

# ILTA Bibliography of PhDs in Language Testing

---

This bibliography is provided by ILTA as a service to students of language testing. In order to ensure that it is as comprehensive as possible, when you complete your PhD, or if you know of any others that have not been included, please send the bibliographic details to [t.brunfaut@lancaster.ac.uk](mailto:t.brunfaut@lancaster.ac.uk) . The bibliography is updated annually.

---

- Abdur-Rahman, Muhibullah (2012) Second language test-takers' writing performance across two rhetorically varied writing prompts: a text-analytic validation study. University of Bristol.
- Abeywickrama, Priyanvada (2007) Measuring the knowledge of textual cohesion and coherence in learners of English as a Second Language. University of California, Los Angeles.
- Abu Bakar, Kaseh (2008) Rasch calibration of Arabic reading diagnostic test items. International Islamic University Malaysia.
- Afitska, Oksana (2008) A formative perspective on teaching and assessment: supporting and promoting linguistic development of young ethnic minority children in English primary schools. University of Bristol.
- Akiyama, T. (2004) Introducing speaking tests into a Japanese Senior High School entrance examination. The University of Melbourne.
- Al Fraidan, Abdullah (2011) Test-taking strategies of EFL Saudi university level learners on two varieties of gap-filling vocabulary achievement tests. University of Essex.
- Alavi, Sayyed Mohammad (1997) An investigation of the usefulness of rhetorical structure theory in testing reading comprehension. Lancaster University.
- Alkhelaiwi, Sahar (2017) Cognitive processes, sub-skills and strategies in academic lecture listening at a Saudi Arabian university: A needs analysis study. Lancaster University.
- Allan, Alistair (1992) EFL reading comprehension test validation: investigating aspects of process approaches. Lancaster University.
- Al-Lawati, Zahra (2014) An investigation of the characteristics of language test specifications and item writer guidelines and their effect on item development. Lancaster University.
- Al-Zadjali, Rima (2008) The integrated assessment of reading and writing: Investigating test products and processes in an EFL context. Lancaster University.
- Ameriks, Yoko S. (2009) Investigating validity across two test forms of the Examination for the Certificate of Proficiency in English (ECPE): A multi-group structural equation modeling approach. Teachers College, Columbia University.
- Ashton, Karen (2009) Comparing proficiency levels in an assessment context: the construct of reading for secondary school learners of German, Japanese and Urdu in England. Cambridge University.
- Assiri, Mohammed (2011) Test-taking strategy use on the reading section of the TOEFL-iBT: A study of Arab ESL learners. Oklahoma State University.
- Baba, Kyoko (2007) Dimensions of lexical proficiency in writing summaries for an English as a foreign language test. Ontario Institute for Studies in Education, University of Toronto.
- Bae, Jungok (2000) The construct validation of certain components of English and Korean writing in children participating in either a two-way immersion program or monolingual classes: a writing assessment and latent variable approach. University of California, Los Angeles.
- Baker, Beverly (2011) In the service of the stakeholder: A critical, mixed-method program of research in high-stakes language assessment. McGill University.
- Banerjee, Jayanti (2003) Interpreting and using proficiency scores. Lancaster University.
- Barati, Hossein (2006) A construct-based approach to the nature of reading: the case of understanding, developing and validating reading skills. University of Bristol.
- Batty, Aaron Olaf (2017) The impact of visual cues on item response in video-mediated tests of foreign language listening comprehension. Lancaster University.

- Benjamin, T. (2013) Signaling trouble: On the linguistic design of other-initiation of repair in English conversation. University of Groningen.
- binti Othman, Normah (2006) Assessment of direct writing in ESL classrooms in selected Malaysian secondary schools. Universiti Putra Malaysia.
- Booth, Dawn (2012). Exploring the washback of the TOEIC in South Korea: A sociocultural perspective on student test activity. The University of Auckland.
- Bosker, Hans Rutger (2014) The processing and evaluation of fluency in native and non-native speech. Utrecht University.
- Bower, Jack (2015) Setting and raising standards: creating and validating institutional standardized English reading and listening tests aligned to the Common European Framework Levels A2-B1 at a Japanese university. Macquarie University.
- Brown, Anne D. (2004) Interviewer variability in oral proficiency interviews. The University of Melbourne.
- Brunfaut, Tineke (2008) Foreign language reading for academic purposes: students of English (native speakers of Dutch) reading English academic texts. University of Antwerp.
- Buck, Gary (1991) The testing of listening comprehension: an introspective study. Lancaster University.
- Carr, Nathan (2003) An investigation onto the structure of text characteristics and reader abilities in a test of second language reading. University of California, Los Angeles.
- Chan, Sathena Hiu Chong (2013) Establishing the validity of reading-into-writing test tasks for the UK academic context. University of Bedfordshire.
- Chapman, Mark (2016) The effect of the prompt on writing product and process: a mixed methods approach. University of Bedfordshire.
- Chen, Ann Shih-yi (2010) The role of learner involvement in the assessment process: A multiple-case study investigating the impact of two approaches to student assessment on adult students' learning of oral English. University of Cambridge.
- Chen, Yu-Hua (2009) Investigating lexical bundles across learner writing development. Lancaster University.
- Cheng, Liying (1998) The washback effect of public examination change on classroom teaching. University of Hong Kong.
- Cheung, K Y F (2014) Understanding the authorial writer: A mixed methods approach to the psychology of authorial identity in relation to plagiarism. University of Derby.
- Cho, Dongwan (1995) The effect of specificity of language test specifications on item construction. University of Illinois at Urbana-Champaign.
- Cho, Yeonsuk (2001) Examining a process-oriented writing assessment in a large-scale ESL testing context. University of Illinois at Urbana-Champaign.
- Clapham, Caroline (1994) The effect of background knowledge on EAP reading test performance. Lancaster University.
- Colby-Kelly, Christian (2011) Using "Assessment for Learning" practices with pre-university level ESL students: A mixed methods study of teacher and student performances and beliefs. McGill University.
- Coombe, Christine (1992) The relationship between self-assessment ratings of functional literacy skills and basic English skills test results in adult refugee ESL learners. The Ohio State University.
- Cordova, Pedro (2001) Learners' input on L2 testing: an issue of washback. Spanish, Italian and Portuguese. University of Illinois at Urbana-Champaign.

- Corrigan, Michael (2015) Construct representation of First Certificate in English (FCE) reading. University of Bedfordshire.
- Cotos, Elena (2010) Automated writing evaluation for non-native speaker English academic writing: The case of IADE and its formative feedback. Iowa State University.
- Csepes, Ildiko (2003) Measuring oral proficiency through paired performance. Eötvös Loránd University.
- Daftarifard, Parisa (2016) Reading comprehension as an interactional ability: Evidence from task based versus response based intervention. Islamic Azad University.
- Dakin, Jee Wha (2010) Investigating the simultaneous growth of and relationship between grammatical knowledge and civics content knowledge of low-proficiency adult ESL learners. Teachers College, Columbia University.
- Dávid, A. Gergely (2000) The use of multitrak items and a small-group oral in the context of the Hungarian higher education. University of Glamorgan.
- De Jong, John H.A.L. (1991) Defining a variable of foreign language ability: an application of item response theory. Twente University.
- Delieza, Xenii (2014). The impact of oral examiners' conduct on candidates' performance: The case of the KPG exam in English. National and Kapodistrian University of Athens.
- Deygers, Bart (2017) Assessing high-stakes assumptions. A longitudinal mixed-methods study of university entrance language tests, and of the policy that relies on them. University of Leuven.
- DiGennaro, K. (2011) An exploration into the writing ability of generation 1.5 and international second language writers: A mixed methods approach. Teachers College, Columbia University.
- Dimitrova-Galaczi, Evelina (2004). Peer-peer Interaction in a paired speaking test: the case of the First Certificate in English. Teachers College, Columbia University.
- Djiwandono, Patrisius (1998) The relationship between EFL learning strategies, degree of extroversion, and oral communication proficiency. State University of Malang.
- Dragemark Oscarson, Anne (2009) Self-assessment of writing in learning English as a foreign language. A study at the upper secondary school level. University of Gothenburg.
- Duan, Y. (2007) The Influence of the Chinese University Entrance Exam (English). University of Kent.
- Ducasse, A. (2009) Assessing interaction in paired oral proficiency assessments in Spanish. The University of Melbourne.
- Dunlea, Jamie (2015) Validating a set of Japanese EFL proficiency tests: demonstrating locally designed tests meet international standards. University of Bedfordshire.
- Dunn, Karen (2014) What makes L2 words difficult to know? Using Explanatory Item Response Theory to model the difficulty of vocabulary test items for learners of English as a Second Language. Lancaster University.
- East, M. (2005) Dictionary use in foreign language writing exams. The University of Auckland.
- Elder, Catherine A. (1997) The background speaker as learner of Italian, Modern Greek and Chinese: implications for foreign language assessment. The University of Melbourne.
- Elham Fruzande (2008) TEFL program evaluation at Master's level in Iran. Shiraz University, Iran.
- Ellerton, A.W. (1997) Considerations in the validation of semi-direct oral testing. University of Reading.

- Eom, M. (2006) An investigation of operationalized constructs of second language listening tests. University of Iowa.
- Erdosy, Usman (2005) Responding to non-native and native writers of English: A history professor's indigenous criteria for grading and feedback in an undergraduate Sinology course. University of Toronto.
- Farnsworth, Timothy (2008) The construct of academic English language in tests of academic achievement and its effects on student performance: a confirmatory factor analytic study. University of California, Los Angeles.
- Feenstra, Hiske M. (2014) Assessing writing ability in primary education: on the evaluation of text quality and text complexity. University of Twente.
- Fernández, Miguel (2007) Methodological proposal for the creation of a new English test for university entrance in Spain. University of Granada. [Spanish]
- Fortus, Ruth (2008) The effect of language and modality on text comprehension. Tel-Aviv University.
- Fox, Janna (2001) It's all about meaning: L2 test validation in and through the landscape of an evolving construct. McGill University.
- Froetscher, Doris (2017) An investigation into the washback of a standardized national exam on the classroom testing of reading. Lancaster University.
- Frost, Kellie (2017) Test impact as dynamic process: individual experiences of the English test requirements for permanent skilled migration in Australia. The University of Melbourne.
- Fulcher, Glenn (1993) The construction and validation of rating scales for oral tests in English as a foreign language. Lancaster University.
- Gebril, Atta (2006) Independent and integrated academic writing tasks: a study in generalizability and test method. The University of Iowa.
- Geranpayeh, Ardeshir (2000) Language proficiency testing: a comparative analysis of IELTS and TOEFL. University of Edinburgh.
- Glover, Philip (2006) Examination influence on how teachers teach: a study of teacher talk. Lancaster University.
- Goodwin, Sarah (2016) Locus of control in L2 English listening assessment. Georgia State University.
- Gosa, Codruta (2004) Investigating washback: A case study using student diaries. Lancaster University.
- Grabowski, Kirby (2009) Investigating the construct validity of a test designed to measure grammatical and pragmatic knowledge in the context of speaking. Teachers College, Columbia University.
- Green, Anthony B. (2003) Test impact and English for academic purposes: A comparative study in backwash between IELTS preparation and university preessional courses. University of Surrey.
- Green, Brent A. (2007) Factors associated with ESL test reform in a local context. University of California, Los Angeles.
- Gruba, Paul A. (1999) The role of digital video media in second language listening comprehension. University of Melbourne.
- Gu, Xiangdong (2005) Positive or negative? An empirical study of CET washback on college English teaching and learning in China. Shanghai Jiaotong University.

- Gui, Min (2011) Investigating the construct validity of the reading comprehension section of the College English Test in China: A structural equation modeling approach. The University of Texas at Austin.
- Gyllstad, Henrik (2007) Testing English collocations: Developing receptive tests for use with advanced Swedish learners. Lund University.
- Harding, Luke W. (2008) The use of speakers with L2 accents in academic English listening assessment: A validation study. The University of Melbourne.
- Hayes, Belinda M. (2003) IELTS preparation in New Zealand: An investigation into the nature of the courses and evidence of washback. Victoria University of Wellington.
- Hill, Kathryn (2009) The role of classroom-based assessment and the issue of continuity in foreign language learning. The University of Melbourne.
- Holmes, J.L. (1996) Studying in two languages: Study summaries in the classroom. Lancaster University.
- Horak, Tania (2012) An investigation into the effect of the Skills for Life Strategy on assessment and classroom practices in ESOL teaching in England. Lancaster University.
- Hoshino, Yuko (2009) Factors affecting performance of Japanese EFL learners in multiple-choice vocabulary tests in sentential context. University of Tsukuba.
- Hosseini Heidari (2008) On the translation evaluation methods practiced in Iranian universities. Shiraz University.
- Huang, Heng-Tsung Danny (2010) Modeling the relationships among topical knowledge, anxiety, and integrated speaking test performance: A structural equation modeling approach. The University of Texas at Austin.
- Huang, Liyan (2009) Washback on teacher beliefs and behaviour: Investigating the process from a social psychology perspective. Lancaster University.
- Hyland, Theresa A. (2005) A study of the practices of EL1 and EL2 students in reporting information from sources in a timed writing proficiency assessment. OISE at University of Toronto.
- Imao, Yasuhiro (2010) Investigating the construct of lexico-grammatical knowledge in an academic ESL writing test. University of California, Los Angeles.
- In'nami, Yo (2007). The effects of task types on listening test performance: a quantitative and qualitative study. University of Tsukuba, Japan.
- Inoue, Chihiro (2011) Task parallelness: Investigating the difficulty of two spoken narrative tasks. Lancaster University.
- Isaacs, Talia (2010) Issues and arguments in the measurement of second language pronunciation. McGill University.
- Jang, Eunice (2005) A validity narrative: effects of reading skills diagnosis on teaching and learning in the context of NG TOEFL. University of Illinois at Urbana-Champaign.
- Janssen, Gerriet (2016) English proficiency and entrance to doctoral studies at one Colombian university: A standard setting study. University of Hawai'i at Mānoa.
- Ji, Yong-Il (2012). Comparing strategic processes in the iBT speaking test and in the academic classroom. University of Leicester.
- Johnson, Jeffrey S. (2002) Construct validation of the Secondary Level English Proficiency (SLEP) Test at a Japanese junior college. Temple University.
- Johnson, Robert (2015) Assessing the assessments: Using an argument-based validity framework to assess the validity and use of an English placement system in a foreign language context. Macquarie University.

- Jones, Neil (1992) An item bank for testing language proficiency - using the Rasch model to construct an objective measure. Edinburgh University.
- Joo, M.-J. (2007) Korean university students' attitudes to and performances on a Face-to-Face Interview (FTFI) and a Computer Administered Oral Test (CAOT). Institute of Education, University of London.
- Jun, Hee Sung (2014) A validity argument for the use of scores from a web-search-permitted and web-source-based integrated writing test. Iowa State University.
- Kadir, Kadeessa (2008) Framing a validity argument for test use and impact: the Malaysian public service experience. University of Illinois at Urbana-Champaign.
- Kamimoto, Tadimitsu. (2008) Nation's Vocabulary Levels Test and its successors: A re-appraisal. Swansea University.
- Kang, Sungwoo (1999) Modeling relationships between the use of English as a second language learning strategies and the test performance of Asian students. University of Illinois at Urbana-Champaign.
- Kangzhou, Peng (2014) A validation study of listening comprehension based on criterion-referenced language assessment which has been done in 2010. Shanghai International Studies University.
- Kasai, Masahiro (1997) Application of the rule space model to the reading comprehension section of The Test of English as a Foreign Language (TOEFL). University of Illinois at Urbana-Champaign.
- Katona, Lucy (1997) Negotiated oral test discourse in Hungarian EFL examinations. Lorand Eotvos University.
- Kattan, Jeanne (1990) The construction and validation of an EAP test for second year english and nursing majors at Bethlehem University. Lancaster University.
- Kawauchi, Chieko (2003) The effect of pre-task activities on L2 oral performance. Lancaster University.
- Kenyon, Dorry (1995) Linking multiple-choice test scores to verbally-defined proficiency levels : an application to Chinese reading proficiency. University of Maryland.
- Khalifa, Hanan (1997) A study in the construct validation of the reading module of an EAP proficiency test battery: Validation from a variety of perspectives. Reading University.
- Kim, A. (2011) Cognitive diagnosis of second language reading ability based on a language ability framework. Teachers College, Columbia University.
- Kim, Jiyoung (2008) Development and validation of an ESL diagnostic reading-to-write test: an effect-driven approach. University of Illinois at Urbana-Champaign.
- Kim, Jungtae (2006) The effectiveness of test-takers' participation in development of an innovative web-based speaking test for international teaching assistants at American colleges. University of Illinois at Urbana-Champaign.
- Kim, Jun-Shik (2010) Developmental patterns of Korean learners corresponding to morphosyntactic items. Lancaster University.
- Kim, Mikyung (2002) Process and product: An investigation of the writing of non-native speakers of English on a computer-based academic English writing test. University of California, Los Angeles.
- Kinoshita, Toru (2000) Understanding the relationships among learners' second language proficiency and selected individual characteristics: A structural equation modeling approach. University of California, Los Angeles.

- Kinzley, Simon (2011) The impact of a university pre-session course on the academic writing behaviours of a group of Chinese undergraduate students studying for a degree in media and cultural studies. Lancaster University.
- Knoch, Ute (2007) Diagnostic writing assessment: The development and validation of a rating scale. University of Auckland.
- Koizumi, Rie (2006) Relationships between productive vocabulary knowledge and speaking performance of Japanese learners of English at the novice level. University of Tsukuba.
- Kollias, Charalambos (2017) Virtual standard setting in language testing: Exploring the feasibility and appropriateness of setting cut scores in synchronous audio and audio-visual environments. Lancaster University.
- Korevaar, Serge (2015) The impact of computer interface design on Saudi students' performance on a L2 reading test. University of Bedfordshire.
- Krausert, S.R. (1991) Determining the usefulness of self-assessment of foreign language skills: post-secondary ESL students' placement contribution. University of Southern California.
- Krishnan S, D. (2010) Validating aspects of a model of academic reading. University of Bedfordshire.
- Kunnan, Antony J. (1991) Test taker characteristics and test performance: a structural modeling approach. University of California, Los Angeles.
- Lallmamode, Sheila P. (2012) The development and argument-based validation of an analytic rubric for L2 writing electronic portfolios. International Islamic University Malaysia
- Lanting, Ashley (2000) An empirical study of a district-wide K-2 performance assessment program: Teacher practices, information gained, and use of assessment results. University of Illinois at Urbana-Champaign.
- Lateigne, Betty (2004) Task descriptions of non-western English language use with cross-cultural comparisons and discussions of language choice. Indiana University of Pennsylvania.
- Lazaraton, Anne (1991) A conversation analysis of structure and interaction in the language interview. University of California at Los Angeles.
- Leaper, David (2015) Consistency in performance in the group oral discussion test: An interactional competence perspective. Macquarie University.
- Lee, Hee-Kyung (2004) Constructing a field-specific integrated writing test for an ESL placement procedure. University of Illinois at Urbana-Champaign.
- Lee, Ho (2005) Collaboration in language testing: Investigating the validity of a new rating procedure for EFL writing assessment in Korea. University of Illinois at Urbana-Champaign.
- Lee, Hyeong-Jong (2008) Testing the measurement invariance of large-scale assessment of second language proficiency using simultaneous multiple group covariance structure and item response analyses. University of Illinois at Urbana-Champaign.
- Lee, King Siong (2012) Investigating student experience of assessment in higher education: A case study. National University of Malaysia.
- Lee, Young-Ju (2005) Construct validation of an integrated, portfolio-based, and computerized EAP placement test: a mixed method approach. University of Illinois at Urbana-Champaign.
- Leontjev, Dmitri (2016) ICAnDoiT: The impact of computerised adaptive corrective feedback on L2 English learners. University of Jyväskylä.

- Leta, Dejenie (1990) Achievement and proficiency in a school leaving examination: A case of the nature of innovation in an Ethiopian setting. Lancaster University.
- Levi, Tziona (2013) The effect of dynamic assessment on the performance of students in oral proficiency tests in English as a foreign language. Tel Aviv University.
- Lewkowicz, Jozefa (1997) Investigating authenticity in language testing. Lancaster University.
- Li, Z. (2015) An argument-based validation study of the English Placement Test (EPT) – Focusing on the inferences of extrapolation and ramification. Iowa State University.
- Liao, Yen-Fen (2009) A construct validation study of the GEPT reading and listening sections: re-examining the models of L2 reading and listening abilities and their relations to lexico-grammatical knowledge. Teachers College, Columbia University.
- Lim, Gad S. (2009) Prompt and rater effects in second language writing performance assessment. University of Michigan.
- Lin, Chih-Kai (2014) Issues and challenges in current generalizability theory applications in rated measurement. University of Illinois at Urbana-Champaign.
- Lin, Shwu-wen (2010) The requirement of English language proficiency for graduation in Taiwanese universities: Its impact on non-English majors and their English curriculum. University of Bristol.
- Liscombe, J. (2007) Prosody and speaker state: Paralinguistics, pragmatics, and proficiency. Columbia University.
- Liu, Hsin-min (2014) Investigating the relationships between a reading test and can-do statements of performance on reading tasks. University of California, Los Angeles.
- Liu, Hui-Chun (Angie) (1997) Constructing and validating parallel forms of performance-based writing prompts in an academic setting. University of Illinois at Urbana-Champaign.
- Llosa, Lorena (2005) Building and supporting a validity argument for a standards-based classroom assessment of English proficiency. University of California, Los Angeles.
- Londe, Zsuzsa (2008) Working memory and English as a Second language listening comprehension tests: a latent variable approach. University of California, Los Angeles.
- Lopez, Alexis A. (2005) Examining the washback of a new large-scale English proficiency test: aligning tests, teaching, and learning. University of Illinois at Urbana-Champaign.
- Lu, Yang (2006) A discourse analytic study of EFL test-takers' spoken discourse competence and its impact on their oral proficiency and spoken grammatical competence. The University of Reading.
- Ludbrook, Geraldine (2012) Investigating the English language needs of CLIL teachers in Italian secondary school science classrooms. Lancaster University.
- Lumley, Thomas J.N. (2001) The process of the assessment of writing performance: the rater's perspective. The University of Melbourne.
- Luoma, Sari (2001) What does your test measure? Construct definition in language test development and validation. University of Jyväskylä.
- Ma, J. (2017) Exploring the effects of the TOEFL preparation courses and materials on Chinese test-takers' TOEFL iBT performance and their English language proficiency. Queen's University at Kingston Ontario.
- MacMillan, Fabiana (2006) Lexical patterns in the reading comprehension section of the TOEFL® Test. Fluminense Federal University.
- Malone, Margaret E. (1999) The development of the English Speaking Test: an investigation of reliability and validity. Georgetown University.

- Manning, Anthony (2013). EAP teacher assessment literacy. University of Leicester.
- May, Lynette (2007) 'Effective interaction' in a paired speaking test: The rater's perspective. The University of Melbourne.
- McCray, Gareth (2014) Statistical modelling of cognitive processing in reading comprehension in the context of language testing. Lancaster University.
- McNamara, Timothy F. (1991) Assessing the second language proficiency of health professionals. The University of Melbourne.
- Moghaddam, Sharif (2010) Students' Preparation for IELTS: Development of Written and Oral Argumentative Texts. University of Adelaide.
- Moradi, Mahmood (2010) – The washback effect of the specialised English test (SPE). Lancaster University.
- Moritz, C. (1995) Self-Assessment of foreign language proficiency: A critical analysis of issues and a study of cognitive orientations of French learners. Cornell University.
- Motteram, Johanna (2016) Reverse engineering language test constructs for Messick's value implications: A sociolinguistic approach. The University of Adelaide.
- Mousavi, Seyyed A. (2007) Development and validation of a multimedia computer package for the assessment of oral proficiency of adult ESL learners: implications for score comparability. Griffith University.
- Nakatsuhara, Fumiyo (2009) Conversational styles in group oral tests: How is the conversation co-constructed? The University of Essex.
- Neumann, Heike (2010) What's in a grade? A mixed methods investigation of teacher assessment of grammatical ability in L2 academic writing. McGill University.
- Nguyen, Thi N.H (2008) An investigation into the validity of two EFL (English as a Foreign Language) Listening Test: IELTS and TOEFL iBT. The University of Melbourne.
- Nkosana, L. (2004) The impact of curriculum and assessment reform on teaching: The case of the BGCSE curriculum in Botswana (2006). The University of Melbourne.
- North, B. (1996) The development of a Common Framework Scale of Descriptors of Language Proficiency based on a theory of measurement. Thames Valley University.
- O'Loughlin, Kieran J. (1997) The comparability of direct and semi-direct speaking tests. The University of Melbourne.
- Ockey, Gary (2006) Making a case for the group oral discussion test: the effects of personality on the group oral's score-based inferences. University of California, Los Angeles.
- Onaiba, Abdulhamid Mustafa El-Murabet (2013). Investigating the washback effect of a revised EFL public examination on teachers' instructional practices, material and curriculum. University of Leicester.
- Orr, Mike (2000) Using rater reports to help interpret test scores in the Cambridge FCE speaking test. University of Valencia.
- Owen, Nathaniel (2017) An evidence-centred approach to Reverse Engineering: Comparative analysis of IELTS and TOEFL iBT reading sections. University of Leicester.
- Papadima-Sophocleous, Salomi (2005) Development, implementation, and evaluation of an online English placement test at college level: a case study. Middlesex University.
- Papageorgiou, Spiros (2007). Setting standards in Europe: the judges' contribution to relating language examinations to the Common European Framework of Reference. Lancaster University.
- Pardo Ballester, Maria Cristina (2007) The development of a web-based Spanish listening placement exam. University of California, Davis.

- Park, Moonyoung (2015) Development and validation of virtual interactive tasks for an aviation English assessment. Iowa State University.
- Park, Sang Bok (2011) Using motivation theories to analyse students' perceptions of an examination and their inclination to study for it. Lancaster University.
- Perrin, Geoffrey (2000) The effect of multiple-choice foreign language tests of listening and reading on teacher behaviour and student attitudes. Lancaster University.
- Perrone, Michael (2010) The impact of the First Certificate of English (FCE) on the EFL classroom: A washback study. Teachers College, Columbia University.
- Phakiti, Aek (2003) An empirical investigation into the relationship of state-trait strategy use to EFL reading comprehension test performance: a structural equation modeling approach. The University of Melbourne.
- Pill, Thomas John Hamilton (2013) What doctors value in consultations and the implications for specific-purpose language testing. The University of Melbourne.
- Purpura, James (1996) Modeling the relationships between test takers' reported cognitive and metacognitive strategy use and performance on language tests. University of California Los Angeles.
- Puvenesvary (2003) A comparative study of the criteria employed by academics and workplace professionals in evaluating business correspondence. The University of Melbourne.
- Pyo, Kyong Hyon (Serena) (2001) Construct validation of an integrated-approach EAP placement test using multi-group structural equation modeling. University of Illinois at Urbana-Champaign.
- Qian, David D. (1998) Depth of vocabulary knowledge: assessing its role in adults' reading comprehension in English as a second language. Ontario Institute for Studies in Education, University of Toronto.
- Raofmoeini, Mohammad (2003) What does C-test Measure? Exploration and validation of strategies and processes used in C-test across L1 & L2. Shiraz University, Iran.
- Rea, Pauline M. (1988) The relationship between grammatical abilities and aspects of communicative competence with special reference to the testing of grammar. Lancaster University.
- Rethinasamy, S. (2006) The effects on rating performance of different training interventions. University of Roehampton.
- Rukthong, Anchana (2016) Investigating the listening construct underlying listening-to-summarize tasks. Lancaster University.
- Sadeghi, Karim (2003) An investigation of cloze procedure as a measure of reading comprehension with reference to educational context in Iran. University of East Anglia.
- Sağlam, Aslı Lidice Göktürk (2016) Washback and instruction sensitivity of a theme-based high-stakes English language proficiency test. Yeditepe University.
- Saif, Shahrzad (2000) Theoretical and empirical considerations in investigating washback: a study of EFL and ESL learners. University of Victoria.
- Sakyi, Alfred A. (2003) A study of the holistic scoring behaviours of experienced and novice ESL instructors. University of Toronto.
- Salisbury, K (2005) The edge of expertise? Towards an understanding of listening test item writing as professional practice. King's College, University of London.
- Salvisberg, JoAnn (2007) Developing a framework for designing a diagnostic instrument for oral skills assessment. Akamai University.

- Sapsarin, Sutthirak (2007) A study of trait factors of oral language ability in a computer-based speaking test for Thai university students. Chulalongkorn University.
- Sasaki, Miyuki (1991) Relationships among second language proficiency, foreign language aptitude, and intelligence: a structural equation modeling approach. University of California, Los Angeles.
- Saville, Nick (2009) Developing a model for investigating the impact of language assessment policy within educational contexts by a public examination provider. University of Bedfordshire.
- Sawaki, Yasuyo (2003) A comparison of summarization and free recall as reading comprehension tasks in web-based assessment of Japanese as a foreign language. University of California, Los Angeles.
- Sawatdirakpong, Sumalee (1993) Native languages and differential dimensionality of English as a second language test proficiency: an exploratory study. University of Illinois at Urbana-Champaign.
- Schoonen, Rob (1991) De evaluatie van schrijfvaardigheidsmetingen. Een empirische studie naar betrouwbaarheid, validiteit en bruikbaarheid van schrijfvaardigheidsmetingen in de achtste groep van het basisonderwijs [The evaluation of writing assessments. An empirical study into the reliability, validity and utility of writing assessments in final grade of elementary school.] University of Amsterdam. [Dutch, summary in English]
- Scott, Catriona (2007) Washback from national curriculum tests on classroom assessment of learners with English as an additional language at key stages 1 and 2. University of Bristol.
- Shahriar Mirshahidi, Shahriar (2017) Predicting international teaching assistants' performance in a domain-specific test: The case of complexity, accuracy, fluency, and compensatory strategies. Oklahoma State University.
- Shin, Dong-il (1999) Construct validation of a diagnostic L2 Listening test: an operational model utilized and multidimensionality issues revisited. University of Illinois at Urbana-Champaign.
- Shin, Sang Kuen (2004) How much is good enough? Setting and validation performance standards and cut scores for the web-based English as a second language placement exam at UCLA. University of California, Los Angeles.
- Shin, Sun-young (2007) Examining the construct validity of a web-based academic listening test: an investigation of the effects of response formats in a web-based listening test. University of California, Los Angeles.
- Shirazi, A. M. (2008). A step towards objectivity in writing assessment. University of Tehran.
- Siddiek, Ahmed Gumaa (2004) Analysis and evaluation of the Sudan School Certificate English Examinations. Omdurman Islamic University.
- Siyari, Masood (2011) On the cross relationship between self/peer-assessment rating accuracy and personality traits, and the contribution to writing performance. Allameh Tabataba'i University.
- Smolik, Marcin (2008) Investigating scoring validity. A study of the 'nowa matura' speaking exam in English at the basic level. Maria Skłodowska Curie University.
- So, Young Jang (2010) The development and evaluation of a systematic training program for increasing both rater reliability and rating accuracy. University of Illinois at Urbana-Champaign.

- Song, Minyoung (2008) The effect of the notetaking format on the quality of test takers' notes and their performance on an academic listening test. University of California, Los Angeles.
- Song, Xiaomei (2014) Test fairness in a large-scale high-stakes languages. Queen's University.
- Spence-Brown, Robyn (2004) Authentic assessment? The implementation of an 'authentic' teaching and assessment task. The University of Melbourne.
- Storey, P. (1995) Investigating construct validity through test taker introspection. University of Reading.
- Suvorov, Ruslan (2013) Interacting with visuals in L2 listening tests: An eye-tracking study. Iowa State University.
- Taylor, Lynda B. (1996) An investigation into the development of the summary completion format as a means of assessing reading comprehension ability. University of Cambridge.
- Tsagari, Dina (2007) Investigating the washback effect of a high-stakes EFL exam in the Greek context: participants' perceptions, material design and classroom applications. Lancaster University.
- Tsai, Constance (2004) Investigating the relationships between ESL writers' strategy use and their second language writing ability. Teachers College, Columbia University.
- Tyacke, Marian (1996) Cognitive diversity in language teaching and testing: a three-phase investigation of the relationships between learning style, learner behaviour and reading task performance. Lancaster University.
- Unaldi, Aylin (2004) Construct validation of the reading subskills of the Bogazici University English Proficiency Test. Bogaziçi University, Turkey.
- Unaldi, Aylin (2010) Investigating reading for academic purposes: sentence, text and multiple texts. University of Bedfordshire.
- Usami, Usami (2012) The application of corpora to language testing – Identifying and improving problematic multiple choice grammar questions in Japanese university entrance exams. Lancaster University.
- Van Maele, Jan (1999) Understanding the assessment of oral foreign language proficiency. an investigation into the rater's judgmental contents structure. Université catholique de Louvain.
- Van Moere, Alistair (2007) Group oral tests: How does task affect candidate performance and test scores? Lancaster University.
- Virkkunen, Anu (1991) Affixation as part of an English proficiency test at the University of Helsinki. University of Minnesota.
- Vongpumivitch, Viphavee (2004) Measuring the knowledge of text structure in academic English as a Second Language (ESL). University of California, Los Angeles.
- Voss, Erik (2012) A validity argument for score meaning of a computer-based ESL collocational ability test based on a corpus-driven approach to test design. Iowa State University.
- Wagner, Matthew Joseph (2006) Utilizing the visual channel: An investigation of the use of videotexts on tests of second language listening ability. Teachers College, Columbia University.
- Wagner, M. (2015) The centrality of cognitively diagnostic assessment for advancing secondary school ESL students' writing: A mixed methods study. Ontario Institute for Studies in Education/University of Toronto.

- Wall, Dianne (1999) The impact of high-stakes examination on classroom teaching: A case study using insights from testing and innovation theory. Lancaster University.
- Waller, Daniel (2015) Investigation into the features of written discourse at levels B2 and C1 of the CEFR. University of Bedfordshire.
- Walters, F. Scott (2004) An application of conversation analysis to the development of a test of second-language pragmatic competence. University of Illinois at Urbana-Champaign.
- Wang, Xiaoying (2014) Investigating assessment practices in three EFL oral language classrooms in China. University of Auckland.
- Watanabe, Yoshinori (1997) The washback effect of the Japanese university examinations in English. Lancaster University.
- Wei Wei (2011) The washback effect and impact of a high-stakes test on classroom teaching – a case study from China. University of Leeds.
- Weigle, S. C. (1994) Effects of training on raters of ESL compositions: An investigation of construct validity. University of California, Los Angeles.
- Wilmes, Carsten (2007) Validation of a German language placement test based on a modified C-test procedure. University of Illinois at Urbana-Champaign.
- Winke, Paula (2005) Individual differences in adult Chinese second language acquisition: The relationships among aptitude, memory and strategies for learning. Georgetown University.
- Wiseman, Cynthia (2008) Investigating selected facets in measuring second language writing using holistic and analytic scoring methods. Teachers College, Columbia University.
- Witton-Davies, Giles (2014) The study of fluency and its development in monologue and dialogue. Lancaster University.
- Woo, Kimberly K. (2014) Examining the construct of oral social language in the English language proficiency assessment of young English Language Learners. New York University.
- Wu, R.W. Jessica (2005) Task difficulty in semi-direct speaking tests. University of Surrey Roehampton.
- Wu, Rachel (2011) Establishing the validity of the General English Proficiency Test reading component through a critical evaluation on alignment with the Common European Framework of Reference. University of Bedfordshire.
- Xi, Xiaoming (2003) Investigating language performance on the graph description task in a semi-direct oral test. University of California, Los Angeles.
- Xie, Qin (2010) Test design and use, preparation, and performance: A structural equation modeling study of consequential validity. University of Hong Kong.
- Xu, Jing (2015) Predicting ESL learners' oral proficiency by measuring the collocations in their spontaneous speech. Iowa State University.
- Yamada, Minako (2002) Mediation by a non-native language: Can cooperative dialogic processes lead to communicative effectiveness? Sophia University, Tokyo.
- Yamashita, Junko (1999) Reading in a first and a foreign language: A study of reading comprehension in Japanese (the L1) and English (the L2). Lancaster University.
- Yang, Rui (2009) A many-facet Rasch analysis of rater effects on an oral English proficiency test. Purdue University.
- Yin, Muchun (2005) A progressively focused qualitative study of teacher thinking in English for academic purposes (EAP) classroom language assessment. University of Bristol.

- Youn, S. J. (2013) Validating task-based assessment of L2 pragmatics in interaction using mixed methods. University of Hawai'i.
- Yu, Guoxing (2005) Towards a model of using summarization tasks as a measure of reading comprehension. University of Bristol.
- Yun, Youngmi (2005) Factors explaining EFL learners' performance in a timed essay writing test: a structural equation modeling approach. University of Illinois at Urbana-Champaign.
- Zahedi, Hameed (2001) A fuzzy-logic approach toward appreciating partial knowledge in MC tests. Shiraz University.
- Zainal Abidin, S.A. (2006) Testing spoken language using computer technology: a comparative validation study of 'live' and computer delivered test versions using Weir's framework. Roehampton University.
- Zhang, W. (1998) The rhetorical patterns found in Chinese EFL student writers' examination essays in English and the influence of these patterns on rater response. Hong Kong Polytechnic University.
- Zhang, Ying (2010) Exploring the native- and nonnative-raters' decision making behavior in the assessment of CET-SET. Monash University.
- Zhao, Cecilia Guanfang (2010) The role of voice in high-stakes second language writing assessment. New York University.
- Zhao, Huihui (2008) Who takes the floor: Peer assessment and teacher assessment? A longitudinal comparative study of peer and teacher assessment in a Chinese university EFL writing class. University of Bristol.
- Zhao, S.-F. (2005) A consideration of the washback effect of English as a Foreign Language Oral Examinations at tertiary level in China: Teaching, testing, and syllabus change. Cardiff University.
- Zhou, Yujia (2008) Effects of computer delivery mode on testing second language speaking. Tokyo University of Foreign Studies.