

English version of the manuscript submitted to the May 2017 issue of *The English Teachers' Magazine*

Japan Language Testing Association (JLTA)
Proposal for Handling English Testing within the 'Prospective University Entrance Scholastic Abilities Evaluation Test [provisional name]'

Japan Language Testing Association's (JLTA's) Committee of the Position Statement on the New University Entrance Examination Policy

As a part of various initiatives targeting the articulation between secondary and tertiary education in Japan, assorted measures have been discussed by the Ministry of Education, Culture, Sports, Science and Technology (MEXT). These measures include the introduction of comprehensive English testing in university entrance examinations and the use of private-sector qualification and certification examinations. The prospect on such measures has also been made available to the public. Since its establishment in 1996, the Japan Language Testing Association (JLTA) has endeavored to acquaint the educational sector and society at large not only with the practical challenges associated with appropriate ways of constructing and using tests, but also with how to verify learning and educational evaluation, while dealing with even broader issues such as university entrance examinations. In the course of a recent series of reforms in particular, a special committee was convened and discussions have been summarized as a position paper, which was subsequently submitted to MEXT as of January 4, 2017. This paper is intended to introduce the full, annotated text of the recommendations to a broader readership.

Proposal for Handling English Testing within the 'Prospective University Entrance Scholastic Abilities Evaluation Test [provisional name]'

1. Preface

A document was issued on August 31, 2016 by MEXT, reporting on the progress of reforms to the articulation between secondary and tertiary education. The document included an announcement with regard to the status of the 'Prospective University Entrance Scholastic Abilities Evaluation Test [provisional name]' (http://www.mext.go.jp/b_menu/houdou/28/08/1376777.htm), offering the prospect of "implementation policies" being formulated by early 2017. While endorsing their general orientation, the JLTA herewith voices our opinions to the range of issues regarding the policy based on our expertise.

2. Comments on English Testing within the 'Prospective University Entrance Scholastic Abilities Evaluation Test [provisional name]'

(1) Transparent Accreditation Procedures

Although it is ultimately the state (i.e., the National Center for University Entrance Examinations) that carries out accreditation when use is made of private-sector qualification and certification examinations, it is necessary to ensure the clarity of accreditation procedures. At such times, we propose implementation of the following two schemes.

- i. Determination for accreditation be made only after sufficient verification of not only the feasibility but also the quality of the test.¹
- ii. Multiple language testing researchers be deployed when formulating accreditation procedures and that such procedures should be highly transparent and reflective of the results of the latest language testing research.²

(2) Disclosure of information

With regard to private-sector qualification and certification examinations, we propose implementation

of the following three points.

- i. Private-sector qualification and certification examinations often differ in terms of their testing purpose, target proficiency, and target test takers respectively.³ This fact should be disseminated to all test users (i.e., examinees, individual universities making use of scores, and anyone else involved with university entrance examinations).
- ii. Testing agencies should publish the kind of detailed information that is essential for selecting tests, including a test's purpose and targeted proficiencies, scoring criteria and methods, as well as how the test was drafted and implemented and methods for its appropriate use. In particular, it should be clearly stipulated how the test complies with the government's official curriculum guidelines.⁴
- iii. Under the direction of MEXT, testing agencies and test users should be clearly requested to thoroughly implement i. and ii. above.

(3) Considerations for Selecting University Entrance Examinations for Selection Purposes

i. Handling Score Conversion Tables

When accrediting multiple private-sector qualification and certification examinations, the current proposal assumes a comparison of scores using a conversion table. However, the target proficiencies measured by each test differ. As a result, direct comparison of scores between tests and interpreting their minute differences as differences in general English proficiency would be inappropriate. MEXT should be explicit about such constraints and request that test users interpret and make use of test results appropriately.⁵

ii. Linking Testing to Each University's Needs and Admissions Policies

Each university, having conducted systematic analysis sufficient for its respective needs and admissions policies, should on that basis select a test and make appropriate use of the test results. MEXT should propose specific procedures for selecting and making use of the results of appropriate tests based on full knowledge of language testing research.⁶

(4) Fairness of Opportunity

Concrete countermeasures should continue to be considered and implemented to prevent examinees from experiencing inequality of opportunity in taking exams, such as the one arising from disabilities or economic or regional disparities. Especially with regard to regional disparities, increased support should be given to provide fairness of opportunity. Tests that can be taken in rural areas are limited, to such an extent that even where taking a test is possible, doing so requires considerable time and investment.⁷

(5) Measures Not Solely Reliant on Tests and Examination System Reforms

An overview of empirical studies of past washback effects leads us to expect that the introduction of comprehensive four-skills testing by itself does not necessarily lead to improvements in English education at the high-school level.⁸ The findings of these studies should be fully considered, so that measures⁹ can be formulated and implemented to produce an intended effect more effectively. Simultaneously, rather than relying solely on testing reforms to improve English-language education, we recommend that the following two points be carried out to strengthen teacher training aimed at improving teachers' leadership skills.

- i. Teacher training be scrutinized in terms of both programs and practical content, so it may fully reflect the results of language testing research.¹⁰
- ii. This content includes the original purpose of testing in English-language education, the role that instructors should play, and basic knowledge and appropriate usage methods regarding testing.¹¹

(6) Utilization of Test Results Leading to Learning

To fully take advantage of the results of comprehensive four-skills testing in high school and university learning, all testing agencies should provide score reports and related materials useful for guidance and learning.¹² Furthermore, procedures and measures should be proposed that are helpful to test users in guidance and learning.¹³

3. Conclusion

The JLTA has spent the past twenty years engaging in educational activities to promote assessment literacy learning (i.e., knowledge and skills concerning assessment and testing) among educators and testing officials. Our association, in addition to continuing such activities, is prepared to actively work together toward reforming university entrance examinations while also cooperating with other academic societies and organizations. Moreover, each and every one of our association's members, as a teacher or as someone otherwise concerned with evaluation and testing, stands ready to utilize our respective expertise and knowledge and to cooperate fully in improving evaluation and educational practices in the classroom and in university entrance examinations.

¹ While it goes without saying that score reliability – the stability of scores such that the same result will be obtained, barring any change in proficiency, even if the test is taken several times – should be taken into consideration when making important decisions, it is still necessary to prioritize whether or not this has a favorable impact on students as the foremost judgment criterion.

² The latest research results should include for example those of Watanabe, Koizumi, Iimura, and Takanami (2016), Kunnan (2013), and Fulcher and Davidson (2012).

³ Examples of testing purpose, target proficiency, target test takers, etc. are summarized on the “Eigo 4 ginō shiken jōhō saito [English 4-Skill Test Information Site]” (<http://4skills.jp/qualification/comparison.html>). However, some details remain unpublished with regard to further disclosure that should be sought, as pointed out in section (2) ii. Examples of necessary information include score reliability and standard error of measurement, training and grading methods employed by assessors, and reports on the results of periodic testing analysis.

⁴ For instance, it is desirable that the vocabulary dealt with in English tests as part of the Academic Proficiency Test for University Applicants (Provisional Name) is consistent with the common core vocabulary covered in the MEXT guidelines. Furthermore, due consideration needs to be given to the relationship between the government's official curriculum guidelines and the number of words to be learned for various external tests.

⁵ Even when we consider TOEFL[®] Tests, TOEIC[®] Tests, and the Practical English Proficiency Test – all widely used in Japan – as examples of private-sector qualification and certification examinations, we find that they all differ in terms of what they seek to measure. Respectively, the tests measure “the ability to pursue studies at institutions of higher education using English,” “English-language communication skills for a wide range of situations, from daily life to business,” and “English language skills necessary for social life in English-speaking countries.” In addition, because the methods used for presenting test results (e.g., how scores are calculated, how pass/fail is determined, etc.) also differ, such comparisons and conversions are extremely difficult.

⁶ The level and quality of English proficiency required when entering university, the English proficiency necessary for taking courses and acquiring credits after entering, and the content of English-language education all vary for different universities, courses, and specializations. For that reason, it is desirable that each university engages in systematic and specific studies, and selects and uses tests found to be appropriate. For further details, see discussions of methods for needs analysis in language education and language test creation (e.g., Koyama, 2016; Bachman & Palmer, 2010).

⁷ Major disparities exist in terms of the nature and frequency of exams that can be taken in metropolitan areas in comparison with depopulated areas. It will be necessary to demonstrate specific measures rectifying such disparities, such as setting examination venues and candidate acceptance numbers in proportion to population size.

⁸ Despite its potential for engineering a positive effect of motivating test-takers to work towards improvement in four skills, it may also happen that the time devoted to the teaching and study for the test may increase even to the degree of narrowing the scope and content of instruction and learning, leaving thereby neglecting those skills, which are important but not covered in the test.

⁹ With regard to the educational effects of testing, empirical studies have shown that if test users employ tests appropriately, they will yield a positive effect, while if users do so inappropriately, they will have a negative effect. In cases where multiple tests are used, users will need to be accommodated by providing them with a summary of details including the objective and sample problems for each test.

¹⁰ We might conceive of the addition of courses for improving assessment literacy in teacher training programs or teachers' practical training. The proper evaluation of comprehensive four-skills English training can be expected to lead to effective and balanced feedback-based instruction.

¹¹ The fact that a variety of tests are widely used means that users will require the proficiency to correctly interpret testing results. Teachers, too, will need to master the fundamentals of test preparation and learn to see this as an opportunity to gain knowledge about basic statistics for the processing, analysis, and interpretation of test results.

¹² Examples include overall English proficiency, scores for each skill, descriptions of what learners who achieve such a score should typically be able to accomplish, and learning advice (for details, see Watanabe et al., 2016, section 3.2.10).

¹³ Methods for improving English proficiency are closely related to the learning methods used for achieving high scores on English-language examinations. For instance, summarizing key points by reading and listening to English while taking notes is an effective way to learn to prepare for any test problem. Rather than being put off by such examinations, instructors in particular are called upon to show an attitude of proactive engagement.

References

- Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice*. Oxford University Press.
- Fulcher, G., & Davidson, F. (Eds.). (2012). *The Routledge handbook of language testing*. Routledge.
- Koyama, Y. (2016). Tokutei mokuteki no tameno eigo (ESP) to needs bunseki [English for specific purposes (ESP) and needs analysis: Theory and practice]. In Y. Ishikawa, S. Ishikawa, Y. Shimizu, T. Tabata, K. Cho, & T. Maeda (Eds.), *Language studies and quantitative methods* (pp. 279–298). Tokyo: Kinseido.
- Kunnan, A. (Ed.). (2013). *The companion to language assessment*. John Wiley & Sons.
- Watanabe, Y., Koizumi, R., Iimura, H., & Takanami, S. (Eds.). *JLTA Journal 2016 Vol. 19 supplementary: 20th anniversary special issue*. Chiba: Japan Language Testing Association. (posted on the JLTA website as of June 15, 2017)

Drafting Committee

Yoshinori WATANABE (Chair; Sophia University), Rie KOIZUMI (Juntendo University), Yukie KOYAMA (Nagoya Institute of Technology), Hidetoshi SAITO (Ibaraki University), Yasuyo SAWAKI (Waseda University), Yuko SHIMIZU (Ritsumeikan University), Kazuhiko KATAGIRI (Senshu University), Makoto FUKAZAWA (University of the Ryukyus), Yuichiro YOKOUCHI (Hiroasaki University)