebildet werden? Das Beispiel TestDaF, in Brandl, H, Duxa, S, Leder, G and Riemer, C (Eds) *Ansätze zur Förderung akademischer Schreibkompetenz an der Hochschule (Materialien Deutsch als Fremdsprache 83)*, Göttingen: Universitätsverlag, 1–10.

Zimmermann, S and Rupprecht, E (2013) Typisch DaZ? Ein Vergleich schriftlicher Leistungen von Studierenden mit Deutsch als Erst-, Zweit- und Fremdsprache, in Brandl, H, Arslan, E, Langelahn, E and Riemer, C (Eds) *Mehrsprachig in Wissenschaft und Gesellschaft: Mehrsprachigkeit, Bildungsbeteiligung und Potenziale von Studierenden mit Migrationshintergrund*, Bielefeld: Universität Bielefeld, 81–89.