

Category 1: General

- Amano, I. (2000). Admission policies in Japanese universities. *JLTA (Japan Language Testing Association) Journal*, 3, 30-45. [Category 1]
- Brown, J. D. (1999). The roles and responsibilities of assessment in foreign language education. *JLTA (Japan Language Testing Association) Journal*, 2, 1-21. [Category 1, 15]
- Chapman, M. (2005). A case study of the need for change in the language testing policies of a Japanese corporation. *JLTA (Japan Language Testing Association) Journal*, 8, 51-67. [Category 1, 14]
- Cho, D. W., & Park, J. Y. (2013). Which grading system, a letter-grading system or an S/U system, is more appropriate for language courses? *JLTA (Japan Language Testing Association) Journal*, 16, 15-42. Retrieved from <http://ci.nii.ac.jp/naid/110009767201/en> [Category 1, 16]
- Fujita, T., & Mayekawa, S. (2012). Puroguramu hyoka no tameno saiteki komoku outou moderu to touka houhou no sentaku. *JLTA (Japan Language Testing Association) Journal*, 15, 153-171. [Category 1]
- Hattori, T. (2006). Komoku tokusei kyokusen o mochiita toka keisu no suitei—Dankai hanno model no baai [Estimating equating coefficients using item characteristic curve—The case of graded response model]. *JLTA (Japan Language Testing Association) Journal*, 9, 1-20. [Category 1]
- Hayashi, H. (2006). Jiko kettei riron ni motozuku daigakusei yo eigo gakushu doki zuke shakudo no sakusei—Kizon shakudo tonon hikaku kosatsu [Development of a scale of motivation for learning English for university students based on self-determination theory—Comparison with existent measurement]. *JLTA (Japan Language Testing Association) Journal*, 9, 117-128. [Category 1]
- Hirai, A. (2006). Enhancing test practicality for in-house English proficiency tests: An efficient item-banking system. *JLTA (Japan Language Testing Association) Journal*, 9, 141-153. [Category 1]
- Hirai, A. (2006). Eigo shutoku ni okeru gengogakuteki kyori [Linguistic distance in English language learning]. Bonds of language: A festschrift for Dr. Yasuaki Fujiwara on the occasion of his sixtieth birthday, 602-616. [Category 1]
- Iitaka, T., & Hirai, A. (2006). Open software wo riyo shita siken shutudai saiten to group kanri no rendou [Using open software for item construction, scoring, and grouping]. *IEICE Technical Report, ET2006-9*, 49-54. [Category 1]
- Ikeda, H. (2000). What we need for research on language testing in Japan—A psychometrician's view—. *JLTA (Japan Language Testing Association) Journal*, 3, 1-29. [Category 1]
- In'nami, Y., & Koizumi, Y. (2010). Can structural equation models in second language testing

- and learning research be successfully replicated? *International Journal of Testing*, 10, 262-273. [Category 1]
- In'nami, Y., & Koizumi, R. (2010). Database selection guidelines for meta-analysis in applied linguistics. *TESOL Quarterly*, 44, 169-184. [Category 1]
- In'nami, Y., & Koizumi, R. (2013). Review of sample size for structural equation models in second language testing and learning research: A Monte Carlo approach. *International Journal of Testing*, 13, 329-353. doi:10.1080/15305058.2013.806925 [Category 1]
- In'nami, Y., Koizumi, R., & Nakamura, K. (2016). Factor structure of the Test of English for Academic Purposes (TEAP®) test in relation to the TOEFL iBT® test. *Language Testing in Asia*, 6(3), 1–23. doi:10.1186/s40468-016-0025-9 Retrieved from <http://www.languagetestingasia.com/content/6/1/3> [Category 1, 13]
- Irie, Y. (2012). Chiiki ninongo kyoshitsu ni okeru gaikoku seki jumin no nihongo noryoku ni kansuru jiko hyoka. *JLTA (Japan Language Testing Association) Journal*, 15, 115-132. [Category 1]
- Ito, A., Shimatani, H., Norizuki, K., & Kinoshita, M. (2013). Development and validation of a self-assessment questionnaire as an indicator of TOEIC® test score. *JLTA (Japan Language Testing Association) Journal*, 16, 87–106. Retrieved from <http://ci.nii.ac.jp/naid/110009767204/en> [Category 1]
- Japan Language Testing Association (JLTA) Steering Committee (Eds.), JLTA Committee of Language Testing Terms. (2006). *The JLTA bilingual list of language testing terms / The JLTA Code of Good Testing Practice*. Nagano: JLTA Secretariat. [Category 1, 15]
- Kanaoka, M. (2005). Creative project-based instruction (CPBI) for the development of ESP-focused language proficiencies. *JLTA (Japan Language Testing Association) Journal*, 8, 86-101. [Category 1]
- Koizumi, R., & Iimura, H. (2010). Neural test riron no tokucho: Kotenteki test riron, Rasch Modeling tonu hikakukara [Characteristics of Neural Test Theory: Comparison with Classical Test Theory and Rasch Modeling]. *JLTA (Japan Language Testing Association) Journal*, 13, 91-109. [Category 1]
- Lee, W.-K. (2015). Creativity-fostering assessment of English at schools: Why and how? *JLTA (Japan Language Testing Association) Journal*, 18, 37–54. [Category 1]
- Lee, S. I., Yoshizawa, K., & Shimabayashi, S. (2006). The content analysis of the TOEIC and its relevancy to language curricula in EFL contexts in Japan. *JLTA (Japan Language Testing Association) Journal*, 9, 154-173. [Category 1]
- Murata, K. (2012). Topikku no sentaku wa nihongo gakushushua ni totte yuri to narieruka. *JLTA (Japan Language Testing Association) Journal*, 15, 173-184. [Category 1]
- Nakamura, Y. (1998). A questionnaire on ALT's usage of authentic English. *JLTA (Japan*

- Language Testing Association) Journal, 1*, 15-34. [Category 1]
- Nakamura, Y. (2007). The necessity for co-operation in language testing in Asia. *JLTA (Japan Language Testing Association) Journal, 10*, 1-11. [Category 1]
- Newfields, T. (2001). Voices in the Field: An Interview with Dan Douglas. *Shiken: JALT Testing and Evaluation SIG Newsletter, 5* (3). Available from www.jalt.org/test/dou_new.htm [Category 1, 14]
- Newfields, T. (2002). Voices in the Field: An Interview with Liz Hamp-Lyons. *Shiken: JALT Testing and Evaluation SIG Newsletter, 6* (2). Available from www.jalt.org/test/ham_new.htm [Category 1, 14]
- Newfields, T. (2003). Voices in the Field: An Interview with J. D. Brown. *Shiken: JALT Testing and Evaluation SIG Newsletter, 7* (1). Available from www.jalt.org/test/bro_new.htm [Category 1, 14]
- Newfields, T. (2005). Voices in the Field: An Interview with Yoshinori Watanabe. *Shiken: JALT Testing and Evaluation SIG Newsletter, 9* (1). Available from www.jalt.org/test/wat_new.htm [Category 1, 14]
- Newfields, T. (2006). Teacher development and assessment literacy. In T. Newfields, I. Gledall, M. Kawate-Mierzejewska, Y. Ishida, M. Chapman, & P. Ross (Eds.), *Proceedings of the 2006 JALT Pan-SIG Conference on Lifelong Learning* (pp. 48-73). [Category 1]
- Newfields, T. (2007). Engendering assessment literacy: Narrowing the gap between teachers and testers. In Y. S. Lee, et al. (Ed). *Assessing Foreign Language Performances: Proceedings of the 2007 Korea English Language Testing Association International Conference* (pp. 22-42). Seoul, Korea: Korea English Language Testing Association. Available from www.tnewfields.info/Articles/testlit.htm [Category 1,15]
- Newfields, T. (2007). Voices in the Field: An Interview with Glenn Fulcher. *Shiken: JALT Testing and Evaluation SIG Newsletter, 11* (2). Available from www.jalt.org/test/ful_new.htm [Category 1, 14]
- Ohtomo, K., Trans. (1972). *Eigo no sokutei to hyoka* [Harris, D. P. (1969). *Testing English as a second language*]. Tokyo: ELEC. [Category 1]
- Ohtomo, K. (1978). Oral Approach. In K. Ohtomo, Y. Ogawa, K. Ito, M. Yoshizawa, N. Kakita, A. Sasaki, et al. *Eigo kyojuho kakuron*. [English Teaching Methods and Approaches] (pp. 48-66). Tokyo: Kenkyusha. [Category 1]
- Ohtomo, K. (1978). Testing overall English language proficiency of Japanese students. In I. Koike et al (Eds.) *The teaching of English in Japan* (pp. 463-477). Tokyo: Eichosha. [Category 1]
- Ohtomo, K. (1982). Hihan: Inshibunseki ni yoru tanitsu gengenoryoku kasetu. [Criticism: Unitary competence hypothesis by factor analysis]. *Language Studies, 4*, 1-27. [Category

1]

- Ohtomo, K. (1989). Komoku hanno riron no kenkyu: Komoku tokusei kyokusen. [Studies in Item Response Theory: Item characteristic curves]. *Studies in Foreign Language Teaching*, 11, 23-40. [Category 1]
- Ohtomo, K. (1991). Komoku hanno riron no kenkyu: PROX ni yoru parameter suitei no monndaiten [Studies in Item Response Theory: Problems in PROX]. *Studies in Foreign Language Teaching*, 13, 1-20. [Category 1]
- Ohtomo, K. (1992) Komoku hanno riron no kenkyu: Test no tohka (1) [Studies in Item Response Theory: Equating (1)]. *Gaikokugo Kyoiku Ronsyu*, 14, 1-20. [Category 1]
- Ohtomo, K. (1993) Komoku hanno riron no kenkyu: Test no tohka (2). *Studies in Foreign Language Teaching*, 15, 15-34. [Category 1]
- Ohtomo, K. (1996). *Komoku ohto riron nyumon* [Introduction to item response theory]. Tokyo: Taishukan. [Category 1]
- Ohtomo, K. (1994). Komoku hanno riron no kenkyu: Logit tokuten to Rasch model [Studies in Item Response Theory: Logit scores and Rasch model]. *Studies in Foreign Language Teaching*, 16, 85-100. [Category 1]
- Ohtomo, K. (1994). Gengo test to daini gengo shutoku [Language tests and second language acquisition]. In I. Koike (Ed.), *Daini gengo shutoku kenkyu ni motozuku saishin no eigokyoiku* [New English teaching based on the second language acquisition research] (pp. 300-310). Tokyo: Taishukan. [Category 1]
- Ohtomo, K., & Ikeda, H., Trans. (1997). *Gengo test no kiso* [Bachman, L. (1990). Fundamental considerations in language testing] . Tokyo: Mikuni. [Category 1]
- Ohtomo, K., & Thrasher, R. (Eds.), Trans. (2002). *Jissen gengo test sakusei ho* [Bachman, L., & Palmer, A. (1996). *Language testing in practice*]. Tokyo: Taishukan. [Category 1, 10]
- Ohtomo, K. (Editorial supervisor) (2003). VII kenkyu to sokutei [Research and measurement]. In I. Koike (Editor in chief), *Ohyo gengogaku jiten* [Handbook of Applied Linguistics] (pp. 721-787). Tokyo: Kenkyusha . [Category 1]
- Ohtomo, K., Asano, H., Hattori, T., & Yoshie, M. (1987). Item difficulty of English language tests for Japanese students: The Rasch model calibration. *JACET (Japan Association of English Teachers) Bulletin*, 18, 109-125. [Category 1]
- Ohtomo, K., Asano, H., Hattori, T., & Yoshie, M. (1989). Test-free person measurement in tests of English for Japanese students. *JACET (Japan Association of English Teachers) Bulletin*, 20, 71-87. [Category 1]
- Ohtomo, K., Carroll, J. B., & Ogawa, Y. (1970, February). Sekai no gengo shinri gakusha: J. B. Carroll no gengokan to kyojuho. [The world famous psycholinguist, Dr. John B. Carroll: His view on languages and language teaching]. *Gendai eigo kyoiku* [Modern English

- Teaching*], 2-8. Tokyo: Kenkyusha. [Category 1]
- Ohtomo, K., Hatori, H., Asano, H., Ohinoue, S, Imura, M. & Kiyokawa, H. (Eds.) (1979). *Eigo shidoho hand book (4): Hyoka hen* [Handbook of Teaching English (4): Assessment]. Tokyo: Taishukan. [Category 1]
- Ohtomo, K., Nakamura, K., & Matsui, T. (Eds.). (1984). *Sengo no eigo kyoiku* [Teaching English in Japan after the World War II]. Tokyo: Sanseido. [Category 1]
- Ohtomo, K., (Editorial supervisor) & Nakamura, Y. (2002). *Test de gengo noryoku ha hakareruka*. [Can you measure my language proficiency by tests?]. Tokyo: Kiriara. [Category 1]
- Saida, C. (2004). Ken kibo eigo gakuryoku test no inshi bunseki teki kosatsu [Factor structures of Prefecture-wide English tests for high school students]. *JLTA (Japan Language Testing Association) Journal*, 6, 92-107. [Category 1]
- Saida, C. (2007). Tachigata IRT model o mochiita eigo gakuryoku test no doji toka—Nichigata IRT model ni yoru toka kekka tonon hikaku [Comparison of concurrent calibration of English achievement tests under dichotomous and polytomous item response models]. *JLTA (Japan Language Testing Association) Journal*, 10, 119-133. [Category 1]
- Saito, H. (2004). Language testing and SLA research: Some issues concerning task difficulty. *JLTA (Japan Language Testing Association) Journal*, 6, 30-39. [Category 1]
- Shimada, K. (1997). Communicative language testing: Principles and problems. *English Review*, 12, 3-24. [Category 1]
- Shimizu, Y. (2002). Survey research on the use of placement tests at four-year universities in Japan. *Ritsumeikan Studies in Language and Culture Studies*, 14, 231-243. [Category 1]
- Shimizu, Y. (2005). Sokutei ni okeru datosei no rikai no tame ni--Gengo test no kihon gainen to shite [Understanding of test validity: As a basic concepts in language testing]. *Ritsumeikan Studies in Language and Culture Studies*, 16, 241-254. [Category 1, 13]
- Shimizu, Y., Sugino, N., & Fraser, S. (2003). The construction of a database of university English entrance examinations in Japan. *Testing International*, 13, 8-11. [Category 1]
- Shizuka, T. (2007). Effects of learning English in elementary school days on the proficiency of, and the attitude towards, the language in high school years (III). *JLTA (Japan Language Testing Association) Journal*, 10, 88-105. [Category 1]
- The JLTA (Japan Language Testing Association) language testing bibliography. (2008). *JLTA (Japan Language Testing Association) Journal*, 11, 219-255. [Category 1]
- Utsunomiya, Y., Maruyama, M., & Ogasawara, S. (2016). Estimating TOEIC scores using G-TELP scores: A Bayesian Model in a Japanese national university. *JLTA (Japan Language Testing Association) Journal*, 19, 27–45. [Category 1, 20]
- Watanabe, Y. (1999). Juken gijutsu no san shurui to kyoshi, test sakusei sha, kenkyu sha no

yakuwari [Three kinds of techniques for entrance examination and the roles of teachers, test-makers and researchers]. *JAFLE (Japan Association of Foreign Language Education) Bulletin*, 2, 22-28. [Category 1]

Watanabe, Y. (2004). Eigo kyoiku kaizen no tame no test no joshiki [Common sense of tests for improving English education]. *Eigo Tenbo*, 110, 20-23. [Category 1]

Watanabe, Y. (2005). Gengo test to hyoka [Language testing and assessment]. In I. Koike (Ed.), *Daini gengo shutoku no genzai [The current state of second language acquisition research]* (pp. 275-294). Tokyo: Taishukan. [Category 1]

Yanase, Y. (2008). Gengo communication ryoku no sanjigen teki rikai [A three dimensional understanding of communicative language ability]. *JLTA (Japan Language Testing Association) Journal*, 11, 77-95. [Category 1]

Yoshimura, O., Shojima, K., Sugino, N., Nozawa T., Shimizu, Y., Saitoh, E., et al. (2005). Daigaku nyushi center shiken kishutsu mondai o riyoshita kyotsu jukensha keikaku niyuru Eigo gakuryokyu no keinen henka chosa. [An research on the trend of academic achievement of English by IRT equating on National Center Test using the common-subject design]. *Japanese Journal for Research on Testing*, 1, 51-58. [Category 1, 11]

Category 2: Testing Listening

- Fujita, R. (2012). Learners' listening strategy use when listening to authentic and inauthentic materials: Based on listening strategy questionnaire. *JLTA (Japan Language Testing Association) Journal*, 15, 133-151. [Category 2]
- Hijikata, Y. (2005). The effects of Japanese EFL learners' rhythm perception on the chunking process during listening. *JLTA (Japan Language Testing Association) Journal*, 8, 68-85. [Category 2]
- Hirai, A. (1999). The relationship between listening and reading rates of Japanese EFL learners. *Modern Language Journal*, 83, 367-384. [Category 2]
- Hirai, A. (2002). The relationship between listening and reading rates: A comparative study of Japanese and international EFL learners. *Special Issue of Studies in Languages and Cultures 1*. Tsukuba: University of Tsukuba. [Category 2, 3]
- Hirai, A. (2004). Variables affecting the listening comprehension of Japanese learners of English. *ELT: The case of Japan*. In V. Makarova & T. Rodgers (Eds.), *LINCOM Studies in Second Language Teaching 02* (pp. 90-113). LINCOM EUROPA. [Category 2]
- Hirai, A. (2005). Factors predicting EFL learners' listening and reading fluency. *JACET (Japan Association of English Teachers) Bulletin*, 41, 19-36. [Category 2, 3]
- Iimura, H. (2007). Measuring listening recall protocols: A comparison of the three types of idea units. *JLTA (Japan Language Testing Association) Journal*, 10, 24-35. [Category 2]
- Iimura, H. (2010). The effects of repeating questions on auditory versions of multiple-choice listening tests: A preliminary study. *JLTA (Japan Language Testing Association) Journal*, 13, 41-54. [Category 2]
- Iimura, H. (2014). Attractiveness of distractors in multiple-choice listening tests. *JLTA (Japan Language Testing Association) Journal*, 16, 17-40. [Category 2]
- Iimura, H. (2016). Are the least frequently chosen distractors the least attractive?: The case of a four-option picture-description listening test. *JLTA (Japan Language Testing Association) Journal*, 19, 46-65. [Category 2]
- In'nami, Y. (2001). The effects of text and task on the listening scores of Japanese university students. *JLTA (Japan Language Testing Association) Journal*, 4, 56-75. [Category 2]
- In'nami, Y. (2002). Psychometric dimensionality of L2 listening construct. *JLTA (Japan Language Testing Association) Journal*, 5, 72-90. [Category 2]
- In'nami, Y. (2005). The effects of task types on listening test performance: A retrospective study. *JLTA (Japan Language Testing Association) Journal*, 8, 1-20. [Category 2]
- In'nami, Y., & Koizumi, R. (2014). Effects of verbalization on L2 listening test performance. *Journal of Shinawatra University*, 1(1), 22-41. [Category 2, 20]
- Katagiri, K. (2013). Progress of vocabulary size and listening, reading, and speaking

proficiencies among Japanese high school EFL students over a three-year period. *JLTA (Japan Language Testing Association) Journal*, 16, 127–146. Retrieved from <http://ci.nii.ac.jp/naid/110009767206/en> [Category 2, 4]

Kinoshita, M., Ikari, Y., Ishii, K., Imamura, T., Ohtsu, A., Kawajiri, M., et al. (1996). Oral communication noryoku o jushi shita listening test no kaihatu [Developing a listening comprehension abilities focusing on oral communication]. *Language Laboratory*, 33, 81-96. [Category 2, 6]

Kinoshita, M., Ishii, K., Otsu, A., Kawajiri, M., Shimatani, H., Takanashi, Y., et al (1998). The formation and analysis of a test of listening comprehension abilities for high school students in Japan and South Korea. *Language Laboratory*, 35, 35-51. [Category 2, 6]

Kinoshita, M., Ishii, K., Ohtsu, A., Kawajiri, M., Shimatani, H., Takanashi, Y., et al. (1998). Oral communication noryoku o jushi shita listening test: Mondai syutsudai joken ga test seiseki ni oyobosu eikyo [Listening comprehension test focusing on oral communication ability: Effects of testing conditions on the test results]. *JLTA (Japan Language Testing Association) Journal*, 1, 1-14. [Category 2, 11]

Norizuki, K., Ito, A., & Shimatani, H. (2011). Exploring Item-Examinee Response Characteristics in Search of Diagnostic Functions of TOEIC® Tests for University Students in Japan. *JLTA (Japan Language Testing Association) Journal*, 14, 1-20. [Category 2, 3]

Ohtomo, K. (1990). Issues in listening comprehension tests. Oyo gengogaku no kenkyu [Studies in applied linguistics]. In JAAL (Japan Association of Applied Linguistics) in JACET (pp. 182-187). Tokyo: Liber Press. [Category 2]

Ohtomo, K., & Strichers, G. (1978). *TOEFL listening comprehension*. Tokyo: Goken. [Category 2]

Sato, R. (2005). Effectiveness of repetitions and pauses on listening comprehension. *JLTA (Japan Language Testing Association) Journal*, 8, 21-30. [Category 2]

Shimada, K. (2000). *Dimensionality in a listening comprehension test*. Unpublished master's thesis Macquarie University, Sydney, Australia. [Category 2]

Shimada, K. (2001). Skills separability in a listening comprehension test. *ARELE (Annual Review of English Language Education in Japan)*, 12, 61-70. [Category 2]

Shimada, K. (2001). A comparison between higher-order questions and lower-order questions: item difficulty and item discriminability in a listening comprehension test. *CELES (Chubu English Language Education Society) Bulletin*, 30, 247-256. [Category 2]

Shimada, M. (2007). Nihongo chokai test ni okeru sentakushi teiji keishiki no eikyo—Noryoku betsu bunseki— [The effects of styles of presenting multiple-choice options in a Japanese listening test—Analysis by proficiency levels--]. *JLTA (Japan Language Testing*

Association) Journal, 10, 12-23. [Category 2]

- Shimatani, H., Kinoshita, M., Laskowski, T., Takanashi, Y., Ohtsu, A., Ishii, K., et al. (1999). Nihon to kankoku no kokosei ni taisuru eigo listening test ni mirareta test bias: kotenteki test riron to komoku oto riron ni motozuku data bunseki [Test biases of a listening comprehension test administered to Japanese and Korean high school students: An analysis based on the classical test theory and item response theory]. *JLTA (Japan Language Testing Association) Journal, 2, 35-54. [Category 2]*
- Shimatani, H., & Laskowski, T. (2001). Washback effect of the introduction of a listening test into a national university entrance examination. *Proceedings of the Foreign Language Education and Technology Conference IV, 738-744. [Category 2, 14]*
- Shimatani, H., & Laskowski, T. (2001). Influences of the introduction of a listening test into the entrance examination at Kumamoto University: Compared with the situation in South Korea. *KASELE (Kyushu Academic Society of English Language Education) Bulletin, 29, 1-9. [Category 2, 14]*
- Shizuka, T. (2010). The influence of adopting different strategies when answering TOEIC listening items. *JLTA (Japan Language Testing Association) Journal, 13, 71-90. [Category 2]*
- Takaki, N., Shimatani, H., Yoshida, M., Rosen, A., Toda, T., Matsuse, K., et al (1999). Eigo kikitori shiken dounyu to kumamoto daigaku nyugakusya no eigo noryoku [Listening testing the entrance examination and English language ability of Kumamoto university students]. *Memoirs of the Faculty of Education Kumamoto University (the Humanities), 48, 39-47. [Category 2, 14]*
- Tani, S. (2004). Nihongo noryoku shiken “chokai” mondai ha genjitsu no nihongo shiyo bamen o dono teido hanei shite iruka: Text no shitsuteki bunseki o toshite [To what degree listening sections in Japanese Language Proficiency Test reflect real Japanese language use: Through qualitative text analysis]. *JLTA (Japan Language Testing Association) Journal, 6, 40-55. [Category 2]*
- Tsuchihira, T. (2014). Examining the dimensionality of an English listening test using TESTFACT and PCAR. *JLTA (Japan Language Testing Association) Journal, 16, 61-82. [Category 2]*
- Ushiro, Y. (2003). An analysis of EFL learners' use of pauses, Japanese translation, and English repetition in reading and listening. *ARELE (Annual Review of English Language Education in Japan), 14, 131-140. [Category 2, 3]*
- Yanagawa, K. (2003). Relative difficulty of three multiple choice listening comprehension item formats for Japanese high school learners of EFL. *The Bulletin of the Graduate School of Education of Waseda University, 10, 1-14. [Category 2, 11]*

- Yanagawa, K. (2004). Tashisentakushiki listening test no mondaibun to sentakushi no teiji jiki ga test performance ni ataeru eikyo [The Effects of provision of item stems and/or answer options on multiple-choice question listening comprehension test performance]. *STEP (Society for Testing English Proficiency) Bulletin*, 16, 30-42. [Category 2, 11]
- Yanagawa, K. (2007). Effects of stems and/or option preview on item difficulty and discrimination. *Japan Journal for Research on Testing*, 3, 3-13. [Category 2, 11]
- Yoshida, H. (2012). Can TOEIC Bridge test scores predict TOEIC test scores?: An investigation of the relationship between TOEIC Bridge and TOEIC tests. *JLTA (Japan Language Testing Association) Journal*, 15, 101-114. [Category 2, 3]

Category 3: Testing Reading

- Beglar, D. (2013). A broad conceptualization of assessing extensive reading. *JLTA (Japan Language Testing Association) Journal*, 16, 1–14. Retrieved from <http://ci.nii.ac.jp/naid/110009767200> [Category 3]
- Hamada, A. (2015). Linguistic variables determining the difficulty of Eiken reading passages. *JLTA (Japan Language Testing Association) Journal*, 18, 57–77. [Category 3, 11]
- Hijikata, Y. (2006). Does processing chunk-by-chunk require an inadequate amount of memory? *JLTA (Japan Language Testing Association) Journal*, 9, 129-140. [Category 3]
- Hirai, A. (2002). The relationship between listening and reading rates: A comparative study of Japanese and international EFL learners. *Special Issue of Studies in Languages and Cultures 1*. Tsukuba: University of Tsukuba. [Category 2, 3]
- Hirai, A. (2005). Factors predicting EFL learners' listening and reading fluency. *JACET (Japan Association of English Teachers) Bulletin*, 41, 19-36. [Category 2, 3]
- Isaji, T. (2005). Test-wisness no kanten kara mita eigo reading jukutatsudo test no datosei no kensho [Examining validity of English reading proficiency test from the viewpoint of test-wisness]. *JLTA (Japan Language Testing Association) Journal*, 7, 125-144. [Category 3, 13]
- Isaji, T. (2006). Nihonjin kokosei eigo gakushusha no eigo reading jukutatsudo test ni okeru test taking strategy nitsuite—Joisha to kaisha no hikaku [Test taking strategy of Japanese EFL learners in high school in taking English reading proficiency tests—Comparison between higher and lower proficiency level learners]. *JLTA (Japan Language Testing Association) Journal*, 9, 36-54. [Category 3]
- Kinoshita, M., & Ohtsu, A. (2002). Nihon to Taiwan ni okeru daigaku eigo nyushi mondai no readability ni kansuru hikaku bunseki—2001 nendo daigaku nyushi center shiken to hachijuku gakunendo daigaku togo shosei gakkakoushi no eigo dokkai mondai o chushin ni [Comparing the readability of the reading sections of the unified exam English test administered in Japan and Taiwan]. *JLTA (Japan Language Testing Association) Journal*, 5, 12-33. [Category 3]
- Koyama, Y., Shimizu, Y., & Kimura, S. (1999). Effect of key word identification tasks on reading Comprehension. *Nagaoka Univeristy of Technology Bulletin of Language Science and Humanities*, 13, 83-93. [Category 3]
- Loucky, J. P. (1994). Teaching and testing English reading skills of Japanese college students. *KASELE (Kyushu Academic Society of English Language Education) Kiyo*, 22, 29-34. [Category 3]
- Loucky, J. P. (1999). Practical CAI language teaching ideas: Reading and writing connections. *Seinan JoGakuin Kiyo*, 46, 1-7. [Category 3, 5]

- Loucky, J. P. (2004). Maximizing vocabulary development using a semantic field keyword approach. *Online in CALLing JAPAN*, 12, 7-20. [Category 3]
- Loucky, J. P. (2005). Combining beneficial strategies from both intensive and extensive reading approaches for more effective and enjoyable language learning. In K. Bradford-Watts, C. Ikeuchi, & M. Swanson (Eds.), *JALT 2004 Conference Proceedings* (pp. 1041-1055). Tokyo: JALT. [Category 3]
- Loucky, J. P. (2005). *Combining beneficial strategies from both intensive and extensive reading approaches for more effective and enjoyable language learning*. In K. Bradford-Watts, C. Ikeuchi, & M. Swanson (Eds.), *JALT 2004 Conference Proceedings: Language Learning for Life* (Conference CD, pp. 1041-1055). Tokyo: JALT. [Category 3]
- Loucky, J. P. (2006). Developing integrated online English courses for enjoyable reading and effective vocabulary learning. *Proceedings of JALT CALL 2005, Globalization: Bringing people together*, 165-169. [Category 3, 6]
- Miyasako, N. (2005). The effects of oral reading practice on English language ability and metacognition of oral reading among Japanese senior high school students. *JLTA (Japan Language Testing Association) Journal*, 7, 92-109. [Category 3]
- Naganuma, N., & Wada, T. (2002). Sokudoku ryoku to text no yomiyasusa niyoru eigo reading noryoku no sokutei [Measurement of English reading ability by reading speed and text readability]. *JLTA (Japan Language Testing Association) Journal*, 5, 34-45. [Category 3]
- Norizuki, K., Ito, A., & Shimatani, H. (2011). Exploring Item-Examinee Response Characteristics in Search of Diagnostic Functions of TOEIC® Tests for University Students in Japan. *JLTA (Japan Language Testing Association) Journal*, 14, 1-20. [Category 2, 3]
- Kimura, Y. (2011). Questions in English Reading Tests for Facilitating EFL Learners' Reading Comprehension. *JLTA (Japan Language Testing Association) Journal*, 14, 21-39. [Category 3]
- Ohtomo, K. (1981). Cloze testing: Analysis and problems. *Language Studies*, 3, 53-81. [Category 3, 6]
- Sakuma, Y., & Ushiro, Y. (2001). The different loads in reading and listening span tests: The relationship between processing and retention. *JLTA (Japan Language Testing Association) Journal*, 4, 21-37. [Category 3]
- Sato, R. (2006). Exploration into correlation of declarative knowledge with procedural knowledge and reading. *JLTA (Japan Language Testing Association) Journal*, 9, 21-35. [Category 3]
- Sawaki, Y. (2005). The generalizability of summarization and free recall ratings in L2 reading assessment. *JLTA (Japan Language Testing Association) Journal*, 7, 21-44. [Category 3]

- Shimada, K. (1993). Dokkai test-ni okeru suirontekihatsumon-no reberu [Levels of inferential questions in reading comprehension tests]. *CELES (Chubu English Language Education Society) Bulletin*, 22, 1-6. [Category 3]
- Shimada, K. (1997). The validity of multiple-choice rational cloze tests. *CELES (Chubu English Language Education Society) Bulletin*, 27, 139-146. [Category 3, 6]
- Shimada, K. (1998). Item analysis of a multiple-choice rational cloze test. *English Review*, 13, 3-16. [Category 3, 6]
- Shimada, K. (2003). Tashisentakugata kurozu test no sentakushi no sentaku [Choice of options in multiple-choice cloze tests]. *CELES (Chubu English Language Education Society) Bulletin*, 32, 283-288. [Category 3, 6]
- Shimada, K. (2004). EAP goi test—Juyogoi-to seigyoteki happyogoi-no kankei— [EAP vocabulary test: The relationship between receptive and controlled productive vocabulary]. *CELES (Chubu English Language Education Society) Bulletin*, 33, 113-120. [Category 3, 6]
- Shimada, K. (2005). Tashisentakushiki-itotekisakujo cloze test no sentakushi-no kaizen [Improvement of options in a multiple-choice rational cloze test]. *English Review*, 19, 39-57. [Category 3, 6]
- Shimada, K. (2007). EAP goichishiki-ni okeru hiroso-to fukasa-no kankei [The relationship between breadth and depth in EAP vocabulary knowledge]. In A. Mochizuki, A. Kubota, H. Iwasaki, & Y. Ushiro (Eds.), *Atarashii eigokyoiku-no tameni* [New perspectives on English language education] (pp. 75-86). Tokyo: Seibido. [Category 3, 6]
- Shimada, S. (2008). The effects of visual aids on item difficulty and item discrimination in a reading test. *JLTA (Japan Language Testing Association) Journal*, 11, 124-138. [Category 3]
- Shimatani, H., Kinoshita, M., Otsu, A., Laskowski, T., Okuda, Y., & Yamamoto, H. (2003). Comparing the reading sections of the national unified English tests of Japan, South Korea, and Taiwan. *The JACET Kyushu-Okinawa Chapter Annual Review of English Learning and Teaching*, 8, 1- 17. [Category 3, 14]
- Shimatani, H., Kinoshita, M., Yamamoto, H., and Laskowski, T. (2004). High school students' impressions and perceptions of some attributes of each reading subtest in the national unified exam English tests of Japan, South Korea, and Taiwan. *Journal of Asia TEFL*, 1, 47-73. [Category 3, 14]
- Shimizu, Y., Kimura, S., & Koyama, Y. (1997). Kakushu reading test ni okeru jiko-tokuten no yosoku [Prediction of scores by examinees in various reading tests]. *SELT (Studies in English Language Teaching)*, 20, 66-71. [Category 3]
- Shizuka, T. (1999). Reading test komoku no nijigen bunruiho [Two-dimensional classification

- of reading test items]. *JLTA (Japan Language Testing Association) Journal*, 2, 85-104. [Category 3]
- Shizuka, T. (2000). Reading test komoku no nijigen bunrui ni okeru konnanten [Difficulties in two-dimensional classification of reading test items: What made them disagree?]. *JLTA (Japan Language Testing Association) Journal*, 3, 47-66. [Category 3]
- Shizuka, T. (2016). How multiple-choice 1-blank partial dictation items behave. *JLTA (Japan Language Testing Association) Journal*, 19, 66–85. [Category 3]
- Takaki, S. (2010). Investigating EFL learners' mental representations with the verb-clustering test. *JLTA (Japan Language Testing Association) Journal*, 13, 163-175. [Category 3]
- Takemura, M. (2001). The correlation between Japanese EFL learners' L1 and L2 reading performance. *JLTA (Japan Language Testing Association) Journal*, 4, 38-55. [Category 3]
- Terada, Y. (2009). Word order anagram test and reading comprehension in EFL: What type of word order awareness should be focused on in teaching. *JLTA (Japan Language Testing Association) Journal*, 12, 1-15. [Category 3]
- Ushiro, Y. (1999). Differences between L1 & L2 readers in processing paragraph. *JLTA (Japan Language Testing Association) Journal*, 2, 65-84. [Category 3]
- Ushiro, Y. (2003). An analysis of EFL learners' use of pauses, Japanese translation, and English repetition in reading and listening. *ARELE (Annual Review of English Language Education in Japan)*, 14, 131-140. [Category 2, 3]
- Ushiro, Y. (2007). *Schema modification processes of Japanese EFL readers*. Unpublished doctoral dissertation, University of Tsukuba. [Category 3]
- Ushiro, Y., Hamada, A., Hasegawa, Y., Dowse, E., Tanaka, N., Suzuki, K., Hosoda, M., & Mori, Y. (2015). A generalizability theory study on the assessment of task-based reading performance. *JLTA (Japan Language Testing Association) Journal*, 18, 92–114. [Category 3, 13]
- Ushiro, Y., Hijikata, Y., Shimizu, M., In'nami, Y., Kasahara, K., Shimoda, A., et al. (2005). Reliability and validity of translation tests as a measure of reading comprehension. *ARELE (Annual Review of English Language Education in Japan)*, 16, 71-80. [Category 3, 13]
- Ushiro, Y., Hijikata, Y., Shimizu, M., Nakagawa, C., Koga, T., Ohno, M., et al. (2007). Relationships between cue types in recall tests and L2 reading proficiency. *ARELE (Annual Review of English Language Education in Japan)*, 18, 31-40. [Category 3]
- Ushiro, Y., Hirai, N., Hoshino, Y., Nakagawa, C., Kai, A., Ide, M., Oki, T., Suzuki, S., Terada, Y., Takaki, S., & Shimizu, H. (2009). Comparing effects of two types of vocabulary knowledge on six question types in reading tests among Japanese EFL learners. *JLTA (Japan Language Testing Association) Journal*, 12, 104-115. [Category 3, 6]

- Ushiro, Y., Koga, T., Shimizu, M., In'nami, Y., Hijikata, Y., & Hamano, Y. (2006). Construct validation of reading sections in communicative tests. *ARELE (Annual Review of English Language Education in Japan)*, *17*, 201-210. [Category 3, 13]
- Ushiro, Y., Koizumi, R., In'nami, Y., Shimizu, M., Hijikata, Y., Kikuchi, K., et al. (2004). The effects of background knowledge on L2 reading comprehension: Special focus on background knowledge test methods. *JACET (Japan Association of College English Teachers) Bulletin*, *38*, 63-77. [Category 3]
- Ushiro, Y., Morimoto, Y., Hijikata, Y., Nakagawa, C., Watanabe, F., Kai, A., et al. (2007). What makes distractors plausible in multiple-choice reading tests? *JLTA (Japan Language Testing Association) Journal*, *10*, 56-67. [Category 3]
- Ushiro, Y., Nakagawa, C., Morimoto, Y., Koga, T., Konno, K., & Nabeta, K. (2007). Examining the characteristics of scoring methods for written recall tests: Focusing on propositional and idea unit analyses. *JACET (Japan Association of College English Teachers) JOURNAL*, *44*, 85-99. [Category 3, 12]
- Ushiro, Y., & Sakuma, Y. (2000). Modifying reading and listening span tests for group testing. *JLTA (Japan Language Testing Association) Journal*, *3*, 67-82. [Category 3]
- Ushiro, Y., Shimizu, H., Hoshino, Y., Kai, A., Shimada, S., Gomi, N., et al. (2008). Comparison of L2 readers' performance in translation, recall, and summary tasks. *JLTA (Japan Language Testing Association) Journal*, *11*, 109-123. [Category 3]
- Ushiro, Y., Shimizu, M., Hijikata, Y., Kasahara, K., Nakagawa, C., Morimoto, Y., et al. (2006). The effects of visual materials on reading comprehension tests. *JACET (Japan Association of College English Teachers) Bulletin*, *42*, 97-112. [Category 3]
- Ushiro, Y., Shimizu, M., Koizumi, R., In'nami, Y., & Hijikata, Y. (2004). Which affects EFL reading comprehension more, background knowledge or reading proficiency? *ARELE (Annual Review of English Language Education in Japan)*, *15*, 149-158. [Category 3]
- Ushiro, Y., Takaki, S., Kobayashi, M., Hasegawa, Y., Nahatame, S., Hamada, A., & Kimura, Y. (2013). Measures of macroproposition construction in EFL reading: Summary writing task vs. the meaning identification technique. *JLTA (Japan Language Testing Association) Journal*, *16*, 185–204. Retrieved from <http://ci.nii.ac.jp/naid/110009767209/en> [Category 3]
- Wada, T. (2004). A study into the reading construct: Validating item types of reading tests. *JLTA (Japan Language Testing Association) Journal*, *6*, 56-75. [Category 3, 13]
- Yoshida, H. (2012). Can TOEIC Bridge test scores predict TOEIC test scores?: An investigation of the relationship between TOEIC Bridge and TOEIC tests. *JLTA (Japan Language Testing Association) Journal*, *15*, 101-114. [Category 2, 3]
- Yoshizawa, K. (2000). Development and field-testing of a reading strategy questionnaire. *JLTA*

(Japan Language Testing Association) Journal, 3, 100-118. [Category 3]

Category 4: Testing Speaking

- Fukazawa, M. (2011). Validity of Self-Assessment of Speech Performance: A Case of Japanese High School Students. *JLTA (Japan Language Testing Association) Journal*, 14, 61-79. [Category 4, 16]
- Hirai, A., Ito, N., & O'Ki, T. (2011). Applicability of Peer Assessment for Classroom Oral Performance. *JLTA (Japan Language Testing Association) Journal*, 14, 41-59. [Category 4, 16]
- Hirai, A., & Koizumi, R. (2013). Validation of empirically derived rating scales for a Story Retelling Speaking Test. *Language Assessment Quarterly*, 10, 398-422. doi:10.1080/15434303.2013.824973 [Category 4, 12]
- Horikawa, Y., & Tokuma, N. (2008). Nihongo group oral test no gakushu heno hakyu koka ni kansuru yobi teki kenkyu: Kyoshitsu nai test no baai [Preliminary study of washback on learning of group oral tests for Japanese language in a classroom environment]. *JLTA (Japan Language Testing Association) Journal*, 11, 139-155. [Category 4, 14]
- Katagiri, K. (1998). *What Is Speaking Evaluation? -Main Factors to Determine Speaking Evaluation-* Unpublished M.A. thesis submitted to Tokyo Gakugei University. [Category 4]
- Katagiri, K. (1998). Specifying main factors that determine speaking evaluation by Japanese English teachers through transcription analysis. *LEO*, 27, 31-56. [Category 4]
- Katagiri, K. (1999). Evaluation of L2 speaking: JTEs versus AETs. *ARELE (Annual Review of English Language Education in Japan)*, 10, 93-102. [Category 4]
- Katagiri, K. (2012). Longitudinal measurements and developmental patterns of receptive vocabulary size: A study of Japanese university EFL students. *JLTA (Japan Language Testing Association) Journal*, 15, 43-57. [Category 6]
- Katagiri, K. (2013). Progress of vocabulary size and listening, reading, and speaking proficiencies among Japanese high school EFL students over a three-year period. *JLTA (Japan Language Testing Association) Journal*, 16, 127-146. Retrieved from <http://ci.nii.ac.jp/naid/110009767206/en> [Category 2, 4]
- Koizumi, R. (2002). The effects of motivation, language anxiety, and test anxiety on English proficiency of Japanese junior high school students. *JLTA (Japan Language Testing Association) Journal*, 5, 91-110. [Category 4, 11]
- Koizumi, R. (2005). Attainment levels of speaking ability of Japanese junior high school students. *JLTA (Japan Language Testing Association) Journal*, 8, 31-50. [Category 4]
- Koizumi, R. (2005). Predicting speaking ability from vocabulary knowledge. *JLTA (Japan Language Testing Association) Journal*, 7, 1-20. [Category 4, 6]
- Koizumi, R. (2005). Speaking performance measures of fluency, accuracy, syntactic complexity,

- and lexical complexity. *JABAET (Japan-Britain Association for English Teaching) Journal*, 9, 5-33. [Category 4, 13]
- Koizumi, R. (2015). Factor structure and four-skill profiles of the TOEIC® tests among Japanese university learners of English. *ARELE (Annual Review of English Language Education in Japan)*, 26, 109–124. [Category 4, 5]
- Koizumi, R., & Fujimori, C. (2010). An exploration of the measures to detect changes in speaking performance: A case study based on picture description. *JACET (Japan Association of College English Teachers) Journal*, 50, 81-91. [Category 4]
- Koizumi, R., & Hirai, A. (2010). Exploring the quality of the Story Retelling Speaking Test: Roles of story length, comprehension questions, keywords, and opinions. *ARELE (Annual Review of English Language Education in Japan)*, 21, 211-220. [Category 4]
- Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, 4, 900-913. Retrieved from <http://ojs.academypublisher.com/index.php/jltr/article/view/jltr0405900913> [Category 4, 6]
- Koizumi, R., & In'nami, Y. (2014). Modeling complexity, accuracy, and fluency of Japanese learners of English: A structural equation modeling approach. *JALT Journal*, 36, 25–46. Retrieved from <http://jalt-publications.org/jj/articles/3728-modeling-complexity-accuracy-and-fluency-japanese-learners-english-structural-equat> [Category 4]
- Koizumi, R., In'nami, Y., & Fukazawa, M. (2016). Development of a paired oral test for Japanese university students. In C. Saida, Y. Hoshino, & J. Dunlea (Eds.), *British Council New Directions in Language Assessment: JASELE Journal Special Edition* (pp. 103–121). Tokyo: British Council Japan. [Category 4, 13]
- Koizumi, R., & Katagiri, K. (2007). Changes in speaking performance of Japanese high school students: The case of an English course at a SELHi. *ARELE (Annual Review of English Language Education in Japan)*, 18, 81-90. [Category 4, 11]
- Koizumi, R., & Katagiri, K. (2009). Changes in speaking performance of Japanese high school students: Longitudinal and cross-sectional studies at a SELHi. *ARELE (Annual Review of English Language Education in Japan)*, 20, 51-60. [Category 4]
- Koizumi, R., & Kurizaki, I. (2002). Developing rating scales for monologues from Japanese junior high school students. *ARELE (Annual Review of English Language Education in Japan)*, 13, 51-60. [Category 4, 12]
- Koizumi, R., & Kurizaki, I. (2002). Nihonjin chugakusei no monologue ni okeru speaking no tokucho [Speaking characteristics of monologues given by Japanese junior high school

- students]. *Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 16, 17-28. [Category 4]
- Koizumi, R., & Yamanouchi, I. (2003). Nihonjin chugakusei no speaking noryoku no hattatsu [Development in speaking ability among Japanese junior high school students: Using self-introduction task]. *Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 17, 33-44. [Category 4, 11]
- Lee, W.-K. (2012). The speaking test of Korea's NEATs 2 & 3: Its usability for university admission qualifications. *JLTA (Japan Language Testing Association) Journal*, 15, 25-42. [Category 4]
- Nekoda, H. (2006). Developing a standard of language proficiency required for English teachers: Using generalizability theory and item response theory. *ARELE (Annual Review of English Language Education in Japan)*, 17, 191-200. [Category 4, 12]
- Nekoda, H., Nekoda, K., & Miura, S. (2007). Developing an oral proficiency scale through the MFRM analysis of teachers' subjective judgements. *ARELE (Annual Review of English Language Education in Japan)*, 18, 221-230. [Category 4, 12]
- Ohtomo, K. (2005). Hanasukoto no shido ni okeru hyoka [Assessment in the teaching of speaking]. PCOLA degital han eigoka kyoiku jugyo jissen shiryō: Riron-hen [PCOLA data for the teaching of English: Theory] (pp. 97-114). Tokyo: Nichibun. [Category 4]
- Sato, T. (2013). The influential features on linguistic laypersons' evaluative judgments of second language oral communication ability. *JLTA (Japan Language Testing Association) Journal*, 16, 107-126. Retrieved from <http://ci.nii.ac.jp/naid/110009767205/en> [Category 4, 12]
- Shimatani, H., Minoda, T., Maemura, Y., & Sawamura, Y. (2004). Eigo communication noryoku no keizokuteki hyoka: Fuzoku cyugakkou ni okeru jissen no Kensyō [Evaluating the continuous assessment of English communication ability at the attached junior high school of the faculty of education, Kumamoto university]. *Memoirs of the Faculty of Education Kumamoto University (the Humanities)*, 54, 149-156. [Category 4, 9]
- Uenishi, K. (2003). A research of students' oral presentation in world studies: Focusing on assessment. *Hiroshima Kenritsu Aifuchu Koto Gakko Kenkyu Kiyō*, 19, 107-117. [Category 4, 9]
- Uenishi, K. (2004). A study of factors contributing to English speaking proficiency: Comparing Japanese high school and university students' speaking factors. *ARELE (Annual Review of English Language Education in Japan)*, 15, 119-128. [Category 4, 11]
- Uenishi, K. (2004). A study of students' oral presentation in international course: With a focus on teachers' assessment. *Kanseigakuin University Kyoshoku Kyōyo Kenkyū*, 9, 81-86. [Category 4, 9]

- Uenishi, K (2006). Nihonjin eigo gakushusha no speaking ryoku ni kansuru jisshoteki kenkyu: Goi hatsuon tonon kanrensei ni shoten o atete [An empirical research on speaking ability of Japanese learners of English: With a focus on the relation to vocabulary and pronunciation]. *Hiroshima Daigaku Gakkokyoiku Gakubu Eigo Kenkyukai Eigo to Eigo kyoiku, Tokubetsugo, Ozasa Toshiaki Taishoku Kinen Ronbunshu* [English and English education: Special number, Society for the English research at Faculty of Education of Hiroshima University, Festschrift for Dr. Toshiaki Ozasa's retirement], 135-144. [Category 4, 11]
- Uenishi, K (2007). *Factors in determining English speaking ability: With a focus on Japanese EFL learners*. Hiroshima: Keisuisha. [Category 4, 11]
- Yokouchi, Y. (2015). Do input text length and presentation mode affect speaking performance in retelling tasks? *JLTA (Japan Language Testing Association) Journal*, 18, 115–133. [Category 4, 11]
- Zhou, Y. J. (2007). Comparison of examinees' reactions toward computer-based test of speaking and face-to-face interview. *The Journal of TUFSS (Tokyo University of Foreign Studies) Language Education Studies*, 2, 73-95. [Category 4, 13]
- Zhou, Y. (2008). A comparison of speech samples of monologic tasks in speaking tests between computer-delivered and face-to-face modes. *JLTA (Japan Language Testing Association) Journal*, 11, 189-208. [Category 4, 11]
- Zhou, Y. (2012). Computerized assessment of second language speaking: A review of tests. *JLTA (Japan Language Testing Association) Journal*, 15, 81-100. [Category 4]

Category 5: Testing Writing

- Koizumi, R. (2015). Factor structure and four-skill profiles of the TOEIC® tests among Japanese university learners of English. *ARELE (Annual Review of English Language Education in Japan)*, 26, 109–124. [Category 4, 5]
- Koizumi, R., In'nami, Y., Asano, K., & Agawa, T. (2016). Validity evidence of Criterion® for assessing L2 writing proficiency in a Japanese university context. *Language Testing in Asia*, 6(5), 1–26. doi:10.1186/s40468-016-0027-7 Retrieved from <http://www.languagetestingasia.com/content/6/1/5> [Category 5, 13]
- Kudo, H. (2000). Jiyusakubun heno hashiwatashi to shiteno Dicto-Comp [Usefulness of Dicto-Comp: Bridge to free compositions]. Unpublished graduation thesis, Tokyo University of Foreign Studies. [Category 5]
- Loucky, J. P. (1999). Practical CAI language teaching ideas: Reading and writing connections. *Seinan JoGakuin Kiyo*, 46, 1-7. [Category 3, 5]
- Matsuno, S. (2005). Feasible writing assessment at Japanese universities. *JLTA (Japan Language Testing Association) Journal*, 7, 45-65. [Category 5]
- Min, H. (2014). Introducing the writing section of the NEAT in Korea. *JLTA (Japan Language Testing Association) Journal*, 16, 1–16. [Category 5, 12]
- Nakamura, Y., & Tobiwatari, H. (2004). An empirical, statistical and comparative study of students' writing abilities in Japanese and English. *JLTA (Japan Language Testing Association) Journal*, 6, 1-12. [Category 5]
- Sugita, Y. (2000). Eisakubun shido ni okeru ayamari no teiseiho ni kansuru kenkyu [Effects of error feedback of Japanese EFL learners]. *JLTA (Japan Language Testing Association) Journal*, 3, 127-135. [Category 5]
- Sugita, Y. (2009). Developing and improving rating scales for a task-based writing performance test. *JLTA (Japan Language Testing Association) Journal*, 12, 85-103. [Category 5]
- Sugita, Y. (2012). Effects of rater training on raters' severity, consistency and biased interactions in a task-based writing assessment. *JLTA (Japan Language Testing Association) Journal*, 15, 61-80. [Category 5]
- Sugita, Y. (2013). Comparability of accuracy and communicability tasks: Are they all equally difficult? *JLTA (Japan Language Testing Association) Journal*, 16, 65–86. Retrieved from <http://ci.nii.ac.jp/naid/110009767203/en> [Category 5, 12]
- Tsuura, M. (2001). *The influence of topics on writing performance*. Unpublished Graduation Thesis, Tokyo University of Foreign Studies. [Category 5]
- Uenishi, K (2006). A study of factors contributing to English writing ability: With a focus on two types of writing task. *ARELE (Annual Review of English Language Education in Japan)*, 17, 71-80. [Category 5, 11]

Category 6: Testing Language Elements

- Amma, K. (2004). Development of multiple-choice grammaticality judgement tests. *JLTA (Japan Language Testing Association) Journal*, 6, 13-29. [Category 6]
- Hasegawa, Y. (2010). Context effects on EFL vocabulary learning: Focusing on word imageability and test formats. *JLTA (Japan Language Testing Association) Journal*, 13, 145-161. [Category 6]
- Hasegawa, Y. (2013). Who will benefit from glossed context in intentional vocabulary learning? Comparing high school and university students. *JLTA (Japan Language Testing Association) Journal*, 16, 43–64. Retrieved from <http://ci.nii.ac.jp/naid/110009767202/en> [Category 6]
- Hirose, K. (1999). Goi test komoku no bunseki [Analyzing test items of vocabulary]. *JLTA (Japan Language Testing Association) Journal*, 2, 55-64. [Category 6]
- Hoshino, Y. (2009). Reactivity of thinking aloud in EFL learners' performance in vocabulary tests. *JLTA (Japan Language Testing Association) Journal*, 12, 16-25. [Category 6]
- Hoshino, Y. (2015). Development of a test to measure knowledge regarding multiple meanings of basic words. *JLTA (Japan Language Testing Association) Journal*, 18, 78–91. [Category 6]
- Ito, A. (2000). Eigo kankeishisetu no sanshutsu ni okeru Hamilton kasetsu no datosei [Validity of Hamilton hypothesis in English relative clause production]. *JLTA (Japan Language Testing Association) Journal*, 3, 136-149. [Category 6]
- Kanatani, K., & English Diagnostic Test (EDiT) Development Group. (2006). Eigo shindan test kaihatsu heno michi—ELPA Eigo Shindan Project no kiseki [Road to the development of the EDiT: Trajectory of the ELPA English Diagnostic Project]. Tokyo: Association for English Language Proficiency Assessment (ELPA). [Category 6, 13]
- Kasahara, K. (2006). Producing a revised version of the Mochizuki Vocabulary Size Test. *JLTA (Japan Language Testing Association) Journal*, 9, 55-72. [Category 6]
- Katagiri, K. (2000). Quick and rough estimates of general English ability, using Mochizuki's Vocabulary Size Test for Japanese EFL learners. *JLTA (Japan Language Testing Association) Journal*, 3, 83-99. [Category 6]
- Katagiri, K. (2001). *Developing the Ten-Minute Vocabulary Tests for Quick and Approximate Estimates of General English Ability of Japanese EFL Learners*. Unpublished Ph.D. thesis submitted to Tokyo Gakugei University. [Category 6]
- Katagiri, K. (2001). The ten-minute vocabulary tests for quick and rough estimates of general English ability of Japanese EFL learners I. *JLTA (Japan Language Testing Association) Journal*, 4, 1-20. [Category 6]
- Katagiri, K. (2001). The ten-minute vocabulary tests for quick and rough estimates of general

- English ability of Japanese EFL learners II. *The Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 15, 39-58. [Category 6]
- Katagiri, K. (2002). The ten-minute vocabulary tests for quick and approximate estimates of general English ability of Japanese EFL learners III. Chiba: *Reitaku University Journal*, 74, 159-182. [Category 6]
- Katagiri, K. (2002). The ten-minute vocabulary tests for quick and approximate estimates of general English ability of Japanese EFL learners IV. *JLTA (Japan Language Testing Association) Journal*, 5, 111-128. [Category 6]
- Katagiri, K. (2002). The ten-minute vocabulary tests for quick and approximate estimates of general English ability of Japanese EFL learners V. *The Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 16, 1-16. [Category 6]
- Katagiri, K. (2004). The rationale for inferring learners' general English ability through measuring their vocabulary knowledge I. *JLTA (Japan Language Testing Association) Journal*, 6, 76-91. [Category 6]
- Katagiri, K. (2005). A practical report: Conducting vocabulary size tests at university every year. *JACET (The Japan Association of College English Teachers) Summer Seminar Proceedings*, 3, 70-76. [Category 6]
- Katagiri, K. (2007). Examining the rationale of vocabulary size tests based on word-frequency levels. *JLTA (Japan Language Testing Association) Journal*, 10, 36-55. [Category 6]
- Kimura, S. (1996). Eitango no rinsetsu kankei ni chumokushita seijo-mondai saiten-ho to eigo noryoku shindan chiryo yobo program no kanousei [Computer assisted scoring of English word sequencing tests with a focus on word relations in the sentence: Its validity and possibility]. *Japan Society for Educational Technology*, 12, 387-388. [Category 6, 13]
- Kinoshita, M., Ikari, Y., Ishii, K., Imamura, T., Ohtsu, A., Kawajiri, M., et al. (1996). Oral communication noryoku o jushi shita listening test no kaihatsu [Developing a listening comprehension abilities focusing on oral communication]. *Language Laboratory*, 33, 81-96. [Category 2, 6]
- Kinoshita, M., Ishii, K., Ohtsu, A., Kawajiri, M., Shimatani, H., Takanashi, Y., et al. (1998). Oral communication noryoku o jushi shita listening test: Mondai shutsudai joken ga test seiseki ni oyobosu eikyo [Listening tests focusing on oral communication ability: Effects of task conditions on test scores]. *JLTA (Japan Language Testing Association) Journal*, 1, 1-14. [Category 2, 6]
- Kinoshita, M., Ishii, K., Otsu, A., Kawajiri, M., Shimatani, H., Takanashi, Y., et al (1998). The formation and analysis of a test of listening comprehension abilities for high school students in Japan and South Korea. *Language Laboratory*, 35, 35-51. [Category 2, 6]
- Koizumi, R. (2003). A productive vocabulary knowledge test for novice Japanese learners of

- English: Validity and its scoring methods. *JABAET (Japan-Britain Association for English Teaching) Journal*, 7, 23-52. [Category 6, 13]
- Koizumi, R. (2005). Predicting speaking ability from vocabulary knowledge. *JLTA (Japan Language Testing Association) Journal*, 7, 1-20. [Category 4, 6]
- Koizumi, R. (2011). Test-Taking Processes of the Lexical Organisation Test: Comparing It With the Word Associates Test. *ARELE (Annual Review of English Language Education in Japan)*, 22, 153-168. [Category 6, 13]
- Koizumi, R., & Mochizuki, M. (2011). Development and validation of the PC version of the Mochizuki Vocabulary Size Test. *JACET (Japan Association of College English Teachers) Journal*, 53, 35-55. [Category 6, 13]
- Koizumi, R., Sakai, H., Ido, T., Ota, H., Hayama, M., Sato, M., & Nemoto, A. (2011). Toward Validity Argument for Test Interpretation and Use Based on Scores of a Diagnostic Grammar Test for Japanese Learners of English. *Japanese Journal for Research on Testing*, 7, 99-119. [Category 6, 13]
- Konno, K., Takanami, S., Okuyama, Y., & Hirai, A. (2009). Examining the effects of involvement load on Japanese EFL learners' vocabulary retention. *JLTA (Japan Language Testing Association) Journal*, 12, 46-64. [Category 6]
- Loucky, J. P. (1996). *Developing and testing vocabulary training methods and materials for Japanese college students studying English as a foreign language*. Unpublished Doctoral Dissertation, Pensacola Christian College, Pensacola, FL. (ERIC Document Reproduction Service No. ED406844) [Category 6]
- Loucky, J. P. (1997). Maximizing vocabulary acquisition: Recommendations for improving English vocabulary learning for foreign language learners. *KASELE (Kyushu Academic Society of English Language Education) Kiyō*, 25, 101-111. [Category 6]
- Loucky, J. P. (1997). Summary of "Developing and testing vocabulary training methods and materials for Japanese college students studying English as a foreign language." *Annual Review of English Learning and Teaching (JACET Kyushu-Okinawa Chapter)*, 2, 15-36. [Category 6]
- Loucky, J. P. (1998). Suggestions for improving ESL/EFL vocabulary instruction. *Seinan JoGakuin Kiyō*, 45, 25-36. [Category 6]
- Loucky, J. P. (2002). Assessing the potential of computerized bilingual dictionaries for enhancing English vocabulary learning. In P. N. D. Lewis (Ed.), *The changing face of CALL: A Japanese perspective* (pp.123-137). Lisse: Swets & Zeitlinger. [Category 6]
- Loucky, J. P. (2002). Choosing your words wisely: Selecting which vocabulary to teach. *Oriō Joshi Keizai Tanki Taigaku Ronshū*, 36, 45-58. [Category 6]
- Loucky, J. P. (2002). Comparing translation software and OCR reading pens. In M. Swanson, D.

- McMurray, & K. Lane (Eds.), *Pan-Asian Conference 3 at 27th International Conference of JALT, National Conference Proceedings CD* (pp. 745-755). Kitakyushu, JAPAN. [Category 6]
- Loucky, J. P. (2002). Improving access to target vocabulary using computerized bilingual dictionaries. *ReCALL*, 14, 3-312. [Category 6]
- Loucky, J. P. (2002). When Eastern Oriental meets Western Occidental language system: Crossing the English vocabulary threshold versus breaking the Kanji Barrier. *Seinan JoGakuin Tandai Kiyō*, 48, 19-38. [Category 6]
- Loucky, J. P. (2003). Enhancing students' English reading and vocabulary skills using CALL innovations. *Seinan Women's University Tandai Kiyō*, 49, 1-15. [Category 6]
- Loucky, J. P. (2003). Testing vocabulary levels in Japan. *Oriō Joshi Keizai Tanki Taigaku Ronshū*, 37, 51-87. [Category 6]
- Loucky, J. P. (2003). Testing vocabulary levels in Japan, Part 2. *The Japanese learner*, 29, 15-20. Oxford University. [Category 6]
- Loucky, J. P. (2003). Using CALL innovations to enhance students' English reading and vocabulary skills. In P. N. D. Lewis, C. Imai, & K. Kitao (Eds.), *Local Decisions, Global Effects: The Proceedings of JALT CALL 2002*, 121-128. [Category 6]
- Loucky, J. P. (2003). Using computerized bilingual dictionaries to help maximize English vocabulary learning at Japanese colleges. *CALICO Journal*, 221, 105-129. [Category 6]
- Loucky, J. P. (2004). Clarifying and resurrecting the Semantic Field Keyword Approach for rapid lexical acquisition: Employing semantic organization, bilingual computerized glosses and pushed output generation to enhance L2 vocabulary learning. *Seinan JoGakuin Tandai Kiyō*, 50, 61-87. [Category 6]
- Loucky, J. P. (2004). Improving cognitive, linguistic and computational processing of new vocabulary using an online Semantic Field Keyword Approach. *Japan Association of Language Teachers CALL Special Interest Group, C@lling Japan*, 12, 7-20. [Category 6]
- Loucky, J. P. (2005). Combining the benefits of electronic and online dictionaries with CALL Web sites to produce effective and enjoyable vocabulary and language learning lessons. *Computer Assisted Language Learning*, 18, 389-416. [Category 6]
- Loucky, J. P. (2005). Systematically employing a trinity of user-friendly vocabulary assessment tools. *JLTA (Japan Language Testing Association) Journal*, 7, 66-91. [Category 6]
- Loucky, J. P. (2006). Developing integrated online English courses for enjoyable reading and effective vocabulary learning. *Proceedings of JALT CALL 2005, Glocalization: Bringing people together*, 165-169. [Category 3, 6]
- Loucky, J. P. (2006). Harvesting CALL websites for enjoyable and effective language learning. *Proceedings of JALT CALL 2005, Glocalization: Bringing people together*, 18-22.

[Category 6]

- Loucky, J. P. (2006). Maximizing vocabulary development by systematically using a depth of lexical processing taxonomy, CALL resources, and effective strategies. *CALICO Journal*, 23, 363-399. [Category 6]
- Miyazaki, A., & Katagiri, K. (2008). Imi no wakaru eitango no nakade kaku kotoga dekiru eitango no wariai chosa [What percentage of English words you receptively know can you spell out correctly?]. *JLTA (Japan Language Testing Association) Journal*, 11, 61-76. [Category 6]
- Morimoto, Y. (2006). Comparison between matching- and supplying-format in multiple choice vocabulary tests. *JLTA (Japan Language Testing Association) Journal*, 9, 73-85. [Category 6]
- Morimoto, Y. (2006). Goi test no keishiki ga goi chishiki to dokkai noryoku no sokutei ni oyobosu eikyo [Effects of vocabulary test formats on measuring vocabulary knowledge and reading ability]. *STEP Bulletin*, 18, 77-91. [Category 6, 13]
- Morimoto, Y. (2007). Test-taking processes of vocabulary tests in context from the perspective of think-aloud analysis. *JLTA (Japan Language Testing Association) Journal*, 10, 68-87. [Category 6]
- Ohtomo, K. (1981). Cloze testing: Analysis and problems. *Language Studies*, 3, 53-81. [Category 3, 6]
- Sakano, E., & Sakai, M. (2006). Rasch model ni yoru kanji placement test no kairyo [Refininig of kanji placement test using the Rasch model]. *Japanese Journal for Research on Testing*, 2, 91-100. [Category 6]
- Sakuma, Y. (2008). Working memory nai no eigo shori ni kansuru bogo washa to nihonjin gaikokugo gakushusha tonon hikaku kenkyu [A comparative study of English processing and storage in working memory between English native speakers and Japanese EFL learners]. *JLTA (Japan Language Testing Association) Journal*, 11, 173-188. [Category 6, 11]
- Shimada, K. (1997). The validity of multiple-choice rational cloze tests. *CELES (Chubu English Language Education Society) Bulletin*, 27, 139-146. [Category 3, 6]
- Shimada, K. (1998). Item analysis of a multiple-choice rational cloze test. *English Review*, 13, 3-16. [Category 3, 6]
- Shimada, K. (2003). Tashisentakugata cloze test no sentakushi no sentaku [Choice of options in multiple-choice cloze tests]. *CELES (Chubu English Language Education Society) Bulletin*, 32, 283-288. [Category 3, 6]
- Shimada, K. (2004). EAP goi test—Juyogoi-to seigyotekihappyogoi-no kankei— [EAP vocabulary test: the relationship between receptive and controlled productive vocabulary].

- CELES (Chubu English Language Education Society) Bulletin*, 33, 113-120. [Category 3, 6]
- Shimada, K. (2005). Tashisentakushiki-itotekisakujo cloze test-no sentakushi-no kaizen [Improvement of options in a multiple-choice rational cloze test]. *English Review*, 19, 39-57. [Category 3, 6]
- Shimada, K. (2007). EAP goichishiki-ni okeru hiroso-to fukasa-no kankei [The relationship between breadth and depth in EAP vocabulary knowledge]. In A. Mochizuki, A. Kubota, H. Iwasaki, & Y. Ushiro (Eds.), *Atarashii eigokyoiku-no tameni* [New perspectives on English language education] (pp. 75-86). Tokyo: Seibido. [Category 3, 6]
- Shimatani, H., Kinoshita, M., Laskowski, T., Takanashi, Y., Otsu, A., Ishii, K., et al (1999). Nihon to kankoku no koukousei ni taisuru eigo listening test ni mirareta test bias: kotennteki test riron to komoku oto riron ni motozuku data bunseki [Test biases of a listening comprehension test administered to Japanese and Korean high school students: An analysis based on the classical test theory and item response theory]. *JLTA (Japan Language Testing Association) Journal*, 2, 35-54. [Category 2, 6]
- Shimizu, Y., Kimura S., Sugino, N., Yamakawa, K., Ohba, M., & Nakano, M. (2003). Eibunpo noryoku hyojun test no shinraisei / datosei no kensho to shin eibunpo noryoku test: MEG [Validity and reliability of standardized English grammar tests and a new grammar test: MEG (Measure of English Grammar)]. *Policy Science Association*, 10, 59-68. [Category 6, 13]
- Sugita, Y. (2010). Reliability and validity of task-based writing performance assessment for Japanese learners of English. *JLTA (Japan Language Testing Association) Journal*, 13, 21-40. [Category 5, 13]
- Takanami, S. (2012). Evaluating Japanese university students' spelling test performance: Focusing on spelling test formats. *JLTA (Japan Language Testing Association) Journal*, 15, 185-204. [Category 6]
- Ushiro, Y., Hirai, N., Hoshino, Y., Nakagawa, C., Kai, A., Ide, M., Oki, T., Suzuki, S., Terada, Y., Takaki, S., & Shimizu, H. (2009). Comparing effects of two types of vocabulary knowledge on six question types in reading tests among Japanese EFL learners. *JLTA (Japan Language Testing Association) Journal*, 12, 104-115. [Category 3, 6]
- Yamazaki, A. (1999). Ways of enhancing incidental vocabulary learning. *JLTA (Japan Language Testing Association) Journal*, 2, 23-34. [Category 6]

Category 7: Purposes for Testing

- Kamimoto, T. (2005). Goi levels test no bunseki: Kotenteki test riron to komoku hanno riron kara [Vocabulary Levels Test--A discussion in terms of Classical Test Theory and Item Response Theory]. *KGU Journal of Language & Literature*, 11 & 12, 61-105. [Category 7, 13]
- Loucky, J. P. (2002). *Testing vocabulary levels in Japan*, Part I. *The Japanese learner*, 15-21. Oxford University. [Category 7]
- Newfields, T. (2007). Game theory approaches to grading: An experiment with two incentive point systems. *Toyo University Keizai Ronshu*, 32, 33-43. Available from [www.tnewfields.info/Articles/ game.htm](http://www.tnewfields.info/Articles/game.htm) [Category 7, 9]
- Ohtomo, K., [Editorial supervisor.], Trans. (1980). *Eigo gakushu totatsu mokuhyo no settei* [Valatte, R. M., & Disick, R. S. (1972). Modern language performance objectives and individualization] Tokyo: Tamagawa University Press. [Category 7]
- Valatte, R. M., & Disick, R. S. (1980). *Eigo gakushu totatsu mokuhyo no settei* [Modern language performance objectives and individualization] (K. Ohtomo, [Editorial supervisor.], Trans.). Tokyo: Tamagawa University Press. (Original work published 1972) [Category 7]

Category 8: Specific Purposes Testing

Kimura, S. (2001). Shogai-gakushu-sha to eigo kyoiku: Shikaku-shogai-sha no juken ni kakawaru mondai to shiken jikan encho sochi nitsuite no kenkyu no gaikan [An overview of English testing and impaired individuals with special focus on extended test times for braille form tests]. *Language and Culture*, 14, 85-97. [Category 8, 11]

Kimura, S. (2002). Shogai-gakushusha to eigo kyoiku: Shikaku-shogai-sha no tame no test mondai no bubun henko sochi nitsuite [An overview of English testing and impaired individuals with special focus on modified test items for braille form tests]. *Language and Culture*, 15, 79-91. [Category 8, 11]

Category 9: Alternative Approaches to Assessment

- Loucky, J. P. (2005) Making a user-friendly vocabulary knowledge scale for Japanese students: Designing and testing a dual assessment vocabulary instructor-evaluator. *Kitakyushu: Orio Aishin Tanki Daigaku Ronshu*, 39, 33-86. [Category 9]
- Murata, Y. (2002). Manabi o unagasu portfolio hyoka—jissen hokoku [Empowering the learner through portfolio assessment]. *JLTA (Japan Language Testing Association) Journal*, 5, 1-11. [Category 9]
- Newfields, T. (2007). Game theory approaches to grading: An experiment with two incentive point systems. *Toyo University Keizai Ronshu*, 32, 33-43. Available from [www.tnewfields.info/Articles/ game.htm](http://www.tnewfields.info/Articles/game.htm) [Category 7, 9]
- Shimatani , H., Minoda, T., Maemura, Y., & Sawamura, Y. (2004). Eigo communication noryoku no keizokuteki hyoka: Fuzoku cyugakkou ni okeru jissen no Kensyou [Evaluating the continuous assessment of English communication ability at the attached junior high school of the faculty of education, Kumamoto university]. *Memoirs of the Faculty of Education Kumamoto University (the Humanites)*, 54, 149-156. [Category 4, 9]
- Uenishi, K (2003). A research of students' oral presentation in world studies: Focusing on assessment. *Hiroshima Kenritsu Aifuchu Koto Gakko Kenkyu Kiyo*, 19, 107-117. [Category 4, 9]
- Uenishi, K (2004). A study of students' oral presentation in international course: With a focus on teachers' assessment. *Kanseigakuin University Kyoshoku Kyoyo Kenkyu*, 9, 81-86. [Category 4, 9]
- Yim, S. Y. (2016). Teachers' perspectives on using peer assessment among young EFL learners. *JLTA (Japan Language Testing Association) Journal*, 19, 15–23. [Category 9, 18]

Category 10: Principles of Test Design

- Bachman, L. F. (2015). Justifying the use of language assessments: Linking test performance with consequences. *JLTA (Japan Language Testing Association) Journal*, 18, 3–22. [Category 10, 13]
- Davidson, F. (2012). Releasability of language test specifications. *JLTA (Japan Language Testing Association) Journal*, 15, 1-23. [Category 10]
- Kamimoto, T. (1993). Tailoring the C-Test to fit: Improvement of C-Test through classical item analysis. *Language Laboratory*, 30, 47-61. [Category 10, 13]
- Nakamura, Y. (2003). Sokutei to hyoka ni kansuru kihontekina kangaekata [Basic concepts in measurement and evaluation]. In T. Watanabe (Ed.), *Eigo ga tsukaeru nihonjin no ikusei* [Cultivating Japanese with English abilities] (pp. 39-47). Tokyo: Sanseido, [Category 10, 13]
- Nakamura, Y. (2006). Gengo test ni okeru test riron to data bunseki [Test theories and data analyses in language testing]. In Kokuritsu Kokugo Kenkyujo [The national institute for Japanese language] (Ed.), *Sekai no Gengo test* [Language tests in the world]. (pp. 15-24). Tokyo: Kuroshio. [Category 10, 13]
- Ohtomo, K., & Thrasher, R. (Eds.), Trans. (2002). Jissen gengo test sakusei ho [Bachman, L., & Palmer, A. (1996). *Language testing in practice*]. Tokyo: Taishukan. [Category 1, 10]
- O'Sullivan, B. (2015). Validity, validation and development: Building and operationalizing a comprehensive model. *JLTA (Japan Language Testing Association) Journal*, 18, 25–33. [Category 10, 13]
- Watanabe, Y. (2003). Test sakusei ha aojashin zukuri kara [The first step to make tests is to make blueprints]. *STEP Eigo Joho*, 5-6, 30-33. [Category 10]
- Watanabe, Y. (2003). Test data bunseki to kaishaku no kiso—Heikin to hyojun hensa [Basic analysis of test data and its interpretation—Mean and standard deviation]. *STEP Eigo Joho*, 9-10, 36-39. [Category 10]
- Watanabe, Y. (2003). Test o test suru—Komoku bunseki [Testing tests—Item analysis]. *STEP Eigo Joho*, 11-12, 34-37. [Category 10]

Category 11: Factors Affecting Test Performance

- Hamada, A. (2015). Linguistic variables determining the difficulty of Eiken reading passages. *JLTA (Japan Language Testing Association) Journal*, 18, 57–77. [Category 3, 11]
- Ito, T., Kawaguchi, K., & Ohta, R. (2008). Gaikokugo noryoku no jiko hyotei ni okeru gengo task keiken no eikyo [The influence of experiential factors on self-assessment of foreign language proficiency]. *JLTA (Japan Language Testing Association) Journal*, 11, 156-172. [Category 11]
- Kamimoto, T. (2003). A comparison of the Vocabulary Levels Test in L1 and L2. *KGU (Kumamoto Gakuen University) Journal of Language & Literature*, 9 & 10, 21-49. [Category 11, 13]
- Kimura, S. (2001). Shogai-gakushu-sha to eigo kyoiku: Shikaku-shogai-sha no juken ni kakawaru mondai to shiken jikan encho sochi nitsuite no kenkyu no gaikan [An overview of English testing and impaired individuals with special focus on extended test times for braille form tests]. *Language and Culture*, 14, 85-97. [Category 8, 11]
- Kimura, S. (2002). Shogai-gakushusha to eigo kyoiku: Shikaku-shogai-sha no tame no test mondai no bubun henko sochi nitsuite [An overview of English testing and impaired individuals with special focus on modified test items for braille form tests]. *Language and Culture*, 15, 79-91. [Category 8, 11]
- Kinoshita, M., Ishii, K., Ohtsu, A., Kawajiri, M., Shimatani, H., Takanashi, Y., et al. (1998). Oral communication noryoku o jushi shita listening test: Mondai syutsudai joken ga test seiseki ni oyobosu eikyo [Listening comprehension test focusing on oral communication ability: Effects of testing conditions on the test results]. *JLTA (Japan Language Testing Association) Journal*, 1, 1-14. [Category 2, 11]
- Koizumi, R. (2002). The effects of motivation, language anxiety, and test anxiety on English proficiency of Japanese junior high school students. *JLTA (Japan Language Testing Association) Journal*, 5, 91-110. [Category 4, 11]
- Koizumi, R., & Katagiri, K. (2007). Changes in speaking performance of Japanese high school students: The case of an English course at a SELHi. *ARELE (Annual Review of English Language Education in Japan)*, 18, 81-90. [Category 4, 11]
- Koizumi, R., & Yamanouchi, I. (2003). Nihonjin chugakusei no speaking noryoku no hattatsu [Development in speaking ability among Japanese junior high school students: Using self-introduction task]. *Bulletin of the Kanto-Koshin-Etsu English Language Education Society*, 17, 33-44. [Category 4, 11]
- Maeda, H. (2001). Nihon no kokosei eigo gakushusha ni yoru eigo gakushu horyaku shiyo—Gakushu horyaku sokutei no tameno kisokenkyu [Japanese high school students' strategy use of English language learning—Foundation study for measuring learning

- strategy]. *JLTA (Japan Language Testing Association) Journal*, 4, 95-106. [Category 11]
- Nakanishi, T. (2008). Learner perceptions and attitudes toward learning styles in Japan and Taiwan. *JLTA (Japan Language Testing Association) Journal*, 11, 21-40. [Category 11]
- Onozuka, W. (2016). Bujinesu nihongo testuto niokeru DIF no bunseki—Seibetsu oyobi kyojuuchi wo kai shuudan toshite [DIF analyses on the Business Japanese Proficiency Test: Comparisons across gender and districts of residence] *JLTA (Japan Language Testing Association) Journal*, 19, 86–108. [Category 11, 20]
- Sakuma, Y. (2008). Working memory nai no eigo shori ni kansuru bogo washa to nihonjin gaikokugo gakushusha tonon hikaku kenkyu [A comparative study of English processing and storage in working memory between English native speakers and Japanese EFL learners]. *JLTA (Japan Language Testing Association) Journal*, 11, 173-188. [Category 6, 11]
- Shimada, M. (2006). Nihongo ni oite sentakushi teiji keishiki ga kaito katei ni ataeru eikyo—Hatsuwa shikoho o mochiite [The effects of styles of presenting multiple-choice options on thinking process of Japanese listening comprehension test: A retrospective study]. *JLTA (Japan Language Testing Association) Journal*, 9, 101-116. [Category 11]
- Shimatani, H. (2004). A study on the English section of the College Scholastic Ability Test and its washback effects on English education in South Korea. *Studies in English Teaching and Learning in East Asia*, 1, 13-23. [Category 11, 14]
- Shimizu, Y., & Zumbo, B. D. (2005). A logistic regression for differential item functioning primer. *JLTA (Japan Language Testing Association) Journal*, 7, 110-124. [Category 11]
- Sugino, N., Saitoh, E., Takahashi, S., Shimizu, Y., Negishi, M., Nozawa, K., et al. (2003). The influence of the test taking strategies on the reliability of test items in English Tests. *ARELE (Annual Review of English Language Education in Japan)*, 14, 221-230. [Category 11, 13]
- Sunada, M., & Suzuki, Y. (2014). Concurrent validity of sentence repetition test: The role of pause and sentence length. *JLTA (Japan Language Testing Association) Journal*, 16, 41–60. [Category 11, 13]
- Takeuchi, N. (2000). Eigo ni yoru test shijibun ga kokosei ni ataeru eikyo ni kansuru chosa [The survey of the effect of test instructions in English given to senior high school students]. *JLTA (Japan Language Testing Association) Journal*, 3, 150-159. [Category 11]
- Tanaka, H. (2005). Donoyoni sureba gakushusha no doki zuke ha takamerareru noka?: Gakushusha no doki zuke o sokushin suru horyaku no koka kensho [How is learners' trait motivation facilitated?: Examination of a motivational strategy in classroom]. *JLTA (Japan Language Testing Association) Journal*, 7, 163-176. [Category 11]
- Tanaka, H. (2008). What facilitates students' active engagement in learning English in the

classroom? Causal analysis between psychological needs, English classroom motivation and engagement. *JLTA (Japan Language Testing Association) Journal*, 11, 96-108. [Category 11]

- Uenishi, K (2004). A study of factors contributing to English speaking proficiency: Comparing Japanese high school and university students' speaking factors. *ARELE (Annual Review of English Language Education in Japan)*, 15, 119-128. [Category 4, 11]
- Uenishi, K (2006). Nihonjin eigo gakushusha no speaking ryoku ni kansuru jissshoteki kenkyu: Goi hatsuon tonon kanrensei ni shoten o atete [An empirical research on speaking ability of Japanese learners of English: With a focus on the relation to vocabulary and pronunciation]. *Hiroshima Daigaku Gakkokyoiku Gakubu Eigo Kenkyukai Eigo to Eigo kyoiku, Tokubetsugo, Ozasa Toshiaki Taishoku Kinen Ronbunshu* [English and English education: Special number, Society for the English research at Faculty of Education of Hiroshima University, Festschrift for Dr. Toshiaki Ozasa's retirement], 135-144. [Category 4, 11]
- Uenishi, K (2006). A study of factors contributing to English writing ability: With a focus on two types of writing task. *ARELE (Annual Review of English Language Education in Japan)*, 17, 71-80. [Category 5, 11]
- Uenishi, K (2007). *Factors in determining English speaking ability: With a focus on Japanese EFL learners*. Hiroshima: Keisuisha. [Category 4, 11]
- Watanabe, Y. (2004). Exploring test-wiseness of Japanese senior high school students. *Akita Daigaku Kyoiku Bunka Gakubu Kenkyu Kiyo, Kyoiku Kagaku*, 59, 27-34. [Category 11]
- Yanagawa, K. (2003). Relative difficulty of three multiple choice listening comprehension item formats for Japanese high school learners of EFL. *The Bulletin of the Graduate School of Education of Waseda University*, 10, 1-14. [Category 2, 11]
- Yanagawa, K. (2004). Tashisentakushiki listening test no mondaibun to sentakushi no teiji jiki ga test performance ni ataeru eikyo [The Effects of provision of item stems and/or answer options on multiple-choice question listening comprehension test performance]. *STEP (Society for Testing English Proficiency) Bulletin*, 16, 30-42. [Category 2, 11]
- Yanagawa, K. (2007). Effects of stems and/or option preview on item difficulty and discrimination. *Japan Journal for Research on Testing*, 3, 3-13. [Category 2, 11]
- Yokouchi, Y. (2015). Do input text length and presentation mode affect speaking performance in retelling tasks? *JLTA (Japan Language Testing Association) Journal*, 18, 115-133. [Category 4, 11]
- Yoshimura, O., Shojima, K., Sugino, N., Nozawa T., Shimizu, Y., Saitoh, E., et al. (2005). Daigaku nyushi center shiken kishutsu mondai o riyoshita kyotsu jukensha keikaku niyoru Eigo gakuryokyu no keinen henka chosa.[An research on the trend of academic

achievement of English by IRT equating on National Center Test using the common-subject design]. *Japanese Journal for Research on Testing*, 1, 51-58. [Category 1, 11]

Zhou, Y. (2008). A comparison of speech samples of monologic tasks in speaking tests between computer-delivered and face-to-face modes. *JLTA (Japan Language Testing Association) Journal*, 11, 189-208. [Category 4, 11]

Category 12: Rating and Rating Scales

- Hirai, A., & Koizumi, R. (2008). Validation of an EBB scale: A Case of the Story Retelling Speaking Test. *JLTA (Japan Language Testing Association) Journal*, 11, 1-20. [Category 12, 13]
- Hirai, A., & Koizumi, R. (2013). Validation of empirically derived rating scales for a Story Retelling Speaking Test. *Language Assessment Quarterly*, 10, 398-422. doi:10.1080/15434303.2013.824973 [Category 4, 12]
- Koizumi, R., & Kurizaki, I. (2002). Developing rating scales for monologues from Japanese junior high school students. *ARELE (Annual Review of English Language Education in Japan)*, 13, 51-60. [Category 4, 12]
- Loucky, J. P. (2005) Making a user-friendly vocabulary knowledge scale for Japanese students: Designing and testing a dual assessment vocabulary instructor-evaluator. *Kitakyushu: Orio Aishin Tanki Daigaku Ronshu*, 39, 33-86. [Category 12, 13]
- Min, H. (2014). Introducing the writing section of the NEAT in Korea. *JLTA (Japan Language Testing Association) Journal*, 16, 1–16. [Category 5, 12]
- Nakamura, Y. (2002). Effectiveness of paired rating in the assessment of English compositions. *JLTA (Japan Language Testing Association) Journal*, 5, 61-71. [Category 12]
- Nekoda, H. (2006). Developing a standard of language proficiency required for English teachers: Using generalizability theory and item response theory. *ARELE (Annual Review of English Language Education in Japan)*, 17, 191-200. [Category 4, 12]
- Nekoda, H., Nekoda, K., & Miura, S. (2007). Developing an oral proficiency scale through the MFRM analysis of teachers' subjective judgements. *ARELE (Annual Review of English Language Education in Japan)*, 18, 221-230. [Category 4, 12]
- Ohtomo, K. (1994). Eigo kentei shiken no tameno Tsukuba eigo noryoku sokutei shakudo no settei. Gaikokugo kentei seido kaihatsu kenkyu houkokusho. [TSUs: Tsukuba English language proficiency measurement scale. A report on the development of revised Tsukuba English language testing system]. *Foreign Language Center, University of Tsukuba*, 25-62. [Category 12]
- Sato, T. (2013). The influential features on linguistic laypersons' evaluative judgments of second language oral communication ability. *JLTA (Japan Language Testing Association) Journal*, 16, 107–126. Retrieved from <http://ci.nii.ac.jp/naid/110009767205/en> [Category 4, 12]
- Sugita, Y. (2013). Comparability of accuracy and communicability tasks: Are they all equally difficult? *JLTA (Japan Language Testing Association) Journal*, 16, 65–86. Retrieved from <http://ci.nii.ac.jp/naid/110009767203/en> [Category 5, 12]
- Ushiro, Y., Koga, T., Sato, T., In'nami, Y., Hijikata, Y., Niklai, G., et al. (2004). What test

constructors should keep in mind: Constructing and scoring a translation test. *JLTA (Japan Language Testing Association) Journal*, 7, 145-162. [Category 12]

Ushiro, Y., Nakagawa, C., Morimoto, Y., Koga, T., Konno, K., & Nabeta, K. (2007). Examining the characteristics of scoring methods for written recall tests: Focusing on propositional and idea unit analyses. *JACET (Japan Association of College English Teachers) JOURNAL*, 44, 85-99. [Category 3, 12]

Category 13: Test Reliability and Validity

- Hirai, A., & Koizumi, R. (2008). Validation of an EBB scale: A Case of the Story Retelling Speaking Test. *JLTA (Japan Language Testing Association) Journal*, 11, 1-20. [Category 12, 13]
- In'nami, Y. (2010). Factor structure of the renewed TOEIC® test. *Proceedings of a symposium co-sponsored by Kobe University Shin Ishikawa Laboratory and Institute for International Business Communication (IIBC), Kyoto, Japan*, 23–33. [Category 13]
- In'nami, Y., Koizumi, R., & Nakamura, K. (2016). Factor structure of the Test of English for Academic Purposes (TEAP®) test in relation to the TOEFL iBT® test. *Language Testing in Asia*, 6(3), 1–23. doi:10.1186/s40468-016-0025-9 Retrieved from <http://www.languagetestingasia.com/content/6/1/3> [Category 1, 13]
- Isaji, T. (2005). Test-wiseness no kanten kara mita eigo reading jukutatsudo test no datosei no kensho [Examining validity of English reading proficiency test from the viewpoint of test-wiseness]. *JLTA (Japan Language Testing Association) Journal*, 7, 125-144. [Category 3, 13]
- Ito, A., Shimatani, H., & Norizuki, K. (2010). Shin kyu TOEIC no yosokuteki datosei no kensho [Two versions of the TOEIC®: A predictive validation study]. *JLTA (Japan Language Testing Association) Journal*, 13, 111-125. [Category 13]
- Ito, A., Shimatani, H., Norizuki, K., & Kinoshita, M. (2009). Kotenteki test riron ni motozuku shin kyu TOEIC® no hikaku bunseki [Two versions of the TOEIC® Test: A comparative study based on the Classical Test Theory]. *JLTA (Japan Language Testing Association) Journal*, 12, 26-45. [Category 13]
- Kamimoto, T. (1993). Tailoring the C-Test to fit: Improvement of C-Test through classical item analysis. *Language Laboratory*, 30, 47-61. [Category 10, 13]
- Kamimoto, T. (2003). A comparison of the Vocabulary Levels Test in L1 and L2. *KGU (Kumamoto Gakuen University) Journal of Language & Literature*, 9 & 10, 21-49. [Category 11, 13]
- Kamimoto, T. (2005). Goi levels test no bunseki: Kotenteki test riron to komoku hanno riron kara [Vocabulary Levels Test--A discussion in terms of Classical Test Theory and Item Response Theory]. *KGU Journal of Language & Literature*, 11 & 12, 61-105. [Category 7, 13]
- Kanatani, K., & English Diagnostic Test (EDiT) Development Group. (2006). *Eigo shindan test kaihatsu heno michi—ELPA Eigo Shindan Project no kiseki* [Road to the development of the EDiT: Trajectory of the ELPA English Diagnostic Project]. Tokyo: Association for English Language Proficiency Assessment (ELPA). [Category 6, 13]
- Kimura, S. (1996). Seijo-mondai no saiten hoho ni tsuite no kosatu [On scoring of sequencing

- tests]. *SELT (Studies in English Language Teaching)*, 19, 72-77. [Category 13]
- Kimura, S. (1996). Eitango no rinsetsu kankei ni chumokushita seijo-mondai saiten-ho to eigo noryoku shindan chiryo yobo program no kanousei [Computer assisted scoring of English word sequencing tests with a focus on word relations in the sentence: Its validity and possibility]. *Japan Society for Educational Technology*, 12, 387-388. [Category 6, 13]
- Kimura, S. (2004). Shukan test no saiten training ni okeru k-keisu, kappa-keisu, taso Rasch model no riyo [The roles of k-coefficient, kappa-coefficient and many-facet Rasch measurement in examiner training for subjective tests]. *Language and Culture*, 17, 27-36. [Category 13]
- Koizumi, R. (2003). A productive vocabulary knowledge test for novice Japanese learners of English: Validity and its scoring methods. *JABAET (Japan-Britain Association for English Teaching) Journal*, 7, 23-52. [Category 6, 13]
- Koizumi, R. (2005). Speaking performance measures of fluency, accuracy, syntactic complexity, and lexical complexity. *JABAET (Japan-Britain Association for English Teaching) Journal*, 9, 5-33. [Category 4, 13]
- Koizumi, R. (2005). Nihonjin chu kosei ni okeru happyo goi chishiki no hirosa to fukasa no kankei [Relationships between breadth and depth of productive vocabulary knowledge among Japanese junior and senior high school students. *STEP (Society for Testing English Proficiency) Bulletin*, 17, 63-80. [Category 6, 13]
- Koizumi, R. (2011). Test-Taking Processes of the Lexical Organisation Test: Comparing It With the Word Associates Test. *ARELE (Annual Review of English Language Education in Japan)*, 22, 153-168. [Category 6, 13]
- Koizumi, R., & In'nami, Y. (2016). Language assessment literacy of examinees: Their perceptions of tests before and after taking a course in language assessment. In Y. Ishikawa, S. Ishikawa, Y. Shimizu, T. Tabata, K. Cho, & T. Maeda (Eds.), *Language studies and quantitative methods* (pp. 100–115). Tokyo: Kinseido. [Category 13]
- Koizumi, R., In'nami, Y., Asano, K., & Agawa, T. (2016). Validity evidence of Criterion® for assessing L2 writing proficiency in a Japanese university context. *Language Testing in Asia*, 6(5), 1–26. doi:10.1186/s40468-016-0027-7 Retrieved from <http://www.language-testing-asia.com/content/6/1/5> [Category 5, 13]
- Koizumi, R., In'nami, Y., & Fukazawa, M. (2016). Development of a paired oral test for Japanese university students. In C. Saida, Y. Hoshino, & J. Dunlea (Eds.), *British Council New Directions in Language Assessment: JASELE Journal Special Edition* (pp. 103–121). Tokyo: British Council Japan. [Category 4, 13]
- Koizumi, R., Sakai, H., Ido, T., Ota, H., Hayama, M., Sato, M., & Nemoto, A. (2011). Toward Validity Argument for Test Interpretation and Use Based on Scores of a Diagnostic

- Grammar Test for Japanese Learners of English. *Japanese Journal for Research on Testing*, 7, 99-119. [Category 6, 13]
- Koizumi, R., & Mochizuki, M. (2011). Development and validation of the PC version of the Mochizuki Vocabulary Size Test. *JACET (Japan Association of College English Teachers) Journal*, 53, 35-55. [Category 6, 13]
- Koyama, Y., Shimizu, Y., & Kimura, S. (1997). Shitsumon-shi niyoru eigo reading test no hyomenteki datosei to hakyu-koka no kenkyu [Face validity and washback of English reading tests]. *Language Laboratory*, 34, 75-92 [Category 13, 14]
- Loucky, J. P. (2005) Making a user-friendly vocabulary knowledge scale for Japanese students: Designing and testing a dual assessment vocabulary instructor-evaluator. *Kitakyushu: Orio Aishin Tanki Daigaku Ronshu*, 39, 33-86. [Category 12, 13]
- Maeda, H. (2000). Koseigainen no datosei no kensho—Nihon no eigo kyoiku gaku kenkyu ni okeru keiko to tenbo [Examining construct validity--Tendency and prospect in English education research in Japan]. *JLTA (Japan Language Testing Association) Journal*, 3, 119-126. [Category 13]
- Maeda, H. (2004). Inga bunseki no datosei no kensho—Nihon no eigo kyoikugaku kenkyu ni okeru keiko to tenbo—[Examining validity of causal analysis—Tendency and prospect in English education research in Japan]. *JLTA (Japan Language Testing Association) Journal*, 6, 140-147. [Category 13]
- Mizumoto, A., & Takeuchi, O. (2009). Comparing frequency and trueness scale descriptors in a Likert scale questionnaire on language learning strategies. *JLTA (Japan Language Testing Association) Journal*, 12, 116-136. [Category 13]
- Morimoto, Y. (2006). Goi test no keishiki ga goi chishiki to dokkai noryoku no sokutei ni oyobosu eikyo [Effects of vocabulary test formats on measuring vocabulary knowledge and reading ability]. *STEP Bulletin*, 18, 77-91. [Category 6, 13]
- Nakamura, Y. (2003). Sokutei to hyoka ni kansuru kihontekina kangaekata [Basic concepts in measurement and evaluation]. In T. Watanabe (Ed.), *Eigo ga tsukaeru nihonjin no ikusei* [Cultivating Japanese with English abilities] (pp. 39-47). Tokyo: Sanseido, [Category 10, 13]
- Nakamura, Y. (2006). Gengo test ni okeru test riron to data bunseki [Test theories and data analyses in language testing]. In Kokuritsu Kokugo Kenkyujo [The national institute for Japanese language] (Ed.), *Sekai no Gengo test* [Language tests in the world]. (pp. 15-24). Tokyo: Kuroshio. [Category 10, 13]
- Nekoda, H. (2010). Eigo koto unyo gino no jukutatsudo ni kansuru hyokashakudo no shitsuteki datoka [Qualitative validation of rating scales on oral English proficiency]. *JLTA (Japan Language Testing Association) Journal*, 13, 55-70. [Category 13]

- Ohtomo, K. (1990). Komoku hanno riron no kenkyu: Komoku jyoho kansu to test jyoho kansu [Studies in Item Response Theory: Item information function and test information function]. *Studies in Foreign Language Teaching*, 12, 1-20. [Category 13]
- Saida, C. (2002). Validation of prefectural English tests for high school students. *JLTA (Japan Language Testing Association) Journal*, 5, 129-146. [Category 13]
- Sato, T. (2010). Validation of the EIKEN can-do statements as a self-assessment measure using Rasch measurement. *JLTA (Japan Language Testing Association) Journal*, 13, 1-20. [Category 13]
- Sugita, Y. (2010). Reliability and validity of task-based writing performance assessment for Japanese learners of English. *JLTA (Japan Language Testing Association) Journal*, 13, 21-40. [Category 5, 13]
- Shimizu, Y. (2005). Sokutei ni okeru datosei no rikai no tame ni--Gengo test no kihon gainen to shite [Understanding of test validity: As a basic concepts in language testing]. *Ritsumeikan Studies in Language and Culture Studies*, 16, 241-254. [Category 1, 13]
- Shimizu, Y., Kimura S., Sugino, N., Yamakawa, K., Ohba, M., & Nakano, M. (2003). Eibunpo noryoku hyojun test no shinraisei / datosei no kensho to shin eibunpo noryoku test: MEG [Validity and reliability of standardized English grammar tests and a new grammar test: MEG (Measure of English Grammar)]. *Policy Science Association*, 10, 59-68. [Category 6, 13]
- Shizuka, T. (2004). Reliability and validity of “invisible-gap filling” items. *JLTA (Japan Language Testing Association) Journal*, 6, 108-127. [Category 13]
- Sugino, N., Saitoh, E., Takahashi, S., Shimizu, Y., Negishi, M., Nozawa, K., et al. (2003). The influence of the test taking strategies on the reliability of test items in English Tests. *ARELE (Annual Review of English Language Education in Japan)*, 14, 221-230. [Category 11, 13]
- Sunada, M., & Suzuki, Y. (2014). Concurrent validity of sentence repetition test: The role of pause and sentence length. *JLTA (Japan Language Testing Association) Journal*, 16, 41-60. [Category 11, 13]
- Tanaka, H., & Maeda, H. (2004). Gaikokugo gakushu doki kenkyu ni okeru koseigainen “amotivation”: Sokutei no datosei kensho to negative na shitsumon komoku no eikyo [Reexamining the construct of amotivation: Validity of measurement and bias from negative items]. *JLTA (Japan Language Testing Association) Journal*, 6, 128-139. [Category 13]
- Ushiro, Y., Hamada, A., Hasegawa, Y., Dowse, E., Tanaka, N., Suzuki, K., Hosoda, M., & Mori, Y. (2015). A generalizability theory study on the assessment of task-based reading performance. *JLTA (Japan Language Testing Association) Journal*, 18, 92-114. [Category

3, 13]

- Ushiro, Y., Hijikata, Y., Shimizu, M., In'nami, Y., Kasahara, K., Shimoda, A., et al. (2005). Reliability and validity of translation tests as a measure of reading comprehension. *ARELE (Annual Review of English Language Education in Japan)*, 16, 71-80. [Category 3, 13]
- Wada, T. (2004). A study into the reading construct: Validating item types of reading tests. *JLTA (Japan Language Testing Association) Journal*, 6, 56-75. [Category 3, 13]
- Watanabe, Y. (2001). TOEFL gain scores: What do they mean to the English language program. *International Christian University Language Research Bulletin*, 15, 117-132. [Category 13]
- Watanabe, Y. (2003). Hyokasha no shinraisei [Rater reliability]. *STEP Eigo Joho*, 7-8, 34-37. [Category 13]
- Watanabe, Y. (2003). Datosei—Yoi test no saijuyo joken [Validity—The most important requirement]. *STEP Eigo Joho*, 1-2, 48-51. [Category 13]
- Watanabe, Y. (2004). Does English ability really matter?--Validating the final grades of university EFL courses. *Annual Research Report on General Education, Akita University*, 6, 19-28. [Category 13]
- Watanabe, Y. (2005). Evaluating a new course in English for academic purposes. *Annual Research Report on General Education, Akita University*, 7, 1-11. [Category 13]
- Wu, J. R. W. (2016). A socio-cognitive approach to assessing speaking and writing: the GEPT experience. *JLTA (Japan Language Testing Association) Journal*, 19, 3–11. [Category 13, 14]
- Yanase, Y. (2001). Is strategic competence ‘competence?’: Reconsidering strategic competence as a pedagogical notion. *JLTA (Japan Language Testing Association) Journal*, 4, 76-94. [Category 13]
- Yoshida, H. (2006). Using generalizability theory to evaluate reliability of a performance-based pronunciation measurement. *JLTA (Japan Language Testing Association) Journal*, 9, 86-100. [Category 13]
- Zhou, Y. J. (2007). Comparison of examinees’ reactions toward computer-based test of speaking and face-to-face interview. *The Journal of TUFS (Tokyo University of Foreign Studies) Language Education Studies*, 2, 73-95. [Category 4, 13]

Category 14: Washback and Impact

- Chapman, M. (2005). A case study of the need for change in the language testing policies of a Japanese corporation. *JLTA (Japan Language Testing Association) Journal*, 8, 51-67. [Category 1, 14]
- Cheng, L., & Watanabe, Y., with Curtis, A. (2004). *Washback in language testing: Research methods and contexts*. NJ: Lawrence Erlbaum. [Category 14]
- Cho, K. (2002). Language aptitude test no riyo to sono kyoikuteki koka ni tsuite—Daigakusei no baai [Usefulness and pedagogical effect of language aptitude test—Case of university students]. *JLTA (Japan Language Testing Association) Journal*, 5, 46-60. [Category 14]
- Horikawa, Y., & Tokuma, N. (2008). Nihongo group oral test no gakushu heno hakyu koka ni kansuru yobi teki kenkyu: Kyoshitsu nai test no baai [Preliminary study of washback on learning of group oral tests for Japanese language in a classroom environment]. *JLTA (Japan Language Testing Association) Journal*, 11, 139-155. [Category 4, 14]
- Koyama, Y., Shimizu, Y., & Kimura, S. (1997). Shitsumon-shi niyoru eigo reading test no hyomenteki datosei to hakyu-koka no kenkyu [Face validity and washback of English reading tests]. *Language Laboratory*, 34, 75-92 [Category 13, 14]
- Newfields, T. (2001). Voices in the Field: An Interview with Dan Douglas. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 5 (3). Available from www.jalt.org/test/dou_new.htm [Category 1, 14]
- Newfields, T. (2002). Voices in the Field: An Interview with Liz Hamp-Lyons. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 6 (2). Available from www.jalt.org/test/ham_new.htm [Category 1, 14]
- Newfields, T. (2003). Voices in the Field: An Interview with J. D. Brown. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 7 (1). Available from www.jalt.org/test/bro_new.htm [Category 1, 14]
- Newfields, T. (2005). TOEIC washback effects on teachers: A pilot study at one university faculty. *Toyo University Keizai Ronshu*, 31, 83-106. Available from www.newfields.info/Articles/washback.htm [Category 14]
- Newfields, T. (2005). Voices in the Field: An Interview with Yoshinori Watanabe. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 9 (1). Available from www.jalt.org/test/wat_new.htm [Category 1, 14]
- Newfields, T. (2007). Voices in the Field: An Interview with Glenn Fulcher. *Shiken: JALT Testing and Evaluation SIG Newsletter*, 11 (2). Available from www.jalt.org/test/ful_new.htm [Category 1, 14]
- Shimatani, H., & Laskowski, T. (2001). Washback effect of the introduction of a listening test into a national university entrance examination. Proceedings of the *Foreign Language*

Education and Technology Conference IV, 738-744. [Category 2, 14]

- Shimatani, H., & Laskowski, T. (2001). Influences of the introduction of a listening test into the entrance examination at Kumamoto University: Compared with the situation in South Korea. *KASELE (Kyushu Academic Society of English Language Education) Bulletin*, 29, 1-9. [Category 2, 14]
- Shimatani, H., Kinoshita, M., Otsu, A., Laskowski, T., Okuda, Y., & Yamamoto, H. (2003). Comparing the reading sections of the national unified English tests of Japan, South Korea, and Taiwan. *The JACET Kyushu-Okinawa Chapter Annual Review of English Learning and Teaching*, 8, 1-17. [Category 3, 14]
- Shimatani, H. (2004). A Study on the English section of the College Scholastic Ability Test and its washback effects on English education in South Korea. *Studies in English Teaching and Learning in East Asia*, 1, 13-23. [Category 11, 14]
- Shimatani, H., Kinoshita, M., Yamamoto, H., and Laskowski, T. (2004). High school students' impressions and perceptions of some attributes of each reading subtest in the national unified exam English tests of Japan, South Korea, and Taiwan. *Journal of Asia TEFL*, 1, 47-73. [Category 3, 14]
- Shizuka, T. (2008). Backwash of performance testing: Auditory and acoustic analyses of an utterance recorded before and after an EFL pronunciation course. *JLTA (Japan Language Testing Association) Journal*, 11, 41-60. [Category 14, 16]
- Takaki, N., Shimatani, H., Yoshida, M., Rosen, A., Toda, T., Matsuse, K., et al (1999). Eigo kikitōri shiken dōnyū to kumamoto daigaku nyūgakusha no eigo nōryōku [Listening testing the entrance examination and English language ability of Kumamoto university students]. *Memoirs of the Faculty of Education Kumamoto University (the Humanities)*, 48, 39-47. [Category 2, 14]
- Watanabe, Y. (1993). Washback effects of college entrance examination on Language learning strategies. *JACET (Japan Association of College English Teachers) Bulletin*, 23, 175-194. [Category 14]
- Watanabe, Y. (1996). Does grammar-translation come from entrance examinations? Preliminary findings from classroom-based research. *Language Testing*, 13, 318-333. [Category 14]
- Watanabe, Y. (1996). Investigating washback in the Japanese EFL classroom--Problems of methodology. *Australian Review of Applied Linguistics (Series S)* (pp. 208-239). [Category 14]
- Watanabe, Y. (1997). Nyūshi ga warui to iu maeni—Nyūshi eigo o ikasu tameno jugyō ni mukete no teian [Suggestion for lessons to make use of English appearing in entrance examination]. *ASTE Newsletter*, 37, 1-8. [Category 14]
- Watanabe, Y. (1997). Nyūsyi kara eigo o hazusu to jugyō ha kawaru ka [Do lessons change

- when we do not conduct English tests in entrance examination?]. *The English Teachers' Magazine TAISHUKAN, special issue*, 30-35. [Category 14]
- Watanabe, Y. (1998). Daigaku nyushi no hakyu koka (Jo)—Jugyo kenkyu kara kosatsu suru [Washback effect of university entrance examination Part I—Discussion from classroom research]. *The English Teacher's Magazine TAISHUKAN, 1*, 68-71. [Category 14]
- Watanabe, Y. (1999). Daigaku nyushi no hakyu koka (Ge)—Jugyo kenkyu kara kosatsu suru [Washback effect of university entrance examination Part 2—Discussion from classroom research]. *The English Teacher's Magazine TAISHUKAN, 2*, 68-71. [Category 14]
- Watanabe, Y. (2000). Nyushi no tayoka de mondai ga meikaku ni nattekita [Problems have been clarified because of variation of entrance examination]. *The English Teacher's Magazine TAISHUKAN, 6*, 16-18. [Category 14]
- Watanabe, Y. (2000). Washback effects of the English section of Japanese university entrance examinations on instruction in pre-college level EFL. *Language Testing Update* (International Language Testing Association Official Newsletter), 27, 42-47. [Category 14]
- Watanabe, Y. (2001). Does the university entrance examination motivate learners?: A case study of learner interviews. *Trans-equator exchanges: A collection of academic papers in honour of professor David E. Ingram* (pp. 100-110). Akita, Japan: Faculty of Education and Human Studies, Akita University. [Category 14]
- Watanabe, Y. (2003). Gengo test no hakyu koka [Washback effects of language tests]. *STEP Eigo Joho, 3-4*, 40-43. [Category 14]
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng & Y. Watanabe, with A. Curtis (Eds.), *Washback in language testing: Research methods and contexts*. (pp. 19-36). NJ: Lawrence Erlbaum. [Category 14]
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng & Y. Watanabe, with A. Curtis (Eds.), *Washback in language testing: Research methods and contexts*. (pp. 129-146). NJ: Lawrence Erlbaum. [Category 14]
- Watanabe, Y. (2004). Test to daigaku eigo kyoiku [Test and English education in University]. *Eigo Seinen, 12*, 20-21.
- Watanabe, Y. (2005). Building a Tower of Babel. *Language Assessment Quarterly, 2*, 69-75. [Category 14]
- Watanabe, Y. (2006). Does teaching to the test have to be decontextualized? *Language Assessment Quarterly, 3*, 375-380. [Category 14]
- Wu, J. R. W. (2016). A socio-cognitive approach to assessing speaking and writing: the GEPT experience. *JLTA (Japan Language Testing Association) Journal, 19*, 3–11. [Category 13, 14]

Category 15: Ethics and Standards

- Association of Japanese language Teachers in Europe. (2005). Europe ni okeru nihongo kyoiku to Common European Framework of Reference for Languages [Japanese language education in Europe and Common European Framework of Reference for Languages]. Tokyo: The Japan Foundation. [Category 15]
- Brown, J. D. (1999). The roles and responsibilities of assessment in foreign language education. *JLTA (Japan Language Testing Association) Journal*, 2, 1-21. [Category 1, 15]
- Japan Language Testing Association (JLTA) Steering Committee (Eds.), JLTA Committee of Language Testing Terms. (2006). *The JLTA bilingual list of language testing terms / The JLTA Code of Good Testing Practice*. Nagano: JLTA Secretariat. [Category 1, 15]
- Newfields, T. (2007). Engendering assessment literacy: Narrowing the gap between teachers and testers. In Y. S. Lee, et al. (Ed). *Assessing Foreign Language Performances: Proceedings of the 2007 Korea English Language Testing Association International Conference* (pp. 22-42). Seoul, Korea: Korea English Language Testing Association. Available from www.tnewfields.info/Articles/testlit.htm [Category 1, 15]
- Noguchi, H., Kumagai, R., Wakita, T., & Wada, A. (2007). Nihongo can-do statements ni okeru DIF komoku no kenshutsu [Detection of DIF items in Japanese can-do-statements]. *JLTA (Japan Language Testing Association) Journal*, 10, 106-118. [Category 15]
- Ohtomo, K. (1988). Eigo test ni okeru komokusentei kijyun no settei [Standard setting for selecting good items in English language tests]. *Studies in Foreign Language Teaching*, 10, 3-25. [Category 15]
- Wada, T., Naganuma, N., & Tanaka, A. (2004). Gengo noryoku no hattatsu dankai kijutsu ni tsuite [Developmental stages of language proficiency]. *Gengo johogaku kenkyu hokoku [Reports on Language Information Research]*, 2, 95-100. [Category 15]

Categories below were added from 2008.

Category 16. Classroom Based Assessment

Cho, D. W., & Park, J. Y. (2013). Which grading system, a letter-grading system or an S/U system, is more appropriate for language courses? *JLTA (Japan Language Testing Association) Journal*, 16, 15–42. Retrieved from <http://ci.nii.ac.jp/naid/110009767201/en> [Category 1, 16]

Fukazawa, M. (2011). Validity of Self-Assessment of Speech Performance: A Case of Japanese High School Students. *JLTA (Japan Language Testing Association) Journal*, 14, 61-79. [Category 4, 16]

Hirai, A., Ito, N., & O'Ki, T. (2011). Applicability of Peer Assessment for Classroom Oral Performance. *JLTA (Japan Language Testing Association) Journal*, 14, 41-59. [Category 4, 16]

Kanaoka, M. (2009). Self-focused written assignments for authentic self-assessment in college EFL classes: Beyond test-based evaluation methods. *JLTA (Japan Language Testing Association) Journal*, 12, 65-84. [Category 16]

Shizuka, T. (2008). Backwash of performance testing: Auditory and acoustic analyses of an utterance recorded before and after an EFL pronunciation course. *JLTA (Japan Language Testing Association) Journal*, 11, 41-60. [Category 14, 16]

Category 17. Technology and Language Testing

Category 18. Testing Young Learners

Sakuma, Y. (2010). A study of Japanese and English versions of Stroop and Reverse-Stroop tests in Japanese elementary school higher grade students' English activities. *JLTA (Japan Language Testing Association) Journal*, 13, 127-143. [Category 18]

Yim, S. Y. (2016). Teachers' perspectives on using peer assessment among young EFL learners. *JLTA (Japan Language Testing Association) Journal*, 19, 15–23. [Category 9, 18]

19. Language Testing Policy

20. Research Methodology

In'nami, Y., & Koizumi, R. (2014). Effects of verbalization on L2 listening test performance.

Journal of Shinawatra University, 1(1), 22–41. [Category 2, 20]

- In'nami, Y., & Koizumi, R. (2014). Research synthesis and meta-analysis in second language learning and testing. *English Teaching & Learning*, 38(3), 1–27. Special Issue: Research synthesis and meta-analysis in second language learning and testing. doi:10.6330/ETL.2014.38.3.01 [Category 20]
- Koizumi, R., In'nami, Y., Azuma, J. Asano, K., Agawa, T., & Eberl, D. (2015). Assessing L2 proficiency growth: Considering regression to the mean and the standard error of difference. *Shiken*, 19(1), 3–15. Retrieved from <http://teval.jalt.org/node/16> [Category 20]
- Onozuka, W. (2016). Bujinesu nihongo testuto niokeru DIF no bunseki—Seibetsu oyobi kyojuuchi wo kai shuudan toshite [DIF analyses on the Business Japanese Proficiency Test: Comparisons across gender and districts of residence] *JLTA (Japan Language Testing Association) Journal*, 19, 86–108. [Category 11, 20]
- Utsunomiya, Y., Maruyama, M., & Ogasawara, S. (2016). Estimating TOEIC scores using G-TELP scores: A Bayesian Model in a Japanese national university. *JLTA (Japan Language Testing Association) Journal*, 19, 27–45. [Category 1, 20]