Enlightenments of Standard-based Academic Assessment System--SBAC in America and SA in Hong Kong

Haiyan Zhang[1], Qi Xu[2]

[1, 2] School of Foreign Studies, Nantong University, Jiangsu, China

Abstract: *As standard-based reform sweeps across the world, many countries and areas are establishing the standard-based assessment system of core courses to enhance their education quality. As the products of this educational reform, the Smarter Balanced Assessment Consortium (SBAC) in America and the Student Assessment (SA) in Hong Kong have earned their reputations for innovative design, powerful development team and advanced supporting technology. In this situation, Mainland China also initiates a project to construct China's standards of English and national English assessment system. Therefore, the study of these two representative academic assessment systems will benefit the construction of national English assessment system in China.*

This paper probes into the development of standard-based assessment reform, as well as the characteristics of SBAC and SA. The differences of the two assessments appear in assessment scope, presentation of specific items and test report. Based on those characteristics and differences, some suggestions for the establishment of national English assessment system in China are proposed include constructing academic standards, improving universal accessibility, developing online assessment system and providing post-assessment teaching and learning support.

Keywords: *standard-based reform; assessment system; Student Assessment; Smarter Balanced Assessment Consortium.*

1. INTRODUCTION

English education in China has undergone several reforms since 1990s. One of the most remarkable outcomes is the National English Curriculum (NEC) issued in 2011 which provides teaching guidance and references for English educators. To some extent, NEC improves English education quality in China. However, it fails to exert an impact on English assessment system as English educators and learners expected. In recent years, the National Education Examinations Authority (NEEA) which administers almost all the nationwide educational tests has realized the importance of English academic assessment system to education quality[1]. In 2014, a conference on constructing national foreign language proficiency assessment system was launched in Beijing. During the conference, the goal and working scheme of developing China's Standards of English (CSE) and the foremost aspects of the CSE Working Manual was introduced and clarified. Based on CSE, the concept of establishing national English academic assessment system which combined formative assessments and summative assessments was initiated in 2015.

As NEEA has been preparing for the establishment of national English academic assessment system, the successful experience of areas and countries where assessment systems have been set up is worth exploring and discussing. In the 1980s, many states in America such as Texas, Maryland and California began to implement standard-based reform. Many scholars attach much attention to America on account of its experience. In 2001, Race to the Top (RTTP) allocated 1.75 billion to develop the Smarter Balanced Assessment Consortium (SBAC) and the assessment based on English core standards is an important component in the consortium. And the new assessment systems will send a powerful signal to schools about the meaning of the CCSS and what students should know and be able to do for college readiness and success in work[2]. Hong Kong's English proficiency ranks medium level in the worldwide non-native English countries and areas. In 2005, its Student Assessment (SA), an innovative web-based assessment system, won a Silver Medal at the prestigious Geneva-based international competition le

Salon International Des Inventions. Therefore, the comparison of these two assessment systems will have some implications on the construction of national English assessment system in China.

2. STANDARD-BASED ASSESSMENT SYSTEM

Standard-based assessment is defined as the evaluation of the students' understanding, also known as "outcomes" or "performance". In essence, there is no measurement on which standards are good or perfect though they are believed to be challenging to prepare students to be productive and remain competitiveness in the future. They should stretch educators' beliefs about what students can learn[3]. Standard-based assessment system is equipped with diverse methods including selected response, written responses, physical constructions, etc. to inform how students get along with their study at a specific level. The assessment results are compared to the descriptors of academic standards rather than a ranking compared to norm. In good practice of standard-based assessment, learners have a clear understanding of the explicit standard and become aware of the achievement with respect to these benchmarks. The awareness, known as post-assessment feedback, assists students in pointing to a specific standard of achievement and then strengthens his weakness. Teachers are also fully engaged in the system by giving meaningful feedback to students about the progress with assessment data.

The history of standard-based reform can be dated back to 1956. In Benjamin Bloom's educational work of Taxonomy of Educational Objectives, students' higher-order thinking skills were discussed to replace rote memorized learning[4]. The reform first gained momentum in the 1980s. In 1983, *A Nation at Risk* was issued, which gained much attention to test scores. As federal interest in reforming education lasted in the 21st century, this era is also known as "Goals 2000". Standard-based reform is one of the most prominent features of the current educational landscape[5]. Although the conceptions of SBR have changed and new ideas have developed over time, the key components it constitutes endure. These core elements include the significance of utilizing information to make instructional decisions; a stress on exploiting the standards to facilitate academically challenging instruction; the imperative of similar standards for students with different social and educational backgrounds; most importantly, the measurement of academic outcomes provided by large-scale assessment impels the policy and practice in an education system. Among most conceptions of standard-based reform, there are mutual features on standards. One of them is that academic standards should be aligned with the core elements of educational system to enhance the accomplishment of these standards. As standard-based movement intensifies in the worldwide, advocates of the standards have realized the necessity of consistent standards.

However, the debate on whether the standard-based assessment system is effective or not ever ceases. Supporters emphasize the transparency of the process facilitates the cooperation between teachers and students, which contributes to learning efficiency. Opponents argue that this assessment is infeasible and increases the working load. As the debate swept across the world, countries and areas are gradually engaged in establishing standard-based assessment system. Consequently, the Smarter Balanced Assessment Consortium (SBAC) in America and Student Assessment (SA) in Hong Kong are the products of standard-based reform.

3. CHARACTERISTICS OF SBAC IN U.S.

3.1. Diversification

SBAC not only calls for deeper learning, but also served for the development of 21st century competencies essential for students' future success[6]. Therefore, it caters for students with different educational background. Thus, the SBAC develops different types of assessment to accomplish the purposes and supported by the advanced computer technology to test.

3.2. Integration

There are three assessment modes in SBAC including formative assessment model, interim assessment model and summative assessment model. The purpose of classroom-based formative assessments is to assist teachers in adjusting ongoing teaching plan by actionable feedback. Simultaneously, a digital library in which educators can comment and share is also produced. The interim assessments are designed to monitor the progress of achievement of summative assessments, recognize the learning gap between students, instruct for teaching and affect the decisions in districts. States and districts have the right to decide administration and re-administration times. Research shows that the interim assessments

can make up the deficiency of summative assessments because they are unable to provide educators with all the students' information in school. There are two online types of interim assessments in SBAC, Interim Comprehensive Assessments (ICA) and Interim Assessment Blocks (IAB). Both of the contents of two assessments are aligned to the CCSS. The items in IAB are short and focused and its results generate more elaborate information while the ICA mirrors the year-end assessments. And most of the items in Interim Assessment are scored online except for some constructed response items and performance task. The summative assessments consist of two components, a computer adaptive test and a performance task. It employs multiple items combined with real world problems to assess student attainment and growth efficiently and precisely.

3.3. Computer-Adaption

Unlike traditional testing ways, a key attribute of Smarter Balanced Assessments is its Computer Adaptive Testing (CAT) which essentially individualizes the items that are administered to students based on their prior responses[7]. Throughout the assessment, the computer will adjust the difficulty of questions. For example, if the questions are answered correctly, the system will present harder questions. And if the question is answered incorrectly, easier questions will be presented. Compared with conventional assessments, students' testing pressure is abated because fewer questions need to be done, instantaneously, more accurate results will be produced.

4. CHARACTERISTICS OF STUDENT ASSESSMENT IN HK

The Student Assessment (SA) is a web-based system in HK which assesses students' listening and speaking ability in knowledge, experience and interpersonal dimensions[8]. The main purpose of SA is to inform the teachers of individual student's ability, thereby teachers and students can make use of the resources in the online bank to tackle students' needs and track their gradual improvement. The Assessment content in English language is selected based on curriculum documents from the Curriculum Development Council.

4.1. Complementary

As an assessment for learning, the Student Assessment is developed as a complementary tool for school to effectively conduct internal assessment. Schools can take advantage of SA to supplement other modes of internal assessments, such as classroom performance assessments, tests and examinations. In this way, students' learning progress and needs can be precisely identified and comprehensively understood. Moreover, the integrated application of various assessment means facilitates the identification of learning problems and the supply of instant support and guidance. In terms of assessment standards, the basic competencies and requirements described in English Curriculum are broader than the basic competencies assessed in SA. For example, some speaking and writing skills are difficult to be assessed and scored due to the limitation of computer technology. Therefore, other modes of internal assessments such as class discussions and projects are demanded for teachers to comprehensively understand students' performance and learning progress.

4.2. Flexibility

Teachers can create and administer assessments according to student individual needs and learning progress. There are four types of assessment in the SA item bank, namely, item pre-test, assessment bank, assessment construction and existing assessment. Students of different levels are invited to attend the items in item pre-test. Data about students' performance are collected for item editing purposes. Pre-test assessments with different testing focuses or themes are available in the assessment bank. Teachers can select different pre-test assessments according to students' needs and learning progress. In assessment construction, teachers can create assessments by selecting different basic competencies, which gives teachers great flexibility to create assessments that suit their teaching plans. As an interactive online system, the SA provides an effective and practical assessment tool that informs teachers about their students learning progress and enhances their effectiveness of teaching and learning.

4.3. Web-based Supporting

The Web-based Learning and Teaching Support is developed by the Education Bureau[9]. The resource bank benefits both teachers and students. On the one hand, possible problems are listed at certain level for teachers' reference and suggested follow-up actions or learning and teaching activities and materials

are offered for teachers. Specifically, the possible problems students may encounter will be classified into listening, reading, writing and speaking of different key stages. The teacher just has to choose the possible problem and then the follow-up materials and related module in the textbook will be shown for download. On the other hand, online self-learning equipped with interactive exercises and games are provided for students to address learning difficulties. Students can download the relevant flash to do exercise.

5. DIFFERENCES BETWEEN SBAC AND SA

5.1. Assessment Purposes and Scope

From the outset, SBAC was created and developed for both formative and summative purposes, so SBAC has a more completed structure than SA due to the fact that it was originally designed only for the enhancement of learning and teaching. The SBAC is designed to meet the needs of all students regardless of disability, language, or subgroup. Now, there are more than 7 million students from 17 states, the US Virgin Islands and the Bureau of Indian Education took the summative assessments. Although SA has a longer history than Smarter Balanced Assessments, it is administered only in Hong Kong and inaccessible for the party outside Hong Kong.

5.2. Specific Items

There are several types of items in Smarter Balanced Assessments, namely, selected-response items, constructed-response items, extended response items, performance tasks, technology-enabled items and technology-enhanced items. The diversity of item types meets the different purposes of assessment. The selected-response items are used to assess students' ability of looking for evidence; constructed-response items are used to assess students' ability of employing logical skills to think and deduce; performance tasks are used to assess students' ability of addressing problems in real life by the knowledge and skills they have mastered. The technology-enabled and technology-enhanced items largely rely on the modern technology to present students videos, cartoons and audios and then set items.

The type of items in SA is not as diversified as those in SA because of different assessment purposes. The purpose of SA is to assess students' basic competencies, therefore, there are no items which aim to assess students' relevantly higher ability such as constructed-response items and performance tasks in SA. What worth to be mentioned are the technology-enabled and technology-enhanced items in SA. Apart from presenting videos and audios, these questions are interactive and can stimulate students' interest. For example, in a listening task, students are asked to complete a motorcycle rider's look. Clothes, caps, hair style are presented in a blank, students just need to move the suitable clothes or caps to the rider to complete the task. Although there are limited types of items in SA, the items in SA are more interactive and more likely to arouse second language learners' interest.

5.3. Test Report

The students' testing results of Smarter Balanced Assessments are mainly reported in two ways: Scale Scores and Achievement Levels. The former one is the number that a student scored while the latter is wider proficiency categories students fall into according to their Scale Scores. The Scale Scores can elucidate students' current level of achievement and their progress over time. Besides, the Scale Scores can be used to reveal changes and gaps among schools and districts. According to the Scale Scores, students split up into four different Achievement Levels defined by Achievement Level Descriptors (ALD). The ALD are the specifications describe what knowledge and skills students master at each level. There are three pages in student report of Smarter Balanced Assessments. The first page presents student's Mathematics report including overall score and student's achievement of different claims. Likewise, page 2 reflects student's overall score of English Arts Literacy and achievement in relation to the four claims for EAL including reading, writing, listening and research and inquiry. Page 3 contains explanations of your students' and individual students report. It provides the purpose of report and how it can be used and contains additional details to help you understand how your student's score. What's more, it describes implications of changes to score summaries and how these data are reported and include a legend to understand colors, symbols and icon bands associated with your student report

Compared to the test report of SBAC, SA is simple because the results are not used for accountability. There are two forms of report in SA including student individual report and class report for teacher. In the student individual report, student's performance of each item and overall performance, possible

problems leading to the wrong answer and the basic competency the part tests will be given. The report for teacher focuses on data analysis to help with identifying and teaching. The number of students who answer correctly, the correct percentage and the Basic Competency the item assesses will be shown on this report. Just as the table 2 shows, the average correct percentage of the four items is 58%, the number of students who answer the item 1 correctly is 26, the correct percentage of item 1 is 63%, and the item 1 assesses Basic Competency 1.

Different report ways are employed on account of the assessment purpose. Two report ways and four ALD are employed in the reporting of Smarter Balanced Assessments because it can be used for further study and the assessment targets are aligned to the CCSS. In contrast, the report in SA is simple or because it only assesses students' basic competencies and is a complementary assessment in the whole academic assessment system.

6. IMPLICATIONS FOR ESTABLISHING NATIONAL ENGLISH ASSESSMENT SYSTEM IN CHINA

6.1. Constructing CSE as a Foundation

The Smarter Balanced Assessments and Student Assessment are both developed in line with consistent English academic standards. The two assessment system covers the range of knowledge and skills stated in CCSS and Basic Competency Descriptors respectively. Each test item is designed on the basis of overall assessment targets. To construct a valid and reliable assessment system, the national common standards which describe what students need to know and be able to do should be established in advance. The high quality standards will undoubtedly provide students, teachers and parents with a range of clear expectations that ensure all the skills and knowledge students grasp which will be valued in their life and career wherever they live. In accordance with the successful experience of CCSS, the construction of CSE must involve education chiefs and teachers from different provinces that can offer an overview of the quo status of English education in China. The specific and constructive feedback from teachers is also imperative.

6.2. Emphasizing Universal Accessibility

Compared with SA, one of the advantages Smarter Balanced Assessment possesses is its consideration of students with different educational background or special needs. SBAC employs universal design in the assessment to ensure the accessibility for all the students. There are embedded universal tools such as calculator, digital notepad and spell check serving for students. For the students with disabilities, Braille calculator, talking calculator, text-to-speech setting, speak tool and American Sign Language tool are outfitted in the system. Besides, the test directions were translated into 19 languages to ensure students can understand the requirements of each item. In China, there is not much consideration for the students with disabilities who participate in English assessments. Since NATIONAL ENGLISH ASSESSMENT SYSTEM is constructed to promote educational equity and attempts to be international, the accessibility for every student should be emphasized and practical methods should be found to tackle those problems of visual, auditory and physical access to tests. Only under this circumstance, will the opportunity for every student to demonstrate what they know and can do be ensured.

6.3. Establishing Online Balanced Assessment System

Unlike paper-based assessment, both SA and SBAC apply cutting-edge technology throughout the whole testing process including web-based testing, scoring, reporting, and data analysis. Manifestly, the assessment purpose of NATIONAL ENGLISH ASSESSMENT SYSTEM is similar to SBAC. Therefore, diversified modes of assessment should be adopted in the whole system. The sole summative assessment apparently cannot make the assessment system balanced and monitor students' learning progress. Thus, the complementary interim and formative assessment should be added into the assessment system to make up the deficiency of summative assessment. Instant and beneficial feedback can help educators learn what students can do and what they have learned. As the subjects of NATIONAL ENGLISH ASSESSMENT SYSTEM are mainly non-native English speakers like SA, the application of multimedia such as sound and video in the assessment can stimulate subjects' interest and enjoy the process, especially for the young learners. The possible problems students may have should be listed for students and teachers identify the learning difficulties and then tailor the courses or exercises for individual to address the problems.

As an assessment also serves for summative purposes, scoring online will be more accurate and efficient. The computer adaptive technology can adjust the difficulty level of the item during the testing process. It provides students, teachers and parents report results within weeks, which means that teachers can resort to interim assessment to track student's progress and prepare for the whole year learning plan instantly. And the technology can also enable the sharing of academic assessment information across areas and the online assessment system work efficiently. The support of computer adaptive technology makes the assessment meet different needs of different students because it feels like an assessment tailored for individual.

6.4. Providing Post-assessment Teaching and Learning Support

Both SBAC and SA produce reliable assessment report for students, teachers and parents to inform students' overall score, achievement of specific item, the possible problems or the draft score summary. In SA, web-based teaching and learning support provides assistance for teachers and enables students to learn by themselves online. A complete academic assessment must include formative or interim assessment. Furthermore, after the assessment for learning, the problems or difficulties students have must be addressed. The post-assessment support also needs to be highly valued. Otherwise, the backwash of the assessment may not be beneficial. Furthermore, the online support offers flexibility. Students can conduct self-learning wherever they are if they have a computer. Compared with the paper-based learning, the multimedia effect can largely arouse their learning interest while tackling the problems. Therefore, a web-based post-assessment resource bank supporting for teaching and learning should also be contained in the whole assessment system.

7. CONCLUSION

Consistent academic standards are the foundation of a standard-based academic assessment system whatever it serves for formative or summative purposes. According to the comparison, the information technology should be employed to facilitate the assessment process, scoring and statistics analysis and engender more precise results. Only when the adequate assessing technology, reasonable assessment results and application of assessing tools are ensured can the large-scale assessment produces beneficial wash back effect. In particular, the sharing of academic information online can transcend time and space. Equally important, the balance of assessment system should be ensured. Like SBAC and SA, national English assessment system should consist of various modes of assessment to reduce the high stake of summative assessment. Beyond that, the attribute formative or interim assessments possess can assist educators in checking learners' progress and assist learners in self-assessing and self-learning. As Ravitch notes, the consequences linked to the assessments will provide the "motivation" for teachers and students to work hard to achieve national standards[10]. Additionally, the work after assessment cannot be ignored.

REFERENCES

[1] Liu,Q. (2009). *The National Education Examinations Authority and its English Language Tests.* New York: Routledge.

[2] Herman, J.L & Linn, R.L (2013). *Assessing Deeper Learning: The Status of Smarter Balanced and PARCC Assessment Consortia.* (CRESST Report 823). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

[3] Resnick, L.B., & Resnick,D. P. (1992). *Assessing the Thinking Curriculum: New Ttools for Educational Reform.* Boston: Kluwer.

[4] Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956).*Taxonomy of Educational Objectives*, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.

[5] Hamilton, L., Stecher, B., & Kun, Y. (2008). *Standard-based Reform in the United States. History, Research, and Future Directions.* Washington, D.C.: Center on Education Policy.

[6] Pellegrino, J. W., & Hilton, M.L. (Eds) (2012). *Education for Life and Work: Developing Transferable Knowledge and skills in the 21ˢᵗ Century.* Washington, DC: The National Academies Press.

[7] Herman, J. L., Matrundola, D., & Wang, J. (2015). *On the Road to Assessing Deeper Learning: What Direction do Test Blueprints Provide?* (CRESST Report 849). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

[8] David, C. (2014). *English Language Education and Assessment*. Hong Kong: Springer.

[9] Education Commission. (2000). *Education Blueprint for the 21st Century: Learning for Life, Learning through Life - Reform Proposals for Education System in Hong Kong*. Hong Kong: Government Printer.

[10] Ravitch,D. (1995). *National Standards in American Education: A Citizen's Guide*. Washington, DC: Brookings.

AUTHORS' BIOGRAPHY

Haiyan Zhang, Ed. D, is an associate professor at Foreign Language College in Nantong University, China. Her research interests are early childhood education and English teaching pedagogy.

Qi Xu, is an undergraduate at Foreign Language College in Nantong University, China. Her major is English Education.