BRITISH COUNCIL

ENGLISH LANGUAGE
ASSESSMENT RESEARCH GROUP

ASSESSMENT FOR LEARNING:
Using Aptis as a diagnostic assessment tool
in the EFL writing classroom

AR-G/2017/4

Jinlan Tang
Beijing Foreign Studies University, China

ABSTRACT

The study investigated the efficacy of using Aptis as a diagnostic assessment tool in the classroom. A mixed-methods research approach, combining both quantitative and qualitative studies, was undertaken with a group of 95 EFL learners and two teachers in two English language training programs (entitled Program A and Program B for the research study) at a university in Beijing. Quasi-experimental research in the form of pre- and post-tests with Aptis was conducted to evaluate the learning gain and the effectiveness of the teaching experiment. Pre- and post-test student questionnaires were designed and administered during the experiment. Teachers' views on the diagnostic assessment feedback and the efficacy of the teaching experiment were gauged through an open questionnaire and interim reports.

The pre- and post-test comparison indicated positive learning gains for the experimental group in Program A. Teachers reported more guided and targeted teaching based on the diagnostic feedback and improved efficiency, and students maintained that tests were authentic, interesting, and motivating. Nevertheless, students also voiced the concern that the tests were not related to their immediate purpose of preparing for IELTS.

The research indicated that, if applied properly, Aptis was a useful instrument for both placement and diagnostic assessment (if supported with diagnostic feedback).

The study offers insights into how Aptis can be used effectively as a diagnostic assessment tool in the EFL classroom and on how the diagnostic assessment functions of Aptis can be improved to serve instructional purposes better.

Author

Jinlan Tang is the Deputy Dean and Professor of English in the School of Online and Continuing Education, Beijing Foreign Studies University. Her research has covered the areas of language assessment, tutor feedback, EFL teaching and learning in the e-learning environment. She has published over 30 research papers in academic journals, two books, and served as the principal investigator on a number of research projects. She has presented papers in major national and international conferences on EFL teaching and learning. She is also the Secretary-General of the China Computer-Assisted Language Learning Association (2016–2020) (ChinaCALL, www.chinacall.org.cn), and a researcher of the China Language Assessment Centre, Beijing Foreign Studies University. She holds a PhD from the University of Nottingham in the area of EFL assessment in the e-learning environment.

Acknowledgements

CONTENTS

Tables

INTRODUCTION

Since the implementation of its reform and open-door policy, China has witnessed an increasing number of students pursuing their education abroad. Nevertheless, a considerable number of them still need to improve their English language competence. The 2012 IELTS global report revealed that Chinese students' overall average performance in IELTS was Band 5.6, which was far from satisfactory considering the amount of time and effort that is expended on English language instruction. Furthermore, English has also posed challenges for those Chinese students who use it to function in the native-speaking environment (Tang, 2003).

On the other hand, the use of technology has been required at all educational levels in China. The China National Educational Reform and Development Plan (2010–2020) stressed efforts to accelerate technology use with the purpose of enhancing teacher application of technology for effective teaching and student utilisation of technology for independent learning. Against this background, using technology is considered a viable means to enhance English language instruction.

In recent years, various English language training programs have been launched to assist trainees with their English and prepare them for studying in an English-speaking environment. Students usually study language courses and then take standardised English language tests, such as a national English test or IELTS or TOEFL at the end of the program. Given the explicit goal of students' taking the standardised tests within a limited amount of time, arguably one of the main problems faced by these training programs is the lack of a proper standardised assessment tool to diagnose students' problems during the course of learning and teaching, which consequently might result in ineffective teaching and support.

Aptis is a flexible and accessible online English assessment tool developed and launched by the British Council in 2012 (http://www.britishcouncil.pk/exam/aptis). Being an English test for adults (16+), Aptis tests English levels from A1 to C on the Common European Framework of Reference for Languages (CEFR). Comprising five components (with a core component of grammar and vocabulary and four other components of listening, speaking, reading and writing), Aptis can assess all four English language skills – speaking, listening, reading and writing. It is flexible in that the core language component (grammar and vocabulary) can be combined with the skills of the user's choice, enhancing accurate and relevant testing. The claimed uses of Aptis, according to the developers, are to benchmark candidates, to identify training needs, to filter employees for interview or promotion, to diagnose strengths and weaknesses of people seeking employment, and to evaluate language development projects. Reading and listening components are marked automatically by the online platform, while a live examiner marks the speaking and writing components. An Aptis candidate will receive a score on a numerical scale (0 – 50) for the grammar and vocabulary section, and a score on a numerical scale (0 – 50) and CEFR level (A1–C) for each skill they take, which will make up the candidate's profile of language proficiency.

Within this background, the current project aims to investigate the use of Aptis as a diagnostic assessment tool to diagnose students' strengths and weakness before an English language training program, with corresponding teaching interventions in between, and the use of Aptis again at the end of the program to evaluate students' learning gains and to examine whether students have made progress in relevant areas.

1. LITERATURE REVIEW

1.1 Previous research

Assessment, if used formatively in the classroom, can enhance and improve learning. An important part of formative assessment is feedback, which has been considered the life blood of learning (Rowntree, 1987). It may be argued that the provision of instant and individualised feedback is impossible without the use of computers. It is noted that online diagnostic assessment can offer timely, individualised and diagnostic feedback to learners during the learning process; in the meanwhile, it can also yield useful and instructive feedback to teachers.

Due to the challenges of developing a diagnostic assessment system, there is a paucity of such systems available. One well-known online diagnostic assessment system (Dialang) is currently used for self-assessment purposes, and the issue of its application in the classroom has not been researched.

Nevertheless, as previously mentioned, lack of proper assessment to diagnose students' strengths and weaknesses before and during the English language training program has hindered the efficacy of the English language instruction, which is less guided and targeted.

As yet, no research has been conducted on the impact of diagnostic assessment or on the classroom use of diagnostic assessment, and empirical research is urgently needed (Alderson, Huhta, Nieminen, Ullakonoja & Haapakangas, 2013). Hence, it is hypothesised that Aptis, designed with diagnostic functions/purposes in mind, can enhance English language instruction by identifying strengths and weaknesses, and by offering timely helpful feedback to inform and guide teaching and learning in the classroom in China.

Previous applied research on Aptis seems so far to have mainly focused on gauging user feedback from piloting the test with candidates of different backgrounds (O'Sullivan 2015a), investigating the potential value of introducing the test into a specific sociocultural context (Douglas & Mbali 2015), examining the cognitive processing of test-takers while completing Aptis reading tasks (Brunfaut & McCray 2015), and exploring the relationship between test-takers' L1, their listening proficiency and their performance in the speaking tests (Jaiyote 2015), etc. It might be argued that the research now needs to progress further to perceiving its place within broader educational processes.

The current study, investigating the use of Aptis as a diagnostic tool in the China EFL classroom to identify students' strengths and weaknesses to improve teaching and learning, might bear greater significance for educational advancement in that it has gone beyond just measuring learning, but more importantly, using the assessment result to improve instruction and learning.

1.2 Theoretical framework

The theoretical basis of the Aptis test system has been elaborated in detail in O' Sullivan (2015b). Additionally, the following two theories, which particularly informed the research design of this study, are briefly reviewed.

1.2.1 Diagnostic assessment

Diagnostic assessment is a type of assessment focusing on learners' strengths and weaknesses via detailed feedback which can be acted upon, e.g. remedial instruction or intervention. Diagnostic assessment is a part of the classroom teaching, occurring whenever possible and feeding back into the curriculum (Alderson, Brunfaut & Harding 2014).

It has been argued that to understand how diagnostic assessment can best contribute to teaching, we need more classroom-based assessment research to investigate the processes of diagnosis (see, for example, Doe 2011; Fox & Hartwick 2011) and the interface between diagnosis and treatment (Alderson et al 2014).

1.2.2 Impact theory

Impact research concentrates on evaluating the effects of interventions (e.g. experimental teaching and tests) on the participants in various ways (Hawkey 2000). It has been argued that, given the formative nature of education, impact studies measure both outcomes (i.e. test results or performance gains) and processes (i.e. the teaching and learning approaches).

It has been pointed out that the efficacy of various treatments or interventions based on diagnostic assessment should be investigated (Alderson et al 2014). Accordingly, in the current project, we intend to examine the impact of an Aptis-integrated teaching experiment on the learning outcomes and the teaching and learning process in an English training program in China.

2. RESEARCH METHODOLOGY

2.1 Research questions

The study aims to address the following research questions.

1. What are students' attitudes toward and experiences with the Aptis test and its use as a diagnostic tool?

2. How do teachers view the efficacy of the diagnostic use of the Aptis test in the classroom?

3. How does the teaching experiment involving the use of the Aptis test as a diagnostic assessment tool impact students' performance?

2.2 Research design

The study was undertaken at a university in China, within two English training programs (entitled Program A and B for the research study) preparing students to study abroad. One of the main purposes of these programs was to assist students with the IELTS or TOEFL test after one or two semesters of instruction. The students studied the courses of IELTS reading, writing, speaking and listening in the program and the contact hours were about 20 hours per week on average.

2.2.1 Research participants

The participants in this study consisted of 95 students of four natural intact classes from the two programs, two teachers and one team leader from Program B. There was an experimental and a control group for Program A (23 students in the experimental group and 24 in the control group). However, due to students' varied language proficiency levels in Program B, there were two experimental classes of 24 students each, and no control group.

All of the participants were high school graduates and the majority of them (76 out of 95, 79.1%) had at least nine years of formal English language learning at school. Regarding their familiarity with the computer-based assessment, only one-third of the students (32, 33.4%) had taken computer-based tests before, the remaining two-thirds (63, 65.6%) had never taken such tests (according to data from Question 2 of the pre-test learner questionnaire, n=95).

The two teachers and the team leader all had MA degrees in English language education. Their teaching experiences varied from three to five years.

2.2.2   Research methods

A mixed-methods research approach, combining both quantitative and qualitative studies, was undertaken. Quasi-experimental research in the form of pre- and post-tests in Aptis was conducted to the experimental and the control groups to evaluate the learning gain and the efficacy of the teaching experiment. To reduce the degree of difficulty of the test as a confounding factor, we used A-B tests cross-over, counter-balanced design; if a student took test A during the pre-test, he/she would take test B in the post-test. In this way, the mean score differences analysis of the control/experimental groups would not be confounded by the difficulty of the tests factor.

The Aptis tests of *Grammar* and *Vocabulary* and *Writing* modules were used as the pre- and post-tests, and both tests were administered online and scored by Aptis examiners. The Aptis *Grammar* and *Vocabulary* module comprises 50 objective questions, with 25 questions for *Grammar* and 25 for *Vocabulary*. The Aptis *Writing* module consists of four writing tasks. Due to the diagnostic nature of the research, detailed feedback on students' writing performance was requested by the research team and was provided by Aptis test scorers.

Pre-, interim-, and post-test student questionnaires (informed by O' Sullivan 2015a) were designed and administered to the students during the experiment. Based on O' Sullivan (2015a), the pre-test questionnaire contained 12 questions (nine objective questions and three subjective questions), covering students' English language learning history, their familiarity with the computer-based assessment, their experiences with Aptis and their perceived performance in the Aptis pre-test (see Appendix 1). Altogether 102 students completed the questionnaires; 95 of the responses were valid.

The interim student questionnaire (see Appendix 2) was administered after students had received scores and writing feedback on their pre-test performance; 93 students completed it in the class.

The post-test student questionnaire contained two parts and aimed to inquire about the students' experiences with the post-test and the teaching experiment throughout the whole semester (see Appendix 3). Altogether 70 valid questionnaires were collected.

The three questionnaires were all prepared in Chinese and in print form. The results of the questionnaires will be presented in Section 3, Results.

An open-ended teacher questionnaire (see Appendix 4), journals (see Appendix 5) and project reports were used to elicit teachers' views on Aptis and on the process of how diagnostic feedback from Aptis helped to guide and enhance their teaching.

2.2.3   The teaching experiment

The teaching experiment comprised two interventions with the Aptis core module of grammar and vocabulary and standard writing tests being introduced as a diagnostic assessment tool for the experimental group to diagnose students' English language proficiency at the beginning and the end of the experiment, and the implementation of a teaching experiment to the experimental group based on the diagnostic feedback provided by Aptis test markers, while everything remained the same with the control group.

The pre-test took place on 16 September 2014 and the teaching experiment extended for 12 weeks for Program A (plus one week for the mid-term exam and one week for APEC holiday) (see Appendix 6 for the teaching experiment plan template). In Program A, the writing instruction for the semester mainly centred on English argumentative essay writing, including an overview of English argumentative essays, how to write an opening paragraph, how to write topic and supporting sentences, how to write a conclusion and various types of argumentative essays in IELTS writing. Students in Program A completed six essays during the term, and the essay topics ranged from education, environment, life, technology to roles of government and other popular IELTS essay topics (see Appendix 7).

For Program B, the first four weeks of writing instruction focused on practicing Part 1 of IELTS writing (which was an expository writing task based on a diagram, a map or a figure), and the latter four weeks on Part 2 of IELTS writing (which was in the form of essay writing). In total, Program B students completed 16 IELTS writing tasks within eight weeks.

Informed by the Aptis writing pre-test diagnostic feedback report, in addition to the regular teaching schedule, the teaching experiment (commencing from late September 2014 to mid-January 2015) mainly consisted of interventions on grammar, vocabulary and writing. The underlying rationale was that good writing is built upon a solid foundation of grammar and vocabulary knowledge and skills. The design of the teaching experiment drew largely on the pedagogical suggestions and online resources links provided by the Aptis developers in the Aptis candidate guide (The British Council 2013) and other writing materials (e.g. Langan 2008).

Specifically, along with teacher instruction, the grammar resources provided by the British Council (http://learnenglish.britishcouncil.org/en) and mobile apps such as *IELTS Word Power* and *LearnEnglish Grammar* were used in the class to enhance students' vocabulary and grammar skills. To tackle students' writing problems revealed through the pre-test, such as sentence structures and use of punctuation marks, the teachers referred to Part 4 "Handbook of Sentence Skills" (including grammar, mechanics, punctuation and word use) of *College Writing Skills with Readings* by Langan (2008). To help students with IELTS writing skills, teachers also recommended the writing website by the British Council (http://learnenglish.britishcouncil.org/en/writing-purpose/essays-structure-1).

It may be argued that the teaching experiment bore two important features as follows. Firstly, as most of the resources used in the experiment were up-to-date, online and designed for self-study, students became very interested in them and actively involved in autonomous learning. Second, through introducing the online independent learning into the class, teacher instruction became more targeted, and the class mode changed from teacher instruction solely to students' autonomous learning assisted with teacher instruction. Accordingly, teacher instruction time accounted for one-third to half of the class time, the remaining two-thirds to one-half were for students' online autonomous practice. The traditional grammar–translation classroom switched into a communicative, task-based classroom, as reported by Teacher A in the project report.

Relevant teacher training was provided along with the project implementation, and the training covered not only Aptis technology training but, more importantly, that on teaching pedagogy, particularly on how to use the diagnostic feedback to improve teaching. The training mainly assumed the form of face-to-face conferences and email communication.

2.3   Data analysis

The quantitative data of students' pre- and post-test performance, and students' responses toward the objective questions in the questionnaires, were analysed with SPSS statistical package, while the qualitative data, such as the open-ended questions, teacher journals and interim project reports, were examined through a content analysis. Content analysis is a research tool used to determine the presence of certain words or concepts within texts or sets of texts. It can be defined as the systematic, quantitative analysis of message characteristics (Neuendorf 2002), and can assist researchers to understand historical documents, newspaper stories, political speeches and open-ended interviews.

3. RESEARCH RESULTS AND DISCUSSION

Research results are reported in the order of the three research questions.

3.1 Students' attitudes toward and experiences with the Aptis test and its diagnostic use

This part answered the first research question, which investigated students' experiences with the Aptis test and its use as a diagnostic tool. The data came from the pre-test learner questionnaire, students' responses toward the feedback on the Aptis writing test, and the post-test learner questionnaire.

3.1.1 Students' attitudes toward and experiences with the Aptis pre-test

Given that two-thirds of the candidates had never taken computer-based tests before, it was interesting to see how they viewed Aptis (Table 1). Accordingly, 61.4% students considered the Aptis test interface to be attractive (Question 3), and 88.5% of students viewed Aptis test instruction as very clear (Question 4).

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
3. The look of the Aptis test was attractive and appealing.	15 (15.6%)	44 (45.8%)	36 (37.5%)	1 (1.0%)	0
4. The Aptis test instruction is very clear and easy to follow.	27 (28.1%)	58 (60.4%)	9 (9.4%)	2 (2.1%)	0

Table 1: Students' views on Aptis

Almost two-thirds of students (60.4%) contended that the Aptis test allowed them to demonstrate their true level of English (Question 5). Nevertheless, students offered a more conservative answer regarding whether they had performed to the best of their ability in the test – only 11.4% considered they did well, while 31.2% reported not doing well, and 57.3% adopted a neutral attitude (Table 2). Modesty might partly account for the response, although computer familiarity may also play a role.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
5. The test gave me an opportunity to show my true level of English.	14 (14.6%)	44 (45.8%)	32 (33.3%)	5 (5.2%)	1 (1.0%)
6. I feel that I have performed to the best of my ability in the Aptis test.	1 (1.0%)	10 (10.4%)	55 (57.3%)	29 (30.2%)	1 (1.0%)

Table 2: Students' perceived performance in the Aptis pre-test

Regarding the factors that affected their performance (Table 3), a majority of students (63.6%) considered unfamiliar vocabulary as the main cause. The next biggest number of students (43.7 %) reported lack of time, which coincided with their responses to Question 2 (for two-thirds of the candidates, taking the Aptis test was their first experience of sitting for computer-based assessment, as reported in Section 2.2.1). Nearly one-third (30.3%) of the students held that unfamiliarity with the topics, especially the writing topics, was also an important factor. More than a quarter of students (26%) reported they had difficulty comprehending some sentences and they experienced technology problems respectively. On factors which did not affect their performance, the largest number of students named "My fear of the test" (56.3%) and "Other language problems" (52.1%).

	A lot	Quite a lot	Not a lot	Not at all
7. What affected your performance in this test？				
7a. Unfamiliar words	30 (31.3%)	31 (32.3%)	28 (29.2%)	7 (7.3%)
7b. Difficult sentences	10 (10.4%)	15 (15.6%)	35 (36.5%)	33 (34.4%)
7c. Other language problems	2 (2.1%)	4 (4.2%)	37 (38.5%)	50 (52.1%)
7d. Difficult questions	8 (8.3%)	8 (8.3%)	51 (53.1%)	29 (30.2%)
7e. Unfamiliarity of topics	11 (11.5%)	18 (18.8%)	30 (31.3%)	36 (37.5%)
7f. Time pressure	13 (13.5%)	29 (30.2%)	29 (30.2%)	24 (25.0%)
7g. My fear of tests	6 (6.3%)	9 (9.4%)	27 (28.1%)	54 (56.3%)
7h. Technology	13 (13.5%)	12 (12.5%)	29 (30.2%)	41 (42.7%)

Table 3: Perceived factors affecting students' performance in the Aptis pre-test

Among the three Aptis tests, a large majority of students (85.4%) reported the writing test as "the most and the second most difficult" (Question 8c, Table 4), followed by the vocabulary test (81.2%) (Question 8a). In contrast, nearly half of the students (45.8%) considered the grammar test as "the easiest" (Question 8b). Researcher observations during the test and after-test informal interviews with a few students indicated that students were not familiar with the ways writing questions were asked nor the writing topics. In one case, students felt at a loss when being asked to name a history figure he/she admired and explain the reason why. In another case, students were asked to write about gardening, which might be a common topic in the West, but which is very new in China, particularly to the students. For vocabulary, in addition to their limited size of vocabulary, another reason may be that students were not familiar with the type of vocabulary questions.

	The most difficult	Second most difficult	The easiest
8. Can you rank the sections of the Aptis test according to how difficult you found them?			
8a. The vocabulary test	39 (40.6%)	39 (40.6%)	17 (17.7%)
8b. The grammar test	9 (9.4%)	43 (44.8%)	44 (45.8%)
8c. The writing test	39 (40.6%)	43 (44.8%)	14 (14.6%)

Table 4: Perceived difficulty of the Aptis pre-test

Regarding the section of the Aptis pre-test that helped identify language problems best, 60.4% of students considered the vocabulary test and 56.3% the writing test, with only 37.5% identifying the grammar test. This might indicate that vocabulary and writing tests were most useful in helping students identify their language problems (Table 5).

	The most helpful	Second most helpful	The least helpful
9. Which test helps with identifying your language problems best?			
9a. The vocabulary test	58 (60.4%)	25 (26.0%)	10 (10.4%)
9b. The grammar test	36 (37.5%)	39 (40.6%)	19 (19.8%)
9c. The writing test	54 (56.3%)	36 (37.5%)	5 (5.2%)

Table 5: Perceived helpfulness in identifying problems of the Aptis pre-test

The next three questions were open-ended, aiming to obtain students' views on this form of test. On the main advantage of the Aptis test (Table 6, Question 10), the largest number of students (25) reported that it could help identify their level and weaknesses, while 15 students considered the computer-based test convenient, efficient, intelligent and a novel means of testing.

10. The main advantage of the test is:	**No. of students**
Identify my level and weaknesses	25
Convenient, efficient, intelligent and novel	15
Questions from easy to difficult and can cater for different levels of students	10
High-brow, formal and official	5
Authentic, related to daily life	5
Identify my writing level	3
Clear instructions and interface	3
The grammar test is very useful	2

Table 6: The main advantage of the Aptis pre-test

Regarding the Aptis test items, it was shown that test items in general were interesting and motivating in both content and question types, and students were impressed by the variety of questions in the vocabulary part. Writing tests were authentic, particularly the writing task in the online discussion chat room. Students felt that they were writing to a real audience in the chat room (Table 6). Students also expressed their positive experiences regarding the gradual progression of the four parts of the writing test in both the topic depth and question types, in that as they moved from one task to another, they naturally tended to write more and better (Table 6).

These responses were encouraging, particularly for this study, as the majority of students reported the diagnostic feature of the test as identifying their level and weakness. Nevertheless, as mentioned previously, for two-thirds of students, this was the first time they had undertaken computer-based tests. Not surprisingly, the largest number of students (20) perceived the main disadvantage of the test as "not being used to this new form of test and experiencing problems in spelling and typing" (Table 7, Question 11). The next largest number of students (12) considered the "test not comprehensive in that there were no speaking or listening sections". The remaining students commented on the limited testing time given the number of questions, challenging vocabulary questions and bright interface.

11. The main disadvantage of the test is:	**No. of students**
Not used to the computer-based test, problems in the typing speed and spelling	20
Test not comprehensive, no spoken test nor listening test	12
Too short time for too many questions	7
Difficult vocabulary questions	6
Interface too bright and can cause eye fatigue	3

Table 7: The main disadvantage of the Aptis pre-test

Students' additional comments on the test were both positive and negative (Table 8). Some considered it to be a good and interesting test which can check wide areas of knowledge; others felt they were not used to computer-based tests, expecting more time for the writing test.

12. Please write any comments you have about the test here.	No. of students
Not used to the computer-based test	2
Longer time for the writing test	2
Good test	2
Very interesting	1
It is much easier than IELTS	1
It can test wider areas of knowledge	1
There is a word count for the writing test, which is very helpful	1
The writing test is most useful, while the vocabulary test is not	1

Table 8: Students' additional comments on the Aptis pre-test

3.1.2 Student responses on their pre-test performance and writing feedback

The Aptis writing test contained four parts of different writing tasks, ranging from blank filling of personal information (Part 1), answering questions in an online chat room (Part 2), to writing two emails (one formal and one informal) (Part 3 and Part 4), with each part accounting for five marks. After students took the pre-test, in addition to the scores for each part, we also asked for detailed feedback for the writing test for each student, based on which, teachers could offer more targeted writing instruction. After the score and detailed feedback were provided to the teachers and students, we invited their views on them via a student questionnaire (Appendix 2) and a teacher questionnaire (Appendix 4).

The student questionnaire contained five questions, of which three were objective and the remaining two subjective questions.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
1. I feel Aptis has given an accurate assessment of my grammar and vocabulary level.	13 (13.5%)	45 (46.9%)	28 (29.2%)	7 (7.3%)	0 (0%)
2. I feel Aptis has given an accurate assessment of my writing level.	7 (7.3%)	46 (47.9%)	31 (32.3%)	8 (8.3%)	1 (1.0%)

Table 9: Perceived accuracy of Aptis assessment

Over 60% of students considered that the Aptis test had offered an accurate assessment of their grammar and vocabulary competence, which verified the usefulness of this test (Table 9). For the writing test, more than half (55.2%) of the candidates held the same views. It is, however, worth noting that nearly one-third of students did not give views on these two questions. In general, the majority of the students considered the Aptis pre-test had offered a precise assessment of their language proficiency. Moreover, the majority of students (82.3%) were satisfied with the writing feedback offered (Table 10).

	Very satisfied	Satisfied	Not satisfied
3. How do you find the feedback on Aptis writing test?	5 (5.2%)	74 (77.1%)	14 (14.6)

Table 10: Students' satisfaction with writing feedback

Nearly half of the students (49.5%) reported their main problems with writing (as revealed through the Aptis test) were their grammar errors and monotony of sentence structures (Table 11), followed by 41.9% of students who reported lack of vocabulary, advanced vocabulary in particular. Five students experienced problems in comprehending test instructions, which was also due to their grammar and lack of vocabulary. In conclusion, it seemed that grammar and vocabulary posed the biggest obstacle in their writing test.

4. What is your main problem with writing as revealed through the Aptis test?	No. of students	Percentage
Grammar errors, monotonous sentence structures	46	49.5%
Lack of vocabulary, advanced vocabulary in particular	39	41.9%
Limited testing time	1	1%
Not used to computer-based tests	2	2%
Incorrect understanding of test instructions	5	5%

Table 11: Main writing problems

Over 70% of students considered the feedback to be very useful in helping them locate their problems, which could be attributed to the detailed and specific comments provided by the Aptis markers on each section of the writing test (Table 12). Over 10% of students (10.2%) held that via practising different genres, they got to know their weaknesses in writing.

5. Can you know your main problems in writing through Aptis feedback and why?	No. of students	Percentage
Feedback indicates my problem clearly, very helpful	62	70.5%
Get to know my weaknesses in writing through practising different genres, somewhat helpful	9	10.2%
Very helpful in identifying my problems, therefore I know what I need to work hard at	1	1.1%
Not very useful as it is too general	1	1.1%
Not very useful as my teacher has already pointed out my problem prior to the test	1	1.1%
It can help me acquire a rough idea of my problem, however the Aptis essay is about letter writing only, which is limited in testing the writing skill	1	1.1%
Not very helpful, as I almost forgot the writing questions, therefore feedback not very helpful	1	1.1%

Table 12: Perceived helpfulness of writing feedback

To summarise, the majority of students considered the Aptis pre-test had given an accurate assessment of their language competence, and the writing feedback had helped them locate their problems and improve.

3.1.3 Student views toward the Aptis post-test

Almost two-thirds of students (64.3%) agreed that they had demonstrated their real English level in the post-test (Table 13), which corresponded with their answer to the pre-test (cf. Table 2). This indicates students' very positive views on Aptis being a good and consistent assessment tool for their language proficiency. A quarter (25.7%) reported they had done well in the post-test, nearly half (48.6%) agreed to some extent that they had done a good job, which contrasted strongly with their views in the pre-test, in which only 11.4% (11) agreed that they did well, while one-third (30, 31.2%) did not do well (cf. Table 2). The result might help to verify the efficacy of the teaching experiment, which was also demonstrated by their writing performance in the post-test (see Tables 23–29).

	Strongly agree	Agree	Somewhat agree	Somewhat disagree	Disagree	Strongly disagree
1. The test gave me an opportunity to show my true level of English.	10 (14.3%)	35 (50.0%)	15 (21.4%)	3 (4.3%)	2 (2.9%)	1 (1.4%)
2. I feel that I have performed to the best of my ability in the Aptis test.	6 (8.6%)	12 (17.1%)	34 (48.6%)	12 (17.1%)	1 (1.4%)	1 (1.4%)

Table 13: Students' perceived performance in the Aptis post-test

Regarding the factors affecting their performance in the post-test, the largest number of students (39, 55.7%) reported "unfamiliar words" posing the biggest hindrance (Table 14), which corresponded with their answer in the pre-test, although there were many more students in the former (61, 63.6%) (cf. Table 3). The next factor was "time pressure" reported by 24 (34.3%) students, which was also selected by 42 (43.7%) students (the second biggest number) in the pre-test.

In contrast to the pre-test in which they chose "unfamiliarity with the topics" (29, 30.3%) (cf. Table 3), the third important factor was "difficult sentences" reported by 17 (24.2%) students. As there was no reading test for the candidates, the difficult sentences reported may have come from the vocabulary and grammar part. Similar to their responses in the pre-test, two factors of "My fear of tests" (39, 55.7%) and "Other language problems" (38, 54.9%) were also considered the least influential factors for their performance.

	A lot	Quite a lot	Not a lot	Not at all
3. What affected your performance in this test？				
3a. Unfamiliar words	19 (27.1%)	20 (28.6%)	23 (32.9%)	4 (5.7%)
3b Difficult sentences	5 (7.1%)	12 (17.1%)	19 (27.1%)	29 (41.4%)
3c. Other language problems	2 (2.9%)	1(1.4%)	23 (32.9%)	38 (54.9%)
3d. Difficult questions	2 (2.9%)	6 (8.6%)	34 (48.6%)	23 (32.9%)
3e. Unfamiliarity of topics	3 (4.3%)	9 (12.9%)	25 (35.9%)	28 (40.0%)
3f. Time pressure	9 (12.9%)	15 (21.4%)	22 (31.4%)	19 (27.1%)
3g. My fear of tests	1 (1.4%)	10 (14.3%)	15 (21.4%)	39 (55.7%)
3h. Technology	4 (5.7%)	8 (11.4%)	24 (34.3%)	28 (40.0%)

Table 14: Perceived factors affecting students' performance in the Aptis post-test

	The most difficult	Second most difficult	The easiest
4. Can you rank the sections of the Aptis test according to how difficult you found them?			
4a. The vocabulary test	34 (48.6%)	18 (25.7%)	15 (21.4%)
4b. The grammar test	8 (11.4%)	29 (41.4%)	30 (42.9%)
4c. The writing test	22 (31.4%)	34(48.6%)	11(15.7%)

Table 15: Perceived difficulty of the Aptis post-test

Similar to their responses in the Aptis pre-tests (cf. Table 4), the largest number of students (80%) also perceived the writing test as "the most difficult and the second most difficult" (Question 4C, Table 15), followed by the vocabulary test (74.3%) (Question 4A, Table 15), although the post-test percentages were lower than the pre-tests, which could be traced to one semester of English instruction. The largest number of students (42.9%) considered the grammar test as "the easiest" (Question 4B, Table 15). Students' perceived difficulty of the vocabulary and writing tests corresponded to their answers to Question 3 on the important factors affecting their performance (see Table 14).

	The most helpful	Second most helpful	The least helpful
5. Which test helps with identifying your language problems best?			
5a. The vocabulary test	43 (61.4%)	14 (20.0%)	10 (14.3%)
5b. The grammar test	18 (25.7%)	30 (42.9%)	16 (22.9%)
5c. The writing test	22 (31.4%)	30 (42.9%)	12 (17.1%)

Table 16: Perceived helpfulness in identifying problems of the Aptis post-test

Among the three Aptis post-tests, the vocabulary test was considered by 61.4% of students as the most helpful in identifying their language problems (Table 16), followed by the writing test, selected by 31.4% of students. The grammar test was selected by only 25.7% of students. Though the ranking of the perceived helpfulness of the post-tests was the same as that of the pre-tests, the number of students was fewer in the latter, particularly for the writing test (54/56.3% the pre-test vs. 22/31.4% the post-test). A possible explanation for this might be that after one semester of the teaching experiment, the pre-test feedback from the Aptis test markers, along with targeted writing instruction, students had gradually resolved their writing problems, and therefore fewer students selected the writing test in the post-test. As the teaching experiment focused on writing instruction mainly with vocabulary learning for self-study, the percentage of students who chose the vocabulary test remained more or less the same.

3.1.4 Student experiences with the teaching experiment

Questions 6–15 of the post-test questionnaire were six-point Likert-scale questions, aiming to elicit students' views on the instruction content, teaching methodology and tutorials. Students' responses are reported in Table 17.

Questions	Strongly Agree	Agree	Moderately agree	Moderately disagree	Disagree	Strongly Disagree
6. I think the Aptis writing test administered at the beginning of the term helped me know my writing level.	8 (11.4%)	29 (41.1%)	20 (28.4%)	3 (4.3%)	3 (4.3%)	1 (1.4%)
7. I get to know my weak points through the Aptis writing test.	8 (11.4%)	27 (38.6%)	20 (28.6%)	5 (7.1%)	3 (4.3%)	1 (1.4%)
8. I worked very hard at my weak points in writing during this term.	4 (5.7%)	33 (47.1%)	19 (27.1%)	7 (10.0%)	1 (1.4%)	0
9. I think the teacher focused a lot on our weak points in writing in her instruction.	18 (25.7%)	32 (45.7%)	10 (14.3%)	3 (4.3%)	1 (1.4%)	0
10. I think the teacher provided very targeted writing instruction in this semester.	16 (22.9%)	26 (37.1%)	18 (25.7%)	3 (4.3%)	1 (1.4%)	0
11. I think the teacher has provided very detailed and helpful writing feedback.	20 (28.6%)	30 (42.9%)	10 (14.3%)	2 (2.9%)	2 (2.9%)	0
12. I think we have practised a lot in our writing classes.	22 (31.4%)	29 (41.4%)	10 (14.3%)	1 (1.4%)	1 (1.4%)	6 (8.6%)
13. I read carefully teacher feedback on my essay and will try to avoid making similar mistakes.	21 (30.0%)	34 (48.6%)	7 (10.0%)	1 (1.4%)	1 (1.4%)	0
14. The teacher has provided very useful and helpful writing resources in this semester.	31 (44.3%)	21 (30.0%)	10 (14.3%)	1 (1.4%)	1 (1.4%)	0
15. I do not read carefully teacher feedback on my essay, anyhow I will not write the same essays next time.	3 (4.3%)	5 (7.1%)	3 (4.3%)	10 (14.3%)	31 (44.3%)	10 (14.3%)

Table 17: Students' experiences with the teaching experiment

More than half of the students (52.5%) agreed that the Aptis pre-test helped them know their writing level and half (50) held that they got to know their weak points through the Aptis writing test (Questions 6–7, Table 17), which pinpointed students' positive attitudes toward Aptis being used to diagnose their writing problems.

Over half of students (37, 52.8%) reported working very hard at their weak points in writing (Question 8), and over two-thirds practised a lot in the writing class (Question 12).

On writing instruction and teacher feedback, more than two-thirds of students (71.4%) agreed that the teacher attended to their weak points in instruction (Question 9, Table 17), about 60% held that the teacher had offered very targeted writing instruction and helpful feedback (Questions 10–11), demonstrating students' very positive attitude to the writing instruction and teacher feedback elements of the teaching experiment. More importantly, nearly three quarters of students (74.3%) reported that the teacher had provided very helpful and useful writing resources (Question 14), which again verified the validity of the teaching experiment design and the usefulness of the teaching experiment.

Regarding students' use of teacher feedback, more than three-quarters of students (78.6%) agreed that they would attend to teacher feedback carefully and try to avoid making similar mistakes (Question 13). Students' response toward Question 13 corresponded with their answer to Question 15, where only 15.7% students reported not attending to teacher feedback.

Questions 16–17 are open-ended questions, aiming to investigate students' perceptions on their biggest progress in writing, area for improvement and underlying reasons. Content analysis method was used to analyse students' answers, and major themes and sub-themes were extracted.

More than a third of students (34.4%) perceived their biggest progress was in "essay structure" (Table 18). Students reported logical thinking and clearer structures in their essays. Nearly 30% (29.5%) of students held "vocabulary" as their biggest progress, which was mainly shown in the expanding vocabulary size and choice of words. Some students (16.4%) reported "grammar", in that fewer grammar mistakes were made. Students (13.1%) reported "sentence structures" being more varied and complex. In addition, students also named "being not nervous" and "being able to write more" as progress in the writing.

Responses	No. of students	Percentage
Essay structure	21	34.4%
Vocabulary	18	29.5%
Grammar	10	16.4%
Sentence structure	8	13.1%
Others	4	6.4%
Total	**61**	**100.0%**

Table 18: The perceived biggest progress in writing

As for the underlying reason for the perceived progress, more than half the students (52.1%) traced it to effective writing instruction for this semester, which might verify the effectiveness of the teaching experiment (Table 19).

Responses	No. of students	Percentage
Effective writing instruction	25	52.1%
Adequate writing practice	13	27.0%
Reciting model essays	4	8.3%
Others	6	12.6%
Total	**48**	**100.0%**

Table 19: The underlying reason for the progress

According to the students' responses, the teachers have offered targeted writing instruction in essay ideas and structures, in language building and enrichment. Moreover, the teachers provided helpful writing resources and detailed feedback on their essays. The following are some of the students' comments:

> *Teaching is targeted. For example, my teacher exemplified the English essay structures with templates and we followed them in our practice, which helps a lot. (Student A, Program A)*

> *The teacher gave us very helpful writing resources, such as teaching handouts, model essays, and many writing resources, which we have memorised. It was very helpful. (Student B, Program A)*

> *The teacher provided very detailed and helpful feedback on our essays. (Student A, Program B)*

> *The teacher taught us how to remember new words and expand our vocabulary size and offered a lot of grammar exercises for us to practice. (Student C, Program A)*

The second largest number of students (27.0%) maintained that the reason for their progress was due to adequate writing practice and continuous assessment.

> *We did a lot of writing practice in the class and as a result, the writing speed is improved. (Student B, Program B)*

> *We did a lot of tests in the class as well. (Student D, Program A)*

The third largest number of students (8.3%) attributed their progress to their reciting many model essays and resources.

> *I studied the model essays and writing resources teachers assigned and tried to memorise them. (Student E, Program A)*

In addition, some students also held the underlying reasons as their having learnt how to think in English, clearer logic in writing, becoming more academic in writing and fluency in language, and fewer wrong sentences.

The largest number of students (43.1%) considered vocabulary, particularly the use of advanced and complex vocabulary, the area they needed to improve the most (Table 20), which also corresponded with their responses on the factors that affected their post-test performance most (cf. Table 14). Students also listed improving sentence complexity, grammar and logic in writing for further areas for improvement (Table 20).

Responses	No. of students	Percentage
Vocabulary	25	43.1%
Sentence	11	19.0%
Grammar	11	19.0%
Logic in writing	4	6.9%
Writing resources	4	6.9%
Writing speed	3	5.1%
Total	**58**	**100.0%**

Table 20: The perceived area for improvement

On the cause for their weak areas, the largest number of students (27.3%) attributed it to lack of effort, particularly to their poor English foundation, not working hard enough, and not reciting model essays (Table 21). This finding corresponds with most studies with Chinese students, who usually attribute failure to lack of effort, rather than other reasons.

Other factors that students named were simple and spoken English sentence, poor grammar and numerous grammar mistakes (Table 21). The third factor (reported by 24.2%) was vocabulary problems: no variety in vocabulary, limited vocabulary size, and lack of advanced vocabulary in particular. The fourth factor was logic in writing: students reporting being slow in organising ideas, using a Chinese way of thinking in English writing, and being not familiar with the English way of writing. Finally, their limited command of the writing resources also caused them to spend more time on writing.

Responses	No. of students	Percentage
Efforts	9	27.3%
Sentence and grammar	9	27.3%
Vocabulary	8	24.2%
Logic in writing	5	15.2%
Writing resources	2	6%
Total	**33**	**100.0%**

Table 21: The underlying reason for improvement

To summarise, learner experiences with the Aptis test revealed that students were very positive on Aptis being a good assessment tool for their language proficiency (Table 9). More importantly, students also maintained that the Aptis test could help to diagnose their weak points in writing (Question 7, Table 17). The efficacy of the teaching experiment was verified by students' positive responses as demonstrated in Tables 17, 18 and 19. Effective writing instruction was named as the main factor contributing to their progress in writing (Table 19). Students in the open-ended questions particularly related teachers' in-class targeted writing instruction based on Aptis feedback, assigned writing practice, provision of useful online writing, vocabulary and grammar resources, and detailed teacher feedback on their writing exercises during the term, which again verified the effectiveness of the teaching experiment.

3.2   Teachers' view on the efficacy of the Aptis-based teaching experiment

To answer Research Question 2: *How do teachers view the efficacy of the diagnostic use of the Aptis test in the classroom?*, teacher experiences with the teaching experiment were collected mainly through a teacher questionnaire, teacher journals and project reports. As mentioned above, two teachers were involved in the teaching experiment. In this paper, we refer to the teacher for Program A as Teacher A, and to the teacher for Program B as Teacher B.

3.2.1   Teacher responses toward Aptis pre-tests and feedback

A teacher questionnaire containing five open-ended questions was designed and administered to augment teachers' views on students' performance in the Aptis pre-test and on the follow-up feedback from the Aptis test markers (see Appendix 4). The two teachers' responses are presented in Table 22.

	Teacher A	Teacher B
1. Have students taken part in other language proficiency test before the Aptis test? If yes, what is the score?	Some students took IELTS before and the score was 5–6.	Only a small number of students took IELTS and the score varied.
2. Are you satisfied with students' performance in Aptis vocabulary and grammar test? Did it meet your expectation?	Basically satisfied and their performance has basically met my expectation.	Basically satisfied and their performance has basically met my expectation.
3. Are you satisfied with students' performance in Aptis writing test? Did it meet your expectation?	Not satisfied, and their performance did not match my expectation. I think the main reason was due to students' unfamiliarity with the writing test item types and with different writing registers.	Not satisfied. Students reported lack of time and also their unfamiliarity with the writing test item types.
Q4. What is the biggest problem in students' writing? Has it been revealed through this test?	The problems were: ideas not convincing, grammar mistakes, lack of cohesion and coherence in structuring ideas. Grammar and punctuation problems were revealed in the test.	Grammar, structure and lack of command of formal writing style, all being somewhat revealed through the test.
5. Are you satisfied with the writing feedback provided by Aptis developers?	Basically satisfied. It would be better if it could be more detailed. For example, if particular grammar items such as attributive clauses, noun clauses, etc. were specified in feedback, teacher instruction would be more targeted. Moreover, students reported if they could view the feedback along with the questions, they could get a clearer idea of their own problems.	Basically satisfied. In the meanwhile I think there should be a unified and explicit marking criteria. Comments like "low word count" and "register" were too general for us to follow, the same problem applied to grammar, the particular grammar problems were not specified.

Table 22: Teachers' views on the Aptis pre-test and feedback

To summarise, both teachers considered their students' performances in the Aptis grammar and vocabulary tests had met their expectations, while some of their writing problems such as grammar, punctuation, style etc. were revealed via the Aptis writing test.

Teachers were basically satisfied with the follow-up feedback, however, they expressed the wish for them to be more specific in pinpointing the problems and there being an explicit assessment criteria.

It is also worth mentioning that students hoped to view the original questions along with the corresponding feedback.

3.2.2 Teacher experiences with the Aptis-based teaching experiment

Though both teachers joined the project from the beginning, participating in all the project meetings and contributing to the teaching experiment design, their teaching approaches varied, as did their experiences.

Teacher A

Teacher A's instruction followed the suggested teaching experiment procedure, and her 12-week teaching was built around instructing students how to write effective essays, commencing from topic sentences, unity, cohesion and coherence to conclusion. Her writing classes were scheduled in the computer lab and an E-learn system was used to assist students' writing. Teacher instruction in the class ranged from one-third to one-half of the class time, with the remaining time for students to do online practice to review and reinforce what they had learned in the class. The underlying teaching philosophy, according to Teacher A, was to change the grammar translation instruction method to task-based class instruction.

In terms of teaching interventions, she first assisted the students with Aptis pre-test result and feedback interpretation, as revealed in the following journal entry.

> *I spent my first lesson assisting student to interpret their Aptis test scores and feedback report, focusing mainly on feedback on their writing problems. As we could not have the test questions on hand, I used the test examples provided on the Aptis Guide to explain and exemplify the four parts of the Aptis writing tests along with the assessment criteria. Most students paid much attention to their performance in the Aptis test, one of the main reasons was that the test was developed by the British Council, who was also the developer of IELTS test. After my interpretation of the assessment criteria, they had gained a better understanding of their scores. Many students also wanted to know how their Aptis scores correlated with IELTS scores, to which I tried to assist by first converting their scores to CEFR bands, then to IELTS scores.*
>
> *(Teacher A, data source: journal 01)*

Via interpreting the assessment criteria to the students, Teacher A was actually developing the students' autonomy. It is also worth noting that this group of students was also very autonomous and goal-driven, judging from the fact that they demanded to see how their Aptis scores correlated with their IELTS scores.

Then Teacher A followed the project team's experimental writing teaching plan, offering extra support to grammar, vocabulary and writing in her instruction based on the Aptis pre-test feedback. In terms of grammar, she introduced the British Council online grammar exercise website to the students, assigned students to do the online grammar exercises in class, and supplemented students' practice with her summary of key grammar points, which resulted in effective grammar instruction.

> *As students' grammar problems lay mainly in verb tenses as revealed in the Aptis pre-test, we used the grammar practice on the BC web site in my class (Grammar: http://learnenglish.britishcouncil.org/en). I summarised the grammar key points and then students practiced online in class. The online grammar exercise was not only interesting, but also meaningful and engaging. Moreover, the online testing form also attracted students. All in all, the online testing form, rich and authentic language learning resources all helped to improve students' grammar.*   *(Teacher A, data source: project report)*

Many students reported their limited vocabulary size after the pre-test, and the pre-test writing feedback also revealed students' problems with the correct use of vocabulary and in spelling. Therefore, how to expand students' vocabulary size, particularly writing vocabulary, and how to improve students' correct use of vocabulary was a problem Teacher A faced in this study. Similar to supplementing the grammar component, Teacher A also incorporated online resources in vocabulary instruction.

> *We introduced two vocabulary learning apps: IELTS Word Power (available in App store) and LearnEnglish Grammar, which includes the graded vocabulary lists of IELTS, TOEFL and various other tests. With rich resources, vocabulary games and many example questions, these two apps were very appealing to the students, to whom, learning vocabulary through the mobile phones is their first experience. The greatest advantage is that students can learn vocabulary at any time at any place, and they are also interesting and engaging compared with print content. Mobile learning via engaging students to practice also develops students' autonomy in learning. Of course, teacher assessment from time to time is also compulsory.*
>
> *(Teacher A, data source: project report)*

For the writing skills intervention, in addition to the original teaching schedule and resources, she incorporated materials such as *Chapter 4 Handbook of Sentence Skills* from the *College English Writing* (Langan, 2008) course book into her instruction to help deal with problems in their sentence writing and use of punctuation marks. She also made active use of online writing resources in the class.

> *More importantly, I also recommended the online writing resource developed by the British Council at http://learnenglish.britishcouncil.org/en/writing-purpose/essays-structure-1, from which students received the most authoritative English writing and IELTS writing guide. Students were very interested in exploring these online websites and in online learning, much more than their interest in reading teacher lecture notes and other familiar methods. However, whether they will go on to study the online materials after class remains a question and it is hard to monitor from my side.*   *(Teacher A, data source: project report)*

All in all, Teacher A made full use of the Aptis writing test diagnostic feedback to make her teaching more targeted, aiming at resolving students' weak points in writing.

> *Feedback on students' Aptis pre-test writing test informed my teaching. For example, the pre-test writing feedback demonstrated that students made the biggest number of errors in the use of punctuation in Parts Two, Three and Four. Accordingly, I incorporated a part on the use of punctuation in my writing instruction, hence students made fewer mistakes in the post-test.*
>
> *(Teacher A, data source: project report)*

With the support of the project team, Teacher A enriched her instruction with many useful resources to assist students to tackle grammar problems as revealed in the pre-test and to expand their vocabulary size and improve their writing, a feature of the intervention program rated very positively by the students under study (see students' response to Question 14 in Table 17).

However, teaching interventions did not always produce the desired effect. Part 4 of the Aptis writing test requires students to distinguish formal and informal styles of writing, which she also taught in the class. However, students still made many errors in this part as shown in the post-test feedback. She explained the reason as follows:

> *Two main factors can be attributed. First, students did not listen carefully in the class. Second, students did not pay attention to this. The main aim for the group of students under study was to pass the IELTS test, therefore they only attended to those related to IELTS. The formal and informal styles of writing in the Aptis writing test Part 4 were not tested in IELTS, as a result, failed to attract students' attention.*   *(Teacher A, data source: project report)*

The students' pragmatic attitude could not be blamed as their main task in this program was to pass the IELTS test, and this attitude was also echoed in Teacher B's report.

Teacher A voiced her views regarding the content of Aptis tests, particularly the writing test and commented on the writing topics and genres.

> *The writing topics and questions on the whole were interesting and they motivated students to write more. After the pre-test, my feeling was that the grammar and vocabulary test was excellent, however, the writing test needed improvement. For example, the topic of gardening might be very popular in the West but very new here in China, as few Chinese families or students had the opportunity to do gardening. I was afraid that the unfamiliarity of the topic might affect students' writing performance. Second, most students were used to writing argumentative essays, and they found it hard doing practical writing as required in Aptis, which was what most students have felt about.* *(Teacher A, data source: project report)*

Judging from her journals and reports, Teacher A seemed to have fully understood and implemented the aims and the content of the project, i.e. using Aptis as a diagnostic tool to identify students' problems and offer feedback, followed by teacher-targeted instruction on the basis of Aptis feedback. She related her thoughts about the project.

> *As a teacher joining in this project, what I felt most was that the Aptis feedback helped me adjust my instruction effectively and offer more specific assistance to students on handling their weak points, particularly in the aspects of grammar, punctuation, and choice of vocabulary. Personally, I felt the post-test feedback was more detailed than the pre-test feedback, in that the former has not only pointed out students' problems but also given relevant examples, while the latter was only spotting out students' problems. Students would get a clearer understanding of their problems after having read the post-test report.* *(Teacher A, data source: project report)*

Teacher A also compared her students' performance in the Aptis pre- and post-tests.

> *The score comparison showed that the two classes had both made progress in the four writing tasks, among which, it seemed that students made the most obvious progress in Part 2, the average score in the post test was nearly the full mark (i.e. 5), indicating most students got full marks for this part.* *(Teacher A, data source: project report)*

Teacher A commented positively on the teaching experiment.

> *The teaching experiment contributed many novel things to the original teaching schedule, the use of online resources and mobile apps added energy to the dull classroom. Teachers and students all felt excited at the new and fresh learning resources and the learning opportunities brought by them. Students made obvious progress, particularly for the experimental group, who had witnessed an obvious growth in the number of students who scored full marks (5) in the post-test. This result would not be possible without the teaching interventions.* *(Teacher A, data source: project report)*

To summarise, Teacher A has fulfilled various roles during the teaching experiment, ranging from Aptis feedback and writing assessment criteria interpreter, learning resources provider and guide user, reflective teaching practitioner (as revealed from her journals), and learning task organiser. As stated before, while engaging students in completing online autonomous learning tasks on grammar, vocabulary and writing, Teacher A switched her class from the traditional grammar translation to more task-based instruction. This might have resulted in improved teaching efficacy as demonstrated in students' improved writing performance in Section 3.3 below.

Teacher B

For Teacher B, her writing instruction during the teaching experiment mainly centred on the two parts of IELTS essays. She admitted the advantage of Aptis in diagnosing grammar problems and noted her students' reduced number of errors in the post-test.

> *Aptis writing feedback specified students' particular problems in writing, such as the subject–verb agreement, the use of pronouns, verb tenses etc., a comparison of students' pre and post scores showed the obvious progress in this aspect, indicating students have made improvement in language accuracy after this semester of instruction.* (Teacher B, data source: project report)

Nevertheless, according to Teacher B, as the Aptis test writing question types differed widely from the IELTS, students did not pay attention to the Aptis writing test report.

> *Some questions in both the Aptis pre- and post-tests seemed to be not helpful or appealing to students. For example, in IELTS, there was no question type such as spelling out the months, therefore, the number of students who made this type of error increased in the post-test. In Aptis writing Part 4, students needed to have knowledge of the difference between formal and informal writing styles. As IELTS writing only required formal writing style and writing in most of their textbooks was between the formal and the informal writing, most students were not familiar with the informal style of writing and scored low in this part.* (Teacher B, data source: project report)

Teacher B voiced her view on Aptis in her report as follows.

> *On the whole, I feel the Aptis test was most useful in helping students locate their grammar problems, particularly to the weak students. However, due to its different testing foci from IELTS, students did not give it due attention.* (Teacher B, data source: project report)

To summarise, it seemed that due to the difference between Aptis and IELTS, Teacher B's instruction seemed solely focused on the two tasks of IELTS writing, which was understandable as the main purpose of the program was to assist students to pass the IELTS test.

Arguably, the teaching experiment seemed to be undertaken fully with Program A with the teaching method changing from the grammar–translation/teacher instruction mode to the task-based learning mode. Changes were also witnessed in the learning process with students in Program A becoming more autonomous, reading and acting on the diagnostic feedback. Moreover, Teacher A's journal revealed the process of the teaching intervention and her reflections along with the process.

3.3 How the teaching experiment involving Aptis impacts students' performance

This part aims to answer Research Question 3: *How does the teaching experiment involving the use of the Aptis test as a diagnostic assessment tool impact students' performance?*

Due to the insufficient number of students from Program B joining the post-test, in the quantitative analysis, only students' performance from Program A (altogether 47 students, with 24 students in the control group, 23 students in the experimental group) was analysed. Students' pre- and post-test scores were inputted into SPSS 16.0. Independent Samples T-test and Paired Samples T-test were conducted to compare students' performance before and after the experiment, between and within groups. The results are presented in Tables 23 to 29.

	Group	N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
pret1	control	24	3.50	1.32	-.33	.41	.24
	experimental	23	3.83	1.37			
pret2	control	24	3.83	.76	.09	.77	.16
	experimental	23	3.74	1.32			
pret3	control	24	2. 67	1.05	-.20	.57	.17
	experimental	23	2.87	1.39			
pret4	control	24	2.08	1.02	-.35	.29	.31
	experimental	23	2.43	1.24			

Table 23: Comparison of two groups' writing scores in the Aptis pre-test

	Group	N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
postt1	control	24	3.92	1.14	-.12	.72	.10
	experimental	23	4.04	1.26			
postt2	control	24	4.17	1.40	-.74	.02	.86
	experimental	23	4.91	.29			
postt3	control	24	3.08	1.47	-.27	.54	.18
	experimental	23	3.35	1.47			
postt4	control	24	2.17	1.17	-.57	.10	.49
	experimental	23	2.74	1.14			

Table 24: Comparison of two groups' writing scores in the Aptis post-test

Tables 23 and 24 show the two groups' writing performance scores before and after the test in the four parts of the Aptis writing test. It seems that the two groups were of the same level before the experiment started with p values all above 0.05 (see Table 23). After the teaching experiment, the experimental group outperformed the control group in all four parts with the mean differences from .12 to .74 (see Table 24), the differences of the two groups in the post-test Part 2 was statistically significant with p value at .02, the effect size at .86.

We also compared two groups' performance in the Aptis grammar/vocabulary pre- and post-tests and the results are shown in Table 25. The independent samples t-test results show that the experimental group outperformed the control group in the pre-test, with the mean difference of 3.54 in favour of the experimental group (the difference was statistically significant with p value at .02). However, after the test, there were no differences between them, suggesting that the same teacher might have introduced the same innovative methods with the control group in the grammar and vocabulary teaching.

	Group	N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
Pregv	control	24	28.42	5.17	-3.54	.02	.70
	experimental	23	31.96	4.90			
Postgv	control	24	31.21	6.93	-.53	.77	.09
	experimental	23	31.74	5.27			

Table 25: Comparison of two groups' G/V scores in the Aptis pre- and post-tests

Students' writing performance within groups in the pre- and post-tests were also compared and analysed (see Table 26). For the experimental group, students' post-test performance outperformed their pre-test performance in the four parts of the writing test, with the mean difference from .21 to 1.17. The experimental group seemed to make the biggest gains in Part 2 of the writing test with 3.74 at the pre- and 4.91 at the post-test; the difference was statistically significant with p value at .00 and effect size at .95. The difference in Part 3 was also statistically significant with p value at .05 and effect size .43. Further research needs to be conducted to explore the reason for these findings.

		N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
Part 1	pret1	23	3.83	1.37	-.21	.45	.16
	postt1	23	4.04	1.26			
Part 2	pret2	23	3.74	1.32	-1.17	.00	.95
	postt2	23	4.91	.29			
Part 3	pret3	23	2.87	1.39	-.48	.05	.43
	postt3	23	3.35	1.47			
Part 4	pret4	23	2.43	1.24	-.31	.25	.25
	postt4	23	2.74	1.14			

Table 26: Comparison of the experimental group writing scores in the Aptis pre- and post-tests

		N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
The experimental group	pregv	23	31.96	4.90	.22	.84	.04
	postgv	23	31.74	5.27			

Table 27: Comparison of the experimental group G/V scores in the Aptis pre- and post-tests

Despite the experimental group students' rapid progress in the post-writing test, their performance in the G/V test remained stable with only the mean difference score at .22, which was not statistically significant (see Table 27).

		N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
Part 1	pret1	24	3.50	1.32	-.42	.17	.29
	postt1	24	3.92	1.14			
Part 2	pret2	24	3.83	.76	-.34	.33	.20
	postt2	24	4.17	1.40			
Part 3	pret3	24	2.67	1.05	-.41	.13	.33
	postt3	24	3.08	1.47			
Part 4	pret4	24	2.08	1.02	-.09	.76	.06
	postt4	24	2.17	1.17			

Table 28: Comparison of the control group writing scores in the Aptis pre- and post-tests

In contrast to the findings of the experimental group (as indicated in Table 26), although the control group students made progress in the post-test with a score gain from .09 to .42, the growth was smaller than the experimental group and the difference was not statistically significant with p values all above .05 (see Table 28).

		N	Mean	Std. deviation	Mean difference	Sig. (2-tailed) P value	Effect size
The control group	pregv	24	28.42	5.17	-2.79	.10	.35
	postgv	24	31.21	6.93			

Table 29: Comparison of the control group G/V scores in the Aptis pre- and post-tests

The control group received higher scores in the G/V post-test with mean difference at 2.79, however, the difference was not statistically significant with p value at .10.

To summarise, in terms of writing performance, the experimental group students in Program A witnessed visible writing gains in the post-test, especially in Part 2 of the Aptis writing test (see Tables 24 and 26). It is also noticeable that the experimental group students in Program A in the post-writing test tended to write more and better, e.g. they could write 30-40 words per answer. However, the control group did not write much in the post-test, which could be partly attributed to the lack of teaching treatment (i.e. lack of teachers' handling the diagnostic feedback and teaching intervention).

4. CONCLUSIONS

4.1 Limitations

A number of limitations exist with this study. First, the length of the experiment was eight weeks, which might not be long enough to witness bigger visible gains. The experiment should be lengthened in a future study.

Second, as revealed through teacher journals and reports, the students in this study were solely concerned with IELTS and would pay attention to IELTS-related resources and instruction. Moreover, some candidates in Program B did not attend the post-test as they had left for home after achieving the required scores in the IELTS test. In future, the study can be undertaken with the on-campus degree program students to ensure the full participation of the candidates through the whole experimental period to examine the efficacy of the teaching interventions.

Although the researcher has endeavoured in training, designing the teaching experiment and organising regular project meetings to handle any emerging issues and offer support during the experiment, it has been demonstrated that variations existed in the two teachers' treatment of the diagnostic feedback and in their implementation of the teaching experiment. Teacher training and support therefore needs to be enhanced in follow-up studies.

4.2 Conclusions

Being the first attempt to use Aptis as a diagnostic assessment tool in Chinese EFL classrooms, the study has yielded useful findings regarding the efficacy of the Aptis tests in evaluating students' language competence and that of the teaching experiment based on Aptis diagnostic feedback. Students and teachers both considered Aptis grammar and vocabulary, and writing tests to be good indicators of candidates' language proficiency; writing test items authentic and motivating; and feedback on the writing tests helping them identify their problems and improve. The teaching experiment in the light of Aptis writing feedback seemed to have produced encouraging results, particularly for Program A students, with the experimental group making progress in both the between-group pre- and post-tests comparison and the within-group pre- and post-tests comparison, and the gain was significant in Part 2 of the writing test (see Section 3.3 for details).

Both teachers and students responded favourably toward the teaching experiment. Teachers viewed the introduction of Aptis as a diagnostic tool in the classroom along with the remedial teaching intervention and the online learning resources to be an effective means of engaging students, of enlivening the classroom atmosphere and of improving teaching efficiency. Students embraced the teaching interventions of teacher in-class targeted writing instruction based on Aptis feedback, assigned writing practice, online writing, vocabulary and grammar resources, and teacher detailed feedback on their writing samples during the term. It may be argued that teacher and students' positive experiences of the Aptis tests and the remedial teaching experiment led to the statistically significant writing gain demonstrated by the experimental group in Program A.

In addition, the study also pinpointed the following recommendations for improvement. Firstly, Aptis test feedback should be provided more quickly if it is to be used for diagnostic assessment purposes. In the current Aptis core test of vocabulary and grammar, only scores for grammar and vocabulary are provided after candidates have finished the test, which might not be helpful for diagnostic teaching and learning purposes. Both teachers and students need to know in which areas students have made mistakes so that they could improve in the future. Given that the vocabulary and grammar test consists of objective questions only, it is suggested that automated diagnostic feedback on grammar can be designed since each grammar question deals with one grammatical item. For the vocabulary part, the automated feedback for each question can be about the particular word along with the vocabulary range that the word falls into, so that students know the word range they should aim for.

For the Aptis writing tests, due to the conflict between the marking workload and the immediacy of feedback demanded, it is suggested that efforts can be made to develop software to offer automated writing feedback in the future. It is also worth noting that both teachers and students in this study considered feedback supported with examples more helpful than feedback alone, which may have implications for Aptis writing test feedback design in the future.

Second, Aptis writing test topics could be more varied to better reflect its international mission. For example, the topic of gardening might be very popular in the West, but is not common in the Chinese context and many students were not familiar with this theme, thus experiencing difficulties in writing about it. Moreover, argumentative writing tasks can be included in the Aptis writing test in the future. The current Aptis writing test mainly consists of practical writing tasks, which might limit its accessibility to a large number of candidates who are preparing for high-stakes exams like IELTS, as indicated in this study.

Third, it is suggested that more targeted teaching and learning resources for the diagnostic instruction purposes could be developed in the future. The research has demonstrated that the online and mobile learning resources suggested by the Aptis developers in the *Aptis Candidate Guide* were very useful references for teachers in designing remedial instruction, and learners found them engaging and helpful as well. Arguably for Aptis tests to become truly a diagnostic assessment tool and an assessment for learning tool in the class, it is envisaged that after students take the test, they would receive not only a score, but also a diagnostic report of their strengths and weakness followed by suggested learning resources targeted at their weak points, as revealed in the test. This is possible with the current development of big data and learning analytics research.

To conclude, Aptis is a useful instrument for both placement and diagnostic assessment (if supported with diagnostic feedback) as demonstrated in this study.

With increasing communication between nations and across borders, English language instruction and assessment has become incredibly important in today's globalised world. Aptis tests, with their accessibility and flexibility, have contributed to producing an accurate and fast assessment of candidate's language proficiency, serving placement, recruitment and educational purposes as claimed. It might be argued that to measure learning is important, but to improve learning is even more important. In this sense, the significance of the current study might not only lie in using Aptis to measure learning, but more importantly, using the Aptis assessment result to design remedial instruction to improve learning. It is hoped that the study has shed light on how to use Aptis effectively as a diagnostic assessment tool in the EFL classroom and on how to improve Aptis diagnostic assessment functions to better serve instructional purposes.

REFERENCES

Alderson, J. C., Brunfaut T. and Harding L. (2014). Towards a Theory of Diagnosis in Second and Foreign Language Assessment: Insights from Professional Practice Across Diverse Fields. *Applied Linguistics* 2014, 1–26.

Alderson, J. C., Huhta, A., Nieminen, L., Ullakonoja R. and E. Haapakangas. (2013). What is the impact of diagnostic language tests? Paper at *10th EALTA Conference 2013.* Istanbul, Turkey, 23-26 May 2013.

British Council. (2013). *The Aptis candidate guide*. www.britishcouncil.org/aptis

Brunfaut, T. and McCray, G. (2015). Looking into test-takers' cognitive processes while completing reading tasks: A mixed-method eye-tracking and stimulated recall study. *ARAGs Research Reports Online, 001.* London: British Council.

Doe, C. (2011). The integration of diagnostic assessment into classroom instruction. In D. Tsagari and I. Csépes (Eds.), *Classroom-Based Language Assessment* (pp. 63–76). Peter Lang.

Douglas, J. and Mbali, C. (2015). Evaluating Aptis Tests with teacher training students and in-service primary school teachers in Durban, South Africa. *ARAGs Research Reports Online, 002.* London: British Council.

Fox, J. and Hartwick P. (2011). Taking a diagnostic turn: Reinventing the portfolio in EAP classrooms. In D. Tsagari and I. Csépes (Eds.), *Classroom-Based Language Assessment* (pp. 47–61). Peter Lang.

Hawkey, R. (2000). *Impact Theory and Practice: Studies of the IELTS Test and Progetto Lingue 2000.* Cambridge: Cambridge University Press.

Jaiyote, S. (2015). The relationship between test-takers' L1, their listening proficiency and performance on paired speaking tests. *ARAGs Research Reports Online, 2.* London: British Council.

Langan, J. (2008). *College Writing Skills with Readings.* Foreign Language Teaching and Research Press.

Neuendorf, K. A. (2002). *The Content Analysis Guidebook.* Sage Publications.

O'Sullivan, B. (2015a). Aptis Formal Trials Feedback Report. *Technical Report 002.* London: British Council.

O'Sullivan, B. (2015b). Aptis Test Development Approach. *Technical Report 001.* London: British Council.

Rowntree, D. (1987).*Assessing Students: How Shall We Know Them?* London: Kogan Page.

Tang, J. (2003). What language courses haven't prepared us for: a qualitative study of Chinese students' language difficulties during a prolonged period of study at a British university. In Hu Wenzhong (Ed.) *Teaching English in China (III)* (pp. 362-395). Beijing: Foreign Language Teaching and Research Press.

APPENDIX 1:
The Pre-test Student Questionnaire

Dear Student,

This questionnaire aims to know your English language learning and your feedback about this test, Please fill in it truthfully, thank you for your support!

Class: ....

Age:...

Gender: a. Male   b. Female

1. When did you start to learn English:
    a. Kindergarten b. Primary school Grade 1 c. Primary school Grade 3
    d. Junior school Grade 1 e. Others, please specify

2. Have you ever taken a computer-based test before?   a. Yes   b. No

3. The look of the Aptis test was attractive and appealing.
    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

4. The Aptis test instruction is very clear and easy to follow.
    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

5. The test gave me an opportunity to show my true level of English.
    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

6. I feel that I have performed to the best of my ability in the Aptis test.
    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

7. What affected your performance in this test? Please tick (√) in the appropriate box.

	A lot	Quite a lot	Not a lot	Not at all
7a. Unfamiliar words				
7b. Difficult sentences				
7c. Other language problems				
7d. Difficult questions				
7e. Unfamiliarity of topics				
7f. Time pressure				
7g. My fear of tests				
7h. Technology				

8. Can you rank the sections of the Aptis test according to how difficult you found them?
    Please tick (√) in the appropriate box.

	The most difficult	Second most difficult	The easiest
8a. The vocabulary test			
8b The grammar test			
8c. The writing test			

9. Which test helps with identifying your language problems best?
    Please tick (√) in the appropriate box.

	The most helpful	Second most helpful	The least helpful
9a. The vocabulary test			
9b. The grammar test			
9c. The writing test			

10. I think the major advantage of the test is:

    ....

    ....

11. I think the major disadvantage of the test is:

    ....

    ....

12. Please write any comments you have about the test here.

    ....

    ....

    ....

APPENDIX 2:
Students' Experience with
the Aptis Pre-test Scores and Writing Feedback

1. I feel Aptis has given an accurate assessment of my grammar and vocabulary level.

    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

2. I feel Aptis has given an accurate assessment of my writing level.

    a. Strongly agree b. Agree c. Neither agree nor disagree d. Disagree e. Strongly disagree

3. How do you find the feedback on Aptis writing test?

    a. Very satisfied b. Satisfied c. Not satisfied

4. What is your main problem with writing as revealed through the Aptis test?

5. Can you know your main problems in writing through Aptis feedback and why?

APPENDIX 3:
The Post-test Student Questionnaire

Dear student,

This questionnaire aims to know your feedback about this test and the writing course in this semester. Please fill in it truthfully, thank you for your support!

Class: ....

Age: ....

Gender: a. male   b. female

Section 1: On the test

1. The test gave me an opportunity to show my true level of English.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree   f. Strongly disagree

2. I feel that I have performed to the best of my ability in the Aptis test.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree   f. Strongly disagree

3. What affected your performance in this test? Please tick (√) in the appropriate box.

	A lot	Quite a lot	Not a lot	Not at all
3a. Unfamiliar words				
3b. Difficult sentences				
3c. Other language problems				
3d. Difficult questions				
3e. Unfamiliarity of topics				
3f. Time pressure				
3g. My fear of tests				
3h. Technology				
3i. Unfamiliarity of writing topics				
3j. Others, please specify.				

4. Can you rank the sections of the Aptis test according to how difficult you found them?
   Please tick (√) in the appropriate box.

	The most difficult	Second most difficult	The easiest
4a. The vocabulary test			
4b. The grammar test			
4c. The writing test			

5. Which test helps with identifying your language problems best?
    Please tick (√) in the appropriate box.

	The most helpful	Second most helpful	The least helpful
5a. The vocabulary test			
5b. The grammar test			
5c. The writing test			

Section 2: On the writing course

6．I think the Aptis writing test administered at the beginning of the term helped me know my writing level.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

7．I get to know my weak points through the Aptis writing test.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

8．I worked very hard at my weak points in writing during this term.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

9．I think the teacher focused a lot on our weak points in writing in her instruction.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

10. I think the teacher provided very targeted writing instruction in this semester.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

11．I think the teacher has provided very detailed and helpful writing feedback.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

12．I think we have practised a lot in our writing classes.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

13．I read carefully teacher feedback on my essay and will try to avoid making similar mistakes.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

14．The teacher has provided very useful and helpful writing resources in this semester.
   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

15．I did not read carefully teacher feedback on my essay, as I would not write the same essay next time.

   a. Strongly agree b. Agree c. Moderately agree d. Moderately disagree e. Disagree f. Strongly disagree

16．I think that I have made the biggest progress in the area (s) of: _____

   The reason(s)：_____

17．I think that I need to improve my writing in the area(s) of: _____

   The reason(s)：_____

APPENDIX 4:
Questionnaire for Teachers

1. Have students taken part in other language proficiency test before the Aptis test?
   If yes, what is the score?

2. Are you satisfied with students' performance in the Aptis vocabulary and grammar test?
   Did it meet your expectation?

3. Are you satisfied with students' performance in Aptis writing test? Did it meet your expectation?

4. What is the biggest problem in students' writing? Has it been revealed through this test?

5. Are you satisfied with the writing feedback provided by Aptis developers?

APPENDIX 5:
The Teacher Journal Template

Program:..

Class time (week):..

Teacher:...

Class:..

Lesson content:...

Observations and Reflections (which can be expanded through the following questions):

1. Students' overall performance in the class（good points and areas for improvement）

2. Teacher reflections（In which area teachers do well, and in which area they need to improve; whether students like to practice writing, what are their writing problems; whether students ask teacher questions and what types of questions they ask）

3. Others

APPENDIX 6:
The Teaching Experiment Plan Template

The Aptis Experimental Teaching Plan (October-December 2014)

Group Name:				
Teacher name:				
No. of weeks	**Teaching objectives**	**Teaching focus**	**Teaching procedures**	**Expected outcome**
1 (Oct 20-24)				
2 (Oct 27-31)				
3 (Nov 3-7)				
4 (Nov 10-14)				
5 (Nov 17-21)				
6 (Nov 24-28)				
7 (Dec 1-5)				
8 (Dec 8-12)				
9 (Dec 16,18)-post test				

APPENDIX 7:
Essay Writing Topics for Program A Students

1	In recent years, studying abroad is becoming increasingly popular among teenagers. What is your idea?
2	It is better for children to grow up in the countryside than in a big city. Do you agree or disagree? Use specific reasons and examples to develop your essay.
3	We have entered a throw-away society and fill the environment with rubbish. What are the causes and what are your solutions?
4	It has been suggested that everyone in the world want to own a car, a TV, and a fridge. Do you think the disadvantages of such a development outweigh the advantages?
5	The government should pay for people's health care and education. To what extent do you agree or disagree?
6	The world is changing rapidly. An increasing number of people change their jobs and the places they live frequently. Is this a negative or positive development?

British Council Assessment Research Awards and Grants

If you're involved or work in research into assessment, then the British Council Assessment Research Awards and Grants might interest you.

These awards recognise achievement and innovation within the field of language assessment and form part of the British Council's extensive support of research activities across the world.