

# **A Standardised Test as a Tool for the Reliable Verification of Acquired Knowledge**

Zdeněk Caha, Jindřiška Šulistová

The Institute of Technology and Business in České Budějovice

## **Abstract**

This article deals with the testing of the achieved language skills of students in a higher education institution. Various forms of testing are reviewed, including the phenomenon of electronic testing (e-testing). A common problem when testing is the validity and the reliability of the tests. It is for this reason that for the research a standardized tool - mock test - was used and converted into electronic form. The aim of the research was to verify, through the aforementioned tests, whether students of English at the Institute of Technology and Business in České Budějovice (ITB), after completing the first two semesters of general English (the course was preceded with English acquisition at primary and secondary schools), actually achieved a B1+ level according to the Common European Framework of Reference for Languages (CEFR) for the written part of the test.

**Key words:** testing, e-testing, validity, reliability, level, B1+, higher education institution, English

## **Introduction**

Many higher education institutions claim that their students have achieved a certain level of English. However, this level is not usually measured through a standardised test and there is often no official record on a candidate's performance. The ITB in České Budějovice has set itself the goal of verifying, through a standardised written mock test, the attained level of English of freshmen attending the first two semesters of general English at the institute.

Testing is in principle a means by which to establish how something works. As far as a learning environment is concerned, testing provides information on the achieved level of knowledge or skills. There are five main types of language testing: placement tests; language aptitude tests; achievement tests; proficiency tests; and diagnostic tests.

The main objective of placement tests is to decide on an appropriate level / class / course for a student. This form of testing is usually used by language schools. Language aptitude tests indicate how well a learner can learn a language under given conditions within a given time framework. Achievement tests can be divided into short-term and long-term tests. These measure how much knowledge from a syllabus has been acquired within a semester / year. Proficiency tests provide information on the knowledge and ability of a student from the broadest possible perspective. This form of testing is independent of a specific textbook and period of time. IELTS, TOEFL or TOEIC are examples of this form of English examination. Diagnostic tests diagnose how much and what a learner knows. It provides both the learner and the teacher alike with information on the weaknesses of a student i.e. they provide valuable feedback on what areas require attention. The key criteria that a test must fulfil to provide thorough and accurate information on a student are reliability, validity, practicality and the backwash effect.

Tests can be taken in either printed form or computer-based. Both forms have their strengths and weaknesses and are suitable for certain subjects and certain learners.

The issue of the differences in learners' performance between computer-based and paper-based tests was investigated by H.K. Lee (2004). Lee states in his comparative study that the analytic components of computer - based tests were assessed considerably better.

Of late, computer-based testing for ESL (English as a Second Language) has been in the spotlight, not only in connection with comparisons to paper-based testing, but also in terms of its assessment, possible self-assessment, accessibility, and the creation of authentic tests. Hunt, Neil, Barnes (2007) state that a vast array of information and communication technology (ICT) is utilised for testing, but its use for assessment purposes is very limited.

Alderson and Huhta (2005) published a paper on computer-based diagnostic tests. They focused on the dialang online assessment system which is based on the Common European Framework of Reference for Languages (CEFR). It is a testing system which gives learners feedback on their language skills rather than on their proficiency.

Hansen and Mislevy (2005) support the idea of accessibility to computer-based testing for students with special needs within a validity framework. However, they go on to express concerns that the extended time frame for the test, larger font size and access to online dictionaries, could undermine the very validity of the tests. Both authors emphasize the need for a framework within which these extra features for learners with special needs would not invalidate the test results.

Douglas and Hegelheimer (2007) deal with the issue of computer-based test assessment. They reveal the threats that are inherent in electronic language testing, such as the wide range of technologies that enable students to cheat in tests, question formulation, feedback, etc. Ockey (2009) also focuses on the developments and challenges facing computer-based test assessments stating that computer-based testing has not met

expectations. He goes on to state that he believes that the growth in its use will expand its evaluation capacity.

Cummins and Davesne (2009) recommend using electronic portfolios for second language assessment. The portfolios offer qualitative and quantitative assessments which can be employed for teaching and learning purposes.

Kuramuto (2013) in his study on reading comprehension level prediction in EFL testing, recommends the use of not only computer-based multiple-choice questions but also written questions in order to assess a learners abilities.

Strother, Fazal, Johnson and Millsap (2008) consider the suitability of language proficiency computer-based testing from the point of view of non-native speakers. The survey they carried out proved that computer-based tests are equivalent to paper-based tests.

In their paper, Fryer, Bovee and Nakao (2014), focus on the motivational orientation of learners. Their findings suggest that not all students are in favour of taking computer-based tests due the low value assigned to the results.

## **Materials and Methods**

The aim of the research was to verify, through a standardized test, whether students of foreign languages at the Institute of Technology and Business in České Budějovice (ITB) achieve a B1+ level for the written part of the test at the end of the second semester. Discussions on the B1+ level centre on the minimum percentage of compliance with the requirements for passing the test. The ITB, taking into account the great heterogeneity of the language skills of students attending higher education, has set this limit at 70 %, whereas for international tests this stands at 50 %. This minimum completion percentage was determined on the basis of many years of experience teaching English as a foreign language. It is considered to be relevant because it corresponds to the minimum requirements for students who wish to continue their studies during which they will be required to study professional language relevant to their specialization.

At the ITB in České Budějovice students can follow full-time and part-time bachelor study programmes (Economics, Mechanical Engineering, Civil Engineering, Transportation, Logistics) of which general language lessons form an integral part. The students of Economics have two 90-minute lessons a week, all the other study directions have one 90-minute lesson a week. The level of English of the freshmen varies from A2 to B2 according to the Common European Framework of Reference for Languages (CEFR). The differences are caused by many factors, such as motivation, foreign exchange stays, age, whether or not a school leaving examination for English was taken at secondary school, etc. The objective within the first year is to equalize this initial level of the learners to B1 level in order to prepare them for studying professionally-oriented language. The preparation is aimed at improving their receptive and productive skills in writing on the basis of the standardised City and Guilds mock test (Caha 2013, 2014, 2015).

Based on the B1 requirements students "*(You) can understand the main points of clear standard communication; convey information, feelings & opinions on familiar topics; engage in discussion to reach shared understanding; produce simple connected text; describe experiences, events, hopes & ambitions.*" (City & Guilds).

### **The research sample and the implementation period of research**

The research took place during the examination period of the spring semester of the 2014/2015 academic year. The research sample consisted of 504 students enrolled in the AJP\_2 course. Of the total research sample, 362 were full-time students (143 men; 219 women) and 142 were part-time students (55 men; 87 women). The full-time students were aged 18 - 25 years, the part-time students 25 - 55 years.

The written part of a standardized B1 level test for the City & Guilds International Examination was used as the diagnostic tool for measuring the level of students' knowledge. The written part takes about 2 hours and 40 minutes and includes listening (approx. 30 minutes), reading and writing (approx. 2 hours and 10 minutes).

The listening part consists of four parts in which students either choose the best answer or fill in missing information.

The reading part also consists of four parts. The tasks focus on looking for missing paragraphs / sentences, appropriate headings and answering questions with a limited number of words.

The test was completed electronically. This was followed by electronic test evaluation with exception to the writing task, which was corrected manually by teachers based on the regulations and methodology of City & Guilds. The results were statistically analysed.

The research analysis should either confirm or refute the following research assumption:

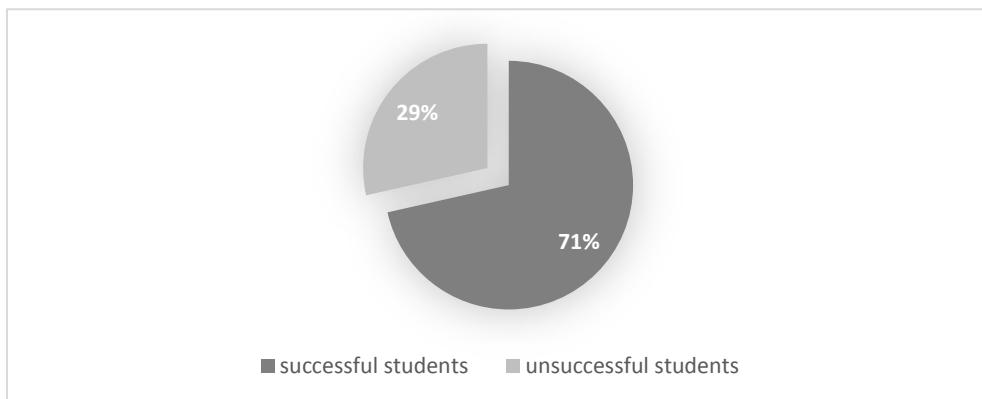
*80 % of the students attending lessons of general English will achieve the B1+ level for the written part of the test (with the minimum limit for passing the test set at 70 %).*

The percentage was elicited from the past students' results which were verified by internal (non-standardized) test. And the research authors supposed that the internal test is at the same level as the standardized one.

## **Results**

As can be seen in Figure 1, the research assumption was refuted because the percentage of students who achieved the B1+ level was only 71 %.

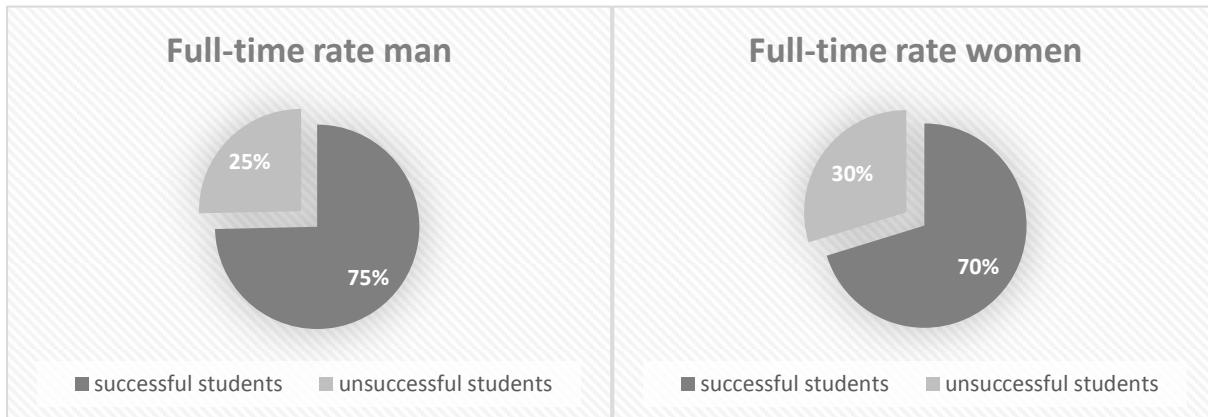
Figure 1: All tested students – overall results



Source: authors

Figure 2: shows the difference between full-time and part-time students. It is clear that, overall, part-time students were more successful when tested.

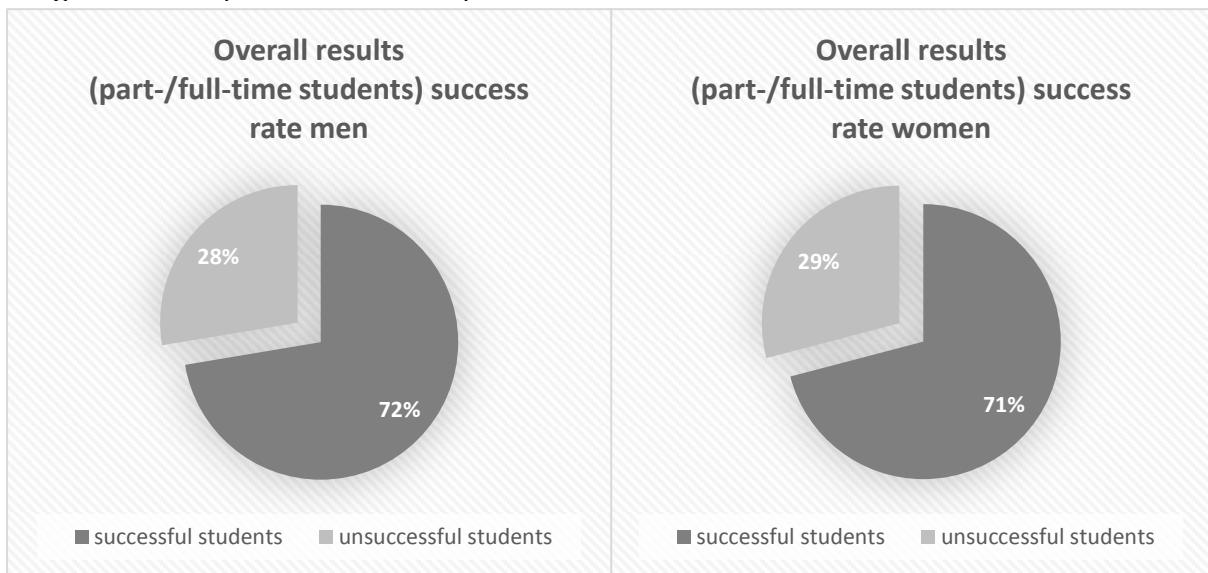
Figure 2: Full-time students – total rate of success



Source: authors

In Figure 3 a comparison is made of the results of full-/part-time students by gender. The graph shows almost no difference in the success rate.

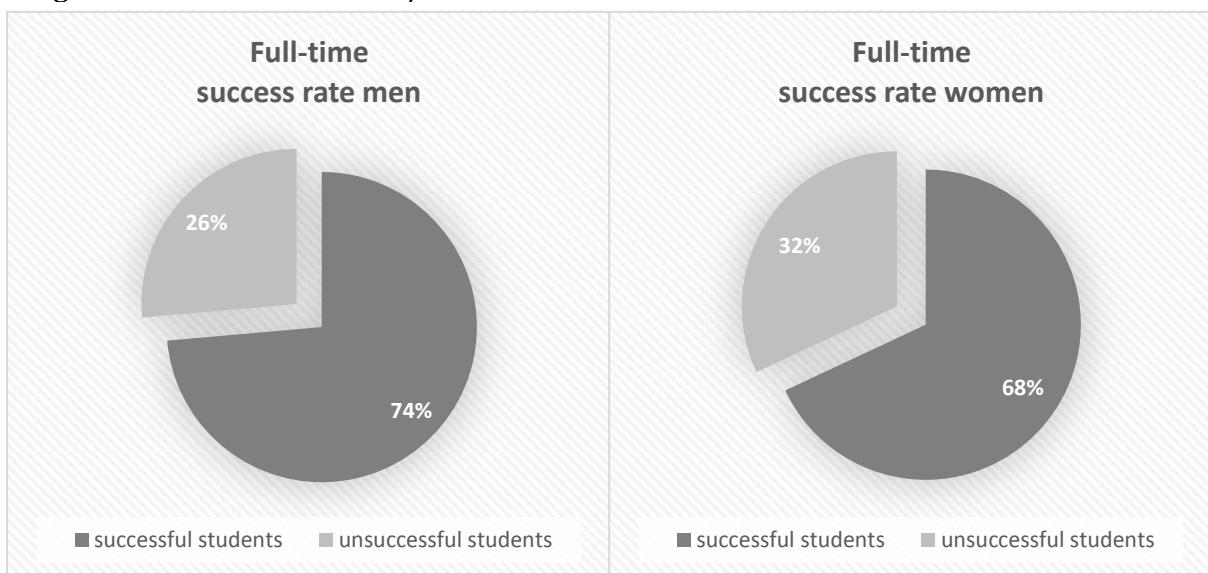
Figure 3: Part-/full-time female / male students – differences in success rate



Source: authors

In Figure 4 a comparison is made of the results of full-time students by gender. The graph shows that men were more successful.

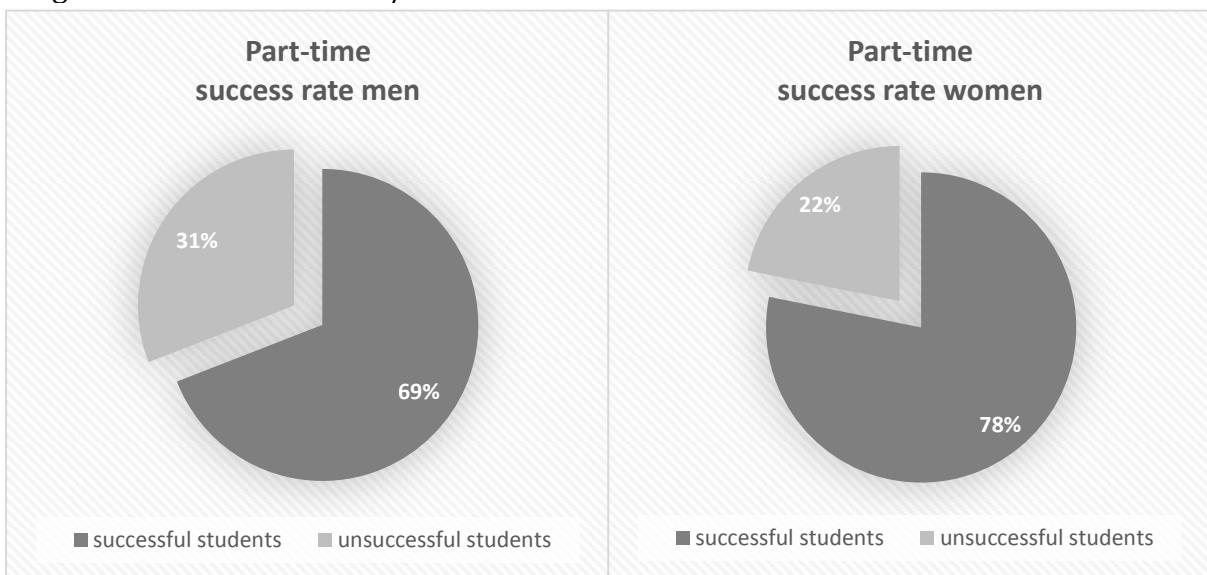
Figure 4: Full-time female / male students – success



Source: authors

Figure 5 compares the results of part-time students by gender. It is evident that female part-time students were almost 10 % more successful than their male counterparts, and that this group (at 78 %) was the closest to confirming the research hypothesis.

Figure 5: Part-time female / male students – success rate

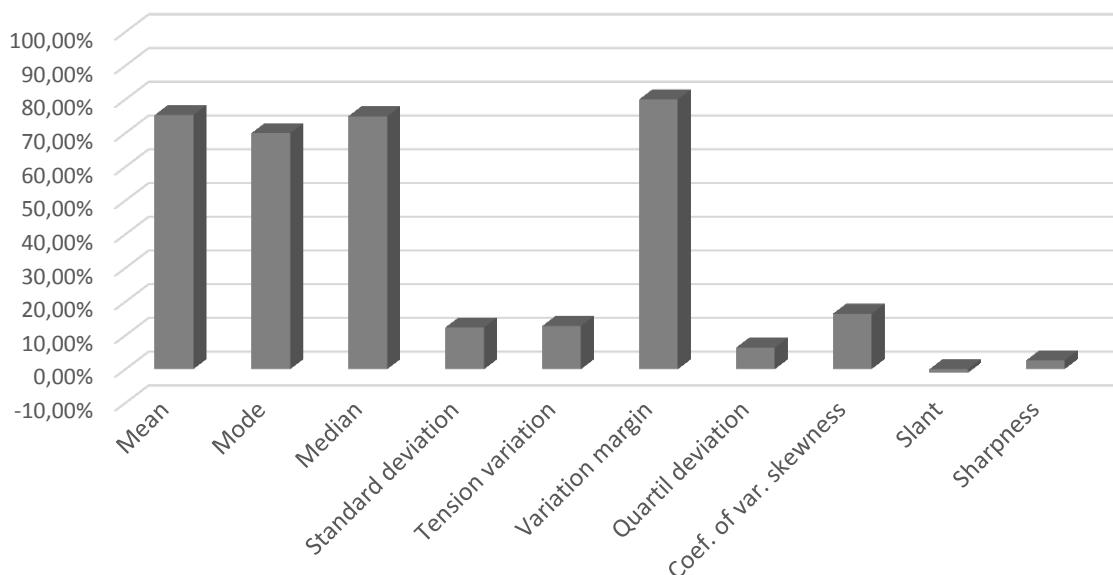


Source: authors

Figure 6 gives a more detailed analysis of the test results. The resulting values, mean, mode and median reveal that the average student is able to take and pass the examination.

The standard deviation and variation margins are not high. This proves that a relatively large number of students are just below or just above the threshold for passing the test.

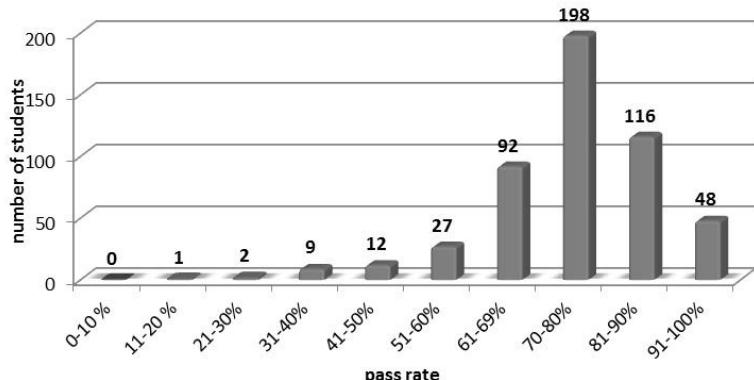
Figure 6: Detailed analysis of overall results – Part, full time students



Source: authors

Figure 7 illustrates the distribution of students. From this it is clear that the distribution largely corresponds with the Gaussian curve.

Figure 7: Assessment of full-time and part-time students



Source: authors

## Discussion and conclusion

Although the results of the research refuted the research assumption, the standardized test proved to be a very convenient tool for checking the outcome of the language skills of students. The standardized test has shown the areas which are needed to be paid the attention to, use of English – collocations, reading and listening comprehension with the focus on the gathering detailed pieces of information.

The use of computer-based testing in combination with a standardised examination tool seemed to motivate students in comparison with the internal tests used before because the former tests lack the variety of tasks. The standardized electronic test reminded of the atmosphere when taking a real test and this fact caused that students' approach was more responsible, in particular the part-time students. Consequence of the testing was an obvious improvement of students' attitude to the language acquisition. This fact has been confirmed by all professional language teachers at ITB. The survey was innovative at the institution in that it pointed to the potential practical use of a mock test for the verification of a declared or achieved level of English, as a result of which a significant number of students were encouraged to take the real examination. As far as level testing is concerned - based on the updated information from other language departments where the language is not the major – it is not very common. However, it is highly recommendable. The other positive aspects of standardized computer-based test application are taking into account the number of students who participate, the related overhead costs, and environmental sustainability.

Further, the research has shown an interesting fact, i.e. the male students in full-time form of study are a bit more successful than the male students in part-time form of study. On the other hand, the findings by the female students show the opposite result. It could be caused by a wide range of aspects, which will be examined in the next study of authors.

Future examination preparation should be focused on the further development and training of language skills. The authors have found out weaknesses and strengths in the written test skills which enables to concentrate on them in detail in the future. It supposed to verify the results at the end of each academic year and based on the findings to take steps. Based on the differences between the study forms and gender the research will be carried out and it is expected to identify the causes. Consequently, the measures will be adopted according to the current situation.

## **References:**

- ABBASIAN, G., A. HARTOONIAN, 2014. Using Self-Regulated Learning Strategies in Enhancing Language Proficiency with a Focus on Reading Comprehension. *English Language Teaching*. **7**(6), 160. ISSN 1916-4742.
- ALDERSON, J. CH., 2000. Technology in testing: The present and the future. *System*. **28**(4), 593-603. ISSN: 0346-251X.
- ALDERSON, J. CH. and A. HUHTA, 2005. The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*. **22**(3), 301-320. ISSN 0265-5322.
- BAX, S., 2013. The cognitive processing of candidates during reading tests: Evidence from eye-tracking. *Language Testing*. **30**(4), 441-465. ISSN 0265-5322.
- BLOOM, B., V. SMIDOWICZ, 2005. *City&Guilds International English for Speakers of Other Languages 3 Achiever Student's Book New Edition + audio CD*. London: City&Guilds. ISBN: 9780851931630.
- BRANTMEIER, C. and R. VANDERPLANK, 2008. Descriptive and criterion-referenced self-assessment with L2 readers. *System*. **36**(3), 456-477. ISSN: 0346-251X.
- BRUNFAUT, T. and A. REVESZ, 2015. The role of task and listener characteristics in second language listening. *TESOL Quarterly*. **49**(1), 141-168. ISSN 1545-7249.
- CAHA, Z., 2013. Aplikace elektronických médií při výuce a testování v cizích jazycích na profesně zaměřené škole. In: *E-konference na téma Využití elektronických médií ve výuce a testování v cizích jazycích na profesně zaměřené škole. 1. vyd*. České Budějovice: Tribun EU s. r. o., 15-20. ISBN 978-80-263-0358-9.
- CAHA, Z., 2014. Sumarizace zkušeností z elektronického testování rámci předmětů cizí jazyk na vysoké škole. In: *Quare 2014 : sborník příspěvku. 1. Vyd*. Hradec Králové: Magnaimitas, 843-848. ISBN 978-80-879552-00-7.
- CAHA, Z., 2015. Language Education System at Institute of Technology and Business in České Budějovice. *Auspicia*. **12**(2), 8-13. ISSN 1214-4967.
- CHAPELLE, C. A. and D. DOUGLAS, 2006. Assessing language using computer technology. *Language Testing*. **23**(4), 544-550. ISSN 0265-5322.
- CHRÁSKA, M., 1999. *Didaktické testy*. Brno: Paido. ISBN 80-85931-68-0.

- CUMMINS, P. W. and C. DAVESNE, 2009. Using electronic portfolios for second language assessment. *The Modern Language Journal*. **93**(1), 848-867. ISSN 1540-4781.
- DE LA FUENTE, M. J., 2014. Learners' attention to input during focus on form listening tasks: the role of mobile technology in the second language classroom. *Computer Assisted Language Learning*. **27**(3), 261-276. ISSN 0958-8221.
- ESOL-LEVELS, 25. 3. 2016. *cityandguildsenglish*. [online]. [accessed: 2016-03-29]. Available at: [http://www.cityandguildsenglish.com/ESOL\\_levels](http://www.cityandguildsenglish.com/ESOL_levels)
- FRYER, L. K., H. N. BOVEE and K. NAKAO, 2014. E-learning: Reasons students in language learning courses don't want to. *Computers & Education*, **74**, 26-36. ISSN 0360-1315.
- HANSEN, E. G. and R. J. MISLEVY, 2005. Accessibility of computer-based testing for individuals with disabilities and English language learners within a validity framework. *Online assessment and measurement: Foundation, challenges, and issues*. Hershey, PA: Idea Group Publishing, Inc.
- HANSEN, E. G. et al., 2005. Accessibility of tests for individuals with disabilities within a validity framework. *System*. **33**(1), 107-133. ISSN: 0346-251X.
- HARMER, J., 2015. *The Practice of English Language Teaching. Longman Handbooks for Language Teaching*. Harlow: Pearson. ISBN 978-1447980254.
- HUNT, M., S. NEILL and A. BARNES, 2007. The use of ICT in the assessment of modern languages: the English context and European viewpoints. *Educational Review*. **59**(2), 195-213. ISSN 0013-1911.
- IN'NAMI, Y. and R. KOIZUMI, 2011. Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly*. **8**(3), 250-276. ISSN 1543-4311.
- JAMIESON, J., 2005. Trends in computer-based second language assessment. *Annual Review of Applied Linguistics*. **25**, 228-242. ISSN 0267-1905.
- JEONG, H. et al., 2011. Testing Second Language Oral Proficiency in Direct and Semidirect Settings: A Social-Cognitive Neuroscience Perspective. *Language Learning*. **61**(3), 675-699. ISSN 0958-8221.
- KARL, M. et al., 2007. Comparison between computer-aided testing and traditional multiple choice: an equivalence study. *European Journal of Dental Education*. **11**(1), 38-41. ISSN 1600-0579.
- KNOCH, U., A. ROUHSHAD and N. STORCH, 2014. Does the writing of undergraduate ESL students develop after one year of study in an English-medium university?. *Assessing Writing*. **21**, 1-17. ISSN 1075-2935.
- KURAMOTO, A., 2013. Computer-based Questions with Considerable Explanatory Power to Predict the English Comprehension Levels of EFL-Learners. IASTED Multiconferences [online]. Innsbruck: ACTA press, [accessed: 2016-03-22]. DOI: 10.2316/P.2013.792-008. Available at: <http://www.actapress.com/Abstract.aspx?paperId=455479>

- LABORDA, J. G., 2009. Interface architecture for testing in foreign language education. *Procedia-Social and Behavioral Sciences*. **1**(1), 2754-2757. ISSN 1877-0428.
- LEE, H. K., 2004. A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing*. **9**(1), 4-26. ISSN 1075-2935.
- LIM, G. S. et al., 2013. Standard setting to an international reference framework: Implications for theory and practice. *International Journal of Testing*. **13**(1), 32-49. ISSN 1530-5058.
- NEWBOLD, D., 2009. Co-certification: a new direction for external assessment?. *ELT journal*. **63**(1), 51-59. ISSN 1916-4742.
- NICOLÁS-CONEZA, F., J. R. DE LARIOS. and Y. COYLE, 2014. Development of EFL students' mental models of writing and their effects on performance. *Journal of Second Language Writing*. **24**, 1-19. ISSN 1060-3743.
- OCKEY, G. J., 2009. Developments and Challenges in the Use of Computer-Based Testing for Assessing Second Language Ability. *The Modern Language Journal*. **93**(1), 836-847. ISSN 1540-4781.
- ROEVER, C., 2001. Web-based language testing. *Language Learning & Technology*. **5**(2), 84-94. ISSN 1094-3501.
- SCRIVENER, J., 2011. Chapter 12.2. In: SCRIVENER, J., *Learning Teaching. Third Edition*. Oxford: Macmillan. ISBN: 978023072984.
- STROTHER, J. B., et al., 2007. Effects of test modality—computer-based versus paper-and-pencil—on nonnatives' English results. In: *Proceedings of the Seventh IASTED International Conference. 2008*. p. 063.

### **Contact address of the author(s):**

Mgr. Zdeněk Caha, MBA, Ph.D., Department of Management, The Institute of Technology and Business in České Budějovice, Okružní 10, 370 01 České Budějovice, Czech Republic, *email: caha@mail.vstecb.cz*

Mgr. Jindřiška Šulistová, Department of Foreign Languages, The Institute of Technology and Business in České Budějovice, Okružní 10, 370 01 České Budějovice, Czech Republic, *email: sulistova@mail.vstecb.cz*

---

CAHA, Z. a J. ŠULISTOVÁ, 2016. A Standardised Test as a Tool for the Reliable Verification of Acquired Knowledge. *Littera Scripta* [online]. České Budějovice: The Institute of Technology and Business in České Budějovice, **9**(1), 21-32 [accessed: 2016-06-27]. ISSN 1805-9112. Available at: [http://journals.vstecb.cz/category/littera-scripta/9-rocnik/1\\_2016/](http://journals.vstecb.cz/category/littera-scripta/9-rocnik/1_2016/)

---