# e-Assessment System as a Positive Tool in the Mastery of Putonghua

Kitty Siu, Y.T. Woo, Kat Leung, Kenneth Wong, Reggie Kwan, and Philip Tsang

Caritas Bianchi College of Careers, Hong Kong
{ksiu,ytwoo,kleung,kwong,rkwan,ptsang}@cihe.edu.hk

**Abstract.** Assessment in whatever format has not been favourably received by students at all levels since the beginning of time. There is a misconception that assessment is always regarded as a means to make judgments about student performance and grade allocation. It is often thought to have little, if anything, to do with learning. This paper introduces the design of a self-regulating and diagnostic e-Assessment system in learning Putonghua as a second language. The beauty of the newly developed system is its embrace of the spirit of "assessment for learning" and its demonstration of how well-planned web-based multiple choice test items can be used to facilitate self-directed learning so as to enable students to learn Putonghua Pronunciation independently outside the classroom. It is also an example of the Computerized Adaptive Tests (CATs), which is based on the Item Response Theory (IRT) model capable of adapting to an individual student learning needs and learning pace. The key distinguished features of such an e-Assessment system are specified and the paper concludes by providing a list of potential uses of the system in the context of language teaching and learning.

**Keywords:** assessment for learning, Putonghua learning, computerized adaptive testing.

## 1 Introduction

The application of Assessment for Learning (AFL) to the education environment is not a new trend [1] [2] [3] [4]. AFL has, in fact, been increasingly adopted during the past decade by various educational institutions as a positive tool to enhance student learning and boost the motivation to learn. Many efforts have been made to explore the effect of AFL on student learning and positive results were reported in face-to-face classroom settings [5] [6] [7]. Though [8] [9] [10] put this to use empirically, research on its learning effectiveness in the online teaching and learning environment is still in its infancy and the current pedagogic evidence base about this assessment tool in language education is scarce.

To embrace the trend of AFL and make use of the benefits of information technology to Putonghua learning in full, the Caritas Bianchi College of Careers (CBCC) has actively participated in the development of various e-learning platforms. Since its introduction in 2001 the e-assessment system for Putonghua using the concept of AFL demonstrated satisfactory learner performance. The aim of this paper

R. Kwan et al. (Eds.): ICT 2011, CCIS 177, pp. 351–365, 2011.

is to give an account of how learners can benefit from the system. In the first part, the background in developing the system will be discussed, followed by a brief introduction of its system design and architecture. The second part provides scenarios in the application of the system with conclusions.

## 2 Literature Review

The types of "assessment" have been widely discussed in the western literature and three models of "assessment" have been identified in the past decades. These are: (i) "assessment of learning", (ii) "assessment as learning" and (iii) "assessment for learning" respectively. The different prepositions used in the above terms between the nouns "assessment" and "learning" explicitly highlight the unlike principles behind these three models and the different purposes for which the assessment is designed.

Assessment in most classrooms, worldwide, is dominated by elements of "Assessment of learning" (AOL). AOL is an event designed merely for the purpose of evaluating student performance, assigning grades, sorting and ranking students. It is also used to collect evidence to evaluate programmes, curricula effectiveness, and school achievement. AOL normally happens at the end of a school term or a school year in order to inform teachers, students and their parents what a student has learned and how much he/she has achieved during the assessment period. In the classroom where AOL is practiced, teachers are the sole assessors, taking full control of how, what, and why students will or will not learned or be assessed, Students play their roles as passive participants both in the learning and assessment process. The spirit of AOL is clearly reflected in the writing of Stiggins and his research team [11] in which they viewed AOL as "a measurer of the impact of instructional interventions". Such an interpretation is similar to that of Chappuis & Stiggins [12]. They described AOL as "an index of school success". In the AOL model, assessment, to a large extent, is equated with the meaning of "measurement" which can be viewed as a process of estimating the amount, extent and level of student learning [13].

As opposed to "Assessment of learning", the practice of "assessment as learning" (AAL) is always completed and carried out by students instead of teachers with the emphasis on the involvement of learners in assessment in the midst of the learning process. Dann's definition of "assessment of learning" is considered as one of the best illustrations of its kind [14]. According to his definition, AAL "is not merely an adjunct to teaching and learning but offers a process through which pupil involvement in assessment can feature as part of learning" (p.153). The concept of AAL also reminds us of the idea of "meta-cognitive process advocated by Brown (1981) focusing on student ownership and responsibility for moving his or her thinking forward and teachers act as facilitator to provide support for students during the learning process [15].

The application of Assessment for learning (AFL) into the day-to-day classrooms has been one of the main concerns in the educational arena, both internationally and locally. The main difference between AFL and the other two assessment models largely lies in its primary purpose of use. The primary purpose of AFL is to use assessment as an instructional tool to enhance and improve student learning. Unlike

AOL, the AFL model is not just the index of change but apparently, the change itself [11]. It occurs throughout the learning process focusing on promoting learning rather than ranking students; boosting student confidence and motivation to learn with special reference to progress rather than enhancing student achievement and performance[12] [17]; getting students involved in the assessment process rather than assigning teachers to play a dominant role; and providing students with descriptive and corrective feedback by emphasizing strengths, identifying challenges and pointing to next steps rather than giving them grades or scores.

Mixed opinions were found in the available literature when it comes to the question whether the practice of AFL equates with that of "formative assessment". On the one hand, AFL and formative assessment are interchangeable in some publications [13] [16]. Similarly, AOL is also called "summative assessment" in the extant literature. On the other hand, some researchers with the opposing view believe that AFL and formative assessment are not the same. As suggested by Stiggins [17] and Stiggin et al. [11], an easy way to distinguish AFL from formative assessment is to ask the following questions: First, "what kind of feedback are teachers using in assessment? Is it descriptive in nature? Second, "Do students get involved in assessment?" If both replies are yes, then the assessment can be called AFL. If not, then the assessment practice is more likely to fall into the category of formative assessment. In short, AFL has a broader meaning than formative assessment [11].

## 2 Background

Since the handover of Hong Kong's sovereignty to the Mainland China, the status of Putonghua has been rising in Hong Kong as it is not only the national language of the Mainland China, but also the national lingua franca among speakers of various dialects. Together with English and Cantonese, Putonghua has already grown to be another important powerful and essential language in local schools and colleges, where regular Putonghua lessons are offered. It is even used as the medium of teaching in place of Cantonese, which has dominated the Chinese language lessons for many years.

Facing the soaring number of students and the congenital limitations of Putonghua learning within fixed bounds of a physical classroom including insufficient time to interact with learners which may further lead to an inability to understand their needs, information technology has offered a good solution with more powerful software and applications, along with mobile devices such as tablet computers, personal digital assistants (PDAs) and laptops. The Internet, being a powerful and affluent resource of learning materials and a convenient means of communication, can be effectively utilised as a desirable learning platform to overcome the limitations with the support of the suitable software and mobile devices so that learning can occur regardless of time and place and according to individual progress.

The idea of developing a self-regulating and diagnostic e-Assessment learning system stems from our interaction with sub-degree students in which severe problems in their self-learning abilities in the mastery of Putonghua after school were noticed. It is apparent that they have difficulty in locating and choosing the appropriate learning materials at a level with which they are comfortable within a self-paced

environment. They are always weak in the identification of their own Putonghua levels, not to mention they are not knowledgeable and confident enough to design an individualised study plan to meet their diverse learning needs and learning pace. This situation can always be found in the learning of tone (聲調) as learners always distinguish the four tones in Putonghua according to those in Cantonese, which have nine tones of inexact equivalence. What is even worse is that most of the self-access Putonghua learning resources in the market are not well-classified according to the language abilities of the learners and the difficulty level of the questions. All the above-mentioned can be seen as barriers to the learning of Putonghua independently outside the classroom or for consolidation of work done inside the classroom.

In light of this, an online assessment learning system (e-assessment system) was developed to enhance the self-learning abilities of learners in the acquisition of Putonghua through a considerable amount of listening exercises regarding tone in the form of multiple choice questions, based on the idea of Computerized Adaptive Testing and Item Response Theory. The system also aims to help individual learners diversify and customize various approaches to learning "how to learn" languages through self-directed, self-controlled and, to some extent, individually packaged "assessment" opportunities. The purpose of doing this is to reshape their ideas concerning "assessment". Assessment in whatever format has long been unwelcomed by learners at all levels as there is a misconception that assessment is merely a means to make judgments about their performance and through which grades are allocated. These elements have little to do with learning. Therefore, the design of the e-assessment system is mainly focused on how learners can understand their weakness through feedback so as to seek apt and specific ways to improve their proficiency in Putonghua, rather than producing grades for teachers. The feedback can also be useful for Putonghua teachers, in gauging individual student ability and therefore providing them with a basis to re-visit topics in the face-to-face classroom and to thereby develop follow-up prescriptive instruction.

## 3 System Design and Architecture

### Development of Data Bank

The main design concept of the data bank is to develop an e-Assessment system that accommodates the diverse learning needs of sub-degree students. The system will not only help solve the problem of teaching a class with considerable variation in language ability, it will also encourage students to improve their language skills on their own, as they would not feel the pressure of learning next to the high achievers or within the fixed time schedule of a lesson.

The system will be designed as multiple choice assessments for Putonghua with item response theory implemented inside. Item bank number for Putonghua will have about 800 sets of questions. The multiple-choice questions are categorized under the curriculum domain, which denotes the "breadth" of the e-Assessment system. They are then further divided into the skill domain, which denotes the "depth" of the system. Four choices will be provided for each question, one of which is the correct

answer. Specifically, there will also be an audio file attached to each of the Putonghua questions (Figure 5).

The item database and the codes of the system will be designed using a 3PL parameter model. However, in actual implementation, a 1PL parameter model will be used ("discrimination" will be set to 1 and "pseudo chance parameter" will be set to 0) i.e. only "difficulty" will be demonstrated. This will leave space for further extensibility when the 3PL model is actually required.

### Development of the Web Interface

When the user enters the system, the login page (Figure 1) will be shown first and if the validation fails, a popup will be displayed. If the validation is successful, the main panel (Figure 2) will be displayed. The system panel will be in Chinese for the subject of Putonghua. An Introduction message will be displayed to describe each of the assessment sets (Figure 3). Users will be able to continue with a previous ability level or begin with a default level. Explanations will be provided for the users after carrying out the assessment.

The result records will include the skill areas needed to be improved upon, the total number of questions answered, the total number of correct answers, ability values, a conclusion on a user's overall ability quotient and the related skill descriptions. Users will be able to access the last five assessment result records they previously undertook.


**Fig. 1.** Login Page


**Fig. 2.** Main Page

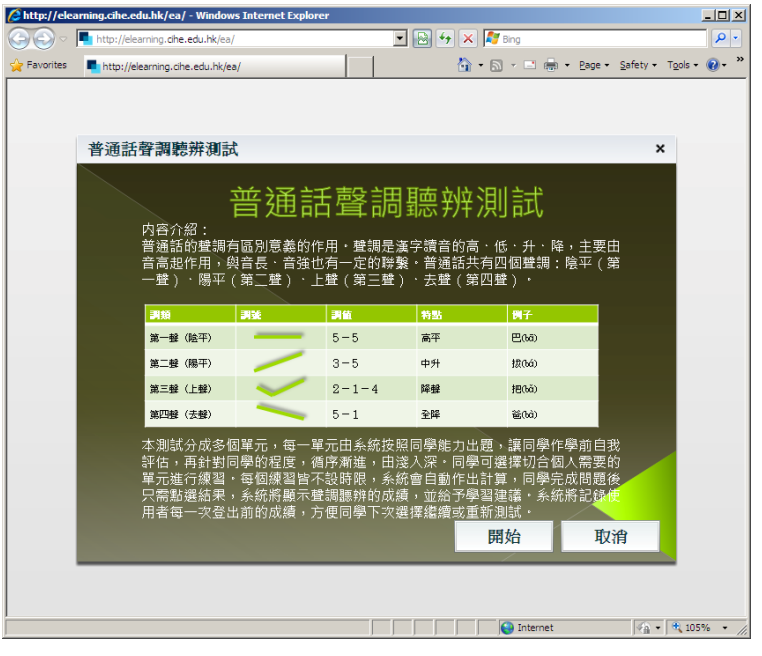


**Fig. 3.** Putonghua Introduction Page

### Curriculum and Skill Design

Based on the common mistakes made by local students, the curriculum of the Putonghua e-Assessment System covers the main areas in Phonetics (Curriculum 1 – C1). Students should have a strong grasp of the concepts in those areas in order to master the speaking skill.

There are various skills involved in the curriculum. Phonetics (C1) has been selected to demonstrate the design of the questions in *Figure 4*. The three main skills involved in C1 are: Tone (聲調) (Skill 1 – S1), Vowels (韻母) (S2) and Consonants (聲母) (S3). To use S1 as an example, different levels of questions with audio files are set to test a user's ability to distinguish between the four tones of Putonghua individually as well as in combinations, such as two characters of the same tone put together (*Figure 5*), the first and second tones, the first and third tones, the first and fourth tones and so on.

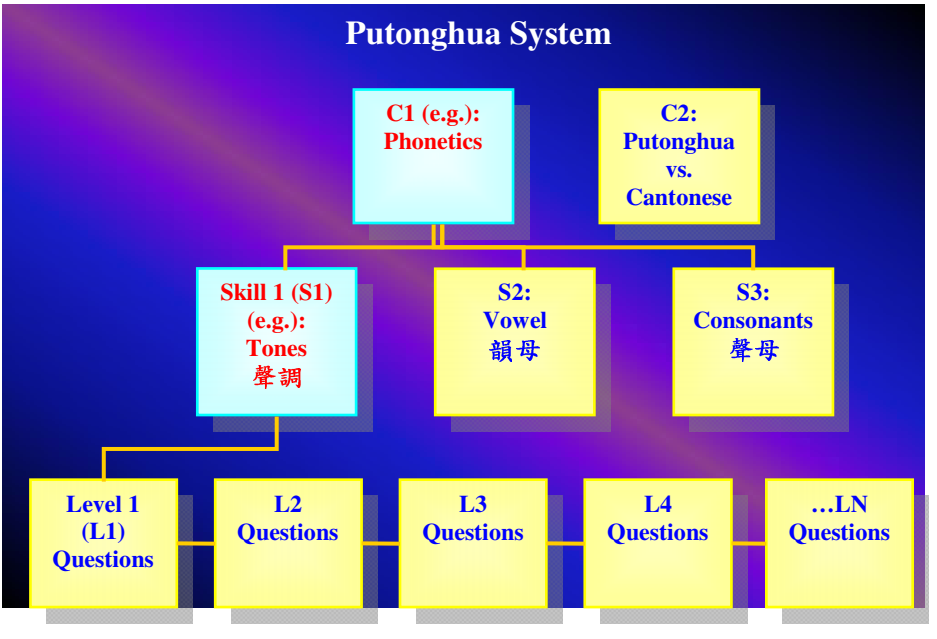


**Fig. 4.** A flowchart outlining the data bank design of the e-Assessment System - Putonghua

### System Rationale

All users will first be treated as equals regardless of the level of their studies. A random question will be generated from the data bank as a reference point of the user's language standard. Every time a user logs in, he/she can choose to begin the assessment at a new level or continue with the previous level if there is history saved. There is no time limit set for each assessment. If the user exits before the end of an assessment, he/she can choose to finish the remaining questions of the previous test or


**Fig. 5.** A question that tests a user's ability to identify the first tone in Putonghua

commence new ones next time. The system will calculate the ability of the users when they have successfully submitted an answer. The difficulty of the next question will be based on the users' answer history. Once the standard error of successive answers calculated is less than a defined value (e.g. 0.01), the system will prompt the user that the assessment has been finished successfully.

Wrong @ Q392

| user_id | quiz_id | question_id | time_spent | answer_history | result_history | difficulty | ability | test_information_value | standard_error | status |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 328 | 30 | C | 1 | -2.20925844 | 3.5 | 0.000176 | 75.361939 | 1 |
| 2 | 2 | 417 | 19 | D | 0 | 3.5 | 0.65 | 0.044423 | 4.744537 | 1 |
| 2 | 2 | 397 | 22 | B | 1 | 0.673924085 | 2.1 | 0.441918 | 1.504282 | 1 |
| 2 | 2 | 430 | 24 | A | 0 | 2.024107789 | 1.3 | 1.129543 | 0.940911 | 1 |
| 2 | 2 | 457 | 10 | C | 0 | 1.294398146 | 0.85 | 1.6864 | 0.770051 | 1 |
| 2 | 2 | 404 | 10 | B | 0 | 0.841754858 | 0.45 | 2.014605 | 0.704539 | 1 |
| 2 | 2 | 392 | 17 | C | 0 | 0.431195234 | 0.1 | 2.201916 | 0.673906 | 1 |
| 2 | 2 | 381 | 43 | C | 0 | 0.121460645 | -0.2 | 2.367092 | 0.649969 | 1 |

Correct @ Q392

| user_id | quiz_id | question_id | time_spent | answer_history | result_history | difficulty | ability | test_information_value | standard_error | status |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 328 | 30 | C | 1 | -2.20925844 | 3.5 | 0.000176 | 75.361939 | 1 |
| 2 | 2 | 417 | 19 | D | 0 | 3.5 | 0.65 | 0.044423 | 4.744537 | 1 |
| 2 | 2 | 397 | 22 | B | 1 | 0.673924085 | 2.1 | 0.441918 | 1.504282 | 1 |
| 2 | 2 | 430 | 24 | C | 0 | 2.024107789 | 1.3 | 1.129543 | 0.940911 | 1 |
| 2 | 2 | 457 | 10 | C | 0 | 1.294398146 | 0.85 | 1.6864 | 0.770051 | 1 |
| 2 | 2 | 404 | 10 | B | 0 | 0.841754858 | 0.45 | 2.014605 | 0.704539 | 1 |
| 2 | 2 | 392 | 20 | D | 1 | 0.431195234 | 0.75 | 3.00877 | 0.576508 | 1 |
| 2 | 2 | 399 | 113 | A | 1 | 0.743514485 | 0.95 | 3.75908 | 0.515774 | 1 |

The above analytical report shows the user's scores and ability levels. The first table indicates a wrong answer to Question 392 with the next question marginally lower in the level of difficulty, and the second table indicates a right answer to Question 392 with the next question marginally higher in the level of difficulty.


**Fig. 6.**


**Fig. 7.**

If the user answers correctly (Figure 6), the next question (Figure 7) will be generated at a higher skill level.


**Fig. 8.**


**Fig. 9.**

3If the user answers wrongly (Figure 8), the next question (Figure 9) will be generated at a lower skill level.

### Assessment Reports with Explanations

After carrying out an assessment, users can navigate to the results window to review their assessment record. Analytical reports on the user's strengths and weaknesses are available and information includes skills that need to be improved upon and related suggestions on further reading will be displayed. Users can click on the advice web link for further information.


**Fig. 10.** Assessment results with the total number of questions answered, the total number of correct answers, ability values and a conclusion on a user's overall ability quotient


**Fig. 11.** A screen capture showing the assessment report and suggestions on further reading

## 4 Scenarios and Potential Use of the e-Assessment System

The **e-Assessment system** is specially designed to provide a supplementary aid to enhance the teaching and learning of Putonghua Phonetics to complement the existing formal assessments during the academic semester. The system can serve the following purposes:

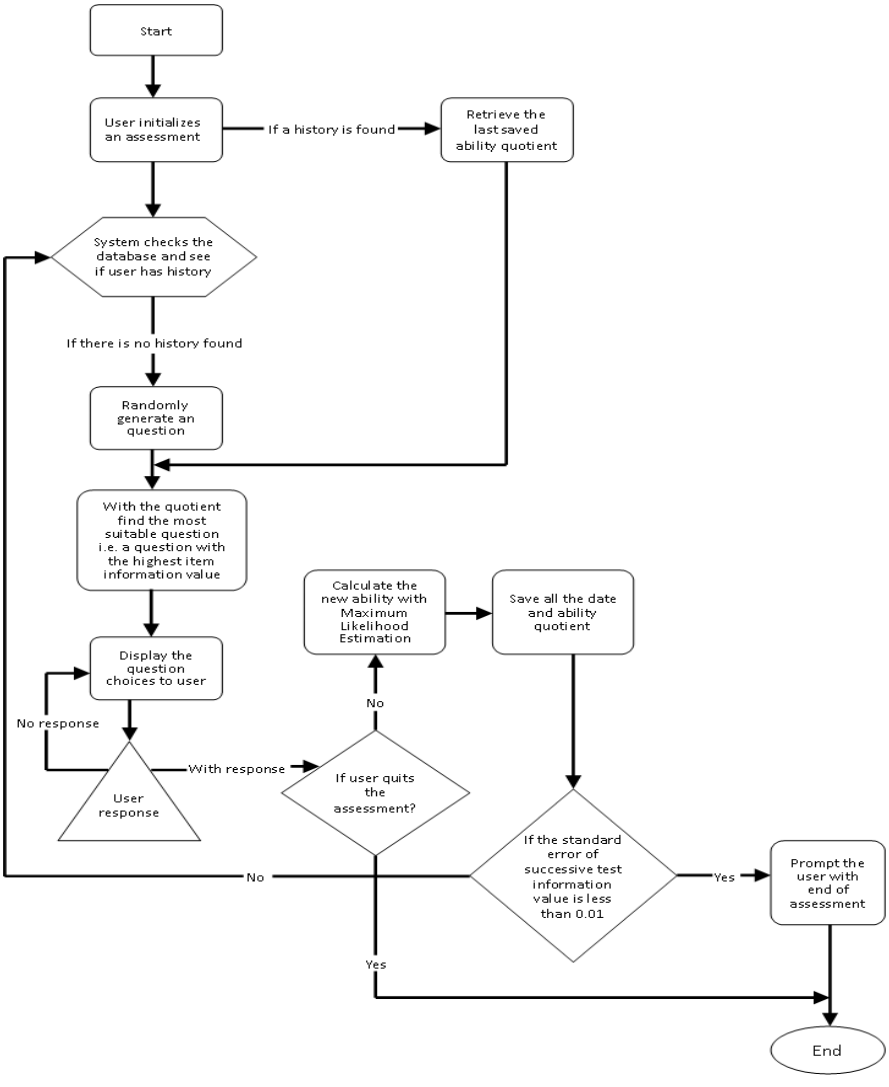


**Fig. 12.** System Flow Chart

## 1. It can be used to boost student motivation to learn

With the e-Assessment system, students become more motivated to learn Putonghua pronunciation and are able to regulate their own learning pace in a step-by-step manner. The e-Assessment system is self-regulating in nature and is capable of adapting to each user's language ability and skill levels. This is different from traditional web-based multiple choice tests with identical question sequence and time limit. The system generates questions according to the user's performance with some random factors to avoid an identical question sequence. The number of questions, the type of questions

and the respective levels of difficulty, will thus vary across different people as well as across time for the same person. As a result, it is difficult for a user to perform well just by memorizing the test answers. The user will only be able to improve his/her subsequent performance through "genuine competence". Every attempt leads to a brand-new learning experience and presents a real challenge to the user.

**2. It can be used to facilitate self-directed learning**

The conventional face-to-face language assessment and classroom learning is a time consuming process. The e-Assessment system is able to offer a diagnostic function to provide an immediate feedback for users which enables them to recognize their strengths and weaknesses in Putonghua phonetics. Detailed study guides with full explanations, follow-up exercises and recommendations for self-improvement, and email addresses of Putonghua teachers for individual tutorial advice are available in the system. Analytical reports on overall performance on the tests will also be generated to give a clear indication of mistakes and deficiencies. Users can plan personalized study for self-improvement by directing their future learning effort to the problem areas. In particular, users are able to select the appropriate learning materials from the library or the Language Centre at a level that is matched with their progress without the need for instructor guidance. Users can repeat this self-learning process in the following cycle: e-Assessment, Diagnosis of weak areas, Improvement, Reassessment.

**3. It can be used to collect evidence to revise classroom instruction**

The performance report generated by the e-Assessment system is also very useful to the instructors. The report contains a detailed record of each student's performance/progress on the online tests, thus providing valuable information for the instructors and helping them decide which topics to be re-visited in subsequent face-to-face classes. Follow-up prescriptive instructions can be developed accordingly. For example, learners will be referred to enroll in a short course in "Putonghua Toning" offered by the Language Centre of the College if they are found to have difficulty in differentiating between the first and the fourth tones in Putonghua on the e-Assessment system.

**4. It can be used to form the basis for curriculum design and development**

The e-Assessment system is able to identify the weak patterns among students with different programmes and at different years of study. By comparing the test performance among students according to their levels of study, types of programme enrolled (Higher Diploma, Associate Degree, Top-up Degree, Bachelor Degree) and nature of programme enrolled (Design, Business, Hospitality Management, Language, Social Work), specific patterns of strengths and weaknesses in terms of programme, course and level of study can be identified eventually through repeated testing. These empirical findings can help the programme/course developers in future curriculum design and development.

## 5 Conclusion

This paper stems from the 24-month project of developing an e-Assessment system in English and Putonghua learning funded by Hong Kong Education Bureau under the Quality Enhancement Grant Scheme, and reports the progress of the project's first year with special reference to Putonghua learning. Since the test item banks and web interface have been fully-developed, effort will be made to test the validity and reliability of the test items in the remaining project period. On top of this, during the coming year, questionnaire surveys and focus-group interviews will be conducted among the students and instructors in order to evaluate the learning effectiveness of the e-assessment system.

## Acknowledgements

This project was funded by the Education Bureau of Hong Kong SAR Government under the Quality Enhancement Grant Scheme.

## References

1. Black, P., Harrison, C., Lee, C.: Working inside the black box: assessment for learning in the classroom. Open University Press, Maidenhead (2003)
2. Black, P.: Assessment for learning: putting it into practice. Open University Press, Maidenhead (2003)
3. Chappuis, S., Stiggins, S., Arter, J., Chappuis, J.: Assessment for learning: an action guide for school leaders. Assessment Training Institute, USA (2004)
4. William, P.: Assessment for learning: why, what and how. University of London, UK (2009)
5. MacLellan, E.: Assessment for learning; the difference perceptions of tutors and students. Assessment & Evaluation in Higher Education 26(4), 307–318 (2001)
6. William, D., Lee, C., Harrison, C., Black, P.: Teachers developing assessment for learning; impact on student achievement. Assessment in Education: Principles, Policy and Practice 11(1), 49–65 (2004)
7. Marshall, B., Drummond, M.: How teachers engage with assessment for learning: lesson from the classroom. Research Papers in Education 21(2), 133–149 (2006)
8. Kwan, R., Wong, K., Yu, C.-w., Tsang, P., Leung, K.: Building an Assessment Learning System on the Web. In: Fong, J., Kwan, R., Wang, F.L. (eds.) ICHL 2008. LNCS, vol. 5169, pp. 91–102. Springer, Heidelberg (2008)
9. Kwan, R., Wong, K.: e-Assessment Learning as a Learning Tool. In: Handbook of Research on Hybrid Learning Models: Tools, Technologies and Applications, pp. 105–114. IGI Publishing (2009)
10. Wong, K., Leung, K., Kwan, R., Tsang, P.: E-Leaning: Developing a Simple Web-based Intelligent Tutoring System Using Cognitive Diagnostic Assessment and Adaptive Testing Technology. In: Tsang, P., et al. (eds.) Hybrid Learning, pp. 23–34. Springer, New York (2010)
11. Stiggins, R.J., Arter, J.A., Chappuis, J., Chappuis, S.: Classroom assessment for student learning: Doing it right using it well (retrieved 30 January 2011 from Google Scholar)

12. Chappuis, S., Stiggins, R.J.: Classroom Assessment for learning. Educational Leadership 60(1), 40–43 (2002)
13. Hargreaves, E.: Assessment for learning? Thinking outside the "black" box. Cambridge Journal of Education 35(2), 213–224 (2007)
14. Dann, R.: Promoting Assessment as learning: Improving the learning process. Routledge Famler, London (2002)
15. Brown, A.L.: Metacogniton: the development of selective attention strategies for learning from texts. In: Kami, M. (ed.) Directions in Reading: Research and Instruction. The national Reading Conference, Washington, DC (1981)
16. Hoover, K.: Open door to learning with keys to assessment. Innovations Perspectives (May 10, 2010)
17. Stiggins, R.J.: Assessment Crisis: the absence of assessment for learning. Phi Delta Kappan 83(10), 758–765 (2002)