UNIVERSITY OF SOUTHERN QUEENSLAND

The effect of web-based formative assessment on students' achievement and attitudes in a large enrolment tertiary English course

A Dissertation submitted by

Somboon Chetchumlong

B.Ed. (English), M.A.T. (Teaching English)

For the award of Doctor of Philosophy
Faculty of Education

2010

ABSTRACT

The study aimed to investigate the effect of Web-based formative assessment on students' achievement and attitudes in a large enrolment tertiary English course at a Thai university. The participants were a stratified random sample of 186 first year undergraduate Thai students who enrolled to study English II (212102) in the Department of Western Languages, Faculty of Humanities and Social Sciences, Burapha University in the second semester of 2004-2005 academic year. The sample was randomly separated to a Web-based formative assessment (WBFA) group as an experimental group and a conventional paper-and-pencil formative assessment (CPFA) group as a control group.

Prior to and after the intervention of the WBFA program, a mixed methods research design with sequential explanatory strategy was utilised to collect two phases of data. First, quantitative data were collected and analysed from objective pre- and post-tests to comprehend the students' achievement in the CPFA and WBFA groups. Then the students' attitudes toward using WBFA in the course were determined from quantitative and qualitative data collected and analysed from pre- and post-WBFA questionnaires and pre- and post-WBFA semi-structured interviews.

Results revealed that the WBFA group outperformed the CPFA group. However, there was no statistical difference between their overall means. With regard to the four language features tested, the WBFA group surpassed the CPFA group. The means of the WBFA group were statistically higher than those of the CPFA group in *reading comprehension* and *vocabulary* sections. In addition, students who completed the WBFA program got higher average score than those who did not. Nonetheless, there was no significant difference. There was no significant correlation between the number of attempts to perform the WBFA and the students' achievement scores. Findings also indicated that after using the program students in the WBFA group had more positive attitudes toward using WBFA in their course. Specifically, the students agreed that WBFA should be integrated in their course because it helped to improve and motivate them to learn English in the course.

CERTIFICATION OF DISSERTATION

I certify that the ideas, experimental work, results, analyses, software and conclusions reported in this dissertation are entirely my own effort, except where otherwise acknowledged. I also certify that the work is original and has not been previously submitted for any other award, except where otherwise acknowledged.

_____ _____

Signature of Candidate Date

ENDORSEMENT

_____ _____

Signature of Supervisor/s Date

_____ _____

Date

ACKNOWLEDGMENTS

I would like express my gratefulness to the Royal Thai Government, the Office of the Higher Education Commission in the Thai Ministry of Education and Faculty of Humanities and Social Sciences in Burapha University for the scholarship and the ongoing support on my study at the Faculty of Education, University of Southern Queensland. My sincere gratitude also goes to my principal supervisor Associate Professor Shirley O'Neill, my associate supervisor Dr. Jeong-Bae Son, and Associate Professor Patrick Danaher (PhD Program Coordinator) for their academic advice and assistance. In addition, I am grateful to Thai undergraduate students at Burapha University and my colleagues at the Department of Western Languages who contributed invaluable data for the study.

I would like to thank Associate Professor Somchai Dechaprompun (former Dean of the Faculty of Humanities and Social Sciences), Assistant Professor Areeluck Harnmontri, Assistant Professor Dr. Malee Nitsaisook, Assistant Professor Dr. Rewat Sangsuriyong, (former Head of HUSO Information Technology Center), CEC's *ChulaOnline*, Thai staff at the Office of Educational Affairs in Canberra, members of postgraduate and early career research (PG & ECR) group, doctoral students at the Faculty of Education at USQ, Dr. Allan Morgan and Dr. Joan Conway for their support, encouragement and empathy.

Finally, I would like to thank my parents Prasarn and Boonyong, my aunt Surangsri and my family Apapak, Juthawasu, Jesada for being my great inspiration and reinforcement.

TABLE OF CONTENTS

List of Figures .....Page

List of Tables .....Page

Chapter 1 Introduction

1.1 Structure of the dissertation

This dissertation is organised into six chapters. Chapter 1 introduces and describes background of the study, research problems, aims of the study, research questions, expected outcomes, and significance of the study. Chapter 2 is concerned with relevant literature reviews and presents a conceptual framework for the use of Web-based assessment, specifically Web-based formative assessment (WBFA), and its effects on students' achievement and attitudes in large enrolment courses reported in previous studies. In Chapter 3, the rationale, research designs, and methods for the study are described. The Chapter also includes details for sample selection, units of statistical analysis, instrumentation, WBFA program, reliability, validity, and methods of data collection. Chapter 4 and Chapter 5 are related to analyses and findings of data collected through a mixed methods research design. These two chapters report results relevant to the effect of WBFA on students' achievement and the effect of WBFA on students' attitudes, respectively. Finally, discussions and conclusions are articulated in Chapter 6.

1.2 Background

The study reported in this dissertation concerns a growing number of students enrolling in compulsory foundation English courses for first year undergraduate students from all disciplines at a Thai university. These students have experienced problems with both obtaining adequate practice with the English language (i.e., completing dialogues, reading comprehension, vocabulary, and grammar) and receiving prompt feedback on language skills to be learned. Most of them have found that they are unable to achieve course objectives in terms of improvement in English language proficiency by the end of the semester. In addition, teachers are confronted with excessive resource demands on increasing time and workloads for marking required by the current conventional paper-and-pencil formative assessment (CPFA). Hence, it is hypothesised that implementing WBFA, instead of CPFA, would be a more efficient and effective approach in terms of providing more flexibility for students to practice the language,

receive immediate feedback, and encouraging students to gain higher achievement scores than utilising CPFA. It is also anticipated that WBFA would be attractive to students while maintaining an appropriate, if not improved, pedagogical focus.

The implementation of innovative computer technologies (e.g. Computer-based test, computer-assisted assessment, Web-based test, Internet-based testing, Online testing, e-assessment) has increased in the area of teaching English as a second/foreign language (ESL/EFL). Language students and teachers in most academic levels are being encouraged to use these technology-based techniques in their learning and teaching. Atkinson and Davies (2000) and Choi, Kim and Boo (2003) noted that, apart from being adopted to reinforce students' motivation, this computer-based testing has been adapted to evaluate students' language achievement. A number of researchers have also examined its impact on students' attitudes; for instance, Cook (2000), Kurubacak (2000), Lin (2003), and Russo (2002) found that students' and teachers' acceptance of this testing system was favourable because of its flexibility, authenticity, instant feedback, and individualisation with regard to assessment. Moreover, Chen, Sullivan, and Savenye (2002) and Tunc and Armstead (2000) discovered support for establishing the use of technology in language classrooms, and highlighted its potential applications to large-size classes in particular.

The World Wide Web (WWW), one of these technology-mediated applications, is the basis so probably should clarify what is meant here as all works off the WWW has already been found to be amenable to both modern learners and visionary teachers. Wong (2003) discovered that more and more face-to-face courses, for example, science, mathematics, and English, are being successfully taught in high schools through the integration of this technology. According to Klinger (2003), young students as well as adult learners at university level, engaging in online learning, demonstrated satisfactory attitudes toward this innovation. Specifically, Chang (2002) argued that the application of technology is particularly relevant to testing and specifically Web-based assessment in improving learners' achievement.

Furthermore, Lewis (2002) indicated that students increased their achievement scores when they were stimulated to practise and undergo frequent tests related to fulfilling the course objectives as with formative assessment. This activates the idea of generating

Web-based assessment for formative purposes, particularly in large enrolment courses. Benefits include substantive feedback to students and reduction of marking workloads for faculty staff (Bennett, Oliver & Pinn, 1999). Cross (2001) also found that both students and teachers regarded formative assessment as beneficial to learning and teaching. Similarly, Bennett (2001) concluded this high technology had the potential to supersede the conventional large-scale assessment, arguing that it surpassed the traditional paper-and-pencil tests, while gaining an increased focus on a current curriculum.

In relation to language acquisition and the relationship between the utilisation of technology and students' achievement, Wright (2002) found that students' reading achievement was improved when technology was used to assist them to develop skills for examination preparation. Zhang (2003) also reported that the use of technology in English language testing had positive effects on students' language achievement.

Blayney and Freeman (2003) and O'Hare (2001) noticed that students experiencing this type of assessment developed favourable attitudes toward computer-based formative assessment. Thus, as noted by Bennett (2002a), there are potential benefits for English language learning via the integration of technology-assisted assessment with traditional face-to-face teaching to create a more powerful learning experience. This supports Ingram's (2003) view that technology should not be implemented to replace face-to-face teaching; rather, it should be treated as another teaching aid that is complementary.

Interestingly, Frankel (2002) and Kobrin and Young (2003) noted that there is evidence that both computerised assessment and paper-and-pencil tests have similar impacts on students even though the former would provide better motivation to students. Nevertheless, important to this study context is that in large enrolment tertiary English courses, WBFA would save more time, reduce teachers' marking loads, provide faster feedback to students, and it can be accessed anywhere and any time (Haycock, 2001). With these potential advantages, there is a strong argument to investigate whether WBFA enhances these first year Thai undergraduate students' achievement and affects their attitudes toward the use of WBFA.

1.3 Research problems

Most EFL classes in Thailand are relatively large, especially at Burapha University, where all first year undergraduate students have to study at least two mandatory foundation English courses during their first two semesters. According to the Australian University Teaching Committee or AUTC (2002), any language class consisting of 50 students or over is considered as a large enrolment class. This study focuses on students studying English II (212102), in which there are approximately 2,000 undergraduates enrolling each semester. These students are separated into 31 classes, each of which contains 50-80 students with diverse backgrounds from different faculties and colleges. Consequently, providing sufficient time to offer an effective English language program to individual students is quite demanding. This also affects teachers' workloads and increases pressure to spend excessive time marking the students' assignments, exercises, and test papers. This situation leads to less time for student-teacher interaction, and low proficiency on students' language resultant competence due to lack of motivation to learn.

A number of studies (e.g., Calhoon, Fuchs & Hamlett, 2000; Gretes & Green, 2000) have shown computerised assessments to be more beneficial than paper-based tests. Positive factors identified include authenticity and ability to improve students' achievement (e.g., Maccini, Gagnon, & Hughes, 2002; Peat & Franklin, 2002), the provision of instant feedback (e.g., Mason, Patry & Bernstein, 2001; Weinberg, 2001), and substantial help to develop both learners' performance and teachers' teaching strategies (e.g., Martinez, 2002). Nonetheless, some studies (e.g., Felzer, 2002; Goldberg, 2000) indicated that students taking paper-based tests received higher average achievement scores whereas their attitudes toward computer-based tests were not as positive as expected. In addition, other studies (e.g., Powers, 2001) revealed that there were no significant differences between the effectiveness of computer-based and paper-and-pencil tests on students' language achievement. Accordingly, it is a strategic and educative exercise to investigate which testing system is more appropriate for evaluating language performance for the possible betterment of both students' language achievement and their perspectives on the testing system, particularly in large enrolment courses.

In modern classrooms, it is highly recommended by Whitworth (2001) that both modes of Web-based and conventional paper-based assessment tools be used. Vonderwell (2003) supports this and notes that students should be familiar with both the face-to-face and online environment to vary their learning experiences. It is also highlighted by Baker (2003) that this integration helps students increase their achievement, while most teachers find the technology beneficial for their teaching as well as being able to provide learning support and independent study or self-access to students. However, there are very few studies, which have examined these issues in a large enrolment foundation English course in higher education, particularly in Asia, and more specifically, in Thailand. Notably, for the university to make a decision to incorporate WBFA as part of the assessment of the English language program there is a great need for research-based evidence.

1.4 Aims of the study and research questions

The primary aim of the study is to examine whether, over a period of one semester, WBFA is able to help improve student's achievement in the English language performance in four language features: *completing dialogues, reading comprehension, vocabulary,* and *grammar*, when compared with CPFA. The secondary aim is to investigate students' attitudes toward WBFA. To achieve the above aims, the study attempts to answer the following questions:

1. Does WBFA serve as a tool for students to elevate their levels of achievement compared with CPFA when measured by objective testing?
2. Do students who use WBFA get higher achievement scores on completing dialogues, reading comprehension, vocabulary and grammar for English than students who use CPFA?
3. How does the frequency of participation in WBFA affect students' learning outcomes?
4. What are students' attitudes toward the use of WBFA and learning English in the course?

1.5 Expected outcomes

It is anticipated that this study will identify the impact of WBFA on Thai undergraduate students' achievement in studying foundation English courses and investigate the extent to which students view this test system positively. The results of the study will hopefully enable English language teachers to be more confident about the implementation of the online formative assessment system. It is also expected that they will help to solve the assessment problems in large EFL classes, which cause many drawbacks to foreign language teaching, learning, and testing. Since EFL teachers have to handle a large number of papers, and marking must be completed as quickly as possible for prompt feedback, using the Web technology will possibly be more effective and feasible in the course management system, as noted by Zin, Darus, Nordin, and Yusoff (2003).

For language learners, it is expected that the interactive capabilities and authentic materials linked to WBFA, which enable students can perform real-world tasks demonstrating meaningful application of essential knowledge and skills (Mueller, 2006), will improve their achievement and inspire positive attitudes. Additionally, for faculty staff, the results of the study using the WBFA assessment, pre-test, post-test, questionnaires, and semi-structured interviews may be used in language test design, test construction, test item delivery, item management, item analysis and interpretation, and score reporting (Fulcher, 2000).

1.6 Significance of the study

The study has the potential to assist EFL teachers to build their confidence in the use of the Web, the broadest and most powerful Internet application, for WBFA. The use of the technology also has the possibility to transform English language assessment, particularly in the Burapha University context and that of many other universities with large enrolment in foundation English courses. It may also be of assistance to test developers who may be searching for a more systematic approach to test design.

In addition, the use of WBFA will introduce and encourage the practice of autonomous and lifelong learning for Thai EFL students at Burapha University and potentially other similar EFL contexts. The testing system may prove to be helpful to passive, low-proficient students who do not like to participate in traditional EFL classes (Inoue & Bell, 2006). This is one of the major problems of most Thai EFL students. WBFA potentially provides a non-threatening environment and may assist students to share more of their ideas among themselves and with teachers via e-communication and hence, enrich their language achievement. Through Web-based language tests, EFL teachers can stimulate students to test themselves for all levels of language skills regardless of time and place. With exciting and authentic features available on the Web, WBFA should provide a pleasing, communicative atmosphere in language learning. With this practical way of online language testing it will also effectively handle large-scale assessment. If successful, the online formative test will be of great help to those creative academic administrators with the vision to make this development nationwide, designating Web-based formative assessment through e-learning as a future educational goal in the Faculty of Humanities and Social Sciences, Burapha University, and possibly for the wider public in Thailand.

1.7 Summary

This chapter justifies the rationale for conducting the research. There are three major problems pertaining to learning and teaching EFL in Burapha University: large classes, heavy marking loads, and students' low achievement and motivation. Hence, WBFA is introduced to investigate whether it would ease the problem. The study focuses on the effects of WBFA on students' achievement and attitudes in a large enrolment tertiary English course. The next chapter aims to describe relevant literature reviews in previous studies.

Chapter 2 Literature review

This chapter provides a review of the literature relevant to the study. It looks at theoretical perspectives underpinning the study, EFL in Thailand before the introduction of the World Wide Web, the introduction of the World Wide Web to Thailand, Web-based assessment, Web-based formative assessment, Web-based formative assessment and students' achievement, Web-based formative assessment and students' attitudes, and Web-based formative assessment in large enrolment courses.

2.1 Theoretical perspectives

There are four theoretical perspectives that underpin and guide this study: second language acquisition, language pedagogy, language assessment, and technology integration (see Figure 2.1).

2.1.1. Second language acquisition and WBFA

Second language acquisition (SLA) relates to the process and sequence that people gain knowledge of other languages in addition to their native language. SLA refers to the similar language development when people acquire their first language. SLA involves acquisition, which is a subconscious process of obtaining a language through exposure to language use and learning, which is a conscious process of studying a language (Ellis, 1994). In general, it is accepted that the process of learning the first language and the second language is not different (Ervin-Tripp, 1974, cited in Brown & Rodgers, 2002, p. 25).

Krashen (1985) described that SLA consists of five hypotheses or generalisations: the acquisition-learning hypothesis, the natural order hypothesis, the monitor hypothesis, the input hypothesis, and the affective filter hypothesis. He emphasised that the input hypothesis is the integral part of SLA. It states that people obtain a language through comprehending messages or receiving understandable input. He argued that comprehensible input is essential, but language students need to be well prepared to receive the input. Moreover, to acquire effectively in SLA, their affective filter or mental block must be lowered. He included that this situation occurs when language

students are motivated, confident, and not anxious because of the threat of possible failure.

Krashen (1987) and Scarino and Liddicoat (2009) suggested that language students could acquire proficiency when the language is applied for communicative purposes. He also suggested that language teachers' responsibility is not simply in a language classroom. Language students need to be encouraged to continue improving their linguistic competence by themselves through provision of sufficient comprehensible input so that they can communicate with the world outside. As noted by Beale (2002), owing to communicative principles, language teaching has to be responsive to learners' needs and interests related to exposure to examples of authentic language from the target language community.

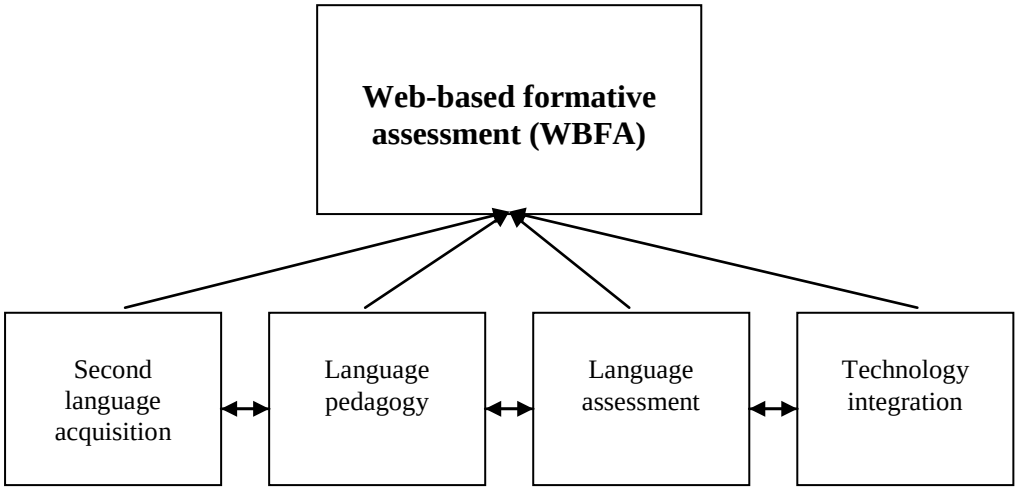


Figure 2.1 Theoretical perspectives underpinning the WBFA program

Consequently, learning experiences should motivate and stimulate language students to learn a language to communicate in a purposeful way. They should be provided real-world and meaningful communicative activities involving integration of different language skills. In relation to a learner-centred philosophy, students should be encouraged to examine, manoeuvre, and investigate answers for themselves (Richards & Rodgers, 2001; Rodgers 2001).

In accordance with this theoretical perspective, WBFA can be used as comprehensible input to link to meaningful communicative activities to support students' language learning. Therefore, students can develop and move on at their own rate, comprehending the input one pace ahead of their present linguistic ability (Schulz, 2007). Along with social constructivist theory, students can be encouraged, by using WBFA, to participate in authentic tasks and real-world environments, as noted by Jonassen (1994, cited in Murphy, 1997). In addition, WBFA promotes student centeredness, in that students could learn anywhere and any time to create knowledge for themselves, independently or socially as they study (Hein, 1991). Vygotsky (1985, cited in Schutz, 2004) indicated that teachers, as facilitators of learning rather than instructors, should supervise and scaffold students' language learning in learning experience to stimulate their students' involvement. In this case, WBFA would be of great assistance to language teachers to motivate students to be more active to improve their language development.

2.1.2 Language pedagogy and WBFA

As noted by Krashen (1985), simply providing comprehensible input to language students is insufficient. It is crucial that language students be ready to obtain the input. Owing to the language pedagogy, students are able to understand what they learn effectively when they are well motivated. Alternatively, they may find it is difficult to comprehend the language when they are in a less motivating atmosphere, for example, in large EFL classes. That takes into account the importance of the language pedagogy theory by providing non-threatening situations along with systematic feedback to students throughout their course (Davies, 1999). Mory (1992) also agreed that such feedback had an influence on learning and it improved learners' performance.

Therefore, with application of electronic feedback on the Web, language students will be accommodated with timely, systematic, non-threatening responses. Moreover, Widmayer and Gray (2000) indicated that the Web provides tools for pedagogical purposes, for instance, discussion boards, online quiz creation, and students' workspace. The system's greatest advantage is that EFL teachers can create communicative activities, providing their students with real-world opportunities for practice in the target language outside the classroom. Since information online is

interesting and attractive (Shuey, 2002), students will be well motivated to learn and EFL students will have more opportunities to interact with real-life materials.

Importantly, WBFA tasks are set at an appropriate level of difficulty so that students could experience success and some challenges. This should reinforce their learning more readily. As a result of repeatedly using WBFA activities and receiving systematic feedback to their responses at appropriately challenging levels, students should also be encouraged to develop positive attitudes toward the use of WBFA. If so, such attitudes, according to Cook (2000), would greatly help stimulate their self-confidence, thus leading to their higher overall language learning achievement. In addition, Dalziel (2000) noted that practising questions with prompt feedback would assist students to consolidate their understanding during the learning process.

Higgins and Tatham (2003) noticed that receiving instant feedback during the learning process through computer-based assessment may increase overall students' learning outcomes. This is in keeping with Skinner's stimulus response learning theory (O'Neill & Gish, 2008) where one would expect immediate systematic feedback to reinforce learning. Due to automated marking and provision of prompt feedback online, WBFA may be viewed as the most suitable solution for solving the problem of the increasing number of students and assessment burdens of language teachers as well.

2.1.3 Language assessment and WBFA

Relevant to the second language acquisition theory and language pedagogy, there are several means to provide comprehensible input in non-threatening circumstances to language students. One potential way is to provide the input through language assessment, specifically formative assessment.

Related to the perspective of formative assessment, Black and Wiliam (1998a) described that it relates to ongoing assessment undertaken for receiving information to be used to amend learning and teaching activities that are engaged. They also stated that there was firm evidence indicating that formative assessment is an essential component of classroom work and its development can raise students' achievement. As noted

earlier, formative assessment is the most important assessment practice because it has the potential to have a powerful effect on students' learning (Black & Wiliam, 1998b). Crooks (1988) supported that any assessment could have a positive impact on students' learning when teachers employ it to support students instead of judging them.

In relation to language assessment, Bachman (1990) described that language tests could be helpful resources for the usefulness of language learning and teaching when they supply valuable evidence in evaluating the effectiveness of different approaches to language teaching. For formative purposes, language teachers can use language assessment to help detect students' strengths and weaknesses to measure their progress and to encourage their linguistic achievement.

In this study, it is anticipated that WBFA would assist students to improve their learning through self-assessment with instantaneous feedback (timely score reporting and correct answers), so they are able to take some control of their ongoing progress and move on to achieve. In addition, WBFA provides feedback to every student more rapidly than conventional paper-and-pencil formative assessment does (CPFA). Therefore, EFL students can identify their weaknesses or strengths in learning in their large enrolment foundation course difficult context. Furthermore, apart from indicating how well or not a student has learned, the formative language assessment can also provide feedback on teaching effectiveness to language teachers. The teachers can use the feedback to trace their students' progress and offer assistance when their students need help.

To construct and develop WBFA, the researcher focused on its qualities based on Bachman and Palmer's (1997) test usefulness: reliability, construct validity, authenticity, interactiveness, impact and practicality. WBFA was trialled for its reliability prior to its implementation to ensure its consistency. Even though multiple-choice questions used in WBFA may seem to make it less authentic and interactive, it would not cause harmful impact since it was considered as low-stakes assessment. For practicality, WBFA would be suitable being implemented in its given situation, a large enrolment foundation course.

2.1.4 Tchnology integration and WBFA

Lao (2000) defined technology integration as "the utilisation, combination, mix, and supplementation of technology tools with instruction to aid and improve learning in the classroom" (p. 1). This definition corresponds to what Jonassen, Howland, Moore, and Marra (2003) concluded that technology integration should engage learners in active, constructive, intentional, authentic, and cooperative learning. Kim, Rich, and Hannafin (2004) argued that it is possible that technology integration will create new educational paradigms shifted from an old paradigm to a new paradigm, as shown in Table 2.1. That is, learning will be launched when students construct their own understanding in meaningful ways. Therefore, students have to change their roles to be proactive learners instead of being only recipients of knowledge. Additionally, having technology integration in their instruction, teachers also have to become active assistants to their students rather than knowledge arbiters.

Table 2.1: Impacts of technology integration on new educational paradigms

OLD PARADIGM	NEW PARADIGM
Teacher-centred instruction	Student-centred learning
Single-sensory stimulation	Multi-sensory stimulation
Single-path progression	Multi-path progression
Single media	Multimedia
Isolated work	Collaborative work
Information delivery	Information exchange
Passive, receptive learning	Active, inquiry-based learning
Factual, knowledge-based	Critical thinking, informed decision making
Reactive response	Proactive, planned
Isolated, artificial context	Authentic, real-world context

Source: Adapted from NETS (2000, p. 2)

According to Roblyer (2003), technology integration in education will stimulate learners' motivation, increase their attention, and link them to modern information sources and learning tools. Interestingly, Wilson and Lowry (2000) supported the idea

of integration technology, especially the Web, for ideal scaffolding in keeping with constructivist learning because this technology can provide access to rich sources of information, encourage meaningful interactions with content, and bring students to challenge, support, or respond to each other.

In foreign language instruction, integration of technology will be based on the assumptions that EFL instruction should be student-centred, and should emphasise combination of different language skills. This is also noted by National Education Technology Standards or NETS (2000), adding that technology provides students with "proactive, authentic, and real-world" (p. 2) activities. It also helps them have better attitudes and self-image. Inoue and Bell (2006) include that "using computers moves students from passive to active mode" (p. 172). They added, "For language learning, repetition is important, and the computer does excellent job providing repetitious drill and relieves the teachers of this often difficult task, especially when students are not dealing with in their first language." (p. 96). Hence, integration of technology in testing second language teaching makes assessment more efficient and serviceable (Chalhoub-Deville, 2001). In addition, Web-based assessment is especially appropriate to indicate a student's level of achievement in a language skill (Dunkel, 1999).

The following section outlines the context and particular circumstances pertaining to EFL classes at Burapha University where this research was carried out. It assists the reader to understand how EFL is taught in Universities like Burapha in Thailand and why an initiative like WBFA is worthy of trial to facilitate EFL learning along with the relevance of the underpinning four theoretical perspectives noted above.

2.2 EFL in Thailand before the introduction of the World Wide Web

Prior to the introduction of the World Wide Web or WWW, the only technology used in EFL classrooms in most state primary and secondary schools in Thailand was a cassette tape player, which was used to increase student's listening and speaking skills. At the tertiary level, some state universities and teacher colleges provided their students with a limited number of language laboratories, containing personal cassette tape recorders and headsets to replace the portable cassette tape player to enhance undergraduate students' listening and speaking skills. At that time, these technologies were considered

the most authentic EFL materials. There were not many state schools or colleges in which students learned the language from native speakers of the English language. The native speakers were volunteers from American Peace Corps and Australia. Studying EFL with these volunteers occurred for a short period and a limited number of students had that opportunity.

Because of text-based instruction and a teacher-centred approach in EFL classes, all contents and activities were used and performed in accordance with those in textbooks and workbooks. Assessments were completely in paper-and-pencil fashions. It seemed that students were not motivated to use the language for communication outside their class because they hardly found someone to listen, talk, or write to. What the students could do were listening and repeating after their Thai EFL teachers, reading only passages in their textbooks, and doing their exercises and homework in their workbooks. Students received the feedback from their EFL teachers the week after or longer depending on when their teachers finished marking. In the middle and at the end of each semester, students were tested on how well and by how much they could memorise English words and grammatical rules related to the textbooks, which were written in both Thai and English. These might be reasons that make students have a low achievement and negative attitudes toward learning the English language.

Later, there were specific English newspapers for students to practice reading authentic materials instead of reading non-updated passages in their textbooks. Up till now, there have been many new technologies to support learning and teaching other language skills in EFL classes. However, many EFL teachers are still using the same approaches in their teaching, even at universities. Hence, most students cannot use the language to communicate effectively. One of the major problems that EFL students in Thailand still find difficult is to develop their English to communicate with the world, as noted by Luanganggoon (2001). She also stated that another vital problem was absence of students' motivation to learn English.

In the next section, additional information of when and why the World Wide Web was introduced to Thailand and eventually to EFL students at Burapha University and how WBFA was created as an intervention in the study are described.

2.3 Introduction of the World Wide Web to Thailand

In Thailand, the Internet was initially used in research before it proliferated into many other fields. This technology has offered many applications to users, but there are two systems which are most well known: electronic mail or e-mail and World Wide Web or WWW. The Internet was introduced to the country by the National Electronics and Computer Technology Center or NECTEC, as shown in Table 2.2. Since the establishment of NECTEC, there have been more Internet organisations and projects founded, for example, the National Science and Technology Development Agency (NSTDA), the Thai Social/Scientific, Academic and Research Network (ThaiSarn), SchoolNet, and UniNet (Malaiwong, 2000).

NECTEC was established in 1986. At that time, it was an initial project under the then Ministry of Science, Technology, and Energy. In 1991, NECTEC was transformed into a specialised national centre under the National Science and Technology Development Agency (NSTDA). NECTEC was the first Web server in Thailand (NECTEC, 2006). However, prior to the establishment of NSTDA, NECTEC instigated the Inter-University Network Project in 1987. Later, the project was developed to UniNet to set up information communication technology infrastructure to network all universities throughout the country (Koanantakool, 1995).

In 1992 the Thai Social/Scientific, Academic, and Research Network (ThaiSarn) was launched to proclaim the second stage of the Inter-University Network Project for scaffolding and broadening multiple applications of the Internet in Thailand. ThaiSarn is presently undertaking its third operation, designated as ThaiSarn3; ThaiSarn 1 was initiated in 1992, ThaiSarn 2 in 1995, and ThaiSarn 3 in 2000 (Charnsripinyo, 2006). When the Ministry of University Affairs, which is now the Commission on Higher Education under the Ministry of Education, developed UniNet, ThaiSarn commenced a new project, SchoolNet, to offer Internet access to state schools under full support from the Ministry of Education (Malaiwong, 2000).

Table 2.2: Introduction of the World Wide Web to Thailand

Years of establishment	Organisations	Objectives
1986	http://www.nectec.or.th/	To undertake, support, and promote the development of electronics and computer technologies through research and development activities (NECTEC, 2006)
1991	http://www.nstda.or.th/	To conduct, support, coordinate, and promote efforts in scientific and technological development between public and private sectors toward maximal benefits for national development (NSTDA, 2007)
1992	http://thaisarn.nectec.or.th	To promote the use of the Internet in Thailand To promote research in networking technologies and applications (ThaiSarn, 2005)
1995	www.school.net.th	To provide Internet access to state schools for awareness creation, promotion of content development, promotion of classroom activities, training of teachers, and provision of network access To empower all schools to access a large pool of information resources using the Internet To serve the goal of universal access for every school nationwide To promote the use of Internet in teaching and learning **(**Koanantakool & Thuvasethakul, 2002)
1997	http://www.uni.net.th	To set up ICT infrastructure connecting all universities, institutions, and campuses across the country To develop self-study centres by providing electronic library databases, Internet, multimedia, video on demand, and other self-study materials To develop social learning and lifelong learning systems by creating multimedia courseware and providing knowledge databases and distance learning systems, and to train teachers and assistants to be able to apply ICT for educational development (AEN, 2003; Thongdhamachart, 2005)
2001	http://www.hrd.e.chiba-u.jp/els-e/main/default.asp	To increase more educational opportunities to people nationwide through e-learning systems To make the course management easier for teachers and students by connecting to the Internet To manage the courses of the CEC's and faculties of Chulalongkorn University as well as other institutions joining the program (CEC, 2006)
2002	http://www.huso.buu.ac.th/els/Main/default.asp	To develop IT system as a tool for efficient learning and teaching management in the Faculty of Humanities and Social Sciences, Burapha University To provide IT trainings, seminars, and workshops to students and the faculty staff (Sangsuriyong, 2007).

In 1995, SchoolNet, Thailand, was established to enable all schools to access and obtain extensive information resources through the Internet (Kiattananan, Koanantakool, Chairatanayut, Kiatisevi, & Beck, 1999). For promotion of the Internet access to the higher education in Thailand, UniNet was established in 1997 to connect all universities and campuses across the country. It was the pioneer of e-learning systems in Thailand. Not only was it an Internet service provider but it was also a knowledge distributor. In 2006, UniNet and ThaiSarn were merged to become Thailand Research and Education Networks or ThaiREN (Tantasanawong, 2007).

Chulalongkorn University was the first Thai university to set up academic Internet connections in 1992 (Koanantakool, 1995). Afterwards, the e-learning system (ELS) was introduced to the Continuing Education Center (CEC) at the university to offer more educational opportunities to the Thai public through the Web: www.chulaonline.com. According to Yoshida (2004), an advisor of CEC, the initial online system of Continuing Education Center was commenced in 2001.

One of the major objectives of CEC is to provide more academic opportunities to all people throughout the country. Consequently, CEC initiated *ChulaOnline*, using an e-learning system called *Chula ELS*. It is also an online course management program, which can be operated using both Thai and English fonts. The program contains a database of courses, examinations, and all information concerning data of students', teachers' and administrators' accessing the program. Since it is an Internet-based program, access can always be available to everyone, regardless of time and place (CEC, 2006). It was available freely at first before it changed a commercial software program (Sangsuriyong, 2007).

2.3.1 WWW in Burapha University

Burapha University Computer Center or BUCC (http://ict.buu.ac.th/) was established in 1992. Its main objectives are to supply computer technology services and supports to students, faculty staff, and communities for research, personnel development, and academic services. The computer centre is also the central network of the university to provide Internet connections to faculties, colleges, and educational institutes, on and off campuses, across the eastern region of Thailand (BUCC, 2007).

In the Faculty of Humanities and Social Sciences (HUSO), afterwards, the Information Technology Center (http://www.huso.buu.ac.th/itc/) or HUSO ITC was established as the faculty integral computer network. According to Sangsuriyong (2007), the former Head of the HUSO ITC, the e-learning program, *Chula ELS*, was initiated for students and faculty staff through cooperation between HUSO and CEC, Chulalongkorn University in 2002. The WBFA program was then created using the program through the faculty Web site: http://www.huso.buu.ac.th/els/Main/default.asp.

For a better understanding of its functions in the study, the description of Web-based assessment together with its benefits and challenges is presented in the following section.

2.4 Web-based assessment

2.4.1 What is Web-based assessment?

Web-based assessment (WBA) is a type of computer-based test that can be implemented by computers connected to the WWW (Olsen, 2000). WBA is processed by Common Gateway Interface (CGI) scripts. Web-based questions delivered through CGI scripts are usually in machine-correctable formats, for example, multiple-choice, fill-in, matching, or true/false questions, using checkboxes, radio buttons, or pull-down menus (Godwin-Jones, 2001). WWW can be an important, academic tool providing information to its users through hypertext. Hypertext can be information, data, still pictures, moving pictures and sounds (Malaiwong, 2000). In addition, Hypertext Markup Language (HTML) in WBA is frequently used to create and perpetuate questions items or tasks, and it can include texts, images, and even sounds. The HTML can be also used to select test topics, test databases, and grading programs (Poulton, 2001). Students can complete WBA at their convenience, regardless of time and place through the Internet access. This type of assessment is inexpensive to produce and maintain; it helps teachers conserve time for marking and provides students with instantaneous feedback (Roever, 2001).

In other words, WBA is a type of online or e-assessments. It is a testing system different from conventional paper-and-pencil assessment because there is no paper or

pencils involved. All question items are transferred into computer programs displayed through a computer screen. Students can perform WBA via the Internet access by reading the items from the screen, using a computer keyboard and mouse (a pointing device) to select or type answers they desire to respond to WBA questions. After completing WBA, students can receive prompt results of their assessment since WBA provides automatic score reporting.

2.4.2 Benefits of Web-based assessment

Researchers have stated several reasons for utilising WBA. Hamilton, Klein, and Lorie (2000) found that WBA provides large question item storage suitable for mass audience assessment. WBA increased flexibility and individual support for a large group of students of different ages, mixed abilities, variable computer experiences and diverse academic backgrounds, as noted by Kendall and Prowse (2005). This is in accordance with the reason that many institutes in academia have been developing their WBA program to cope with the increasing number of students because WBA can allocate a large number of students with flexibility of any time and anywhere access.

WBA also makes it easy to distribute questions, and provides flexibility for teachers to construct and edit test items (Dalziel, 2000). WBA provides for enormous item banks, and validity and reliability of the questions can be potentially secured (Maughan, Peet & Willmott, 2001). Questions on WBA can be scheduled automatically within seconds together with graphics and multimedia. Its electronic scoring reduces human marking errors. It is possible to reduce students' dishonesty in performing WBA through randomisation of question databases. Students can repeat doing the assessment as frequently as desired to achieve their academic target. According to pedagogical issues, students can create their own ways to attain the desired goals. This also encourages the idea of student-centred learning (Polyson, Saltzberg & Godwin-Jones, 2002).

Bull (1999) concluded that the most valuable benefit of WBA is its ability to provide prompt feedback which is an important motivation. Students can receive the response to their performance once they complete WBA, rather than waiting so long to receive the feedback from their teachers in conventional paper-and-pencil assessment. This is supported by Alagumalai, Toh, and Wong (2003) who emphasised that any assessment

is more effective when immediate feedback is presented to students. They concluded that WBA tenders great savings in not only test management time but also the vast amount of paper used, an important consideration in today's societies' needs to conserve.

For EFL learners, WBA is useful in terms of accessibility, renewability and adaptability because it can provide 24-hour independent access, be modified easily and often, and be easily adapted to support students at different English language proficiency levels or needs. Moreover, WBA can be applied to assess several kinds of language skills, for instance, vocabulary, grammar, and reading comprehension (Morrison, 2002).

Brown (1997) added that using computers in language testing has several advantages. He pointed out that computers are more accurate at marking and reporting scores than people. In addition, computers can report test scores more quickly to each student than teachers. Students can work independently at their own pace when using computers. They may feel less stressful than doing conventional paper-and-pencil tests because questions are presented one at a time on the computer screen. In addition, according to Aggarwal and Bento (2003), using the Web still sustains two-way interactions, both synchronous and asynchronous, for language students and teachers. Chat room and Web board provided on WBA are best examples of these interactions.

2.4.3 Challenges of Web-Based assessment

There are some problems associated with WBA. One is its security (Luecht, 2001). As they are not invigilated, it is unable to identify whether the students perform the assessment by themselves. Rovai (2000) noted that this problem may lead to academic dishonesty. McMurty (2001) suggested that an appropriate way to combat e-cheating be that teachers take time to explain and discuss academic honesty policy with their students because most universities have academic integrity policies to discourage cheating.

In addition, there are many strategies that teachers can use to measure student performance online to minimise students' inducement to cheat. Suggestions include having a login system using user name and password, utilising several short tests, or

creating large question pools for randomisation of questions (Olt, 2002; University of Illinois, 2007). Roever (2001) noted that online cheating is less applicable when WBA is used as low-stakes assessment. Low-stakes assessment is generated to give students feedback on their performance and progress in reaching their learning target. This assessment can be integrated with conventional classroom instruction. Low-stakes assessment can also be appropriate for summative test preparation because students can identify their strengths and weaknesses for further improvement.

2.4.4 Multiple-choice questions in Web-based assessment

There is apprehension related to the format of multiple-choice questions (MCQ) frequently used in WBA, especially in EFL courses because the focus of authentic language assessment should ideally be on real-life communicative tasks. Thus, researchers have argued that MCQ technique is less authentic when compared to other question formats used in language testing. They are concerned that MCQ may cause harmful backwash or impact negatively on communicative EFL learning and teaching. According to Hughes (2003), backwash will be detrimental when the test content and testing techniques are at odds with the purposes of the course. The test will also be harmful when it does not assess accurately. Inaccuracy occurs when the test lacks its reliability which may be induced by the test's feature and/or the way it may be scored. However, Hughes also argued that MCQ technique can be relevant to real-life tasks. For example, he described, in a reading comprehension test, that a shop assistant may identify which one of the four dresses a customer is describing. This can be found in a real-life situation, and it is in the multiple-choice format. He concluded that the most apparent benefit of MCQ is the fact that scoring can be trustworthy and therefore contributes to the reliability of the test. Scoring can also be done quickly and inexpensively. In addition, the MCQ technique is the most appropriate for intermittent testing of a great number of students. Bachman (1990) also supported that MCQ format is practical to construct and is also able to be an authentic form of assessment for EFL.

Objective tests using MCQ, according to Bull (2003), are appropriate for WBA which has a great ability of automatic marking. As a result, marking a large number of students can be done swiftly. MCQ diminishes the need for double or triple marking which usually occurs in essay-type assessment. In addition, teachers can test students

with a wide array of topics in a single assessment. Bull emphasised that these tests have been used in education for over 40 years. For instance, there have been used for large-scale testing in the American College Testing examination (ACT), the Scholastic Aptitude Tests (SAT), the Graduate Record Examinations (GRE) and the Law Society Admissions Test (LSAT) in the United States. Another good example is the Test of English as a Foreign Language (TOEFL). In the United Kingdom, the use of objective testing for formative assessments is accepted in various institutions, including the Open University, University of Derby and University of Loughborough. Relevant to pedagogical issues, the use of WBA in the delivery of objective tests enables the provision of automatic feedback to students. Objective tests are good at examining recall of facts, knowledge, application of terms, and questions that require short text or numerical responses.

To confirm the benefit of MCQ in WBA, MacKay and Emerson (2000) proved that the operation of a Web-based testing using MCQ was a positive influence on students' writing skills. Carneson, Delpierre, and Masters (2002) supported that MCQ format has many advantages. MCQ can be marked electronically, and the scoring can be both accurate and objective. Additionally, MCQ can be focused at every cognitive level: knowledge, comprehension, application, analysis, synthesis, and evaluation, on Bloom's Taxonomy. It is also statistically easy to validate the question items in terms of their difficulty and discrimination indexes appropriate for students of differing competence. It can be administered frequently and easily to provide regular and systematic information on students' performance. These questions can provide a better coverage of content than essay-type questions. It also provides a quick diagnostic feedback on students' progress. Although there are several forms of questions available, McKenna and Hesketh (2000) indicated that the most common form of assessment on the Web is the use of MCQ. This is in agreement with Khan, Davies, and Gupta (2001), and Ryan, Scott, Freeman, and Patel (2000).

Jefferies, Constable, Kiely, Richardson, and Abraham (2000) found that using MCQ in WBA is also beneficial to teachers whose academic workloads have been increasing in all aspects of the learning and teaching process. The potential for using WBA is attracting, an increasing level of interest of teachers teaching a large number of students. In their study, participants were 80 first-year undergraduates in one foundation

course. MCQ was utilised in a graded assignment to measure students' skill pertaining to factual content in the course. WBA was chosen because it provided automatic marking, rapid feedback and quick statistical analysis of results. The WBA, installed on a university server on campus, was generated from randomisation of 25 pools of five MCQ questions in each pool. Students had to complete the WBA within 50 minutes and had only one attempt to perform the assessment. Their names and scores obtained were recorded. Most of the questions were taken and adjusted from Teachers' Supplement which was used to support the course text. The assessment was trialled a week before its implementation for validation purposes. All participants took the assessment during one of their computer laboratory sessions. They were required to complete the WBA in one week and most of them obtained relatively high scores in the assessment. After the completion of the WBA which involved some technical problems due to slow response times from the server, participants were asked to respond a questionnaire to capture their attitudes toward the WBA experience. The return rate was 48.75% (*n = 39*). The results revealed that using WBA was a valuable experience as preparation of summative MCQ assessment with a positive rate of 87%. The majority of students (85%) accepted that the assessment provided them very useful benefits of technology in assessment, especially in the large foundation course.

However, the literature (Higgins & Tatham, 2003) identified some problems associated with MCQ; for instance, constructing good MCQ may be time-consuming, and students' creativity cannot easily be tested with this question format. There are also some disadvantages which relate to the difficulty of designing alternative wrong answers and dealing with guessing. Such a phenomenon is considered less significant when this question type is used for formative purposes as opposed to summative and high-stakes assessments. MCQ in WBA is selected as appropriate for use with large, first year foundation courses. The ability to provide immediate feedback (e.g., prompt numerical score reporting) to students is its key benefit (Benson, 2003).

Authenticity is another concern when multiple-choice questions are utilised in the assessment. However, WBFA would be used as student self-assessment, according to this, O'Malley and Valdez Pierce (1996) stated that:

> *Self-assessment promotes direct involvement in learning and the integration of cognitive abilities with motivation and attitude toward learning. In becoming self-regulated learners, students make choices, select learning activities, and plan how to use their time and resources. They have the freedom to choose challenging activities, take risks, advance their own learning, and accomplish desired goals.* (p. 5)

To sum up, assessment is the essential factor of the educational process since without it we are unable to know how much our students learn. There are at least six categories of assessment: achievement, placement, formative, diagnostic, summative, and self-assessment (Born, 2003). Objective testing can be the fundamental approach to perform these types of assessment. Cucchiarelli, Panti and Vanlenti (2000) argued that objective testing can be the fundamental approach to perform these types of assessments. However, as noted by Astin et al. (2003), assessment works best when it is ongoing. That is formative assessment. The use of objective testing in WBA for formative purposes is the focus of the following section.

2.5 Web-based formative assessment

2.5.1 What is Web-based formative assessment?

Web-based formative assessment (WBFA) is WBA used by teachers for tracking students' progress to inform and adjust the educational process and to improve students' learning outcomes (Born, 2003). Formative assessment applies to all activities attempted by both teachers to assess their work to adapt the teaching and students, to meet their needs (Black & Wiliam, 1998b). It also refers to its use for diagnostic purposes to provide feedback to teachers and students over a course of instruction (Boston, 2002; Robles & Braathen, 2002) and to promote further improvement of students' attainment (Crooks, 2001). Formative assessment is well recognised as a powerful means to improve learning and assist students to learn better (Harper, 2001; OECD, 2005), thus justifying it as a most important assessment practice for supporting learning. Quality formative assessment has been shown by FairTest (2007), and Maughan, Peet and Willmott (2001) to have a crucial influence on students' learning when placed online, thus contributing to a richer educational experience. It is also argued that WBFA can be authentic since it can be designed to require students to work

in real-life contexts through a variety of interactive activities that may involve multimedia.

Hanson, Millington and Freewood (2001) and Mulligan (1999) noted that computerised tests, when used as objective tests with formative purposes, provide an efficient approach to learning since the use of frequent continuous feedback through such formative assessment has a positive impact on students' work. Moreover, the potential use of a multimedia assessment system, with automated grading and electronic monitoring of students' performance, can assist in improving traditional approaches to course materials as well as supply additional personal attention to students while recording the students' behaviour on the computer system (Caygill & Eley, 2001; Wong, Wong & Yeung, 2001). In EFL, by providing students with greater autonomy with networked environments with WBFA, it is expected that they will develop a self-directed approach to their language learning (Macdonald, 2002).

2.5.2 Benefits of Web-based formative assessment

Blayney and Freeman (2003) and Dow (2003) stated that an important advantage of Web-based formative assessment is its timely and effective feedback (e.g., automatic scoring) reported to students. In relation to the feedback, Emberger (2002) recommended that simply offering right or wrong answers to students has a negative effect on achievement and giving students with correct answers has a moderately positive effect on students' learning. In addition, explaining what is correct and what is incorrect has a greater effect. However, allowing students to continue working on a task until they are successful has the greatest effect. Emberger concluded that the more immediate the feedback is provided, the more learning improvement occurs.

Another benefit was that a combination between objective tests and WBFA might overcome the challenges of large enrolment courses involving several thousand first year university students. For instance, Henly (2003) created an objective item bank for formative assessment to assist students to develop the skills necessary to become lifelong learners. Questions included short-answer, multiple-choice and matching questions. The result showed that eighty per cent of students regarded the assessment as

helpful for their study. In addition, Cross (2001) introduced a computer-based formative assessment providing immediate feedback to several hundred first year university students. The objectives were to determine students' perceptions toward the assessment and to improve the availability of teaching staff. Participants were students who were randomly allocated to two groups. The first group used paper-based assessment and wrote their answers to submit to their lecturers while the other used computer-based assessment, which was similar to the first group, but they used computer to submit their answers, for the first four weeks. Both groups reversed their roles in the fifth week. Then for the last four weeks, both groups independently used either form of the assessment. The response rate from questionnaires to participants returned for analysis after the fifth week was 81% (42 returns), and after the ninth week was 88% (48 returns). Results revealed that students viewed computer-based formative assessment as helpful in learning. It also enhanced feedback to staff for improving teaching and offering even-handed treatment to students. The study concluded that the assessment reduced frustration which students experienced when waiting for feedback from conventional paper-based formative assessment. At Burapha University the lecturer student ratio is one lecturer to 65 students making it impossible to have a fast formative assessment turn-around-time.

In EFL, Finch (2000) utilised WBFA to promote communicative competence in English for a large number of first year university students. Qualitative and quantitative results, obtained through questionnaires, interviews, learner journals and self-assessment revealed that the language learning is positively affected by promotion of learner autonomy and formative assessment is an effective method of language program development.

Similarly, Hayden (2002) concluded that WBFA was optimistically accepted by both teachers and students. Teachers viewed that the assessment was time-saving and inexpensive to implement statistical analysis of the assessment questions. It provided more feedback than paper-based assessment did. In addition, students viewed the assessment as providing prompt feedback and a good support for learning. It allowed them to take the time they needed for the assessment and with good access, students prefer WBFA to paper-based assessment in the future.

2.5.3 Challenges of Web-based formative assessment

According to Ellis and Ratcliffe (2004), there are four problems related to using WBFA. Firstly, it is difficult for teachers to use a limited number of questions again after the correct answers are provided to students without the risk of students' recognising the answers even though those questions are randomised. Secondly, students using formative assessment may not try their best effort to perform the assessment as they do to summative assessment because they think that WBFA does not have vital influence on their course when the formative scores are not counted toward the final grading at the end of their semester. Thirdly, students may not receive the feedback that they need. If they do not receive explanations why their answers are incorrect, it may be difficult for students to understand by themselves. Finally, teachers may not obtain the feedback that they require from the students. If the testing for summative assessment is not supervised, teachers may have difficulty being convinced on how well their students are learning the materials in their course.

2.6 Web-based formative assessment and students' achievement

The relationship between WBFA and students' achievement has been the focus of researchers. For example, Dufresne, Mester, Hart and Rath (2002) investigated Web-based homework (WBH) for its effect on students' academic achievement at the University of Massachusetts at Amherst. The system was used for a class of 290 students and available 24 hours a day, seven days a week. Students had to log on through a Web-based system to do regular homework assignments, which were graded automatically. For the purposes of assessment security, the variable values were randomly assigned by the system each time the online problem was generated. Overall results, measured by exam performance, revealed that students utilising WBH had higher achievement. In another example, MacKay and Emerson (2000) used a Web-based testing facility with multiple-choice questions stored in databases in a compulsory course for first year university students. The results showed that this Web-based testing had a positive effect on students' learning, particularly in a large class of first year university students.

In addition, Buchanan (2000) used WBFA in an undergraduate course. He aimed to provide prompt and useful feedback to students since it was difficult to do so with a large number of students in conventional classrooms. Multiple-choice tests were used in the assessment as an efficient way to serve the purpose because they were easily administered and scored by computer programs. With WBFA, students could repeat taking the same test to measure their performance for themselves after class during the semester. The participants were 232 undergraduate students of mixed ability and background. They had to enter their identification number, which was recorded, when using the WBFA. All participants had to perform three exercises, each comprising 11-15 multiple-choice questions assessing factual knowledge of topics covered in the module. They also had to do two additional revision exercises, in which 10 questions were randomly selected from a large pool. The exercises could be accessed from any networked computer on or off campus. Access to each question was permitted only once. Participants received instant feedback once they completed their tests. There were no correct answers given. Instead, participants were provided with appropriate sections of textbooks to find out correct answers for the incorrect ones they did. It was suggested they repeat the test after studying the relevant learning materials. Results revealed that for the remaining and available 148 participants, the number of their WBFA uses had a positive correlation with their final exam performance at the end of the semester.

In the teaching of EFL, Kilickaya (2007) examined how computer-assisted language learning had impacts on learners' TOEFL scores through a quasi-experimental study. It emphasised on investigation whether there was a significant difference of students' achievement in a computer-assisted language learning group and a traditional group in preparation for the TOEFL exam, related to structure, reading and listening sections. The participants, aged between 18 and 20, were 34 second year undergraduate students (29 females and 5 males). They were recruited through convenience sampling. They were randomly assigned to the experimental and control groups using a table of random numbers. Each group consisted of 17 participants. The experimental group used computer-assisted instruction in a language laboratory while the control group employed traditional classroom setting. Both groups were taught by the same instructor during the eight-week study. Pre- and post-tests were used as research instruments to collect the data. The questions in both tests, consisting of 140 multiple-choice items, were taken from the book, TOEFL Test Preparation Kit Workbook. The pre- and post-

tests were in identical paper-based version. Test takers received 1 point for each correct answer and 0 for each incorrect one. On the first day of classes, after participants signed student consent forms, the lecturer administered the pre-test to the control and experimental groups in the same class. Students in the experimental group worked on a computer using the provided programs and learnt by themselves. The lecturer's role was to make sure that they worked alone on their computers. Students in the control group attended a traditional lecture class following the lecturer's schedule. Students in both groups studied in this way for three hours a week. For students in the experimental group, English Grammar in Use and Cambridge Advanced Learner's Dictionary were used together with the CD versions of these materials and preparation for the TOEFL test software by ETS. For the control group, practice tests on the CD were converted to paper-based tests. On the last day of classes, the lecturer administered the post-test. The scores obtained from the pre- and post-tests were statistically analysed through an independent samples *t*-test. Results revealed that there was no statistically significant difference, according to overall gain scores, between the control and experimental groups. There was also no significant difference between the control and experimental groups' scores on the structure section of the TOEFL. However, there were statistical differences between the control and experimental groups' scores on the reading and listening sections of the TOEFL.

2.7 Web-based formative assessment and students' attitudes

Students' attitudes toward the use of WBFA are further considered relevant to success in testing through this mode. It has been shown that students' attitudes toward WBFA may influence their learning outcomes. Cook (2000) found that the positive perspectives on the use of technology could greatly increase students' confidence and would lead to higher learning achievement. The evidence to date (e.g., Al-Amri, 2007; Higgins, Russell & Hoffman, 2004) has suggested that the majority of students have positive experience toward multiple-choice, online computer-based assessment, specifically in modern language modules. Accessibility and immediate feedback were reported as the crucial factors contributing to success. Findings indicated that students found this online assessment was less stressful than the traditional paper-based one (Beverly, Beverly, Clarke & White, 2001; O'Hare, 2001).

Arkkelin (2001) investigated students' attitudes toward online testing before and after they experienced online testing and found that most students liked the immediate feedback of their test results and the ability to review the individual questions. Findings revealed that students had significantly more favourable attitudes toward online exams after than before experiencing online exams. Students had greater knowledge, less fear and a more favourable attitude toward computers after experiencing online testing than before. In addition, Walker and Delius (2004) employed online assessment to motivate 182 undergraduate students and to solve problems experienced with slow feedback to students in traditional assessment. They stated that the delay in receiving feedback discouraged students to learn from their mistakes. Two surveys, five-point Likert scales with open-ended questions, were employed. The first was administered after the training session at the beginning of the academic year, and the second at the end of the semester along with a follow-up interview. Results revealed that the majority of respondents (85%) liked the automatic feedback and chances to reattempt the assessment.

Cheung (2004) investigated the perceptions and attitudes of ESL student teachers toward the learning and use of media technologies. A total of twelve second-year undergraduate students were randomly selected to complete the questionnaires and to be interviewed with questions aimed at drawing some in-depth reflections from their learning process. Each of the respondents was asked multiple-choice, ranking, and open-ended questions. The results revealed that 84% of the respondents believed that media technologies could help stimulate learners' motivation and interest in language learning. Also, 75% of the respondents stated that technologies encouraged shy or passive learners to learn more actively while 92% of the respondents stated that they enjoyed using a computer. The majority (92%) found communicating online an enjoyable experience while 58% of the respondents stated that they tended to participate more actively in communication in an online environment than in a face-to-face mode. Similarly, Lin (2003) indicated that 80% of EFL learners in her study appreciated opportunities to practice and extend their language abilities by surfing the Internet to develop reading skills when she examined attitudes of 46 first year junior college students.

In the same study mentioned in the previous section, 2.6, Kilickaya (2007) also examined learners' perceptions related to the use of computer-based assessment. Seventeen participants were interviewed one by one in English after using the assessment for eight weeks. Forty-seven per cent of the participants stated that they did not feel comfortable working with computers while 53% felt comfortable using computers. Eighty-two per cent of the participants agreed that the assessment provided great motivation because they could study as much as they desired. Seventy-one of the participants liked to practise listening skills through the computer-based assessment since they could spend more time on the activity compared with a teacher testing listening.. Listening skills were also found to be ignored or less important in traditional classrooms where students listened to a passage or a dialogue for a few minutes. Twenty-nine of them responded that they liked to practise English structures on the computer-based assessment. However, all participants indicated difficulties in practise reading skills on the computer-based assessment. They complained that they could neither underline the important points nor see a reading passage as a whole because they did not like scrolling up and down to read on computer screens. According to pedagogical issues, some students mentioned that they preferred to have a lecturer to assist them to answer specific questions. At the end of the program, 29% of the respondents accepted that they still had negative feelings toward language computer-based assessment while 71% stated that they had positive feelings.

2.8 Web-based formative assessment in large enrolment courses

The Australian University Teaching Committee or AUTC (2001) defined a class of 80 or more students as a large class. However, in language instruction, a class with 50 or more students is referred to as a large class. AUTC also indicated several problems related to teaching and learning in large classes. For instance, large classes limit students' opportunities to receive instant feedback and to interact with lecturers. In assessment, teachers in large classes encounter heavy marking and feedback loads. One potential solution suggested is to promote self-assessment to students, especially in courses which aim to improve students' skills. Large classes also increase complexity of teaching since teachers have to deal with a more diverse group of students, have more problems in communicating with students, suffer greater managerial burdens, and face difficulties in promoting active participation and scrutinising student progress.

As a result, in large enrolment EFL classes, the opportunity for individual interaction between students and lecturers is very difficult. Hence, the quality of individual learner feedback and support is largely ineffectual. Moreover, lecturers cannot effectively handle instant and interactive feedback, which is considered essential to students in language assessment. As a result, WBFA should be utilised to administer multiple-choice assignments to reduce the faculty marking workloads and deliver immediate feedback to individual students. This system was proved to be a time saver because automatic scoring systems were more practical in both quantity and frequency (Bonham, Titus, Beichner & Martin, 2000).

Arising from a pilot study that found frequent computer-based assessments influenced the work rates of students, Wong, Wong and Yeung (2001) developed Web-based formative assessment as an intervention in large classes of approximately 800 students and investigated students' achievement through their final examination performance. Participants were first-year computer science students at Hong Kong Baptist University. They created WBFA called *Quadrille* to assist lecturers with heavy workloads in marking conventional paper-and-pencil assessment and to support students with instant feedback. The program was supposed to help students to measure their understanding of general knowledge that was taught in lecture classes so they could identify students' areas of weaknesses and intervene on their problems. The program also aimed to examine students' behaviour, performance and perspectives toward the system. To perform the assignment, each student had to provide a user ID and a password to the system. The students were able to use the system to complete their work anytime they desired. The assignment scores were not counted for the course grade. In addition, completing the assignment was not compulsory because it was employed as formative assessment. On completion of the assignments, the students were provided with immediate feedback including correct answers and explanations or hints to engage their understanding of the questions. Instant feedback regarding performance of individual students and the whole class with the average number of attempts and the average time used on each question was also provided to the lecturers. The questions were randomly selected from the question bank of eight types of question formats, for example, multiple-choice, true/false, and fill in. The content in those questions was related to the students' textbook used in the course.

The system was introduced at the commencement of the semester; however, it was the fifth week that the assignments were released after the students had studied the relevant lectures. The students were able to do their assignments until after their final examination. Since completing the assignment was voluntary, the students' performance rate was low during the semester. The frequency increased before the final examination. Most students did these online assignments between 11 a.m. and 8 p.m. Seventy-seven per cent of the students preferred to perform the assignments on campus while the rest did them outside the university. Results revealed that there was a correlation between the number of assignments completed and the students' achievement scores in their final examination. In other words, students who did more assignments got higher scores in their final examination. With regard to students' attitudes toward the system shown in the questionnaire delivered to each student at the end of the semester, more than 85% of 445 students who returned the questionnaire agreed that the system was user friendly. Around 78% of these respondents accepted that the WBFA helped motivate them to do the assignments while 79% mentioned that the WBFA was useful or very useful. However, some students complained that it took long time to download images. Some stated that they did not like to use computers.

In another review, Eustace (2003) created a project called SLATE (Supporting Learners using Audio Tutorial e-learning) to investigate, implement and evaluate an online assessment for a large group. More than 200 first-year students at the University of Salford studying an introductory engineering course were participants. For formative purposes, both short and long formative assessments were taken immediately after the students had studied related matters. The WBFA provided to the students were in different formats, for example, multiple-choice and drag and drop. These question formats were similar to those at the end of the module to allow the students to become familiar with the summative assessment. The students received feedback when they submitted their answers to the system and that feedback provided them with correct answers with relevant explanations. Feedback from students was generally good. Students in large classes felt that computer-based assessment was less threatening than traditional written examination and they felt much more relaxed taking the test. Automatic marking of the assessment provided a significant time saving. This assessment approach is certainly worth taking for large groups of students and changing

from traditional written examination to a multiple-choice format reduced the marking time.

In short, the use of WBFA, with its capabilities of automatic marking, prompt score reporting, all-time accessibility, encouraging self-efficacy and real-world assessment via multimedia systems, appears to be a positive strategy to improve students' academic achievement and attitudes toward the use of technology in learning English as well as a promising approach to help reduce heavy marking workloads for language instructors.

2.9 Summary

This chapter has looked at the four theoretical perspectives of second language acquisition, language pedagogy, language assessment and technology integration, which underpin and guide this research. Then it has described the background of the EFL teaching situation before the introduction of the WWW to Thailand and how the WBFA was introduced to educational institutions in the country, and more specifically how it was integrated into large EFL classes. Later it has presented a review of the literature relating to issues in language testing and specifically related to Web-based assessment, the concept of Web-based formative assessment, and Web-based formative assessment and students' achievement, Web-based formative assessment and students' attitudes, and Web-based formative assessment in large enrolment classes.

In Chapter 3, the details in methodology employed in the study are described.

Chapter 3 Methodology

This chapter describes the research method design employed to collect data for the study. It is categorised into five sections. The first section (3.1) is a brief description of the design of the study. The second section (3.2) is relevant to participants recruited for the study. The third section (3.3) is related to research instruments designed to apply for mixed methods data collection. The fourth section (3.4) is pertaining to the procedure for gathering the data. The final section summaries the chapter.

3.1 Research design

The mixed methods research design, which focuses on collaboration of a quantitative method and a follow-up qualitative method, is utilised to conduct the study. The former method is exploited to collect numerical data to investigate students' achievement mostly through a pre-test, a post-test and the WBFA program. The latter method is applied to collect qualitative data from questionnaires and semi-structured interviews to determine students' attitudes toward WBFA. Rationales for selecting the mixed methods design used in the study are described in the following subsection.

3.1.1 The mixed methods design

In order to investigate both students' achievement and attitudes influenced by the WBFA program, the mixed methods research design was implemented. This research design provides several benefits. As noted by Greene and Caracelli (1997), the main purpose of the mixed methods design is to seek a wider range of interests and perceptions. Utilising this research design may provide more supports to the findings. Because all methods have drawbacks, applying the mixed methods research design may help reduce any weaknesses. In addition, according to Tashakkori and Teddlie (2003a), this research design provides simultaneous, confirmatory and exploratory answers to questions in the same study. It supplies better and stronger inferences and a greater diversity of views. Moreover, Frechtling and Sharp (1997) viewed that the mixed methods design yields richer, more valid, and more reliable findings. This is likely to increase the acceptance of findings and conclusions by the diverse groups of stakeholders in the evaluation because it provides both the ability to generalise and the

perceptions and reactions of the target population. Johnson and Turner (2003) concluded that the fundamental principles of mixed methods research are "(1) to obtain convergence or corroboration of findings, (2) to eliminate or minimise key plausible alternatives explanation for any conclusions drawn from the research data, and (3) to elucidate the divergent aspects of a phenomenon" (p. 299).

With reference to the two methods in the design, Maxwell and Loomis (2003, p. 241) defined that the quantitative method involves experimental planning, deductive approaches, control of extraneous variable, formal hypothesis testing, and theory confirmation whereas the qualitative method relates to naturalistic setting, inductive approaches, detailed description, concentration to context, and the rigorous analysis of particular cases. Regarding mixed methods research, Creswell (2005) viewed that:

> *Quantitative research is a type of educational research in which the researcher decides what to study, asks specific, narrow questions, collects numeric (numbered) data from participants, analyses these numbers using statistics, and conducts the inquiry in an unbiased, objective manner. Qualitative research is a type of educational research in which the researcher relies on the views of participants, asks broad, general questions, collects data consisting largely of words (or text) from participants, describes and analyses these words for themes, and conducts the inquiry in a subjective, biased manner.* (p. 39)

However, the mixed methods design also has some challenges. Tashakkori and Teddlie (2003a) expressed their concern in using this approach that there may be "confusion between the quality of data and the quality of inferences that are made on the basis of the analysis of such data" and that there are some "controversies regarding standards for evaluating inferences" (p. 38). Additionally, Johnson and Christensen (2004) notified that in mixed methods research design "the researcher has to learn about multiple methods and approaches and understand how to appropriately mix them" (p. 414) and "it is more time consuming" (p. 414). Conversely, they accepted that mixed methods research offers the ability to draw valid and reliable conclusions through convergence and confirmation of findings.

In the study, an explanatory mixed methods design with a sequential explanatory strategy (Creswell, 2002) is employed. The explanatory mixed methods design is straightforward and easy to implement. In the sequential explanatory strategy for data

collection, the priority is on the quantitative data collection and analysis, and then the smaller qualitative data are gathered and analysed to explain the quantitative results. This aims to use the results from the qualitative data analysis to help explain and interpret the unexpected findings of the primarily quantitative method. As a result, the two methods are incorporated or mixed in the interpretation phase of the study. In addition, the sequential explanatory strategy provides clear, separate stages for data collection making it easy to describe and to report. The samples engaged in this research design are usually homogeneous and gender is not a main concern. However, it takes long time because of more participants, more research instruments, more data collection and analyses demanded in the two phases (Creswell, 2005).

3.2 Participants

Kemper, Stringfield and Teddlie (2003, pp. 283-284) suggested that mixed methods sampling strategies can greatly fortify the research design of most studies in order to increase both inference quality (internal validity and trustworthiness) and generalisability or transferability. Two types of samples, a probability sample and a non-probability or purposive sample, are advised to fulfil this opportunity. Probability sampling techniques are to enhance inference quality, for example, simple random sampling, systematic random sampling, stratified random sampling, and cluster random sampling. Non-probability or purposive sampling techniques are to confirm generalisability, for instance, convenience sampling, extreme case sampling, homogeneous case sampling, and opportunistic sampling.

Hence, the recruitment of the samples in the study was operated in two phases, according to the mixed methods research design and its sampling techniques. The first phase was for the sampling of quantitative data, and the second phase for qualitative data. Stratified random sampling, one of the probability sampling techniques, was used to select samples from nine faculties and colleges for the quantitative method. In addition, opportunistic sampling, which is one of the non-probability or purposive sampling techniques, was employed to recruit samples for the qualitative method.

3.2.1 Quantitative method sampling

In the phase of the quantitative method data collection, the stratified random sampling was applied because representatives from every discipline in the population were needed in the study. Moreover, Denscombe (2003) underlined the benefits of this sampling technique that:

> *A stratified sample can be defined as one in which every member of the population has an equal chance of being selected in relation to their proportion within the total population. The significant advantage of stratified sampling over pure random sampling is that the social researcher can assert (claim, stress) some control over the selection of the sample in order to guarantee that crucial people or crucial factors are covered by it, and in proportion to the way they exist in the wider population. This obviously helps the researcher when it comes to generalising from the findings of the research.* (p. 13)

3.2.1.1 Quantitative sample size

The participants in the study were 186 first year undergraduate students recruited from nine disciplines through the stratified random sampling technique. The number of participants randomly selected from each discipline is illustrated in Table 3.1. They were all Thai students with the average age of 18. These students enrolled to study English II (212102) at the Department of Western Languages, Faculty of Humanities and Social Sciences, Burapha University, Bang Saen campus, Chon Buri, in the second semester of 2004-2005 academic year (i.e., November 2004 – March 2005). The course was a three-credit compulsory foundation subject for first year undergraduate students.

Prior to their enrolment, all first year undergraduate students were classified into three groups due to their scores in the English test taken earlier in the 2004 national university entrance examination. In the first group, there were students who obtained no more than 45% of the total English score in the entrance examination, and they were required to study English 1 (212101) and English 2 (212102). The second group consisted of students who received 46-64% of the total English score in the entrance examination, and they were required to study English 2 (212102) and English 3 (212103). Students who got higher than 65% of the total English score in the entrance examination were in the last group, and they were required to study English 3 (212103) and English 4 (212104) in their first and second semesters, consecutively. The students

in the first group were the majority of the first year undergraduate students and were randomly selected as the sample for the study.

Table 3.1 A stratified random sample based on student faculties and colleges

Faculties and Colleges	Number of students in each discipline	The sample (10% of the population)	Percentage of samples in each discipline
1. Faculty of Education	239	24	12.9%
2. Faculty of Engineering	230	23	12.4%
3. Faculty of Fine and Applied Arts	115	12	6.4%
4. Faculty of Humanities and Social Sciences	538	54	29.0%
5. Faculty of Nursing	83	8	4.4%
6. Faculty of Public Health	132	13	6.9%
7. Faculty of Science	402	40	21.5%
8. College of Sport Science	72	7	3.8%
9. Maritime College	51	5	2.7%
Total	1862	186	100%

After their enrolment to study the course, the students in the first group were divided into 35 classes, approximately 50-80 students from several disciplines in each class. Typically, the students had to attend three face-to-face classes on two different days a week. Each class lasted 50 minutes. The first two periods were scheduled on the same day for the students to attend a lecture class. The last period was on a different day scheduled for the conventional paper-and-pencil assessment (CPFA) where students completed dialogues, read passages, and did vocabulary and grammar tasks in their workbook. Occasionally, the CPFA was taken as homework. The CPFA was corrected by lecturers who taught the course in each class. The students were usually able to receive individual feedback about a week after the assignment although this constituted a very large, time consuming task for lecturers. The students had to purchase their textbooks and workbooks at the beginning of each academic year from the department which were commercial books previously and carefully selected for them. The books were subject to change each academic year with the approval of the University's Department of Western Languages.

In relation to the assessment during the 16-week semester, there were two summative assessments to evaluate the students' language achievement—the midterm and the final examinations. The content of both examinations was in accordance with the course

objectives, comprising 70 multiple-choice questions, categorised into four main parts: *completing dialogues, reading comprehension* from passages in the textbook, *vocabulary, grammar,* and *reading comprehension* from unseen passages. Both examinations were based on conventional paper-and-pencil assessment, and weighted 35 % each which comprised a total of 70% of the course grade in total. The CPFA required 5 %, while in-class activities were counted 10% of the course grade. Another 10% was for external reading tests administered outside the summative assessment schedule. Therefore, the students had to complete their external reading assignments in the reading room for the last 5% of the course grade.

With regard to the selection of participants for the study, the researcher contacted the Faculty of Humanities and Social Sciences at Burapha University and then received ethics clearance from the University of Southern Queensland for permission to collect the data (see Appendix G). All students in the first group were randomly selected through the applications of the Table of Random Number (Fisher & Yates, 1963). To provide an equal opportunity to every student to be a representative of their discipline in the study, a stratified random sampling technique was applied. A 10% stratified random sample of 186 first year undergraduates selected from a total of 1,862 students was then recruited. The participants were informed of the objectives of the study and student consent forms (see Appendix H) were signed by them. The participants were equally separated into two groups on the basis of student faculties and colleges. That is from each discipline group names were placed in a box and drawn out one by one taking it in turn to allocate to each form of assessment. Therefore, every participant selected in a discipline through the random stratified process had an equal chance to go to either the WBFA group or the CPFA group. As a result, the number of students representing each faculty and college in both WBFA and CPFA groups were identical. It should be noted that one more student was randomly added to each of the following disciplines, the Faculty of Engineering, the Faculty of Public Health, the College of Sport Science and the Maritime College in order to make even numbers of students in WBFA and CPFA groups in these disciplines. Therefore, the actual total number of the participants was 190 (i.e., 95 for each group). Students in the WBFA group were assigned to use the WBFA program, instead of the CPFA, for eight weeks during the release of WBFA.

3.2.2 Qualitative method sampling

To strengthen the mixed methods research design, Rudestam and Newton (2001) explained that, for qualitative data collection, researchers are more interested in understanding what is going on in a specific phenomenon. Hence, sampling issues for qualitative purposes focused on the selection of particular participants, settings, clusters, or circumstances to be questioned or interviewed, rather than the quantity of the samples. The benefit is that only a few cases may be needed to substantiate that a definite quality or attribute exists. The drawback is that it is difficult to control the sample composition to be able to corroborate generalisability. In the phase for the qualitative method, non-probability sampling or purposive sampling techniques were involved. This is quite different from the sampling techniques for the quantitative method because it is definitely not a random selection (Denscombe, 2003). Participants were chosen because they suited the criteria demanded by the researchers that were distinctive to the group under investigation (Ellis & Barkhuizen, 2005). Hence, opportunistic sampling, one of the purposive sampling techniques, was employed to recruit samples for the semi-structured interviews in the study. The researcher selected some specific students whom he thought would provide the most valuable data pertaining to attitudes toward using the WBFA program in the course. The students included those who used the WBFA program regularly and those who did not.

3.2.2.1 Qualitative sample size

Participants in the pre-WBFA semi-structured interview were six students (7.89%) from the WBFA group. The interviews were conducted at the time of the WBFA tryout in the first semester. In addition, another ten students (13.16%) from the WBFA group were recruited for the post-WBFA semi-structured interview in the second semester. Those participants were selected from the same WBFA group, but they were not the same students. These participants in both interviews were selected regardless of their gender and discipline. The participants were well informed of the objectives of the interviews and they all accepted to sign the student consent forms prior to both interviews.

3.3 Materials

A total of five instruments were created to collect the mixed methods data: WBFA, a pre-test, a post-test, a questionnaire, and a semi-structured interview. Details of these materials are described in the following subsections.

3.3.1 WBFA

The WBFA program aims to assist the students to make a succession toward learning English in their course and their summative assessment based on the course objectives. This type of assessment, as noted by Hughes (2003), can be served as an investigation of the students' improvement in order to keep their learning on the right track. In addition, WBFA is not high stakes testing so the process of construction is less time consuming and the items can be used over and over again. WBFA should be seen as evaluating progress toward the objectives on which the summative assessment is based.

3.3.1.1 Development of WBFA

Prior to the WBFA construction, there was consultation with Burapha University and the researcher contacted the Department of Western Languages and Faculty of Humanities and Social Sciences at the University for the preparation of the Web setting for WBFA (see Appendix E). Afterwards, it was related to the designing of the WBFA. The English language content of WBFA was designed by the researcher in line with the English II (212102) course syllabus. Related connections for communicative and authentic language activities were linked to WBFA as well as instant feedback to facilitate and stimulate students' language learning and outcomes.

The software program used for creating WBFA was installed on the faculty Web site and generated through the Web site: http://www.huso.buu.ac.th/els/main/default.asp. The Web page of the software program is illustrated in Figure 3.1. The Web is under the cooperation of the Continuing Education Center (CEC), Chulalongkorn University and the Faculty of Humanities and Social Sciences, Burapha University (see Appendix E). The program is the Internet course management system, which is able to serve both Thai and English languages and fonts. In WBFA, only the English language and fonts

were used. The system includes databases for students, teachers, system administers, online courses and examinations, which run with Active Server Pages (ASP). The system can be used with several databases, for example, Oracle, SQL Server 2000, and Microsoft Access. It also provides both synchronous and asynchronous communication through its Web board and chat room (see Appendix A).

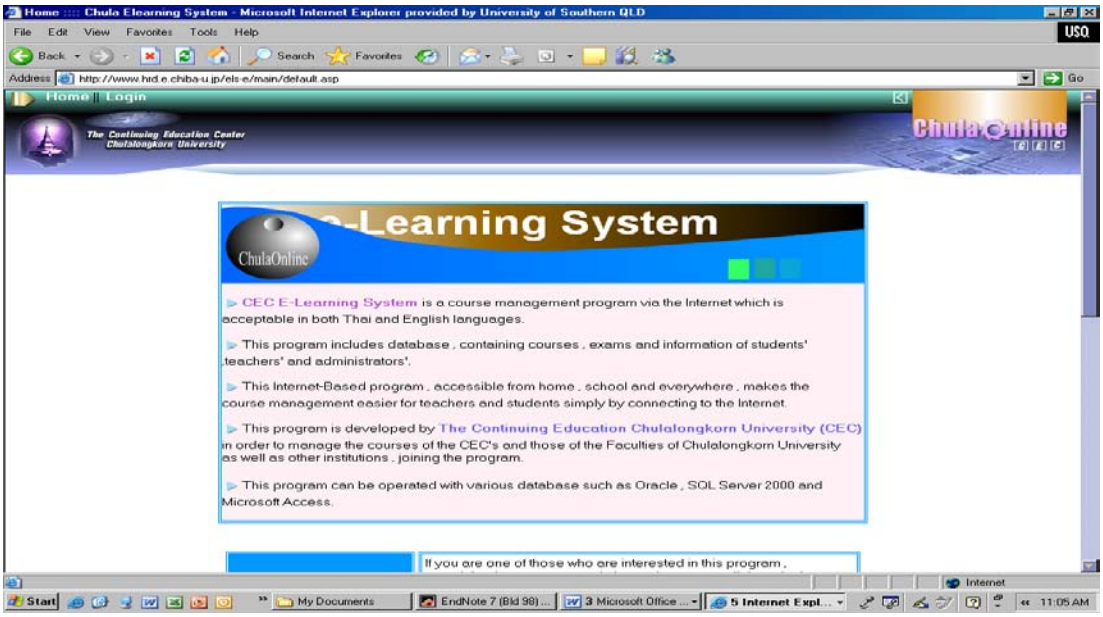


Figure 3.1 ChulaOnline home page

When constructing WBFA, the researcher was able to download and upload the test questions via a variety of applications (see Figure 3.2), including links to other useful Web sites to emphasise the provision of authenticity, which encourages students to be familiar with using the language in real-world situations. This should be another important advantage of WBFA. The students were able to take this benefit when they logged in to the program and clicked to the connected Web sites relevant to their modules in the class before they were ready to perform WBFA. However, there was no online record for this activity. The system was able to record only students' performance on WBFA.

The content of WBFA was designed in accordance with the content of the students' textbook and the objectives of the course. The objectives of the course were: (1) to encourage students to practice English listening, speaking, reading and writing skills,

(2) to enable students to learn basic English vocabulary and structure, (3) to assist students to understand contextual meanings, and (4) to support students to identify the main ideas of selected readings.

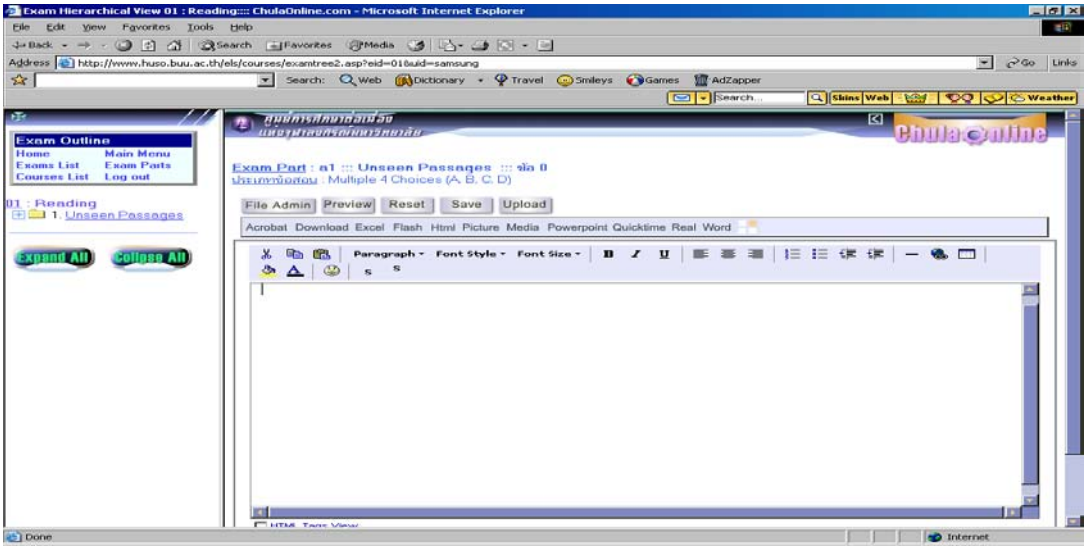


Figure 3.2 WBFA question item constructing page

WBFA was based on objective test questions, which were classified into three types: multiple-choice, fill-in and true/false questions. The question format of WBFA was equivalent to that of CPFA. WBFA contained 12 test sets, with ten multiple-choice questions each, for students to practice in the *completing dialogues, reading comprehension*, *vocabulary*, and *grammar* sections. In each section, the students were able to perform 30 multiple-choice questions to evaluate their own strengths and weaknesses. To complete the program, students had to answer all 120 questions. In the sections of *completing dialogues* and *grammar*, 60 questions were all multiple-choice questions. In the *reading comprehension* section, there were 10 multiple-choice, 10 fill-in and 10 true/false questions. In the *vocabulary* section, there were 10 multiple-choice and 20 fill-in questions. Hence, there were 80 multiple-choice with four alternatives (a, b, c and d), 30 fill-in and 10 true/false questions in the WBFA program. The multiple-choice format was designed in WBFA to make the participants familiar with the later summative assessment. However, the students had chances to type short answers in the boxes provided in the fill-in and true/false questions in WBFA. Students' participation in the WBFA was also weighted 5% of the course grade in the same way as that of the

students in the CPFA group. The WBFA program was available twenty-four hours a day, seven days a week.

3.3.1.2 Trial of WBFA and statitiscal processes

After its construction, WBFA was trialled in the first semester with a group of students similar to those in the WBFA group to verify its index values of item difficulty, item discrimination, reliability, and validity. The item analysis from the upper and lower 27% of the participants in the tryout were employed to ensure that the index of difficulty or *p* value and the index of discrimination (*D*) of all question items in WBFA provided reliable statistics (see Appendix D). As a result, the index of difficulty of WBFA was between 0.20 and 0.80, and the index of discrimination was equal to or greater than 0.20 (Aiken, 1996, pp. 60-61; Bachman, 2004, pp. 123, 138).

In addition, the Cronbach's alpha was also calculated to see the reliability coefficient of WBFA. The WBFA program was a kind of teacher-made test or classroom assessment. Its score was merged with other test scores to assign a course grade for a semester work. Hence, this kind of assessment does not need very high reliability coefficients. A reliability coefficient of .50 to .65 may suffice (Ebel & Frisbie, 1991, pp. 85-86; Ruder & Schafer, 2001; Tuckman, 1988, p. 188). However, the internal consistency reliability (Cronbach's alpha) for the *completing dialogues*, *reading comprehension*, *vocabulary* and *grammar* sections in WBFA were .682, .718, .679, and .725, respectively (see Appendix I). The applications of computer software, Statistical Package for the Social Sciences (SPSS) version 14.0 for Windows, were employed for the statistical analyses.

3.3.1.3 How to use the WBFA program

To enter the WBFA program, each student was required to choose a user name through the WBFA login homepage. Once the user name was accepted, the student had to complete the New Students Registration form online, as shown in Figure 3.3 where they had to fill in their user ID, passwords, names, and other personal details, either in Thai or English fonts. When their registration form was completed and submitted to the system, the students were allowed to use their accepted user ID and password to login to the Web until the end of the program. For general security reasons, each student had

to use this personal information to login any time they entered the WBFA program. Prior to their first login to performing WBFA, all WBFA students enrolling in the course had to receive online permission from the researcher to use the WBFA program for additional security of the assessment system. Moreover, the questions in the WBFA were randomised each time the questions were generated to avoid or reduce cheating.

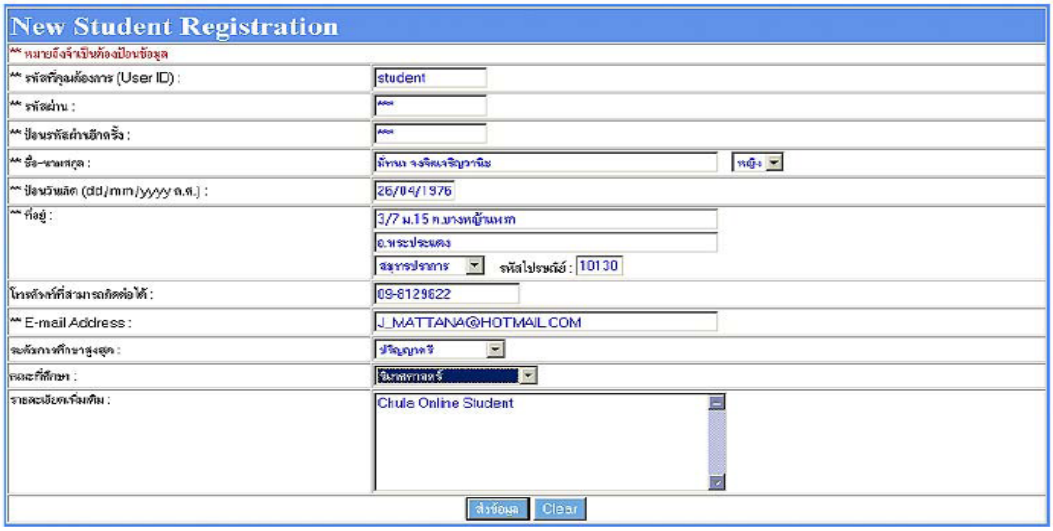


Figure 3.3 New student registration page

Before students were able to log in to the WBFA program, the researcher offered a training session for all students in the WBFA group. Students were trained how to register in the program by filling in the student registration page online. After that, they could apply their registration to the WBFA program. The researcher made a final approval online for students who enrolled in the course and whose name lists were received from the Department of Western Languages after consultation with the lecturers of the course at that time. Once they were approved to use the WBFA program, the students could log in to the program at their convenience regardless of time and places for at least one year or until the end of their study course of study.

The WBFA program was released in January 2005 after the pre-WBFA test. Due to formative purposes, students in the WBFA group were able to use the WBFA program immediately after they finished each module in their lecture class, which was scheduled in the course syllabus. The students' names, answers to each question, the number of attempts to perform WBFA, login dates, scores, and amount of time used during WBFA performance were automatically recorded in order to be analysed regarding the impact

of frequent attempts to perform WBFA on the students' achievement and to identify how well and how often they performed WBFA. The records were saved and accessible by the lecturers and the researcher only.

The WBFA questions were processed on the basis of one item on screen at a time. For reading parts, students could scroll down to continue their reading before performing tests. Basically, the reading text was one-page long. This is to avoid the problem of reading a long text on the computer screen. Students could choose the answers they wanted from the four alternatives. At the bottom right corner of the computer screen, there was a time indicator to allow students to realise how much time remained for each test set, as demonstrated in Figure 3.4. The program ended in February 2005, eight weeks after its commencement.

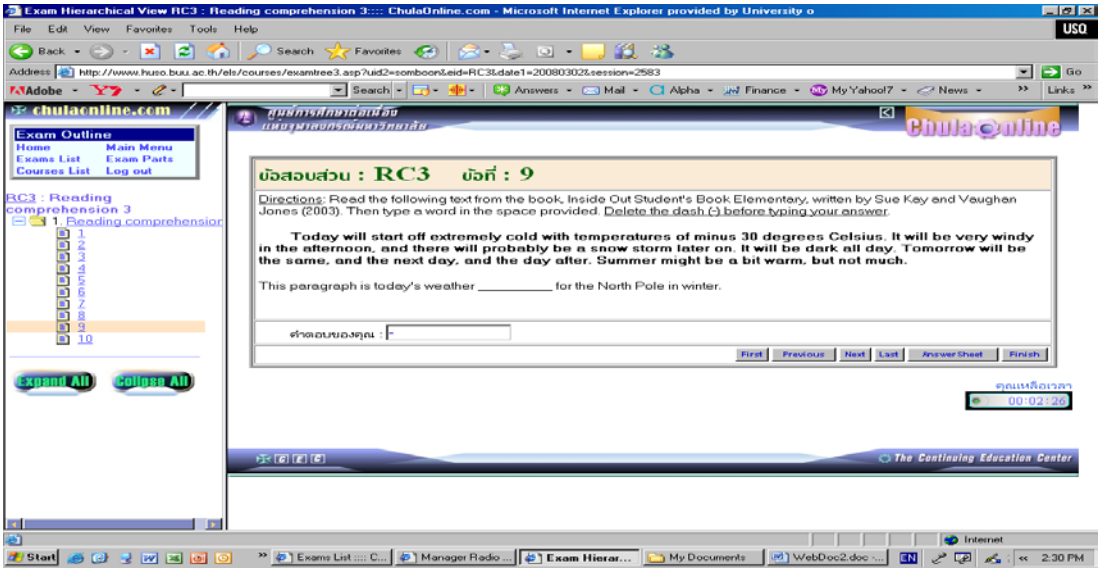


Figure 3.4 A sample question on the WBFA

The students received their feedback and scores reporting on completion of each of 12 test sets, as shown in Figure 3.5, and they could repeat their performance on WBFA as often as they wished. The score reporting was instantly provided to the students once they complete each test set. The students were allowed to see the correct answers when they finished each test set after automatic score reporting. When they needed further explanation, they could contact their lecturers or the researcher any time through the e-contacts provided. During the WBFA program, students in the WBFA group were able to communicate with other students, their lecturers, and/or the researcher through Web

board, chat room, and e-mail instead of attending their face-to-face class. All e-contacts and students' participation levels were also recorded for the study.

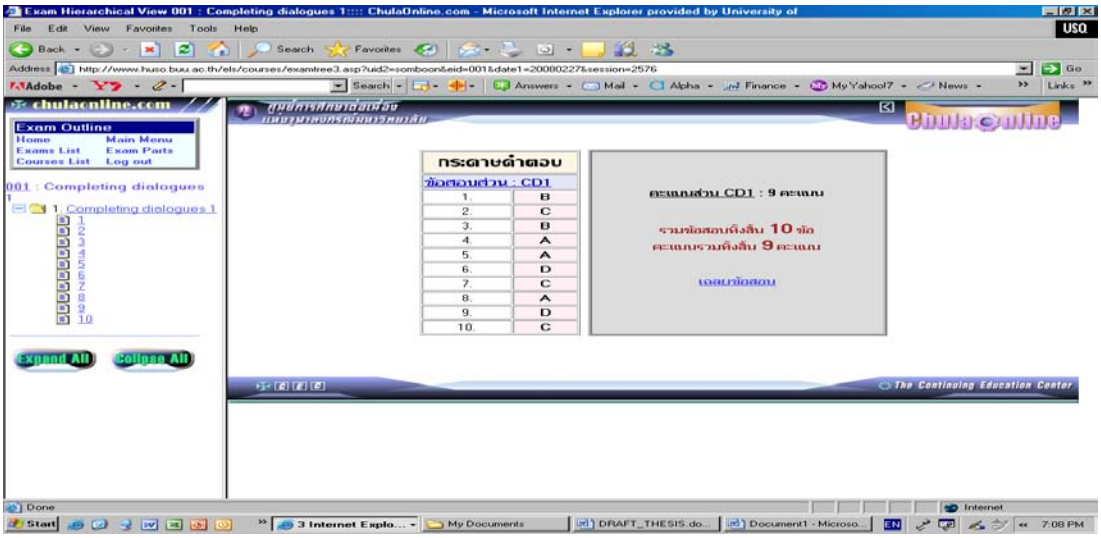


Figure 3.5 A sample of score reporting on WBFA

3.3.2 Pre-test

3.3.2.1 Creation of the pre-test

The pre-test was conducted to measure the participants' achievement in both CPFA and WBFA groups prior to the beginning of the WBFA program. The pre-test mean was compared with that of the post-test to determine whether there was improvement among the participants. The content of the pre-test was based on the course content and course objectives. The pre-test consisted of 45 multiple-choice question items (see Appendix D). Each item had four predetermined options, a, b, c, and d, in which one was the correct answer. The distribution of scores in the test was in the form of a dichotomous scale. Each correct answer for each item was scored 1, while 0 was for every incorrect answer or omission. The test was divided into four sections, *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar*. There were 10 multiple-choice question items for each of the first three sections while there were 15 items for the last section. The pre-test was in conventional paper-based version.

The test was created by the lectures in the Department of Western Languages who taught the course. The researcher was not involved in any stage of the test procedure.

The test was taken on campus and was invigilated by lecturers responsible for each of the 35 classes. The pre-test was marked by the lecturer of each group. Students could request to see their marked answer sheets at any time before they were graded. Then the raw scores were delivered to the researcher for statistical analysis for the study.

3.3.2.2 Trial of the pre-test and statistical processes

For the reliability and validity, the pre-test was trialled with a group of students similar to the participants. The test item analysis was applied to ensure that each item in the test had the index of facility or index of difficulty between .20 and .80, and the index of discrimination of .20 or above (see Appendix D), according to Brown, Bull and Pendlebury (1997). In addition, the Cronbach's alpha or coefficient alpha in the pre-test was .735 for *completing dialogues*, .782 for *reading comprehension*, .714 for *vocabulary*, and .758 for *grammar* sections (see Appendix D). This was considered appropriate as Jacobs (1991) suggested that the test reliability coefficient should be at least .70. The pre-test was conducted in two forms, Form A and Form B, with identical content, but the question items were ordered differently to avoid or reduce cheating. The participants in both CPFA and WBFA groups took the pre-test at the same time on the same day in a similar environment.

3.3.3 Post-test

3.3.3.1 Creation of the post-test

At the end of the WBFA program, the participants in both CPFA and WBFA groups were required to take the paper-based post-test linked directly to the course content and course objectives. It was equivalent to that of the pre-test; however, the content in both tests was not identical. The post-test also contained 45 multiple-choice questions. Each question had four preset alternatives, a, b, c, and d, in which one was the correct answer. The simple allocation of scores in the test is to provide 1 point for each correct response and 0 for each incorrect response or error. The test was separated into four sections, *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar*. Each of the first three sections was consisted of 10 questions whereas 15 questions were in the last section.

The post-test was also constructed by the lectures in the Department of Western Languages who were responsible for the course. The researcher had no involvement in any stage of creating, proctoring and marking the test. The post-test was also taken on campus. The participants in both CPFA and WBFA groups were under the supervision of their lectures responsible for each class while taking the test. The post-test was marked by the lecturers of each group. All participants in both groups were able to see their marked answer sheets any time they wanted before grading. The raw scores were then sent to the researcher for statistical analyses.

3.3.3.2 Trial of the post-test and statistical processes

Prior to its implementation to collect the quantitative data, the post-test was also trialled in a similar procedure as in the pre-test. For its reliability and validity, the procedure of item analysis was applied to certify that each question item in the test had the index of facility or difficulty between .20 and .80, and the index of discrimination of .20 or above (see Appendix D). Additionally, the Cronbach's alpha or coefficient alpha of each section in the test was over .70, which was considerably equivalent to that in the pre-test. The alpha level for the post-test was .773 for the *completing dialogues*, .745 for the *reading comprehension*, .722 for the *vocabulary*, and .753 for the *grammar* sections (see Appendix D). The post-test was conducted in three forms, Form A, Form B, and Form C to avoid or reduce the possibility of any possible order effect and reduce the possibility of cheating. The three forms contained identical content, but provided the question items in different orders.

3.3.4 Questionnaire

The questionnaire aims to gather the mixed methods data regarding students' attitudes toward the use of the WBFA program in their course. This research instrument is considered to be the most appropriate for measuring different viewpoints of participants (Brown & Rodgers, 2002; Johnson & Christensen, 2004). As noted by Johnson and Turner (2003), the self-report instrument including "a mixture of completely open-ended and closed-ended items", which is suitable for the mixed methods research design, is known as "intramethod mixing" (p. 304) due to its mixed methods items.

According to Balnaves and Caputi (2001), the questionnaire is also suitable for the analysis of a large group of participants. The Likert scales are the most universally used approach to measuring attitudes. The questionnaire is used when it is implausible to have direct observation on participants as in the situations of this study.

3.3.4.1 Development of the questionnaire

The participants undertaking the WBFA program were administered both pre- and post-questionnaires, which were designed in the format of Likert scales. The questionnaires were constructed by the researcher under the supervision of experts in the field which included trial of items to make judgements about its content validity. The content in both questionnaires was identical. The inventory consisted of 25 closed-ended items containing details relevant to investigating the students' perceptions of the implementation of the WBFA in relation to the four theoretical perspectives described in Chapter 2. The questionnaire also included eight open-ended questions for students to report independently on their opinions about WBFA (see Appendix G). Therefore, students could have the opportunity to express their own ideas concerning provision of feedback, anxiety toward computer skills needed for the assessment, and use of communication through e-contacts among students and the researcher. Items also covered how students reacted to the implementation of WBFA compared to CPFA, which they experienced in their first semester prior to the study. In addition, all students' answers to the both questionnaires were not considered in relation to the students' course grade in anyway. The questionnaire was written in plain English with translation in Thai at the end of each item.

3.3.4.2 Trial of the questionnaire and statistical processes

The questionnaire was trialled in the first semester prior to the mixed methods data collection for the rigidity of its reliability with a group of 30 students similar to the stratified random sample participated in the study. The questionnaire was available on trial for ease of administration and rapid responses for the improvement of question applicability and question performance (Frary, 2002). Generally, attitude scales should have Cronbach's alpha of reliabilities at least .70 or .75 (Dornyei, 2003; Mehrens & Lehmann, 1973). SPSS (version 14.0 for Windows) was employed for the statistical

analyses. In relation to the reliability, the Cronbach's alpha for the questionnaire utilised in the study was .803 (see Appendix D), which was a relatively strong result.

3.3.4.3 Return rates

With reference to the return or response rates for the survey questionnaire, Babbie (1986) recommended that "a 50 per cent response rate for a questionnaire is adequate, 60 per cent is good, and 70 per cent is very good" (p. 22). The return rate for the pre-WBFA questionnaire in the study was 59.21 per cent, which was a rather good return rate. In addition, the response rate for the post-WBFA questionnaire was 79.98 per cent, which was very good. Additional details for the return rate from each discipline of both questionnaires are described in Chapter 5.

3.3.5 The semi-structured interview

The semi-structured interview is one of the major research instruments in the qualitative method. The verbal transaction utilised in the conduct of interviews gives greater flexibility and is able to gather more applicable responses from the participants' on their perceptions of reality compared with the closed-ended items in the questionnaire (Burns 1999). It diminishes misinterpretation and provides higher response rates for the study (Barnett, 2002). Other benefits of the interview are that the order of answering can be controlled and it yields relatively richer data than the open-ended items in the questionnaire (Brown, 2001).

3.3.5.1 Creation of the semi-structured interview

The semi-structured interview was employed to capture students' more in-depth perspectives on using WBFA. The use of the qualitative data collection method relates to the assumption that the candidates' perceptions are meaningful, knowable, and able to be made explicit. The students' attitudes would also affect the overall success of this assessment as an option for all students and teachers of English in the future. This research instrument was selected due to the fact that interpersonal contact is important, and opportunities for follow-up of emerging issues, interesting and in-depth comments are necessary to fully understand the phenomenon (Frechtling & Sharp, 1997). The

interview was administered before and after the implementation of the WBFA program with a small group of participants, involving an approximate 10% random subsample of the WBFA group. The interview was created by the researcher, containing a set of eight open-ended questions which were identical to those in the questionnaire. The pre- and post-WBFA interviews would take into account issues emerging from the pre- and post-WBFA questionnaires as well. Thus, there was a great deal of flexibility to pursue emerging issues in greater depth. This would make the mixed methods data collection more systematic and comprehensive. In addition, all participants' answers to both semi-structured interviews were not considered in relation to their course grade. Course instructors were not involved and students were fully aware that the research activity was not in anyway linked to their overall course results.

The permutation of the questionnaire and the interview would lead to more complete and appealing interpretation of the differences across the student disciplines. The use of the interview helps researchers to better understand their quantitative findings. As noted by Johnson and Turner (2003), the sequential explanatory strategy using the questionnaire followed by the interview in the mixed methods research design is known as "intermethod mixing" (p. 304).

3.3.5.2 Trial of the semi-structured interview

The semi-structured interview was trialled with three students for content validity in the first semester prior to the mixed methods data collection. These students had performed some examples of WBFA. The trial also aimed for testing the quality of recording techniques and devices to record the interview when gathering the actual data. Both pre- and post-WBFA semi-structured interviews were conducted in Thai.

3.4 Procedure

The procedure to collect the mixed methods data was taken in Thailand and was separated into three stages: the pre-WBFA stage, the WBFA stage, and the post-WBFA stage (see Figure 3.6). In the pre- and post-WBFA stages, there were two phases of data collection. The first phase was related to the quantitative method. The second phase was for the follow-up qualitative method. With regard to the explanatory mixed methods

research design with the sequential explanatory strategy utilised for the study, the priority and majority of the data collection was placed on the quantitative phase. The data collection plan is displayed in Table 3.2.

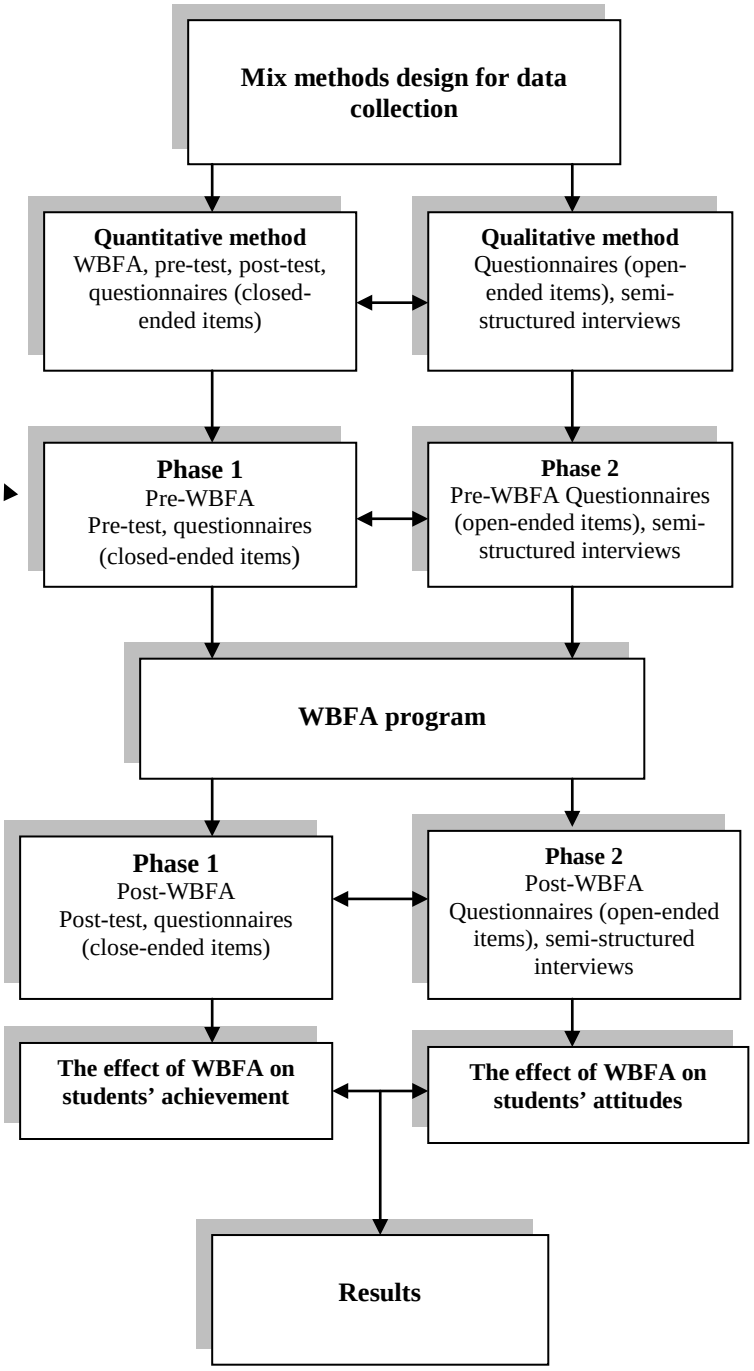


Figure 3.6 Stages of data collection in the mixed methods design in the study

Table 3.2 Research questions and the data collection plan

Research questions	Data collection plan
1. Does WBFA serve as a tool for students to elevate their levels of achievement compared with CPFA when measured by objective testing? Hypothesis: The WBFA English skills' overall mean scores in the post-test-only design would be higher than those of the CPFA group with the alpha level set at .05.	1. Pre-tests of English language for students in WBFA and CPFA groups 2. Post-tests of English language for students in WBFA and CPFA groups
2. Do students who use WBFA get higher achievement scores on completing dialogues, reading comprehension, vocabulary, and grammar for English than students who use CPFA? Hypothesis: The WBFA means of the four language features in the pre-test-post-test design would be higher than those of the CPFA with the alpha level set at .05.	1. Pre-tests of English language for students in WBFA and CPFA groups 2. Post-tests of English language for students in WBFA and CPFA groups
3. How does the frequency of participation in WBFA affect students' learning outcomes? Hypothesis: Students with above average use of the WBFA would have better learning outcomes than those with below average use of WBFA with the alpha level being set at .05.	1. Pre-tests of English language for students in WBFA group 2. Post-tests of English language for students in WBFA Group 3. Records of students' frequency of participation during WBFA program
4. What are students' attitudes toward the use of WBFA and learning English in the course?	1. Pre- and post WBFA program attitudinal questionnaires 2. Pre- and post WBFA program semi-structured interviews

3.4.1 Pre-WBFA stage

After the tryouts of the research instruments in the first semester, participants in the CPFA and the WBFA groups were recruited in keeping with the University ethics clearance and the students' consent forms were completed. Then the author began to collect the data in the second semester of 2004-2005 academic year. This stage consisted of two phases. The first phase was the major quantitative data collection. The initial data collection was relevant to the numerical data through the pre-test. All participants in both CPFA and WBFA groups had to take the pre-test synchronously on campus.

One week after the pre-test, the researcher provided a training session to the participants in the WBFA group. The session occurred in the air-conditioned computer laboratory at the Faculty of Humanities and Social Sciences, Burapha University. When their access to the program was approved, the participants were able to complete some examples of the program and online communication. The instructions regarding how to use the program from login to logout were demonstrated and distributed to every trainee in hard copies in the Thai language. The session took two hours and all information was explained in the Thai language. There were two training sessions on two different days owing to the limited number of the computers available. In each session, there were 100 computers for use. The second phase of data collection in the stage, which was mainly qualitative, took place at the end of the session. The pre-WBFA questionnaire and the pre-WBFA semi-structured interview were administered to participants in the WBFA group.

The raw scores from the pre-test were manipulated for the mean comparisons with that in the post-test to investigate the effect of the WBFA on the students' achievement within and between both CPFA and WBFA groups. The data from both the pre-WBFA questionnaire and the pre-WBFA semi-structured interview were also analysed to determine the students' attitudes toward the use of the WBFA program in the course. Additional details of the data analysis are explained in Chapter 4 and Chapter 5.

3.4.2 WBFA stage

The WBFA program was accessible from 1 January 2005 to 26 February 2005. That was approximately eight weeks (57 days). At that stage, participants from the WBFA group were able to log in to the WBFA program any time and anywhere after they had studied the related module in their lecture class. They did not have to attend the period for CPFA each week. However, the participants in the CPFA still had to attend their conventional class. Once the students logged in, the system would automatically record their login date and time. Every participant was able to view their login history on the student page available in the program. In addition, they were able to communicate with other students or the researcher any time and anywhere when they logged in to the program. There were a Web-board and a chat room for asynchronous and synchronous communication, respectively, on the course page.

When the students needed to perform the WBFA, they could click the "Exam" button after they selected the test set. There were 12 test sets of the four sections of *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar*. Each section was divided into three test sets containing ten multiple-choice, fill in, or true/false questions. The students could click the line of buttons displayed on each page of each question. There were 'First', 'Previous', 'Next', 'Last', 'Answersheet' and 'Finish' buttons on the line placed at the bottom of the page for them to start the test set, go back to previous questions, go to the next ones, go to the last ones, to see the answer sheet, and to submit their answers at the end of each test set, consecutively. The students were asked to confirm their answer submission after clicking the 'Finish' button. Then the system would be generated numerical feedback to report their scores immediately. After the score reporting, the students were able to see the correct answers provided as another feedback if they wanted to. The system would automatically stop the test set when the time set for each test set was over. There was a timer on each question page. However, the students could reattempt as many times as they desired.

It should be noted that five minutes was set for each set of ten questions and there were no correct answers provided during the training session. During the WBFA stage, the time set was expanded to ten minutes for each test set because most students could not finish it in five minutes previously. Additionally, the correct answers were provided at this stage to give some hints to the students and to encourage them to learn from their mistakes. This also aimed to stimulate the students to discuss among themselves and to ask their lecturers or the researcher for more explanations to promote individual student-teacher interaction, instead of providing them with everything.

The system recorded the students' performance on the WBFA program once they clicked the 'Exam' button. The recorded information was displayed on the teacher's page for tracking the students' progress during their study.

3.4.3. Post-WBFA stage

At the end of the WBFA program, participants from the CPFA and the WBFA groups were required to take the post-test. They had to take the test for the quantitative data

collection at the same time on campus. In the second phase, the post-WBFA questionnaire and the post-WBFA semi-structured interview were also administered to the participants in the WBFA group for the follow-up qualitative data collection.

The raw scores from the pre- and post-tests were used to compare the means of students' scores in both CPFA and WBFA groups (see Appendix E). The standard error of the two means was estimated to calculate a confidence interval around each mean. As a result, statistical procedures of the dependent or correlated and independent or uncorrelated *t*-tests were employed through SPSS (version 14.0 for Windows) to test for any statistically significant difference between the means of the pre-test and the post-test. The level of confidence was set at $p < .05$.

The pre-and-post-test design in the same group yielded scores that were dependent because the same group of students had taken both tests. In this case, the dependent *t*-tests were applied to find the significant differences between those means. When comparing CPFA with WBFA, their mean scores were independent because two different groups of students had taken the tests. Then the independent *t*-tests were employed.

SPSS was also employed to calculate correlation between the pre- and post-tests. The correlation between the pre- and post-tests of the CPFA group on the *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* sections was significant at the .01 level (2-tailed). In addition, the correlation between the pre-and post-tests of the WBFA group on the *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* sections was also significant at the .01 level (2-tailed).

3.5 Summary

This chapter has described the mixed methods research design utilised to conduct the research in this study. The sequential explanatory strategy was employed to collect two phases of data. The recruitment of participants using both probability and non-probability sampling techniques was undertaken in the approach to sample selection. Details relevant to research instruments to collect both quantitative and qualitative data

have been also described. In addition, the procedure to gather the mixed methods data has been presented.

In the following chapter, quantitative data analyses and results are reported with regard to the investigation of the effect of WBFA on students' achievement scores.

Chapter 4 Data analyses and results: Students' achievement

Chapter 3 has justified the methodology employed to collect the mixed methods data in this study. The mixed methods research design was utilised to gather a substantial amount of data through the use of four data collection instruments: the WBFA program, the pre-test, the post-test, the questionnaire, and the semi-structured interview, from a large stratified random sample of students. This chapter reports the results of the quantitative data analyses in accordance with the methodology explained in the previous chapter. The results are reported in four sections: *Between-group comparisons* (4.1); *Within-group comparisons* (4.2); *Comparisons of above average performers with below average performers (ABA versus BLA)* (4.3) and (4.4) the chapter summary.

The first section (4.1) reports the results of comparisons of the students' performance in the conventional paper-and-pencil formative assessment (CPFA) group with the students' performance in the Web-based formative assessment (WBFA) group. This analysis aims to detect whether WBFA serves as a tool for students to elevate their levels of achievement compared with CPFA when measured by objective tests at the end of the WBFA program. The comparisons of overall means obtained from the post-test-only design between the control group (i.e., CPFA) and the treatment group (i.e., WBFA) are presented in the section of Between-group comparisons (4.1). The post-test-only design was utilised to avoid threats that may change the participants' perception of the experimental treatment that they were encouraged to perform better provided they had not been rendered the pre-test. The post-test-only design could as well reduce the threats of testing and instrumentation caused by the pre-test, especially in a short-term research study (Berg & Latin, 2008; Creswell, 2008; McMillan, 2004; McMillan & Schumacher, 2008).

In Section 4.1, comparisons between overall mean scores of the CPFA and WBFA groups within each discipline are presented. Comparisons of students' performance in the four language features, *completing dialogues (CD), reading comprehension (RC), vocabulary (VO),* and *grammar (GR)* between the CPFA and WBFA groups within each student's discipline are reported under the subsections of *Between-subgroup*

comparisons (4.1.1- 4.1.10). The comparisons were made to detect whether WBFA serves as a tool for students to elevate their levels of achievement compared with CPFA when measured by objective tests at the end of the WBFA program (see Appendix D). Figure 4.1 shows the data analysis plan for the between-group comparisons in which independent samples *t*-tests were used. In the between-subgroup comparisons the Mann-Whitney *U* tests were included to test independent samples with small sample sizes (less than five) to avoid any violations that may occur when using the *t*-tests due to small sample sizes (Davies, 2007; Jackson, 2009).

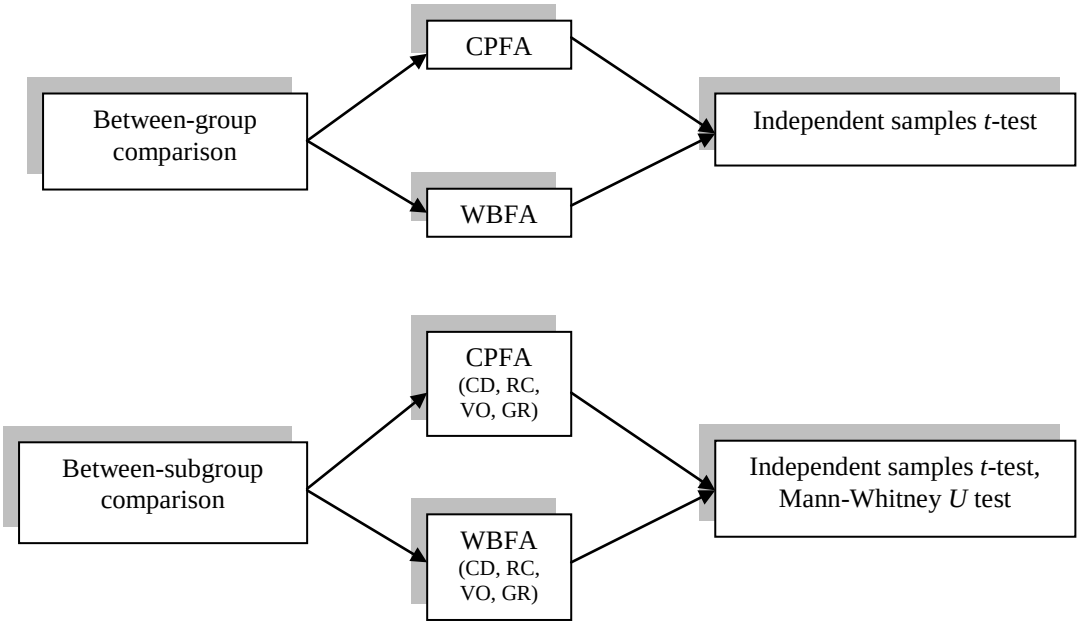


Figure 4.1 Statistical tests used in the post-test-only design

The second section (4.2) reports students' performance in the CPFA and WBFA groups on the pre-test-post-test design within each discipline. The analysis aims to determine whether students who use WBFA get higher achievement scores in the four language features than students who use CPFA. Comparisons of pre- and post-test overall means between both groups in each discipline are presented in the section of *Within-group comparisons* (4.2). Comparisons on means of the four language features between the CPFA and WBFA groups within each discipline are reported under the subsections of *Within-subgroup comparisons* (4.2.1.-4.2.9). The analysis aims to determine whether students who use WBFA get higher achievement scores in the four language features than students who use CPFA. Figure 4.2 shows the data analysis plan for the within-

group comparisons in which paired samples *t*-tests were used. In the within-subgroup comparisons the Wilcoxon signed ranks tests were included to test paired samples with small sample sizes (less than five) to avoid any violations that may occur when using the *t*-tests due to small sample sizes (Ferguson & Takane, 1989).

The third section (4.3) is relevant to analyses and results related to comparisons on the mean scores of two subgroups within the entire WBFA group. The analysis aims to investigate how frequency of participation in WBFA affects students' learning outcomes. It reports comparisons on overall means between students with above average (ABA) use of WBFA and students with below average (BLA) use of WBFA. The overall means across the four language features of the ABA and BLA subgroups are presented in the section of *Comparisons of ABA and BLA* (4.3). In addition, results of comparisons on means of the four language features between ABA and BLA subgroups within each student discipline are reported in subsections that follow, 4.3.1-4.3.6, (see Appendix C). These comparisons are based on students' performance on the pre-test-post-test design in both ABA and BLA subgroups. The independent samples *t*-test and the Mann-Whitney *U* test were employed to locate statistical significance between the performances of both groups. The analysis aims to investigate how frequency of participation in WBFA affects students' learning outcomes. The final section (4.4) summarises the chapter.

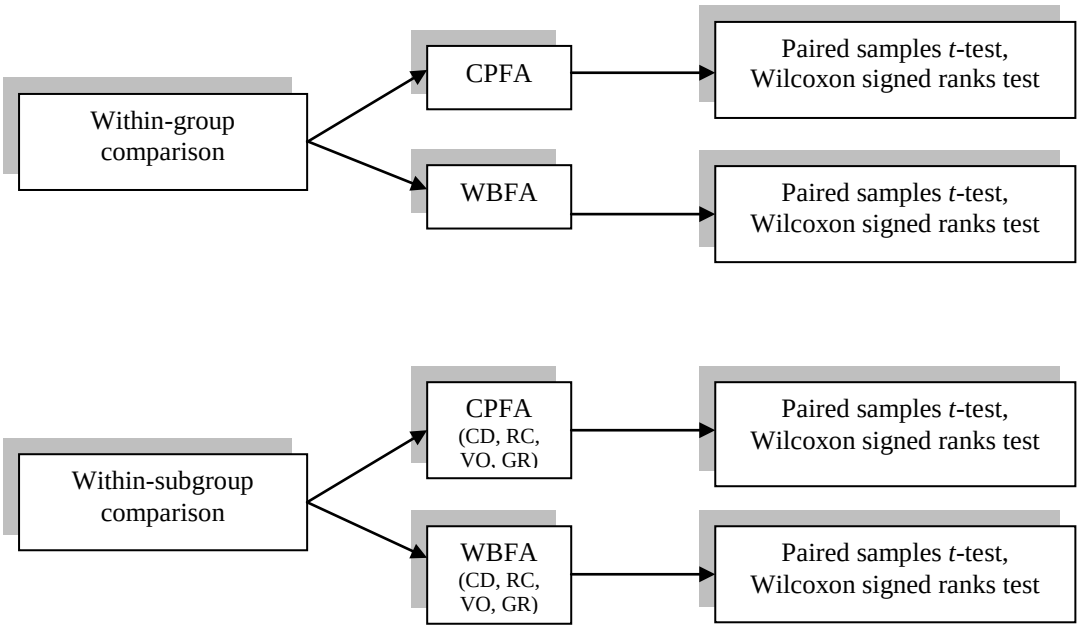


Figure 4.2 Statistical tests used in the pre-post-test design

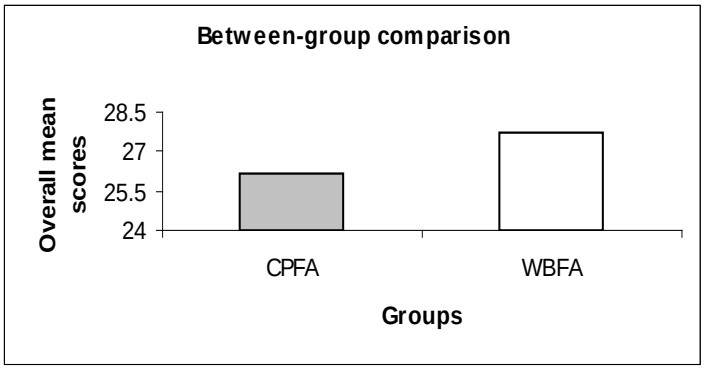
4.1 Between-group comparisons

In this section, overall mean scores of the CPFA group's (*n = 95*) and WBFA (*n = 76*) group's post-test English skills results were compared. The overall mean scores of the CPFA group were calculated from all CPFA students' scores and the overall WBFA means scores came from all WBFA students'. While the study set out to investigate all data from the original equal numbers of students in the two groups of stratified random samples, there were some non-performing students in the WBFA group. It was uncertain why they did not perform the WBFA since they logged into the WBFA program, and then logged out without doing the WBFA. These students were encouraged to participate at various intervals during the WBFA program by their lecturers through personal contacts and by the researcher via electronic contacts, for instance, e-mails and Web-board. They were all content to sign a student consent form prior to the commencement of the WBFA program, and at that time they did not show any sign of dropping out of the WBFA group, so no additional students were randomly recruited to replace them. This study aimed to collect and analyse evidence from students who had performances on WBFA to compare with those in the CPFA group. As a result, only data collected from 76 students (i.e., 80% of students in the WBFA group) who actually performed on WBFA were used in the data analysis with reference to any comparison of students' performance in both groups. The independent or uncorrelated *t*-test was utilised to test for statistical significance between the two groups' performance.

Prior to the implementation of the *t*-test, three assumptions were checked and met to avoid the test power being violated. The first assumption was that the CPFA and WBFA samples were normally distributed. The second assumption referred to the homogeneity of variance. Since the two sample sizes were unequal, the *F*-test was utilised to ensure that the homogeneity-of-variance assumption was not violated (Bachman, 2004; Glass and Hopkins, 1996). The last assumption was that the individuals' performance in CPFA and WBFA groups was evaluated independently. The applications of independent or uncorrelated *t*-test were used where the means of independent samples were compared for statistical significance. In addition, on the basis of the importance of systematic feedback for learning mentioned earlier**,** it was

hypothesised that the post-test of English skills' overall mean scores of the WBFA group would be higher than those of the CPFA group with the alpha level set at .05.

Figure 4.3 shows the result of the comparison of the overall mean scores between students' performance on the post-test of English skills in the CPFA and WBFA groups. The total score of the post-test was 45. The overall mean score of students in the CPFA group was 26.13 while the overall mean score of students in the WBFA group was 27.68. The result reveals that descriptively the overall mean score of the WBFA group was a little higher than the overall mean score of the CPFA group. However, the result of the independent or uncorrelated *t*-test showed that there was no statistically significant difference between the two groups' overall performance on the post-test of English skills with $t = 1.480$, NS at $p > .05$, as shown in Table 4.1.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.3 Comparison on CPFA and WBFA overall means

Following the above comparison the data for each group were broken down into students' disciplines area and comparisons were made between the post-test of English skills overall performance of the WBFA group by discipline with that of the CPFA like-discipline group. Figure 4.4 shows these comparisons across the eight disciplines. The comparison between the CPFA and WBFA groups within the College of Sport Science (SS) was excluded due to the inadequate number of students who performed WBFA for the calculation of the group mean score. The means shown in Figure 4.4 were based on the CPFA and WBFA students' performance on the post-test of English skills which were the combined results of the CD, RC, VO and GR. The total

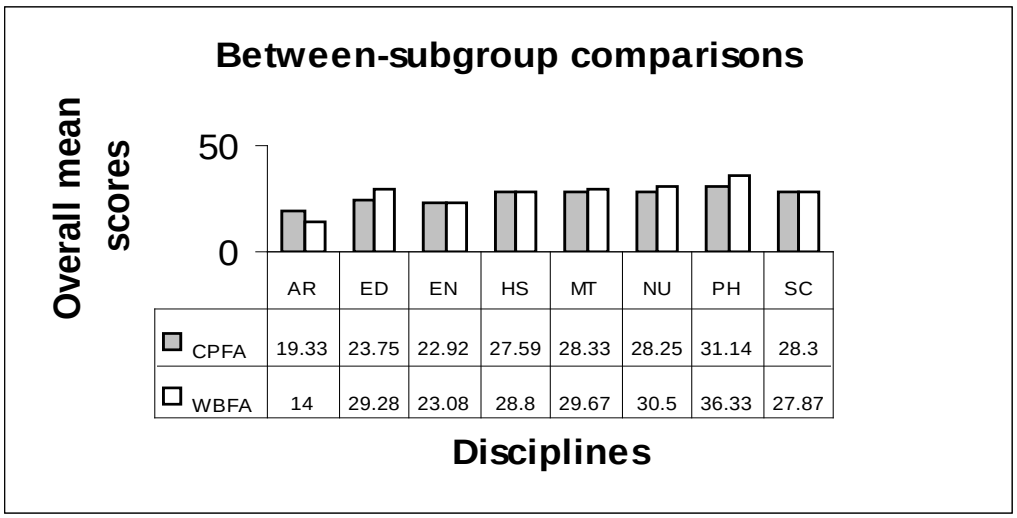
score for each section of the CD, RC and VO was 10 and the total score for the GR section was 15.

Table 4.1 T-tests on CPFA and WBFA overall means

Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
WBFA	27.6842	76	6.55075	.75142	1.480
CPFA	26.1263	95	7.18930	.73761	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

In relation to subgroup comparisons, Figure 4.4 shows that based on descriptive statistics the overall means of two of the CPFA discipline groups were higher than those of the WBFA groups in the same two disciplines; these were the Faculty of Fine and Applied Arts (AR) and the Faculty of Science (SC). The overall mean score of AR students was 19.33 for the CPFA group and 14.00 for the WBFA group (a difference of 5.33), whereas the overall mean scores of SC students in the CPFA group (28.30) and the WBFA group (27.87) were relatively similar (a difference of 0.43). However, in the remaining six discipline groups the overall post-test English skills means for the WBFA group were larger than those in the CPFA group. The overall mean score of students from the Faculty of Education (ED) was 23.75 for the CPFA group, which was lower than 29.28 of the WBFA group (a difference of 5.53). The overall mean score of students from the Faculty of Engineering (EN) was 22.92 for the CPFA group, which differed very little from that of the WBFA group, 23.08 (a difference of 0.16). The overall mean scores of students from the Faculty of Humanities and Social Sciences (HS) were 27.59 for the CPFA group and 28.76 for the WBFA group (a difference of 1.17). Similarly, the overall mean score of students from Maritime College (MT) was 28.33 for the CPFA group and 29.67 for the WBFA group (a difference of 1.34). The overall mean score of students from the Faculty of Nursing (NU) was 28.25 for the CPFA group and 30.50 for the WBFA group (a difference of 2.25). Finally, the overall mean score of students from the Faculty of Public Health (PH) was 31.14 for the CPFA group and 36.33 for the WBFA group (a difference of 5.19).



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
AR = Faculty of Fine and Applied Arts ED = Faculty of Education
EN = Faculty of Engineering HS = Faculty of Humanities and Social Sciences
MT = Maritime College NU = Faculty of Nursing
PH = Faculty of Public Health SC = Faculty of Science

Figure 4.4 Comparisons on CPFA and WBFA overall means between subgroups

It should be noted that AR students had the lowest overall means in both CPFA and WBFA groups with second greatest difference in performance (5.33) but the CPFA performed best. The PH students had the highest overall means in both CPFA and WBFA groups. For each of the remaining subgroups the mean score of the WBFA group was relatively similar to that of the CPFA group, except for the ED students showing the greatest difference of all (5.53).

At the end of the WBFA program, *t*-tests for independent samples were carried out to test if there was any statistically significant difference between the post-test means of the CPFA and the WBFA groups in each discipline . The results revealed that the ED and PH students had statistically higher WBFA overall means than their CPFA counterparts, $t = 1.846$, sig. at $p < .05$, and $t = 1.828$, sig. at $p < .05$, respectively. In addition, the Mann-Whitney *U* tests were employed to locate statistically significant differences between the CPFA and WBFA groups' performance for AR, MT and NU subgroups where the sample sizes were less than five. The *U* tests indicated that there was marginally significant difference only in AR subgroup as $U = 17.5$, $z = 2.195$, at $p < .05$. There were no significant differences in MT and NU subgroups with $U = 5.0$, $z = 0.218$ and $U = 10.5$, $z = 0.722$, at $p > .05$, respectively.

Table 4.2 T-tests and Mann-Whitney U tests on CPFA and WBFA overall means in each subgroup

Subgroup	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
ED	WBFA	29.2857	7	3.72891	1.40940	1.846*
	CPFA	23.7500	12	9.16639	2.64611	
EN	WBFA	23.0833	12	5.91544	1.70764	0.068
	CPFA	22.9167	12	6.15642	1.77721	
HS	WBFA	28.7600	25	5.79713	1.15943	0.678
	CPFA	27.5926	27	6.60570	1.27127	
PH	WBFA	36.3333	6	2.87518	1.17379	1.828*
	CPFA	31.1429	7	6.84175	2.58594	
SC	WBFA	27.8667	15	4.30725	1.11213	-0.237
	CPFA	28.3000	20	6.47343	1.44750	

Subgroup	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
AR	WBFA	14.0000	3	1.00000	.57735	17.5*
	CPFA	19.3333	6	4.50185	1.83787	
MT	WBFA	29.6667	3	3.21455	1.85592	5.0
	CPFA	28.3333	3	5.68624	3.28295	
NU	WBFA	30.5000	4	5.06623	2.53311	10.5
	CPFA	28.2500	4	3.30404	1.65202	

*Sig. at $p < .05$
CPFA = Conventional paper-and-pencil formative assessment
WBFA = Web-based formative assessment
AR = Faculty of Fine and Applied Arts
ED = Faculty of Education
EN = Faculty of Engineering
HS = Faculty of Humanities and Social Sciences
MT = Maritime College
NU = Faculty of Nursing
PH = Faculty of Public Health
SC = Faculty of Science

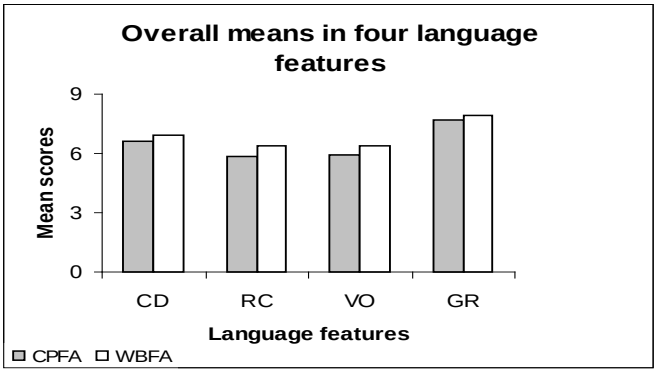
In brief, the WBFA overall mean scores of students from two faculties, Education and Public Health, were statistically higher than the overall mean scores of students in the CPFA group. While there were no significant differences between the overall mean scores of the CPFA and WBFA groups in the five disciplines of EN, HS, MT, NU and SC, in the AR discipline the CPFA group performed statistically significantly better on

the post-test of English skills ($p < .05$) than the WBFA group. However, attention is drawn to the fact that there were only 3 students in the WBFA AR discipline group.

In the following subsections, results from between-subgroup comparisons relevant to students' performance in the CPFA group compared with that in the WBFA group related to the four language features tested, *completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*, within each discipline are reported.

4.1.1 Between-subgroup comparison in four language features

Students' performance on the post-test of English skills was investigated further by breaking down the overall test results into the four language features of *completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*. Then WBFA overall group of students' performance is compared with that of the CPFA overall group on these four language features. This is followed by further exploration on the basis of comparative performance within the eight discipline groups (WBFA versus CPFA). The results are reported in the following subsections.



CPFA = Conventional paper-and-pencil formative assessment
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Figure 4.5 Comparisons on CPFA and WBFA overall means in four language features

Figure 4.5 shows the overall mean scores of the two overall groups' (CPFA and WBFA) performance on the four language features. The results reveal that in the WBFA group the overall mean score for each language feature, while slightly higher than that of the CPFA group, there was very little difference according to the descriptive statistics. The comparisons demonstrate that for *completing dialogues*, the

overall mean score of the CPFA group was 6.65 while 6.89 was the overall mean of the WBFA group. In the section of *reading comprehension*, the overall mean in the CPFA group was 5.83 while that of the WBFA group was 6.41. For the *vocabulary* section, the overall mean in the CPFA group was 5.96 and 6.41 in the WBFA group. For *grammar* the overall mean of the CPFA group was 7.67 while the overall mean for the WBFA group was 7.93.

Although based on descriptive statistics there appears little difference between the WBFA group's performance on the four language features compared with that of the CPFA group when *t*-tests for independent samples were applied the results indicated that the WBFA group's performance on *reading comprehension* was statistically significantly better than that of CPFA group, $t = 1.745$, sig. at $p < .05$. Similarly, the WBFA group's performance on *vocabulary* was statistically significantly better than that of CPFA group, $t = 1.705$, sig. at $p < .05$, as shown in Table 4.3. There was no statistically significant difference between the WBFA group's performance on and the CPFA group's performance on the language features of *completing dialogues* and *grammar*.

Table 4.3 T-tests on CPFA and WBFA overall means in four language features

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	6.8947	76	1.74798	.20051	0.846
	CPFA	6.6526	95	1.99344	.20452	
RC	WBFA	6.4079	76	2.04729	.23484	1.745*
	CPFA	5.8316	95	2.26750	.23264	
VO	WBFA	6.4605	76	1.93522	.22199	1.705*
	CPFA	5.9579	95	1.89015	.19393	
GR	WBFA	7.9342	76	2.62976	.30165	0.607
	CPFA	7.6737	95	2.95873	.30356	

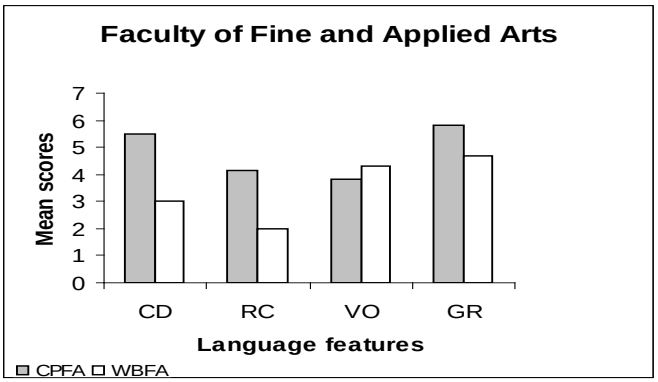
*Sig. at $p < .05$
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading Comprehension
VO = Vocabulary GR = Grammar

In the following subsections, the results of comparisons of mean scores in the four language features between the CPFA and WBFA groups in each discipline are reported.

This aims to find the impact of WBFA on students' learning English in the course through the comparative analysis between the means of both groups related to *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* of English in each discipline.

4.1.2 Faculty of Fine and Applied Arts

Figure 4.6 shows that the WBFA mean score of the stratified random sample from the Faculty of Fine and Applied Arts (AR) was higher than the CPFA mean score in the section of *vocabulary* only. The mean of the WBFA group was 4.33 while the mean of the CPFA group was 3.83. In the last three sections, the WBFA means were lower than the CPFA mean scores. For the section of *completing dialogues*, the mean of the WBFA was 3.00 while the mean of the CPFA group was 5.50. In the section of *reading comprehension*, the mean of the WBFA group was 2.00 while the mean of the CPFA group was 4.17, and in the section of *grammar*, the mean of the WBFA group was 4.67 while the mean of the CPFA group was 5.83.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.6 Comparisons on CPFA and WBFA means in four language features of AR students

The comparisons between the performance of AR students in the CPFA *(n = 6)* and WBFA *(n = 3)* groups are shown in Table 4.4. Their scores in the four language features are reported together with the results of Mann-Whitney *U* tests. The results reveal that there were no significant differences between the performance of AR

students in the CPFA and the WBFA groups in relation to the four language features at the end of the WBFA program. The *U* test results of the four sections were: *completing dialogues U* = 13.0, *z* = 1.032; *reading comprehension U* = 15.0, *z* = 1.549; *vocabulary U* = 10.5, *z* = 0.387; and *grammar U* =12.0, *z* = 0.775 at $p > .05$, respectively.

Table 4.4 Mann-Whitney U tests on four language features between CPFA and WBFA groups of AR students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	WBFA	3.0000	3	1.73205	1.00000	13.0
	CPFA	5.5000	6	2.81069	1.14746	
RC	WBFA	2.0000	3	1.00000	.57735	15.0
	CPFA	4.1667	6	1.83485	.74907	
VO	WBFA	4.3333	3	1.52753	.88192	10.5
	CPFA	3.8333	6	1.32916	.54263	
GR	WBFA	4.6667	3	1.52753	.88192	12.0
	CPFA	5.8333	6	1.94079	.79232	

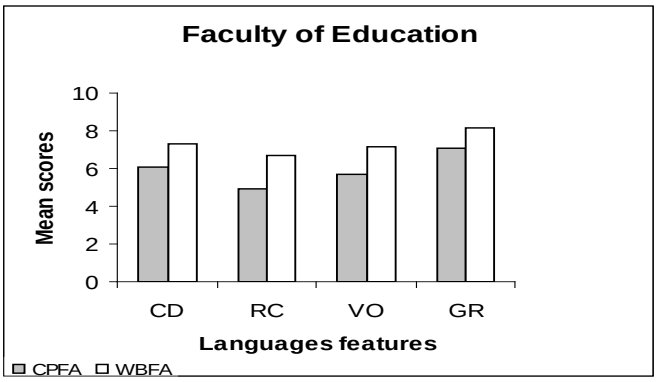
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

To sum up, there were no significant differences between the performance of the CPFA and WBFA groups of AR students in all sections of the language features tested in the post-test of English skills at the end of the WBFA program.

4.1.3 Faculty of Education

Figure 4.7 shows the results related to the stratified random sample from the Faculty of Education (ED). The ED students in the WBFA group (*n = 7*) had higher mean scores than those in the CPFA group (*n = 12*) in all language features. In the *completing dialogues* section, the mean of the CPFA group was 6.08 whereas 7.29 was the mean of the WBFA group. ED students also had a higher WBFA average score, 6.71, than the CPFA mean, 4.92, in the section of *reading comprehension*. In addition, they got higher WBFA mean, 7.14, than CPFA mean, 5.67, in the section of *vocabulary*. In the final

section of *grammar*, the mean of the WBFA group was 8.14 while the mean score of the CPFA group was 7.08.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.7 Comparisons on CPFA and WBFA means in four language features of ED students

Table 4.5 shows the results of independent *t*-tests between ED students in the CPFA and WBFA groups involving the four language features. The results show that the *reading comprehension* mean of the WBFA group was significantly higher than that of the CPFA group, $t = 2.168$, sig. at $p < .05$. In the *completing dialogues*, *vocabulary*, and *grammar* sections, their WBFA means were higher than those of the CPFA group as well; however, there were no significant differences between those means.

In short, the WBFA group of ED students had higher means than the CPFA group in all sections. However, there was only in the section of *reading comprehension* that the mean of WBFA group was significantly higher than the mean of the CPFA group.

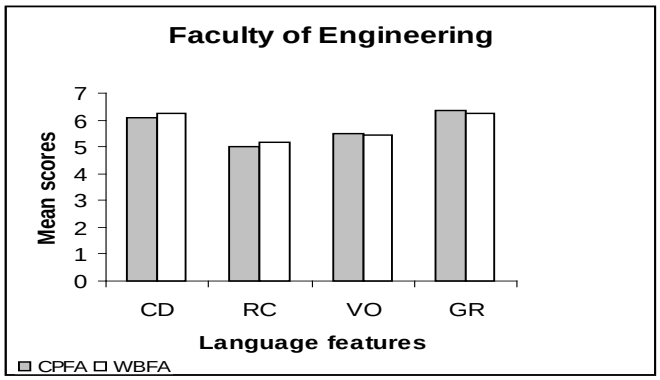
Table 4.5 T-tests on four language features between CPFA and WBFA groups of ED students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	7.2857	7	1.88982	.71429	1.329
	CPFA	6.0833	12	1.92865	.55675	
RC	WBFA	6.7143	7	1.38013	.52164	2.168*
	CPFA	4.9167	12	2.23437	.64501	
VO	WBFA	7.1429	7	2.03540	.76931	1.508
	CPFA	5.6667	12	2.10339	.60720	
GR	WBFA	8.1429	7	1.46385	.55328	0.777
	CPFA	7.0833	12	4.31611	1.24595	

*Sig. at $p < .05$
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.1.4 Faculty of Engineering

Figure 4.8 shows the results of students in the WBFA group (*n = 12*) from the Faculty of Engineering (EN) whose mean scores were slightly higher than those of the CPFA group (*n = 12*) in the two sections of *completing dialogues* and *reading comprehension*. Their average scores in the *completing dialogues* section were 6.08 for the CPFA group and 6.25 for the WBFA group. In the *reading comprehension* section, their CPFA mean was 5.00 while their WBFA mean was 5.17. However, in the last two sections of *vocabulary* and *grammar*, the mean scores of the WBFA group were lower than the mean scores in the CPFA group. In the *vocabulary* section, the mean of the CPFA group was 5.50 while the mean of the WBFA group was 5.42. In the *grammar* section, the mean of the CPFA group was 6.33 while the mean of the WBFA group was 6.25.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.8 Comparisons of CPFA and WBFA means in four language features of EN students

Table 4.6 indicates the results of independent *t*-tests from comparisons between the means of EN students in the CPFA and WBFA groups on the four language features. The results show that even the means of *completing dialogues* and *reading comprehension* sections in the WBFA group were higher than those in the CPFA group, there were no significant differences, $t = 0.285$, NS, at $p > .05$, and $t = 0.192$, NS, at $p > .05$, consecutively.

Table 4.6 T-tests on four language features between CPFA and WBFA groups of EN students

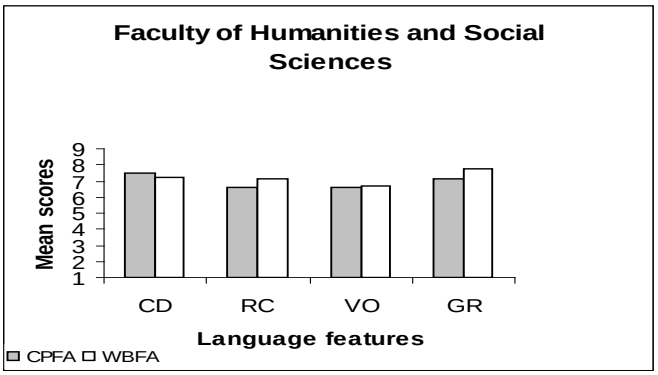
Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	6.2500	12	1.35680	.39167	0.285
	CPFA	6.0833	12	1.50504	.43447	
RC	WBFA	5.1667	12	2.03753	.58818	0.192
	CPFA	5.0000	12	2.21565	.63960	
VO	WBFA	5.4167	12	2.27470	.65665	-0.089
	CPFA	5.5000	12	2.27636	.65713	
GR	WBFA	6.2500	12	2.49089	.71906	-0.096
	CPFA	6.3333	12	1.66969	.48200	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

To put it briefly, there were no significant differences between the means of the four language features in the CPFA and WBFA groups of EN students.

4.1.5 Faculty of Humanities and Social Sciences

Figure 4.9 shows the results of the stratified random sample from the Faculty of Humanities and Social Sciences (HS), whose mean in the WBFA group (*n = 25*) was somewhat lower than that of the CPFA group (*n = 27*) only in the section of *completing dialogues.* In the section, the mean of the CPFA group was 7.48 while the mean of the WBFA group was 7.20. However, in the last three sections, their mean scores in the WBFA group were higher than the mean scores in the CPFA group. In the section of *reading comprehension,* the mean of the CPFA group was 6.59 while the mean of the WBFA group was 7.16. In the *vocabulary* section, the CPFA mean was 6.33 and the WBFA mean was 6.68. In the section of *grammar* the mean of the CPFA group was 7.15 while the average score of the WBFA group was 7.76.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.9 Comparisons of CPFA and WBFA means in four language features of HS students

In Table 4.7, the results of independent *t*-tests from the comparisons between the means of HS students in the CPFA and WBFA groups with reference to the four language features are reported. The results show that there were no significant differences between the means of both groups although the WBFA means were higher than those of the CPFA in the three sections of *reading comprehension*, $t = 0.899$, NS, at $p > .05$, *vocabulary*, $t = 0.718$, NS, at $p > .05$, and *gramma*, $t = 0.913$, NS, at $p > .05$. In the

completing dialogues section, the CPFA mean was higher than the WBFA mean. In short, there were no significant differences between all means of both groups relating to the four language features.

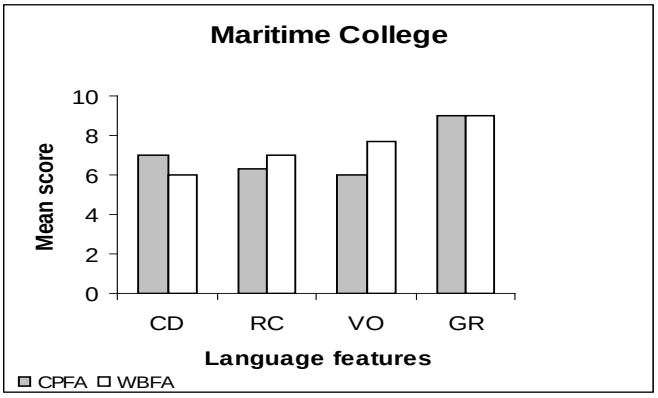
Table 4.7 T-tests on four language features between CPFA and WBFA groups of HS students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	7.2000	25	1.47196	.29439	-0.621
	CPFA	7.4815	27	1.78391	.34331	
RC	WBFA	7.1600	25	2.17332	.43466	0.899
	CPFA	6.5926	27	2.37388	.45685	
VO	WBFA	6.6800	25	1.57374	.31475	0.718
	CPFA	6.3333	27	1.90142	.36593	
GR	WBFA	7.7600	25	2.33238	.46648	0.913
	CPFA	7.1481	27	2.50697	.48247	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.1.6 Maritime College

Figure 4.10 shows the results of the stratified random sample from the Maritime College (MT), whose mean scores in the WBFA cohort (*n =3*) were higher than the mean scores of the CPFA group (*n = 3*) in two sections. The mean of the *reading comprehension* in the CPFA group was 6.33 while the mean of the WBFA group was 7.00. For the section of *vocabulary*, the mean of the CPFA group was 6.00 while the mean of the WBFA group was 7.67. However, in the section of *completing dialogues*, the mean of the CPFA group, 7.00, was a little higher than the mean of the WBFA group, 6.00. The mean scores of MT students in both CPFA and WBFA groups were the same at 9.00 in the *grammar* section.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.10 Comparisons of CPFA and WBFA means in four language features of MT students

Table 4.8 shows the results of Mann-Whitney *U* tests from the comparisons between the performance of MT students in the CPFA and WBFA groups on the four language features. The results indicate that there were no statistical differences between the performance of both groups in all sections. The results in each section were: *completing dialogues U* = 6.5, *z* = 0.873*; reading comprehension U* = 6.0, *z* = 0.655; *vocabulary U* = 7.0, *z* =1.091*;* and *grammar U* = 4.5, *z* = 0, at *p* > .05, respectively.

Table 4.8 Mann-Whitney U tests on four language features between CPFA and WBFA groups of MT students

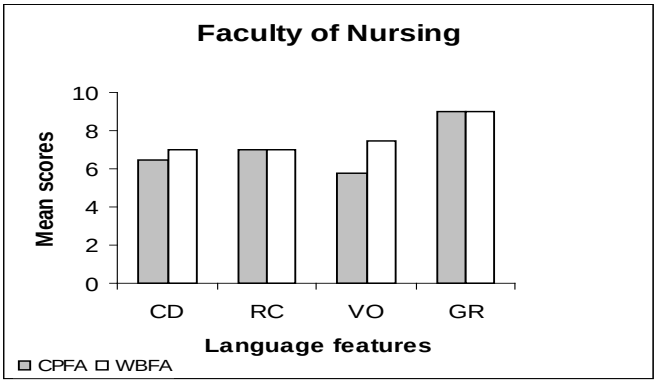
Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	WBFA	6.0000	3	1.00000	.57735	6.5
	CPFA	7.0000	3	1.73205	1.00000	
RC	WBFA	7.0000	3	1.00000	.57735	6.0
	CPFA	6.3333	3	1.52753	.88192	
VO	WBFA	7.6667	3	1.52753	.88192	7.0
	CPFA	6.0000	3	1.73205	1.00000	
GR	WBFA	9.0000	3	2.00000	1.15470	4.5
	CPFA	9.0000	3	1.00000	.57735	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Therefore, it can be summarised that there were no significant differences between the CPFA and WBFA groups' performance of MT students on the four language features.

4.1.7 Faculty of Nursing

Figure 4.11 discloses that the mean scores of the WBFA group (*n = 4*) of the stratified random sample from the Faculty of Nursing (NU) were higher than those of the CPFA group (*n = 4*) in two sections regarding the four language features tested. In the section of *completing dialogues*, the mean of the CPFA group was 6.50, which was slightly lower the mean of the WBFA group, 7.00. In addition, in the section of *vocabulary*, the mean of the CPFA group was 5.75 while the mean of the WBFA was 7.50. However, the NU students had the same mean scores in the CPFA and WBFA groups in the other two sections. In the section of *reading comprehension*, both CPFA and WBFA means were 7.00 while in the *grammar* section their means were the same at 9.00. With reference to the findings, there is verification that NU students in the WBFA group had better development than those in the CPFA group in terms of average scores in *completing dialogues* and *vocabulary* sections.



CPFA = Conventional paper-and-pencil formative assessment
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Figure 4.11 Comparisons of CPFA and WBFA means in four language features of NU students

Table 4.9 shows the results of Mann-Whitney *U*-tests from the comparisons between the performance of NU students in the CPFA and WBFA groups on the four language features. The results indicate that there were no significant differences between the performance of both groups in all sections. The results in each section were: *completing*

dialogues U = 10.5, *z* = 0.577*; reading comprehension U* = 8.5, *z* = 0.144; *vocabulary U* = 14.0, *z* =1.732*;* and *grammar U* = 8.5, *z* = 0.144, at *p* > .05, respectively.

Table 4.9 Mann-Whitney U tests on four language features between CPFA and WBFA groups of NU students

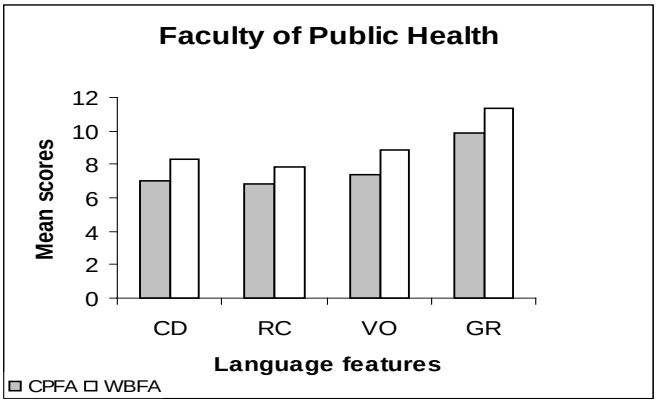
Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	WBFA	7.0000	4	1.41421	.70711	10.0
	CPFA	6.5000	4	1.29099	.64550	
RC	WBFA	7.0000	4	1.41421	.70711	8.5
	CPFA	7.0000	4	1.41421	.70711	
VO	WBFA	7.5000	4	1.29099	.64550	14.0
	CPFA	5.7500	4	.95743	.47871	
GR	WBFA	9.0000	4	2.44949	1.22474	8.5
	CPFA	9.0000	4	.81650	.40825	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In summary, there were no significant differences between the CPFA and WBFA groups' performance of NU students on the four language features.

4.1.8 Faculty of Public Health

Figure 4.12 shows results of the stratified random sample from the Faculty of Public Health (PH) whose mean scores in the WBFA group (*n = 6)* were higher than those in the CPFA group (*n = 7*) in all language features tested. In the section of *completing dialogues*, the mean of the CPFA group was 7.00 while the mean of the WBFA was 8.33. In the section of *reading comprehension*, the mean of the CPFA group was 6.86 while the mean of the WBFA group was 7.83. For *vocabulary* section, the mean of the CPFA group was 7.43 and the mean of the WBFA was 8.83. In addition, in the section of *grammar*, the mean of the CPFA was 9.86 while that of the WBFA was 11.33.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.12 Comparisons of CPFA and WBFA means in four language features of PH students

In Table 4.10, the result of independent *t*-tests from the comparisons between the means of PH students in the CPFA and WBFA groups involving the four language features are presented. The results reveal that there were significant differences between means of both groups in the *reading comprehension* and *vocabulary* sections. In the section of *reading comprehension*, the mean of the WBFA group was significantly higher than that of the CPFA group, $t = 1.820$, sig. at $p < .05$. In the section of *vocabulary*, the WBFA mean was also significantly higher than the CPFA mean, $t = 1.769$, sig. at $p < .05$. In the *completing dialogues* and *grammar* sections, although their means in the WBFA group were higher than those means of the CPFA group, there were no significant differences between them.

In conclusion, for PH students, the WBFA means in the four language features were higher than those in the CPFA group. However, there were statistical differences in two sections of *reading comprehension* and *vocabulary* where the WBFA group outperformed the CPFA group.

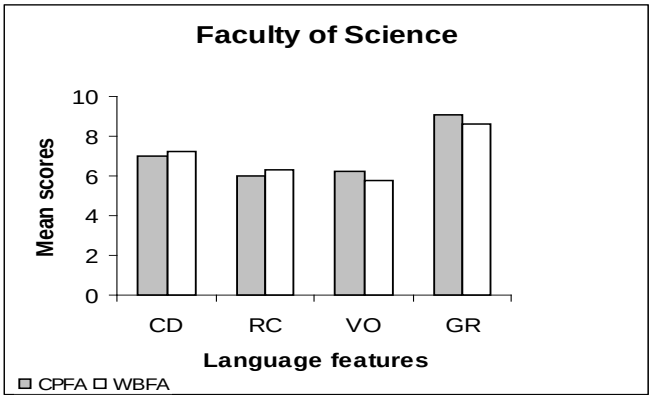
Table 4.10 T-tests on four language features between CPFA and WBFA groups of PH students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	8.3333	6	1.03280	.42164	1.493
	CPFA	7.0000	7	2.08167	.78680	
RC	WBFA	7.8333	6	.40825	.16667	1.820*
	CPFA	6.8571	7	1.34519	.50843	
VO	WBFA	8.8333	6	.98319	.40139	1.769*
	CPFA	7.4286	7	1.81265	.68512	
GR	WBFA	11.3333	6	1.63299	.66667	1.059
	CPFA	9.8571	7	3.23669	1.22336	

*Sig. at $p < .05$
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.1.9 Faculty of Science

In Figure 4.13, on account of the results of the stratified random sample from the Faculty of Science (SC), there were two sections of *completing dialogues* and *reading comprehension* that the means of the WBFA group (*n = 15*) were slightly higher than the means of the CPFA group (*n = 20*). In the former section, the mean of the CPFA group was 7.00 while the mean of the WBFA group was 7.20. In the latter section, the mean of the CPFA group was 6.10 while the mean of the WBFA was 6.27. On the contrary, in the last two sections, the mean scores of the WBFA group were lower than the mean scores of the CPFA group. In the section of *vocabulary*, the mean of the CPFA group was 6.20 while the mean of the WBFA group was 5.80. In addition, in the *grammar* section, the mean of the CPFA group was 9.10 while the mean of the WBFA group was 8.60.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.13 Comparisons of CPFA and WBFA means in four language features of SC students

Table 4.11 shows the results of independent *t*-tests from the comparisons between the means of SC students in the CPFA and WBFA groups concerning the four language features. The findings reveal that there were no significant differences between the means of both groups in the four sections tested. The means of the WBFA group were higher than those in the CPFA group in two sections of *completing dialogues*, $t = 0.324$, NS at $p > .05$ and *reading comprehension*, $t = 0.462$, NS at $p > 0.5$. However, the CPFA means were higher, with no significant differences, than the WBFA means in the last two sections of *vocabulary* $t = -0.849$, NS at $p > .05$ and *grammar*, $t = -0.539$, NS at $p > .05$. In respect of the comparisons, there is evidence that SC students in the WBFA group outperformed those in the CPFA group in *completing dialogues* and *reading comprehension* sections in terms of average scores. Nonetheless, they did not gain better significant improvement than those in the CPFA group at the end of the WBFA program. Hence, the intervention of the WBFA program might not have any impacts to students in the experimental group as it was previously hypothesised.

In brief, there were no significant differences between means of the CPFA and WBFA groups of SC students in relation to the four language features tested.

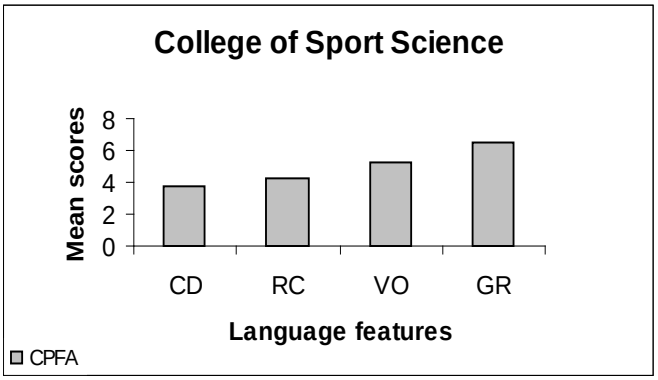
Table 4.11 T-tests on four language features between CPFA and WBFA groups of SC students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	WBFA	7.2000	15	1.74028	.44934	0.324
	CPFA	7.0000	20	1.89181	.42302	
RC	WBFA	6.2667	15	1.03280	.26667	0.462
	CPFA	6.0000	20	2.29416	.51299	
VO	WBFA	5.8000	15	1.42428	.36775	-0.849
	CPFA	6.2000	20	1.32188	.29558	
GR	WBFA	8.6000	15	2.50143	.64587	-0.539
	CPFA	9.1000	20	2.97180	.66451	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.1.10 College of Sport Science

Figure 4.14 displays the results of the mean scores pertaining to the four language features tested in the CPFA group (*n = 4*) of the stratified random sample from the College of Sport Science (SS). Their average scores in the sections of *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* were 6.75, 4.25, 4.00, and 4.75, consecutively. Owing to the inadequate number of students who really performed in the WBFA program in the WBFA group (*n = 4*), the means of the four language features in the group could not be calculated. Actually, there was one student in the experimental group who did perform in the WBFA program while the other three students logged in without WBFA performance. Their login times were recorded; however, there was no information in connection with their participation in the assessment. As a result, there were no comparisons of those means between the CPFA and WBFA groups.



CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.14 Comparisons of CPFA means in four language features of SS students

Table 4.12 shows the results of the means of the SS students in the CPFA group involving the four language features. The comparison between the means of the CPFA and WBFA groups could not be concluded due to the inadequate number of students in the WBFA group to be calculated to sort out the their means.

Table 4.12 Mean scores on four language features in CPFA group of SS students

Language feature	Group	Mean	N	Std. Deviation	Std. Error Mean
CD	CPFA	3.7500	4	1.70783	.85391
RC	CPFA	4.2500	4	2.21736	1.10868
VO	CPFA	5.2500	4	2.06155	1.03078
GR	CPFA	6.5000	4	2.88675	1.44338

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

According to the data analysis from the post-test-only design, there was no significant difference between the overall performance of students in the CPFA and WBFA groups. In relation to the comparisons of the means between the CPFA and WBFA groups in each discipline, there were significant differences in two disciplines, the

Faculty of Education and the Faculty of Public Heath, where the means of WBFA group were significantly higher than the means of the CPFA group.

In relation to the comparisons of the four language features, the overall means indicate that there were significant differences in the *reading comprehension* and *vocabulary* sections where the means of the WBFA group were significantly higher than the means of the CPFA group. In each discipline, the results reveal that there were significant differences in the Faculty of Education and the Faculty of Public Health. In the Faculty of Education, the mean of the WBFA group was significantly higher than the mean of the CPFA group in the *reading comprehension* section. In the Faculty of Public Health, the means of the *reading comprehension* and *vocabulary* sections in the WBFA group were significantly higher than the means of the same sections in the CPFA group.

In the next section, the results from the comparisons of means in the CPFA and WBFA groups through the pre-and post-test design within each discipline are presented. In addition, results from the comparisons between the pre- and post test means on the four language features in the CPFA and WBFA groups in each discipline are reported.

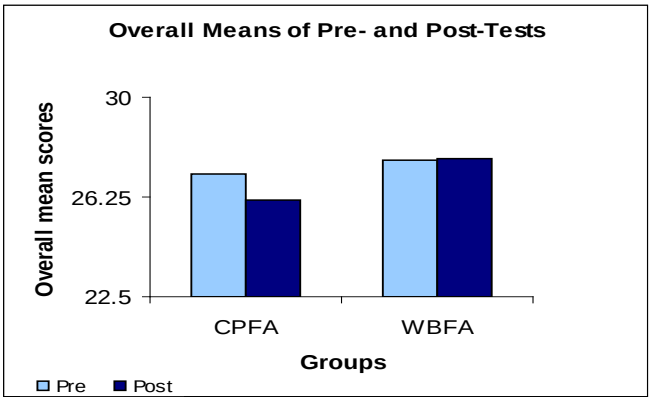
4.2 Within-group comparisons

To investigate whether the intervention of WBFA would elevate students' achievement scores in the WBFA group at the end of their semester when compared with the CPFA group, the pre-and post-test design was employed. All participants in both groups took the pre-test prior to the commencement of the WBFA program and took the post-test at the end of the WBFA program. Both tests were taken in the same atmosphere and environment on the same campus. The pre-test and post-test were equivalent and conducted in traditional paper-and-pencil test format under the same course content and course objectives. Both pre-test and post-test were written, marked, and invigilated by lecturers from the Department of Western Languages, Faculty of Humanities and Social Sciences, Burapha University, who taught the course, English II (212102) at the time of the study. The researcher was not involved in any stage of creating, marking, and proctoring the tests. After the tests, all participants' raw scores were delivered to the researcher for the data analysis. It was hypothesised that, in the WBFA group, the overall mean and means of the four language features in the post-test would be higher

than those means in the pre-test with the alpha level set at .05. The comparisons of means in the CPFA group were also manipulated to compare with the results in the WBFA group.

In this section, *Within-group comparisons* (4.2), the results from the comparisons between overall means of the pre-test and post-test in the CPFA and WBFA groups are reported to find out whether there was any improvement in both groups during the time of the study. This is followed, in the subsections (4.2.1-4.2.9), by the results from the comparisons of overall means of the pre- and post-tests concerning the four language features between the CPFA and WBFA groups in each student discipline to find out whether there was any development in the four language features during the time of the study.

In Figure 4.15, the results reveal that, in the WBFA group (*n = 76*), the overall average score of the post-test was on a par with the overall average score of the pre-test. The WBFA overall post-test mean was 27.68 while that of the pre-test was 27.64. However, in the CPFA group (*n = 95*), the overall mean score of the post-test was lower than the overall mean score of the pre-test with the post-test mean being 26.13 and the pre-test 27.09.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.15 Comparison on overall pre-test and post-test means within CPFA and WBFA groups

Table 4.13 shows the *t*-tests resulting comparisons between the overall pre- and post-tests of English skills within the CPFA and WBFA groups. It is apparent that the overall

mean of the post-test in the CPFA group (26.13) was significantly lower than that of the pre-test (27.09), t = -2.039, sig. at p < .05. In contrast, the overall post-test mean in the WBFA group was 27.68 and its pre-test mean was 27.64. However, there was no significant difference between both means of the WBFA group, t = 0.067, NS at p >.05. With regard to the evidence, it demonstrates that the outcome of students in the CPFA group was not as good as those in the WBFA group and interestingly they did not do as well at the end of the course as they did at the start on the pre-test.

Table 4.13 T-tests on overall means of pre-test and post-test in CPFA and WBFA groups

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	26.1263	95	7.18930	.73761	-2.039*
	Pre	27.0947	95	7.09071	.72749	
WBFA	Post	27.6842	76	6.55075	.75142	0.067
	Pre	27.6447	76	5.52740	.63404	

*Sig. at p <.05
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

In keeping with the comparisons of pre- and post-test means pertaining to the four language features in the CPFA group, Table 4.14 illustrates that there were significant differences in the two sections of *reading comprehension* and *vocabulary*. In the *reading comprehension* section, the post-test mean (5.83) was significantly higher than that of the pre-test (5.17), t = 2.957, sig. at p < .05. In the *vocabulary* section, the post-test mean (5.96) was significantly higher than that of the pre-test (4.95), t = 4.168, sig. at p < .05. However, in the other two sections, their post-test means were significantly lower than the pre-test means. The post-test means in the sections of *completing dialogues* (6.65) and *grammar* (7.67) were lower than the means of the *completing dialogues* (7.67) and *grammar* (9.31) sections in the pre-test. Their *t*-test results were t = - 4.950, sig. at p < .05 and t = -5.007, sig. at p < .05, consecutively. The findings confirm that students in the CPFA group had significant improvement in *reading comprehension* and *vocabulary* sections, but there was no statistical progress in the sections of *completing dialogues* and *grammar*.

Table 4.14 T-tests on four language features of pre-test and post-test in CPFA group

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	6.6526	95	1.99344	.20452	-4.950*
	Pre	7.6737	95	2.01823	.20707	
RC	Post	5.8316	95	2.26750	.23264	2.957*
	Pre	5.1684	95	1.99282	.20446	
VO	Post	5.9579	95	1.89015	.19393	4.168*
	Pre	4.9474	95	2.30803	.23680	
GR	Post	7.6737	95	2.95873	.30356	-5.007*
	Pre	9.3053	95	3.16589	.32481	

*Sig. at p <.05
CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Table 4.15 shows the results from the dependent *t*-tests from the comparisons of the pre- and post-test means with regard to the four language features tested in the WBFA group. There were significant differences in two sections. In the *reading comprehension* section, the mean of the post-test (6.41) was higher than that of the pre-test (5.39), t = 4.352, sig. at $p < .05$. In addition, in the section of *vocabulary*, the post-test mean (6.46) was larger than that of the pre-test (4.83), t = 7.586, sig. at $p < .05$. However, the post-test means of the *completing dialogues* (6.89) and *grammar* (7.93) sections were significantly smaller than the means of the *completing dialogues* (7.80) and *grammar* (9.62) sections in the pre-test. Their *t*-scores were t = -4.216, sig. at $p < 0.5$ and t = -5.673, sig. at $p < .05$, respectively.

Table 4.15 T-tests on four language features of pre-test and post-test in WBFA group

WPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	6.8947	76	1.74798	.20051	-4.216*
	Pre	7.8026	76	1.75134	.20089	
RC	Post	6.4079	76	2.04729	.23484	4.352*
	Pre	5.3947	76	1.47041	.16867	
VO	Post	6.4605	76	1.93522	.22199	7.586*
	Pre	4.8289	76	1.70010	.19502	
GR	Post	7.9342	76	2.62976	.30165	-5.673*
	Pre	9.6184	76	2.52437	.28957	

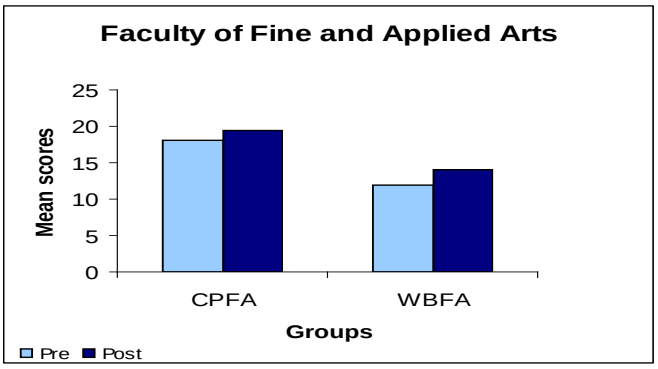
*Sig. at p <.05
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

According to the comparisons of the overall pre- and post-test means of students in the CPFA group, there was no improvement on their overall average score at the end of the course. Actually, it decreased (0.97 point) with significance. On the contrary, the overall mean of students in the WBFA group shows that they had their average score increased (0.04 point) even though it was not large enough for indicating significant difference. Regarding the four language features, students in both CPFA and WBFA groups gained significant improvement in the *reading comprehension* and *vocabulary* sections. It should be noted that the overall means of the CPFA group regarding the section of *reading comprehension* increased 0.66 while that of the WBFA group gained 1.01 points. Related to the section of *vocabulary*, the overall mean of CPFA group increased 1.01 while that of the WBFA group gained 1.63 points.

In the following subsections (4.2.1-4.2.9), the results from the comparisons of the pre- and the post-test means in both CPFA and WBFA groups within each discipline are reported. This is followed by the results from the comparisons of the pre- and post-test means in connection with the four language features in the CPFA and WBFA groups within each discipline. The analysis aims to examine whether students who use WBFA in each subgroup get higher achievement scores on *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* than students who use CPFA during the time of the study.

4.2.1 Faculty of Fine and Applied Arts

Figure 4.16 shows the results in relation to the comparisons of the overall pre- and post-test means of the stratified random sample from the Faculty of Fine and Applied Arts (AR). With reference to the CPFA group ($n = 6$), the results disclose that the overall mean of the post-test (19.33) was higher than that of the pre-test (18.17). In addition, regarding the WBFA group ($n = 3$), its overall post-test mean was 14.00 while the overall mean of the pre-test was 12.00. As a result, it is apparent that both groups of AR students did increase their achievement in terms of overall means at the end of the semester.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.16 Comparisons on pre-test and post-test means within both groups of AR students

Table 4.16 shows the *t*-tests from the comparisons between the pre- and post-tests within the CPFA and WBFA groups concerning AR students. The results inform that even the overall mean of the post-test in the CPFA group was higher than that of the pre-test, there was not a significant difference between those means, Wilcoxon signed ranks test z = -.542, NS at $p > .05$. It also indicates that even though the overall post-test mean in the WBFA group was higher than that of the pre-test, there was no significant difference between them, z = -1.604, NS at $p > .05$. In other words, there was no significant improvement in terms of overall means of both groups in this discipline.

Table 4.16 Wilcoxon signed ranks tests on means of pre-test and post-test in CPFA and WBFA groups of AR students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CPFA	Post	19.3333	6	4.50185	1.83787	-.542
	Pre	18.1667	6	6.43169	2.62573	
WBFA	Post	14.0000	3	1.00000	.57735	1.604
	Pre	12.0000	3	1.73205	1.00000	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Table 4.17 shows the *t*-tests indicating whether there were significant differences between the pre- and post-test means in the four language features tested in both groups of AR students. Pertaining to the CPFA group, the results reveal that the post-test means of the *reading comprehension* (4.17) and *vocabulary* (3.83) sections were higher

than those of the same sections in their pre-test (3.50 and 2.67, consecutively). However, there were no significant differences between those means. The *t*-score for the *reading comprehension* section was $t = 1.581$, NS at $p > .05$ while $t = 1.659$, NS at $p > .05$ was for the *vocabulary* section. The post-test means of *completing dialogues* (5.50) and *grammar* (5.83) sections were lower, with no significance, than those in the pre-test (5.67 and 6.33, respectively). Therefore, it is manifest that there was no significant progress with reference to the four language features in the CPFA group in this discipline at the end of the semester.

Table 4.17 T-tests on four language features of pre-test and post-test in CPFA group of AR students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	5.5000	6	2.81069	1.14746	-0.123
	Pre	5.6667	6	2.87518	1.17379	
RC	Post	4.1667	6	1.83485	.74907	1.581
	Pre	3.5000	6	1.37840	.56273	
VO	Post	3.8333	6	1.32916	.54263	1.659
	Pre	2.6667	6	.81650	.33333	
GR	Post	5.8333	6	1.94079	.79232	-0.311
	Pre	6.3333	6	3.14113	1.28236	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In relation to the WBFA group, the results reveal that the post-test means in the *vocabulary* (4.33) and *grammar* (4.67) sections were higher than the *vocabulary* (2.00) and *grammar* (2.33) pre-test means. However, there were no significant differences between those means. The result of Wilcoxon signed ranks test in the *vocabulary* section was $z = 1.633$, NS at $p > .05$ and that in the *grammar* section was $z = 1.633$, NS at $p > .05$. In the *completing dialogues* and *reading comprehension* sections, the post-test means were lower than those in the pre-test, as shown in Table 4.18. The result of Wilcoxon signed ranks test in the *completing dialogues* section was $z = -1.342$, NS at $p > .05$, while that of the *reading comprehension* section was $z = -1.633$, NS at $p > .05$. Hence, there were no significant differences between pre- and post-test performance of AR WBFA students related to the four language features.

Table 4.18 Wilcoxon signed ranks tests on four language features of pre-test and post-test in WBFA group of AR students

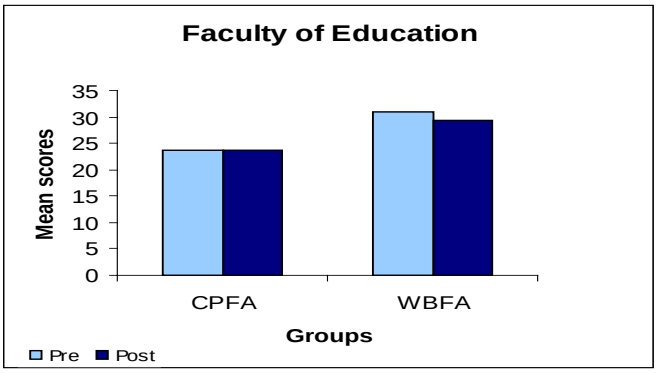
WPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	3.0000	3	1.73205	1.00000	-1.342
	Pre	4.0000	3	1.00000	.57735	
RC	Post	2.0000	3	1.00000	.57735	-1.633
	Pre	3.6667	3	.57735	.33333	
VO	Post	4.3333	3	1.52753	.88192	1.633
	Pre	2.0000	3	1.00000	.57735	
GR	Post	4.6667	3	1.52753	.88192	1.633
	Pre	2.3333	3	.57735	.33333	

WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In brief, in the CPFA group of AR students, their overall average score increased 1.17 points while those in the WBFA group had their overall average score gained 2.00 points. However, there were no significant differences. Regarding the four language features tested in both CPFA and WBFA groups, there were no significant differences between the means of the pre- and post-tests in the three sections of *completing dialogues*, *vocabulary* and *grammar*. However, there was a significant difference in the section of *reading comprehension* in the WBFA group where the pre-test mean was larger than that of the post-test.

4.2.2 Faculty of Education

Figure 4.17 shows the results of the comparisons between the means of the pre-test and post-test within the CPFA (*n = 12*) and WBFA (*n = 7*) groups of students from the Faculty of Education (ED). With reference to the CPFA group, the overall mean of the pre-test was as the same as the post-test overall mean (23.75). With regard to the WBFA group, the overall mean of the post-test (29.28) was lower than the overall pre-test mean (31.00). Consequently, it can be verified that both groups of students in this discipline did not have their average scores increased at the end of the semester. It seemed that the intervention of the WBFA program did not elevate ED students in the WBFA group to gain their overall achievement.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.17 Comparisons on pre-test and post-test means within both groups of ED students

The results from the *t*-test, as shown in Table 4.19, reveal that were no significant differences between the means of the pre- and post-tests in both CPFA and WBFA groups of the ED students. In addition, there was no significant improvement in terms of mean scores in both groups since the pre- and post-test means in the CPFA group were the same while, in the WBFA group, their average score decreased 1.71 points at the end of the WBFA program.

Table 4.19 T-tests on means of pre-test and post-test in CPFA and WBFA groups of ED students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	23.7500	12	9.16639	2.64611	.000
	Pre	23.7500	12	6.32635	1.82626	
WBFA	Post	29.2857	7	3.72891	1.40940	-1.247
	Pre	31.0000	7	2.08167	.78680	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

In Table 4.20, the results of dependent *t*-tests on the four language features in the CPFA group of ED students are reported. It shows that the post-test means were higher than the pre-test means in the *reading comprehension* and *vocabulary* sections. However, there were no significant differences between those means. The *t*-score in *reading comprehension* section was $t = 0.140$, NS at $p > .05$, and in *vocabulary* section, $t = 1.666$, NS at $p > .05$. For *completing dialogues* and *grammar* sections, their post-test means were smaller than their pre-test means. It should be noticed that the pre-test mean

in the section of *completing dialogues* was significantly higher than that of the post-test mean, t = -2.385, sig. at $p < .05$.

Table 4.20 T-tests on four language features of pre-test and post-test in CPFA group of ED students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	6.0833	12	1.92865	.55675	-2.385*
	Pre	7.3333	12	2.34844	.67794	
RC	Post	4.9167	12	2.23437	.64501	0.140
	Pre	4.8333	12	1.94625	.56183	
VO	Post	5.6667	12	2.10339	.60720	1.666
	Pre	4.1667	12	1.80067	.51981	
GR	Post	7.0833	12	4.31611	1.24595	-0.371
	Pre	7.4167	12	3.23218	.93305	

*Sig. at p <.05
CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Table 4.21 illustrates the results of dependent *t*-tests in the WBFA group of ED students involving the four language features. It indicates that the post-test means of the *reading comprehension* and *vocabulary* sections were higher, with no significance, than the means of their pre-test. The *t*-test results in the *reading comprehension* and *vocabulary* sections were t = 0.956, NS at $p > .05$, and t = 1.469, NS at $p > 0.5$, consecutively. In the *completing dialogues* and *grammar* sections, their post-test means were smaller than their pre-test means. It should be noted that in the section of *grammar*, the pre-test mean was significantly larger than the post-test mean, t = -3.333, sig. at $p < .05$.

In brief, the results regarding the stratified random sample from the Faculty of Education reveal that there were no significant differences in terms of overall average scores between the pre-test and post-test in both CPFA and WBFA groups. In relation to the four language features tested, there were also no significant differences between the pre- and post-test means in CPFA and WBFA groups concerning the *reading comprehension* and *vocabulary* sections. However, there were significant differences in the *completing dialogues* section of the CPFA group and in the*grammar* section of the WBFA group where their post-test means were smaller than those in the pre-test. The means in those sections reduced 1.25 and 2.86 points, respectively.

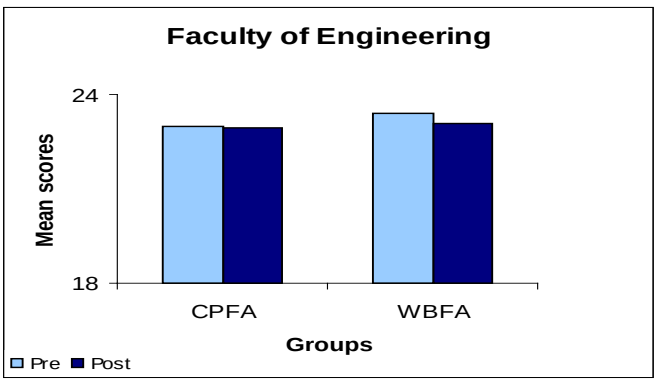
Table 4.21 T-tests on four language features of pre-test and post-test in WBFA group of ED students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.2857	7	1.88982	.71429	-1.871
	Pre	8.2857	7	1.38013	.52164	
RC	Post	6.7143	7	1.38013	.52164	0.956
	Pre	6.0000	7	1.52753	.57735	
VO	Post	7.1429	7	2.03540	.76931	1.469
	Pre	5.7143	7	2.36039	.89214	
GR	Post	8.1429	7	1.46385	.55328	-3.333*
	Pre	11.0000	7	1.82574	.69007	

*Sig. at p <.05
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

4.2.3 Faculty of Engineering

Figure 4.18 demonstrates the comparisons between the pre- and post-test means of students from the Faculty of Engineering (EN). The results reveal that, in the CPFA group (*n =12*), the overall post-test mean (22.92) was slightly lower than the overall pre-test mean (23.00). In addition, in the WBFA group (*n = 12*), the overall post-test mean (23.08) was also slightly lower than the overall pre-test mean (23.42).



CPFA = Conventional paper-and-pencil formative assessment
WBFA = Web-based formative assessment

Figure 4.18 Comparisons on pre-test and post-test means within both groups of EN students

Table 4.22 shows the results of dependent *t*-tests from investigating whether there was any significant difference between the overall pre- and post-tests in both CPFA and WBFA groups of EN students. The results indicate that there were no significant

differences between the overall means of the pre- and post-tests in both groups. The overall average score of the CPFA group decreased 0.08 point, while that of the WBFA group reduced 0.33 point at the end of the WBFA program.

Table 4.22 T-tests on means of pre-test and post-test in CPFA and WBFA groups of EN students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	22.9167	12	6.15642	1.77721	-0.061
	Pre	23.0000	12	6.63325	1.91485	
WBFA	Post	23.0833	12	5.91544	1.70764	-0.247
	Pre	23.4167	12	4.73782	1.36769	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

The results of the *t*-tests in the four language features within the CPFA group of EN students are shown in Table 4.23. The post-test means in the sections *of reading comprehension* and *vocabulary* were higher than their pre-test means. However, there was a significant difference in only the *vocabulary* section, $t = 2.960$, sig. at $p < .05$. In *reading comprehension* section, there was no significant difference, $t = 1.149$, NS at $p > .05$. For *completing dialogues* and *grammar* sections, the post-test means were smaller than the pre-test means. It should be noticed that, in the section of *grammar*, the pre-test mean was significantly larger than the post-test mean, $t = -2.862$, sig. at $p < 0.5$.

Table 4.23 T-tests on four language features of pre-test and post-test in CPFA group of EN students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	6.0833	12	1.50504	.43447	-1.765
	Pre	7.2500	12	2.70101	.77971	
RC	Post	5.0000	12	2.21565	.63960	1.149
	Pre	4.0000	12	1.75810	.50752	
VO	Post	5.5000	12	2.27636	.65713	2.960*
	Pre	3.2500	12	1.76455	.50938	
GR	Post	6.3333	12	1.66969	.48200	-2.862*
	Pre	8.5000	12	3.11886	.90034	

*Sig. at $p < .05$ CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Table 4.24 shows that the post-test means in the sections of *reading comprehension* and *vocabulary* in the WBFA group of EN students were higher than the pre-test means. According to the results of dependent *t*-tests, there were significant differences between those means. The *t*-test result of the *reading comprehension* section was $t = 2.068$, sig. at $p < .05$ while that of the *vocabulary* section was $t = 3.464$, sig. at $p < .05$. The post-test means of the *completing dialogues* and *grammar* sections were significantly smaller than their pre-test means, $t = -3.000$, sig. at $p < .05$ and $t = -3.742$, sig. at $p < 0.5$, consecutively.

Table 4.24 T-tests on four language features of pre-test and post-test in WBFA group of EN students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	6.2500	12	1.35680	.39167	-3.000*
	Pre	7.7500	12	1.81534	.52404	
RC	Post	5.1667	12	2.03753	.58818	2.068*
	Pre	3.9167	12	1.44338	.41667	
VO	Post	5.4167	12	2.27470	.65665	3.464*
	Pre	3.4167	12	1.72986	.49937	
GR	Post	6.2500	12	2.49089	.71906	-3.742*
	Pre	8.3333	12	2.22928	.64354	

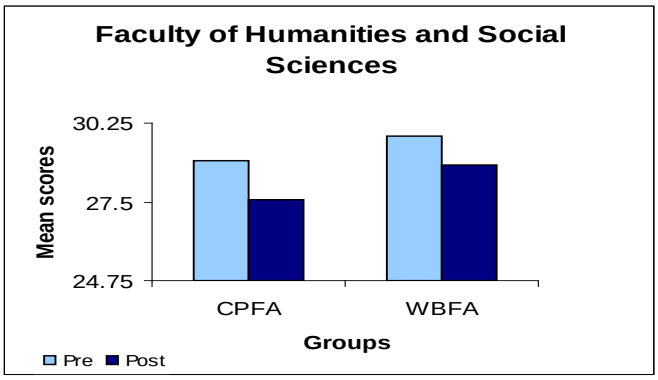
*Sig. at $p < .05$ WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In summary, the results of EN students reveal that there were no significant differences between overall means of the pre- and post-tests in both CPFA and WBFA groups. In relation to the four language features, in the CPFA group, there was a significant difference between the pre-test and post-test means in the *vocabulary* section. In the WBFA group, there were significant differences between means of the pre- and post-tests in the two sections of *reading comprehension* and *vocabulary* where the post-test means were higher than their pre-test means.

4.2.4 Faculty of Humanities and Social Sciences

Figure 4.19 shows results from the stratified random sample of the Faculty of Humanities and Social Sciences (HS), which is the largest sample group in the study,

concerning the comparisons of the overall means between the pre- and post-test are presented in Figure 4.17. In the CPFA group (*n = 27*) the result reveals that the overall post-test mean (27.56) was lower than the overall pre-test mean (28.93). In addition, in the WBFA group (*n = 25*), the overall post-test mean (28.80) was also lower than the overall pre-test mean (29.80).



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.19 Comparisons on pre-test and post-test means within both groups of HS students

The results of dependent *t*-tests in Table 4.25 indicate that there were no significant differences between the means of the overall pre-test and overall post-test within both CPFA and WBFA groups. At the end of the WBFA program, the results specify that the overall average score of the CPFA group decreased 1.37 points while the overall average score of the WBFA group also decreased 1.00 point.

Table 4.25 T-tests on means of pre-test and post-test in CPFA and WBFA groups of HS students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	27.5556	27	6.60031	1.27023	-1.660
	Pre	28.9259	27	6.12640	1.17903	
WBFA	Post	28.8000	25	5.78792	1.15758	-1.015
	Pre	29.8000	25	5.29150	1.05830	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Table 4.26 shows that there were significant differences between means of the pre- and post-tests in two sections of the four language features tested in the CPFA group. The post-test mean of the *reading comprehension* section was higher with significance than the pre-test mean, $t = 3.602$, sig. at $p < .05$. In the *vocabulary* section, the post-test mean was also higher with significance than the pre-test mean, $t = 5.577$, sig. at $p < .05$. However, in *completing dialogues* and *grammar* sections, their post-test means were smaller with significance than their pre-test means.

Table 4.26 T-tests on four language features of pre-test and post-test in CPFA group of HS students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.4815	27	1.78391	.34331	-4.721*
	Pre	8.8148	27	1.44214	.27754	
RC	Post	6.5926	27	2.37388	.45685	3.602*
	Pre	5.0741	27	1.66239	.31993	
VO	Post	6.3333	27	1.90142	.36593	5.577*
	Pre	4.4815	27	2.06380	.39718	
GR	Post	7.1481	27	2.50697	.48247	-6.107*
	Pre	10.5556	27	2.62141	.50449	

*Sig. at p <.05
CD = Completing dialogues
VO = Vocabulary
CPFA = Conventional paper-and-pencil formative assessment
RC = Reading comprehension
GR = Grammar

In Table 4.27, the results disclose that there were significant differences between means of the pre-test and post-test in the WBFA group of HS students in relation to the four language features. The post-test mean of the *reading comprehension* section was significantly higher than that of the pre-test, $t = 3.468$, sig. at $p < .05$. In the *vocabulary* section, the post-test mean was significantly higher than the pre-test mean, $t = 5.850$, sig. at $p < .05$. However, in *completing dialogues* and *grammar* sections, the post-test means were lower with significance than their pre-test means.

In summary, there were no significant differences between overall means of the pre- and post-tests in the stratified random sample of HS students in both CPFA and WBFA groups at the end of the WBFA program. Regarding to the four language features tested the average points in *reading comprehension* and *vocabulary* sections increased with significance in both CPFA and WBFA groups in this discipline.

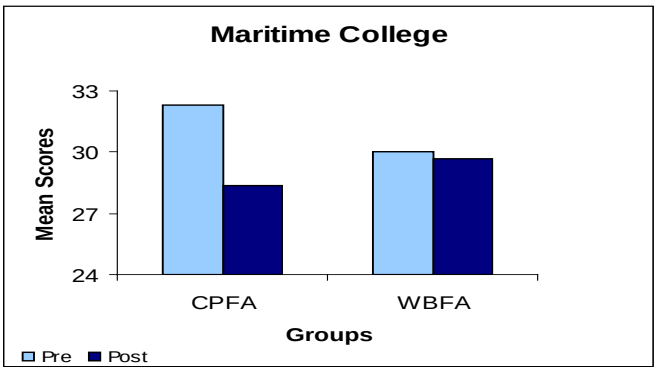
Table 4.27 T-tests on four language features of pre-test and post-test in WBFA group of HS students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.2000	25	1.47196	.29439	-5.938*
	Pre	8.8400	25	1.54596	.30919	
RC	Post	7.1600	25	2.17332	.31475	3.468*
	Pre	5.6000	25	1.41421	.29822	
VO	Post	6.6800	25	1.57374	.31475	5.850*
	Pre	4.8400	25	1.49108	.29822	
GR	Post	7.7600	25	2.33238	.46648	-5.963*
	Pre	10.5200	25	2.48529	.49706	

*Sig. at p <.05
CD = Completing dialogues
VO = Vocabulary
WBFA = Web-based formative assessment
RC = Reading comprehension
GR = Grammar

4.2.5 Maritime College

In Figure 4.20, the results from the comparisons between means of the overall pre- and post-tests analysed through the data collected from stratified random samples in the Maritime College (MT) are reported. In the CPFA group ($n = 3$), the overall post-test mean (28.33) was lower than the overall pre-test mean (32.33). In addition, in the WBFA group ($n = 3$), the overall post-test mean (29.67) is slightly lower than the overall pre-test mean (30.00).



CPFA = Conventional paper-and-pencil formative assessment
WBFA = Web-based formative assessment

Figure 4.20 Comparisons on pre-test and post-test means within both groups of MT students

The Wilcoxon signed ranks test results shown in Table 4.28 indicate that there were no significant differences between means of the overall pre- and post-tests in both CPFA

and WBFA groups. In the CPFA group, the post-test mean was smaller, with no significance, than that of the pre-test. The overall average point of the group decreased 4.00 points. In the WBFA group, the overall average point also decreased 0.33 point when the post-test mean was lower than that of the pre-test at the end of the WBFA program.

Table 4.28 Wilcoxon signed ranks tests on means of pre-test and post-test in CPFA and WBFA groups of MT students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CPFA	Post	28.3333	3	5.68624	3.28295	-1.604
	Pre	32.3333	3	3.21455	1.85592	
WBFA	Post	29.6667	3	3.21455	1.85592	-.447
	Pre	30.0000	3	2.00000	1.15470	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

With regard to the four language features tested in the CPFA group of MT students, the Wilcoxon signed ranks test results displayed in Table 4.29 confirm that there were no significant differences between MT CPFA students' performance in pre- and post-tests. The findings reveal that there was no improvement in terms of average scores in the four language features tested. In the *completing dialogues* section, the pre- and post-test means were equal. For the last three sections of *reading comprehension, vocabulary* and *grammar*, their post-test means were smaller than their pre-test means.

Table 4.29 Wilcoxon signed ranks tests on four language features of pre-test and post-test in CPFA group of MT students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	7.0000	3	1.73205	1.00000	.000
	Pre	7.0000	3	1.00000	.57735	
RC	Post	6.3333	3	1.52753	.88192	-1.000
	Pre	6.6667	3	1.15470	.66667	
VO	Post	6.0000	3	1.73205	1.00000	-1.633
	Pre	7.3333	3	1.15470	.66667	
GR	Post	9.0000	3	2.00000	1.15470	-1.633
	Pre	11.0000	3	1.00000	.57735	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

The results in Table 4.30 shows that, in the WBFA group of MT students, the post-test means in the sections of *reading comprehension* and *vocabulary* were higher than the pre-test means. However, they were not large enough to indicate statistical differences. The post-test means in *completing dialogues* and *grammar* sections were smaller than their pre-test means with no statistical differences.

Table 4.30 Wilcoxon signed ranks tests on four language features of pre-test and post-test in WBFA group of MT students

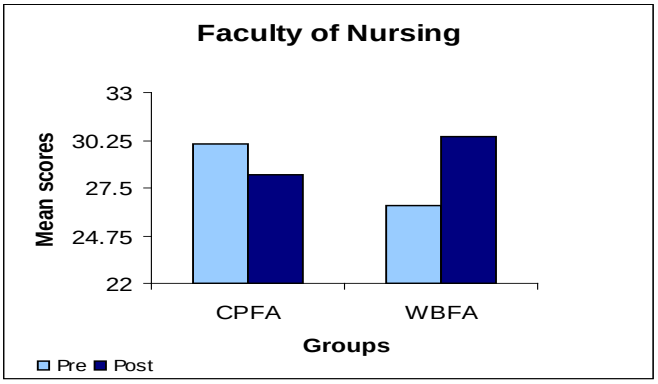
WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	6.0000	3	1.00000	.57735	-1.732
	Pre	7.0000	3	1.00000	.57735	
RC	Post	7.0000	3	1.00000	.57735	1.633
	Pre	5.6667	3	.57735	.33333	
VO	Post	7.6667	3	1.52753	.88192	1.414
	Pre	6.3333	3	.57735	.33333	
GR	Post	9.0000	3	2.00000	1.15470	-1.604
	Pre	11.0000	3	1.00000	.57735	

WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In summary, the results relevant to the stratified random sample in the College of Maritime reveal that there were no significant differences between overall means of the pre- and post-tests in both CPFA and WBFA groups. In addition, there were not significant differences between means of the pre- and post-tests in relation to all sections of *completing dialogues, reading comprehension*, *vocabulary* and *grammar*.

4.2.6 Faculty of Nursing

Figure 4.21 shows the results obtained from the stratified random sample in the Faculty of Nursing (NU). In the CPFA group ($n = 4$), it indicates that the overall post-test mean (28.25) was smaller than that of the pre-test (30.00). In contrast, the overall post-test mean (30.50) was higher than that of the pre-test mean (26.50) in the WBFA group ($n = 4$). Hence, it is evident that the students' outcome in the WBFA group surpasses that in the CPFA group in this discipline.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.21 Comparisons on pre-test and post-test means within both groups of NU students

From the results illustrated in Table 4.31, it is obvious that the average scores of the CPFA group of NU students decreased 1.75 points while that of the WBFA group increased 4.00 points. However, the Wilcoxon signed ranks test results revealed that there were no significant differences between the NU students' performance of pre- and post-tests in both CPFA and WBFA groups ($p > .05$).

Table 4.31 Wilcoxon signed ranks tests on means of pre-test and post-test in CPFA and WBFA groups of NU students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CPFA	Post	28.2500	4	3.30404	1.65202	-1.841
	Pre	30.0000	4	4.24264	2.12132	
WBFA	Post	30.5000	4	5.06623	2.53311	-1.473
	Pre	26.5000	4	1.91485	.95743	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

From the results of the Wilcoxon signed ranks tests in Table 4.32, there were no significant differences between means of the pre- and post-tests in the CPFA group of NU students in relation to the four language features. In the sections of *completing dialogues* and *grammar*, the post-test means were smaller than the pre-test means. In the *reading comprehension* section, the post-test mean was higher than the pre-test means with no significance, z = -1.633, NS at $p > 0.5$. In the *vocabulary* section, the means of both pre- and post-tests were the same.

Table 4.32 Wilcoxon signed ranks tests on four language features of pre-test and post-test in CPFA group of NU students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	6.5000	4	1.29099	.64550	-1.633
	Pre	7.7500	4	1.50000	.75000	
RC	Post	7.0000	4	1.41421	.70711	1.134
	Pre	6.2500	4	2.62996	1.31498	
VO	Post	5.7500	4	.95743	.47871	.000
	Pre	5.7500	4	1.70783	.85391	
GR	Post	9.0000	4	.81650	.40825	-1.890
	Pre	10.2500	4	.95743	.47871	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Table 4.33 shows the results the Wilcoxon signed ranks tests indicating differences between the means of both pre- and post-tests in the four language features in the WBFA group of NU students. It reveals that in all sections of *completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*, the means of the post-test were larger than the pre-test means. However, there were no significant differences between NU WBFA students' performance in both pre- and post-tests.

Table 4.33 Wilcoxon signed ranks tests on four language features of pre-test and post-test in WBFA group of NU students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	7.0000	4	1.41421	.70711	.577
	Pre	6.7500	4	.50000	.25000	
RC	Post	7.0000	4	1.41421	.70711	1.857
	Pre	5.5000	4	1.29099	.64550	
VO	Post	7.5000	4	1.29099	.64550	1.841
	Pre	5.2500	4	.50000	.25000	
GR	Post	9.0000	4	2.44949	1.22474	.000
	Pre	9.0000	4	1.41421	.70711	

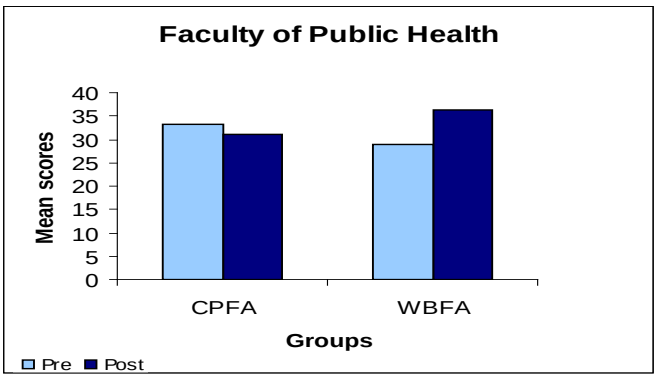
WBFA = Web-based formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In brief, the results reveal that there were not significant differences between the overall means of pre- and post-tests within both CPFA and WBFA groups of NU students, at

the end of the WBFA program. With regard to the four language features tested, there were also no significant differences between NU WBFA students' performance in pre- and post-tests.

4.2.7 Faculty of Public Health

Figure 4.22 shows the results obtained from the stratified random sample in the Faculty of Public Health (PH). In the CPFA group (*n = 6*), it reveals that the overall post-test mean (31.14) was lower than the overall pre-test mean (33.14). On the contrary, in the WBFA group (*n = 6*), the overall post-test mean (36.33) was higher than the overall pre-test mean (29.00). Hence, there is verification that the overall mean of the WBFA group obviously improved at the end of the WBFA program.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.22 Comparisons on pre-test and post-test means within both groups of PH students

The results of *t*-tests shown in Table 4.34 indicate that there were not significant differences between the overall pre- and post-test means in the CPFA group. Actually, their overall average score decreased 2.00 points at the end of the semester. However, the average score of the WBFA group increased 7.33 points at the end of the WBFA program. In addition, the WBFA overall post-test mean was significantly higher than the overall pre-test mean, $t = 4.400$, sig. at $p < .05$. As a result, there is evidence that the intervention of the WBFA program did have positive impacts on students' achievement in the WBFA group.

Table 4.34 T-tests on means of pre-test and post-test in CPFA and WBFA groups of PH students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	31.1429	7	6.84175	2.58594	-0.741
	Pre	33.1429	7	7.05759	2.66752	
WBFA	Post	36.3333	6	2.87518	1.17379	4.400*
	Pre	29.0000	6	3.94968	1.61245	

*Sig. at $p < .05$
CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

With reference to Table 4.35, PH students who used conventional paper-based formative assessment had little improvement in the *reading comprehension* section. Their post-test mean (6.86) was higher, with no significance, than that of the pre-test (6.29), $t = 1.188$, NS at $p > .05$. However, they failed to get better average scores in the *completing dialogues, vocabulary*, and *grammar* sections. Their post-test means in those three sections were smaller than their pre-test means.

Table 4.35 T-tests on four language features of pre-test and post-test in CPFA group of PH students

CPFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.0000	7	2.08167	.78680	-1.400
	Pre	8.2857	7	1.60357	.60609	
RC	Post	6.8571	7	1.34519	.50843	1.188
	Pre	6.2857	7	1.49603	.56544	
VO	Post	7.4286	7	1.81265	.68512	-0.956
	Pre	8.1429	7	1.57359	.59476	
GR	Post	9.8571	7	3.23669	1.22336	-0.471
	Pre	10.4286	7	3.20713	1.21218	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

In the WBFA group, PH students did very well in all sections of the language features tested, as shown in Table 4.36. Their post-test means indicate higher scores outperforming their previous means in the pre-test. In the *completing dialogues* section, they had 8.33 as their post-test average score while the average score of the pre-test in the same section was 7.33. However, there was no significance between the means, $t = 1.225$, NS at $p > .05$. The post-test mean in the *reading comprehension* section (7.83)

was significantly higher than that of the pre-test (5.83), t = 3.873, sig. at p < .05. In addition, in the *vocabulary* section, the post-test mean (8.83) was higher with significance than the pre-test mean (5.83), t = 5.196, sig. at p < .05. In the last section of *grammar*, the post-test mean (11.33) was also larger than the pre-test mean (10.00), but there was no significant difference, t = 1.865, NS at p > .05.

Table 4.36 T-tests on four language features of pre-test and post-test in WBFA group of PH students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	8.3333	6	1.03280	.42164	1.225
	Pre	7.3333	6	1.36626	.55777	
RC	Post	7.8333	6	.40825	.16667	3.873*
	Pre	5.8333	6	1.47196	.60093	
VO	Post	8.8333	6	.98319	.40139	5.196*
	Pre	5.8333	6	1.16905	.47726	
GR	Post	11.3333	6	1.63299	.66667	1.865
	Pre	10.0000	6	1.09545	.44721	

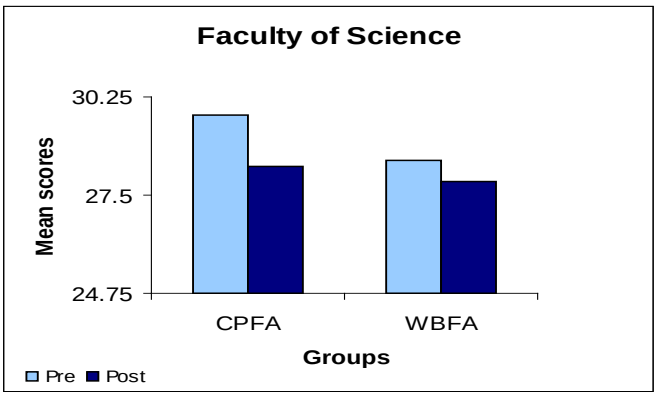
*Sig. at p <.05
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

In brief, the overall average score of PH students in the CPFA group indicates no improvement at the end of the course. In contrast, the overall average score of PH students in the WBFA group increased significantly at the end of the WBFA program. With reference to the four language features tested, there were not significant differences in the CPFA group. On the contrary, there was significantly improvement in the sections of *reading comprehension* and *vocabulary* in the WBFA group where their post-test means were higher than those of the pre-test.

4.2.8 Faculty of Science

According to the results obtained from the stratified random sample in the Faculty of Science (SC), the overall post-test average score (28.30) of the CPFA group (n= *20*) was lower than that of the pre-test (29.75). In the WBFA group (n = *15*), the overall post-test mean (27.87) was also smaller than the overall pre-test mean (28.47). The

results designate that there was no improvement in terms of overall average scores within both CPFA and WBFA group of students in this discipline, as in Figure 4.23.



CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Figure 4.23 Comparisons on pre-test and post-test means within both groups of SC students

On account of the *t*-test results shown in Table 4.37, it is obvious that there were no significant differences within both CPFA and WBFA groups of SC students. The *t*-scores resulted in both groups were t = -1.669 and t = -0.530, NS at $p > .05$, respectively. Actually, their overall average scores decreased 1.45 points for students in the CPFA group and 0.60 point for those in the WBFA group.

Table 4.37 T-tests on means of pre-test and post-test in CPFA and WBFA groups of SC students

Group	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CPFA	Post	28.3000	20	6.47343	1.44750	-1.669
	Pre	29.7500	20	5.63705	1.26048	
WBFA	Post	27.8667	15	4.30725	1.11213	-0.530
	Pre	28.4667	15	1.88478	.48665	

CPFA = Conventional paper-and-pencil formative assessment WBFA = Web-based formative assessment

Table 4.38 shows the result of dependent *t*-tests relating to the four language features tested in the CPFA group. It indicates that there were no significant differences between the pre- and post-test means in all sections. It should be noticed that the post-test means of all sections were smaller than their pre-test means. As a result, it is visible that

students in the group did not have improvement in all sections of *completing dialogues, reading comprehension, vocabulary* and *grammar*, in terms of their average scores.

Table 4.38 T-tests on four language features of pre-test and post-test in CPFA group of SC students

CBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.0000	20	1.89181	.42302	-0.546
	Pre	7.2500	20	1.55174	.34698	
RC	Post	6.1000	20	2.42574	.54241	-1.905
	Pre	7.2500	20	1.55174	.34698	
VO	Post	6.2000	20	1.32188	.29558	-0.244
	Pre	6.3000	20	1.94936	.43589	
GR	Post	9.1000	20	2.97180	.66451	-1.530
	Pre	10.1500	20	2.81490	.62943	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

From the dependent *t*-test results displayed in Table 4.39, it indicates that there were no significant differences in the four language features tested in the WBFA group of SC students. The post-test means of the *reading comprehension* (6.27) and *vocabulary,* (5.80) sections were slightly higher than the means in their pre-test (6.00 and 5.40, respectively). Their *t*-scores in the *reading comprehension* and *vocabulary* sections were $t = 0.718$, and $t = 0.823$, NS at $p > .05$, consecutively. The post-means of *completing dialogues* and *grammar* sections were lower than their pre-means. It should be noticed that, in the section of *grammar*, the pre-test mean (9.80) was significantly higher than that of the post-test (8.60), $t = -1.790$, sig. at $p < .05$.

In summary, the results from the stratified random sample in the Faculty of Science reveal that there was no improvement in terms of overall average scores within both CPFA and WBFA groups at the end of the WBFA program. In addition, there was no significant improvement in relation to the four language features tested in the CPFA group. There was a little improvement in *reading comprehension* and *vocabulary* sections in the WBFA group; however, there were no significant differences.

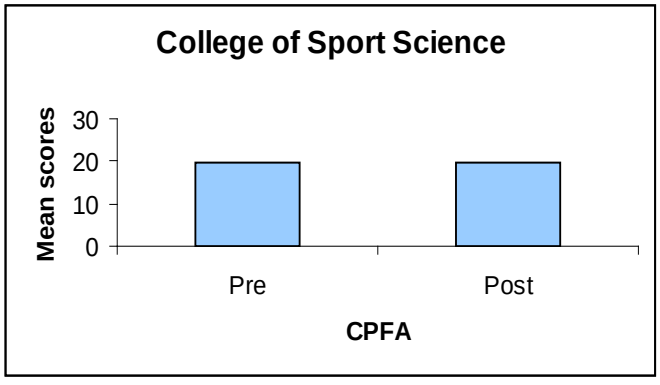
Table 4.39 T-tests on four language features of pre-test and post-test in WBFA group of SC students

WBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*T*
CD	Post	7.2000	15	1.74028	.44934	-0.115
	Pre	7.2667	15	1.38701	.35813	
RC	Post	6.2667	15	1.03280	.26667	0.718
	Pre	6.0000	15	1.00000	.25820	
VO	Post	5.8000	15	1.42428	.36775	0.823
	Pre	5.4000	15	.91026	.23503	
GR	Post	8.6000	15	2.50143	.64587	-1.790*
	Pre	9.8000	15	1.08233	.27946	

*Sig. at p <.05
WBFA = Web-based formative assessment
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

4.2.9 College of Sport Science

The results regarding the stratified random sample in the College of Sport Science (SS) are reported in Figure 4.24. The findings obtained from the CPFA group of SS students are presented solely due to the inadequate number of SS students in the WBFA group for computing their overall pre-and post-test means. The CPFA overall pre- and post-test means were the same at 19.75.



CPFA = Conventional paper-and-pencil formative assessment

Figure 4.24 Comparisons on pre-test and post-test means within CPFA group of SS students

Table 4.40 shows the results of the *t*-test with regard to the four language features in the CPFA group of SS students. It indicates that the post-test means of the *vocabulary* and

grammar sections were higher than those of the pre-test. However, there were no significant differences. Results of the Wilcoxon signed ranks tests for *vocabulary* and *grammar* sections were the same at $z = 1.633$, NS at $p > .05$. The pre- and post-test means relating to the *reading comprehension* section were the same while the pre-test mean was higher than that of the post-test in the *completing dialogues* section with no significant differences,

Table 4.40 Wilcoxon signed ranks tests on four language features of pre-test and post-test in CPFA group of SS students

CBFA	Test	Mean	N	Std. Deviation	Std. Error Mean	*Z*
CD	Post	3.7500	4	1.70783	.85391	-1.841
	Pre	6.7500	4	.95743	.47871	
RC	Post	4.2500	4	2.21736	1.10868	.000
	Pre	4.2500	4	2.62996	1.31498	
VO	Post	5.2500	4	2.06155	1.03078	1.633
	Pre	4.0000	4	1.63299	.81650	
GR	Post	6.5000	4	2.88675	1.44338	1.633
	Pre	4.7500	4	1.70783	.85391	

CPFA = Conventional paper-and-pencil formative assessment
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

To sum up, for SS students, there were no significant differences between SS CPFA students' performance in the pre- and post-tests with reference to the four language features.

With reference to the results of the pre-test and post-test design, this entire section reports that there were no significant differences between overall pre and post-means in the WBFA group. However, there was a statistical difference in the CPFA group where the overall mean of the pre-test was higher than that of the post-test. In relation to the four language features, there were significant increases in *reading comprehension* and *vocabulary* sections in CPFA and WBFA groups. However, students' overall means in the *completing dialogues* and *grammar* sections reduced significantly in both CPFA and WBFA groups.

Regarding the results of the pre- and post-tests in the CPFA and WBFA groups in each subgroup, there was significant overall improvement in the WBFA group of the students from the Faculty of Public Health (PH). In relation to each language feature tested in each discipline, relating to the CPFA group, there was significant improvement in *vocabulary* section of the students in the Faculty of Engineering (EN), and in the *reading comprehension* and *vocabulary* sections in the Faculty of Humanities and Social Sciences (HS). Concerning the WBFA group, there was significant improvement in the two sections of *reading comprehension* and *vocabulary* in three disciplines: the Faculty of Engineering, the Faculty of Humanities and Social Sciences, and the Faculty of Public Health.

In the following section, the results obtained from all participants in the WBFA group are used to investigate how the frequency of participation in the WBFA program may have affected students' learning outcomes

4.3 Comparisons of ABA and BLA groups

To investigate whether the frequency of participation in the WBFA program affected students' learning outcomes, the data collected from the post-test of English skills in the WBFA group (*n = 76*) were analysed. In addition, the records of the students' login times to perform WBFA were tallied and calculated to find out the mean of their frequency of (a) participation in and (b) performance on the WBFA. The number of attempts that these students logged in to perform WBFA was considered when the performance on WBFA was actually occurring. The number of times students logged in without doing the WBFA tasks was not included in the comparisons. This was not considered necessary because the study aimed to investigate the effect of the WBFA program on students who were actually influenced by performing the WBFA tasks.

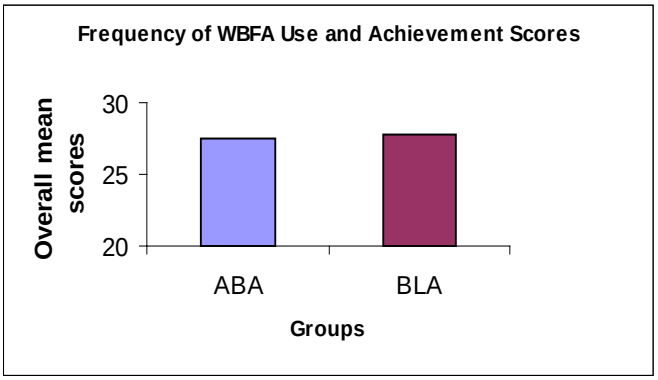
Results showed that the overall average frequency of students' participation for performing WBFA was 15.24. Then, the students in the WBFA group were divided into two groups in relation to the number of times they logged in and performed WBFA, regardless of the length of time they spent and the scores they received once they completed each test set. Students who used WBFA more frequently than the average frequency for the total WBFA group were identified as "above average" and referred to

as the ABA group, while the students who used WBFA at a lower frequency rate than the average of the total group were placed in the below average referred to as the BLA group. As a result, there were 27 students (35.53%) who had above average use of WBFA (16-63 times) and 49 students (64.47%) with below average use of WBFA (1-15 times). Finally, the means of the ABA and BLA groups obtained from the post-test-only design were compared.

The *t*-test for independent or uncorrelated samples was employed to investigate the statistical significance of the post-test of English skills. In addition, the Mann-Whitney *U* test was used to locate significance in the subgroups having small sample size (less than five). The *F*-test was also utilised to ensure that the homogeneity-of-variance assumption was not violated due to the unequal sample sizes. It was hypothesised that students in the ABA group would have better learning outcomes than those in the BLA group with the alpha level being set at .05.

This section (4.3) reports the results of statistical comparison of the ABA and BLA groups overall performance on the post-test of English skills. Then, these two groups are compared on the basis of their performance on each of the four language features of *completing dialogues, reading comprehension, vocabulary* and *grammar.* In the subsections (4.3.1 - 4.3.6), the results of comparisons between the ABA and BLA groups performance on the post-test of English skills within each discipline are reported. This is followed by comparisons between the ABA and BLA groups' performance on each of the four language features within each discipline.

Figure 4.25 shows the results of the comparison on overall mean scores between the ABA (*n = 27*) and BLA (*n = 49*) groups. It reveals that the overall mean of the ABA group was 27.44 while that of the BLA group was 27.84. It indicates very little difference between the overall mean of the ABA group and that of the BLA group.



ABA = Above average group BLA = Below average group

Figure 4.25 Comparison on overall means between ABA and BLA groups

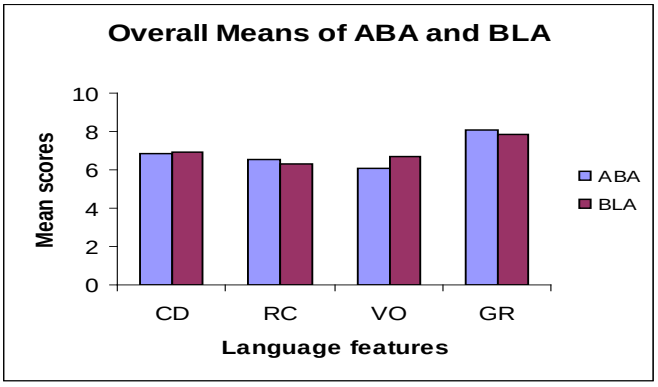
In Table 4.41, the result indicates that there was no statistical difference between the ABA and BLA groups' performance on the post-test of English skills (t = -0.258, NS, at $p > .05$).

Table 4.41 T-tests on overall means between ABA and BLA groups

Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
ABA	27.4444	27	6.29611	1.21169	-0.258
BLA	27.8367	49	6.44640	.92091	

ABA = Above average group BLA = Below average group

Figure 4.26 shows the results of overall mean comparisons between the performance of students in relation to the four language features within the ABA and BLA groups. It shows that the students in the ABA group performed almost the same as the BLA group in the sections of *reading comprehension* and *grammar*. The mean of the ABA group in the *reading comprehension* section (6.56) was a little higher than that of the BLA (6.33). In addition, the ABA mean in the *grammar* section (8.07) was also slightly higher than that of the BLA (7.86). In the sections of *completing dialogues* and *vocabulary*, the BLA group performed with negligible difference from the ABA group. The BLA mean for *completing dialogues* was 6.94 while that of the ABA was 6.81. In the *vocabulary* section, the BLA mean was 6.67 while that of the ABA group was 6.07.



ABA = Above average group
CD = Completing dialogues
VO = Vocabulary
BLA = Below average group
RC = Reading comprehension
GR = Grammar

Figure 4.26 Comparisons on means of four language features between ABA and BLA groups

Table 4.42 shows that at the end of the WBFA program the mean scores of students in the ABA group had a negligible difference from those in the BLA group in the *reading comprehension* and *grammar* sections. Meanwhile, the BLA means indicated insignificant difference when compared with the ABA means in the sections of *completing dialogues* and *vocabulary*.

Table 4.42 T-tests on overall means of four language features between ABA and BLA groups

Language features	Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
CD	ABA	6.8148	27	1.86129	.35820	-0.301
	BLA	6.9388	49	1.70059	.24294	
RC	ABA	6.5556	27	1.69464	.32613	0.504
	BLA	6.3265	49	2.23036	.31862	
VO	ABA	6.0741	27	2.09259	.40272	-1.253
	BLA	6.6735	49	1.83016	.26145	
GR	ABA	8.0741	27	2.36848	.45581	0.359
	BLA	7.8571	49	2.78388	.39770	

ABA = Above average group
CD = Completing dialogues
VO = Vocabulary
BLA = Below average group
RC = Reading comprehension
GR = Grammar

In summary, there was no significant difference between the overall means of both ABA and BLA groups at the end of the WBFA program. In addition, there were no significant differences between overall means of the four language features tested in

both groups. Hence, the frequency of participation in using WBFA during the time of the study did not influence different students' learning outcomes. In other words, the statistical results reported in the section support the argument that more attempts on performing WBFA than the average was not an indicator that students would yield better learning outcomes.

In the following subsections (4.3.1-4.3.6), the results from comparisons of the mean scores, obtained from the post-test-only design, between the ABA and BLA groups within each student discipline are reported. It is noted that there are no such comparisons for the Faculty of Fine and Applied Arts (AR), the Maritime College (MT), and College of Sport Science (SS) because there were no students with above average performance on WBFA in these subgroups.

4.3.1 Faculty of Education

Table 4.43 shows the results of the comparison of the mean scores of ABA (*n = 4*) and BLA (*n = 3*) groups of the Faculty of Education (ED) on their overall performance on the post-test of English skills. The mean of the BLA group (29.67) and the mean of the ABA group (29.00) had a negligible difference. Subsequently, the result of the Mann-Whitney *U* tests revealed that there was no significant difference between both groups' performance on the post-test of English skills ($U = 7.0$, NS at $p > .05$).

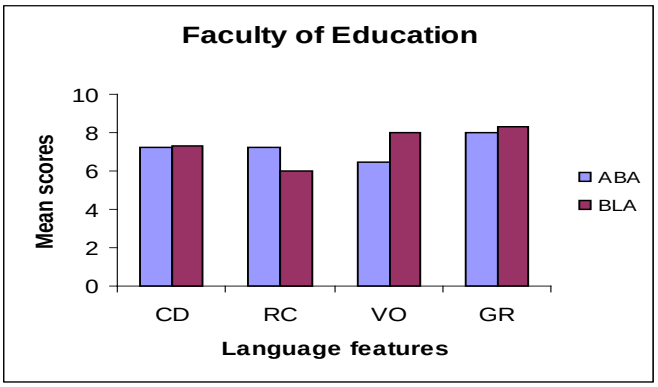
Table 4.43 Mann-Whitney U tests on overall means between ABA and BLA groups of ED students

Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
ABA	29.0000	4	3.16228	1.58114	7.0
BLA	29.6667	3	5.13160	2.96273	

ABA = Average and above average group BLA = Below average group

Figure 4.27 shows the results of the mean comparisons regarding the four language features between the ABA and BLA groups of ED students. It specifies that the students in the ABA group did better than those in the BLA group only in the section of *reading comprehension*. The means of the ABA group on the section was 7.25, while that of the BLA was 6.00. However, the means of the ABA group in the *completing dialogues*

(7.25), *vocabulary* (6.50), and *grammar* (8.00) sections were all smaller than the means of the *completing dialogues* (7.33), *vocabulary* (8.00), and *grammar* (8.33) sections in the BLA group, respectively.



ABA = Above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Figure 4.27 Comparisons on means of four language features between ABA and BLA of ED students

Tables 4.44 illustrates that ED students in the ABA mean was slightly above that of the BLA group in the section of *reading comprehension*; however, there was no significant difference ($p > .05$.). The BLA means also had negligible differences from those of the ABA in the three sections of *completing dialogues*, *vocabulary*, and *grammar*. According to the results of the Mann-Whitney *U* tests, there were no statistical differences between those means.

In relation to the results concerning the ABA and BLA groups of ED students in the WBFA group, it demonstrates that frequency in the participation of the WBFA program did not affect the students' learning outcomes. Students with above and below average use of WBFA in the Faculty of Education had no significant difference in their learning outcomes at the end of the WBFA program.

Table 4.44 Mann-Whitney U tests on four language features in ABA and BLA groups of ED students

Language features	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	ABA	7.2500	4	2.06155	1.03078	6.5
	BLA	7.3333	3	2.08167	1.20185	
RC	ABA	7.2500	4	.95743	.47871	9.0
	BLA	6.0000	3	1.73205	1.00000	
VO	ABA	6.5000	4	2.08167	1.04083	8.5
	BLA	8.0000	3	2.00000	1.15470	
GR	ABA	8.0000	4	1.15470	.57735	7.0
	BLA	8.3333	3	2.08167	1.20185	

ABA = Above average group BLA = Below average group
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.3.2 Faculty of Engineering

Table 4.45 shows the result of the Mann-Whitney *U* tests independent *t*-test from the comparison between the overall mean scores of ABA (*n = 2*) and BLA (*n = 10*) groups obtained from the stratified random sample in the Faculty of Engineering (EN). It indicates that the mean of the BLA group (24.00) was significantly higher than the mean of the ABA group (18.50), $U = 14.5$, NS at $p > .05$.

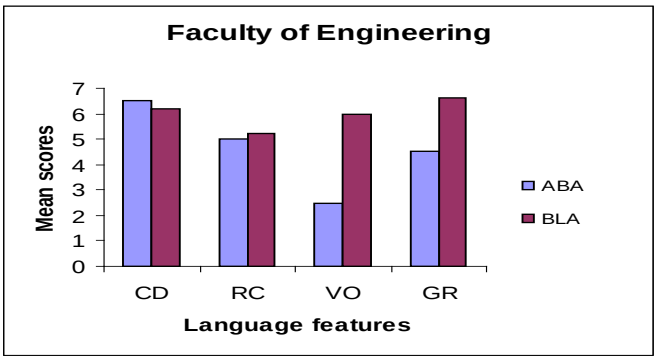
Table 4.45 Mann-Whitney U tests on overall means between ABA and BLA groups of EN students

Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
ABA	18.5000	2	.70711	.50000	14.5
BLA	24.0000	10	6.09189	1.92642	

ABA = Above average group BLA = Below average group

Figure 4.28 shows results of the mean comparisons from the stratified random samples in the Faculty of Engineering regarding the four language features between the ABA and BLA groups. It indicates that the EN students in the ABA group had higher mean score than that in the BLA group only in the section of *completing dialogues*. The mean of the ABA group in the *completing dialogues* section was 6.50 while that of the BLA was 6.20. However, the means of the ABA group in the *reading comprehension* (5.00), *vocabulary* (2.50), and *grammar* (4.50) sections were all lower than the means of

reading comprehension (5.20), *vocabulary* (6.00), and *grammar* (6.60) sections in the BLA group, consecutively.



ABA = Above average group
CD = Completing dialogues
VO = Vocabulary
BLA = Below average group
RC = Reading comprehension
GR = Grammar

Figure 4.28 Comparisons on means of four language features between ABA and BLA of EN students

Tables 4.46 illustrates that at the end of the WBFA program, EN students in the ABA group performed better than those in the BLA group in the section of *completing dialogues* while the BLA group outperformed the ABA group in the three sections of *reading comprehension, vocabulary*, and *grammar*. However, according to the results of the Mann-Whitney *U* tests, there were no significant differences between those means ($p > .05$).

Table 4.46 Mann-Whitney U tests on four language features in ABA and BLA groups of EN students

Language features	Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	ABA	6.5000	2	3.53553	2.50000	10.0
	BLA	6.2000	10	.91894	.29059	
RC	ABA	5.0000	2	1.41421	1.00000	10.0
	BLA	5.2000	10	2.20101	.69602	
VO	ABA	2.5000	2	2.12132	1.50000	18.0
	BLA	6.0000	10	1.88562	.59628	
GR	ABA	4.5000	2	2.12132	1.50000	15.0
	BLA	6.6000	10	2.50333	.79162	

ABA = Above average group
CD = Completing dialogues
VO = Vocabulary
BLA = Below average group
RC = Reading comprehension
GR = Grammar

In summary, for EN students, the overall mean of BLA group was significantly higher than that of the ABA group. In addition, with reference to the language features, the mean in the *vocabulary* section of the BLA group was higher, with significance, than that in the ABA group.

4.3.3 Faculty of Humanities and Social Sciences

Table 4.47 shows the results from the comparison between the overall mean scores of ABA (*n = 8*) and BLA (*n = 17*) groups obtained from the stratified random sample in the Faculty of Humanities and Social Sciences (HS). It indicates that the mean of the BLA group (29.41) was higher than the mean of the ABA group (27.38). However, the result of the independent *t*-test reveals that there was no significant difference between both means, $t = -0.673$, NS at $p > .05$.

It should be noticed that the results of independent *t*-test demonstrated in Table 4.47 might not be robust, due to violation of the homogeneity-of-variance assumption. This might be caused by the unequal sample sizes and the difference between the standard deviations of the means in both groups. As a result, the result from the *F*–test to ensure that both means had equal variance was rejected (Bachman, 2004; Glass and Hopkins, 1996). Therefore, the means in both groups might not be good representatives to the population means and the *t*-test result might not be vigorous. This is the only problematic *t*-test result occurred in the study.

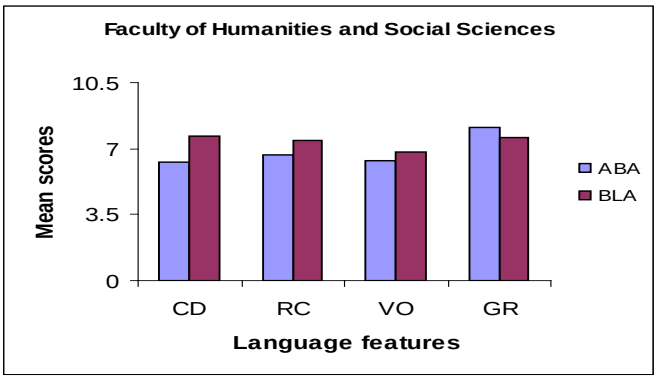
Table 4.47 T-tests on overall means between ABA and BLA groups of HS students

Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
ABA	27.3750	8	7.61460	2.69217	-0.673
BLA	29.4118	17	4.86131	1.17904	

ABA = Above average group BLA = Below average group

Figure 4.29 shows results of the mean comparisons from the stratified random samples in the Faculty of Humanities and Social Sciences pertaining to the four language features between the ABA and BLA groups. It notifies that the HS students in the ABA group outperformed those in the BLA group only in the section of *grammar*. The mean

of the ABA group in the *grammar* section was 8.13 while that of the BLA was 7.59. However, the means of the ABA group in the *completing dialogues* (6.25), *reading comprehension* (6.63), and *vocabulary* (6.38) sections were all smaller than the means of the *completing dialogues* (7.65), *reading comprehension* (7.41), and *vocabulary* (6.82) sections in the BLA group, respectively.



ABA = Above average group
CD = Completing dialogues
VO = Vocabulary
BLA = Below average group
RC = Reading comprehension
GR = Grammar

Figure 4.29 Comparisons on means of four language features between ABA and BLA of HS students

Table 4.48 shows the results of the independent *t*-tests relevant to the four language features between the ABA and BLA groups in the Faculty of Humanities and Social Sciences. For the first three sections of *completing dialogues, reading comprehension,* and *vocabulary,* the results reveal that the BLA group had higher means than the ABA group. Specifically, the BLA mean in *completing dialogues* section was significantly higher than the ABA mean, t = -1.946, sig. at $p < 0.5$. However, in the last section of *grammar*, the mean score of the ABA group was higher with no significance than the BLA mean, t = 0.580, NS at $p > .05$.

In short, there might be no significance between the overall means of the ABA and BLA groups of HS students. Nonetheless, there was a significant difference in the section of *completing dialogues* between the means of both groups where the BLA mean was higher than that of the ABA.

Table 4.48 T-tests on four language features in ABA and BLA groups of HS students

Language features	Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
CD	ABA	6.2500	8	1.90863	.67480	-1.946*
	BLA	7.6471	17	.99632	.24164	
RC	ABA	6.6250	8	2.77424	.98084	-0.728
	BLA	7.4118	17	1.87279	.45422	
VO	ABA	6.3750	8	1.92261	.67975	-0.588
	BLA	6.8235	17	1.42457	.34551	
GR	ABA	8.1250	8	1.95941	.69276	0.580
	BLA	7.5882	17	2.52633	.61273	

*Sig at $p < .05$
ABA = Above average group BLA = Below average group
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

4.3.4 Faculty of Nursing

Table 4.49 shows the result of the Mann-Whitney *U* tests from the comparison between the overall performance of ABA (*n = 2)* and BLA (*n = 2*) groups obtained from the stratified random sample in the Faculty of Nursing (NU). It indicates that the overall mean of the BLA group (32.50) was higher than the overall mean of the ABA group (28.50). However, the result of the Mann-Whitney *U* tests reveals that there was no significant difference between both means, $U = 2.0$, NS at $p > .05$.

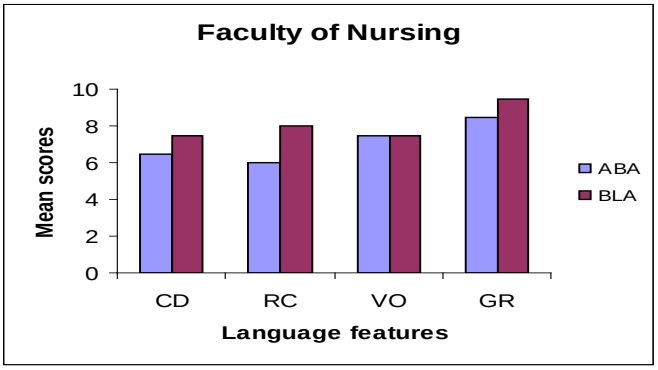
Table 4.49 Mann-Whitney U tests on overall means between ABA and BLA groups of NU students

Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
ABA	28.5000	2	7.77817	5.50000	2.0
BLA	32.5000	2	.70711	.50000	

ABA = Average and above average group BLA = Below average group

Figure 4.30 shows results of the mean comparisons from NU students concerning the language features between the ABA and BLA groups. It notifies that the NU students in the ABA group had the same average score as that in the BLA group in the section of *vocabulary*. The mean of the *vocabulary* section in both groups was 7.50. However, the means of the ABA group in the *completing dialogues* (6.50), *reading comprehension* (6.00), and *grammar* (8.50) sections were all lower than the means of the *completing*

dialogues (7.50), *reading comprehension* (8.00), and *grammar* (9.50) sections in the BLA group, consecutively.



ABA = Average and above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Figure 4.30 Comparisons on means of four language features between ABA and BLA of NU students

Table 4.50 shows the results of the Mann-Whitney *U* tests regarding the four language features between the ABA and BLA group in the Faculty of Nursing. The results demonstrate that the means in *completing dialogues*, *reading comprehension*, and *grammar* sections of the ABA group were smaller that those in the BLA group. However, there were no significant differences between those means. For the section of *vocabulary*, the means in both groups were similar.

Table 4.50 Mann-Whitney U tests on four language features in ABA and BLA groups of NU students

Language features	Groups	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	ABA	6.5000	2	2.12132	1.50000	2.5
	BLA	7.5000	2	.70711	.50000	
RC	ABA	6.0000	2	1.41421	1.00000	4.0
	BLA	8.0000	2	.00000	.00000	
VO	ABA	7.5000	2	.70711	.50000	2.0
	BLA	7.5000	2	2.12132	1.50000	
GR	ABA	8.5000	2	3.53553	2.50000	2.5
	BLA	9.5000	2	2.12132	1.50000	

ABA = Average and above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

In a nutshell, there was no significant difference between overall means in the ABA and BLA groups of the NU students. In addition, there were no significant differences in relation to the four language features of both groups.

4.3.5 Faculty of Public Health

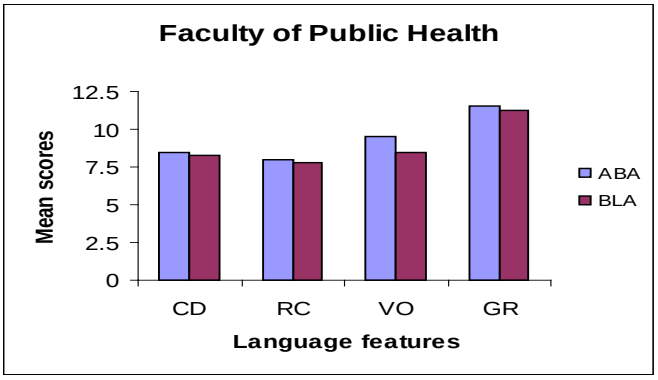
Table 4.51 shows the results from the comparison between the mean scores of ABA (*n = 2*) and BLA (*n = 4*) groups obtained from the stratified random sample in the Faculty of Public Health (PH). It indicates that the mean of the ABA group (37.50) was higher than the mean of the BLA group (35.75). However, the result of the Mann-Whitney *U* tests reveals that there was no significant difference between both means ($p > .05$).

Table 4.51 Mann-Whitney U tests on overall means between ABA and BLA groups of PH students

Group	Mean	N	Std. Deviation	Std. Error Mean	*U*
ABA	37.5000	2	2.12132	1.50000	5.5
BLA	35.7500	4	3.30404	1.65202	

ABA = Average and above average group BLA = Below average group

Figure 4.31 shows the results of the mean comparisons from the PH students with reference to the four language features between the ABA and BLA groups. It indicates that the PH students in the ABA group outperformed those in the BLA group in all four sections of *completing dialogues, reading comprehension, vocabulary*, and *grammar*. The means of the ABA group in the *completing dialogues* (8.50), *reading comprehension* (8.00), *vocabulary* (9.50), and *grammar* (11.50) sections were higher than the means of the *completing dialogues* (8.25), *reading comprehension* (7.75), *vocabulary* (8.50), and *grammar* (11.25) sections in the BLA group.



ABA = Average and above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

Figure 4.31 Comparisons on means of four language features between ABA and BLA of PH students

Table 4.52 shows the results of the Mann-Whitney *U* tests to compare the PH students' performance between ABA and BLA groups with regard to the four language features. The results reveal that students in the ABA group had higher means than the BLA group in all sections. However, there were no significant differences between those means.

Table 4.52 Mann-Whitney U tests on four language features in ABA and BLA groups of PH students

Language features	Groups	Mean	N	Std. Deviation	Std. Error Mean	*U*
CD	ABA	8.5000	2	.70711	.50000	5.0
	BLA	8.2500	4	1.25831	.62915	
RC	ABA	8.0000	2	.00000	.00000	5.0
	BLA	7.7500	4	.50000	.25000	
VO	ABA	9.5000	2	.70711	.50000	6.5
	BLA	8.5000	4	1.00000	.50000	
GR	ABA	11.5000	2	2.12132	1.50000	4.5
	BLA	11.2500	4	1.70783	.85391	

ABA = Average and above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

In brief, there was no significant difference between the overall mean of the ABA and BLA groups among the PH students. In addition, there were no significant differences between the means of the four language features in both groups.

4.3.6 Faculty of Science

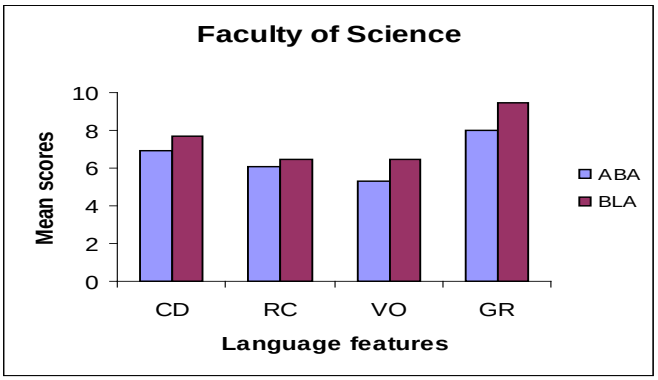
Table 4.53 shows the results from the comparison between the mean scores of ABA (*n = 9*) and BLA (*n = 6*) groups obtained from the stratified random samples in the Faculty of Science (SC). It indicates that the mean of the BLA group (30.17) was significantly higher than the mean of the ABA group (26.33), *t* = -1.895, sig. at *p* < .05.

Table 4.53 T-tests on overall means between ABA and BLA groups of SC students

Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
ABA	26.3333	9	4.24264	1.41421	-1.895*
BLA	30.1667	6	3.54495	1.44722	

*Sig. at *p* < .05
ABA = Average and above average group BLA = Below average group

Figure 4.32 shows descriptive results relating to the four language features between the ABA and BLA groups. It was obvious that SC students in the BLA group outperformed those in the ABA group in sections of *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar*. The means of the BLA group in the *completing dialogues* (7.67), *reading comprehension* (6.5), *vocabulary* (6.5), and *grammar* (9.50) sections were higher than the means of the *completing dialogues* (6.89), *reading comprehension* (6.11), *vocabulary* (5.33), and *grammar* (8.00) sections in the ABA group.



ABA = Average and above average group BLA = Below average group
CD = Completing dialogues RC = Reading comprehension
VO = Vocabulary GR = Grammar

Figure 4.32 Comparisons on means of four language features between ABA and BLA of SC students

From the results of the independent *t*-tests to comprehend statistical significances between the means of the four language features in both ABA and BLA groups, as shown in Table 4.54, the average scores in all language features of the BLA group of SC students were negligibly different from those of the ABA group. Hence, there were no significant differences between those means.

Table 4.54 T-tests on four language features in ABA and BLA groups of SC students

Language features	Groups	Mean	N	Std. Deviation	Std. Error Mean	*t*
CD	ABA	6.8889	9	1.76383	.58794	-0.840
	BLA	7.6667	6	1.75119	.71492	
RC	ABA	6.1111	9	.92796	.30932	-0.664
	BLA	6.5000	6	1.22474	.50000	
VO	ABA	5.3333	9	1.11803	.37268	-1.525
	BLA	6.5000	6	1.64317	.67082	
GR	ABA	8.0000	9	2.34521	.78174	-1.119
	BLA	9.5000	6	2.66458	1.08781	

ABA = Average and above average group
BLA = Below average group
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

In brief, according to the data analysis through the post-test only design at the end of the WBFA program, there were no significant differences between the overall means of the ABA and BLA groups. In addition, there were no statistical differences between the means of the four language features in both groups of SC students.

In addition, it should be noticed as illustrated in Table 4.55 that there were 29 students who completed their performance on the 12 test sets of the WBFA program and there were 47 students who did not. To comprehend whether there was a significant difference between their achievement at the end of the program, their means from the post-test-only design was employed along with the applications of the independent *t*-test. The result reveals that there was not a statistical difference between both means, $t = 0.678$, NS at $p > .05$.

Table 4.55 T-tests on mean comparison between COM and INCOM groups

Group	Mean	N	Std. Deviation	Std. Error Mean	*t*
COM	28.3103	29	6.21118	1.15339	0.678
INCOM	27.2766	47	6.82929	.99615	

COM = Students who completed the WBFA program
INCOM = Students who did not complete the WBFA program

There were 29 students in the WBFA group who complete 12 test sets of the WBFA program. Ten students were from the Faculty of Humanities and Social Sciences, eight from the Faculty of Science, four from the Faculty of Education, another four from the Faculty of Engineering, two from the Faculty of Public Health, and one from the Faculty of Nursing. Twenty-two of them were in the ABA group while seven were in the BLA group. In addition, there were 47 students in the WBFA group who did not complete all 12 test sets in the WBFA program. Four of these students were in the ABA group while 43 of them were in the BLA group.

In summary, this section reports that there were no statistical differences between the overall performance of ABA and BLA groups. In addition, there were no significant differences between both groups in the means. In each discipline, there were significant differences in the Faculty of Engineering and the Faculty of Science where BLA means were higher than those of the ABA group. In relation to the four language features, there were also significant differences in *vocabulary* section of the Faculty of Engineering and the *completing dialogues* section in the Faculty of Humanities and Social Sciences where the BLA means were higher than those of the ABA group. Additional details regarding the number of attempts on the sections of *completing dialogues, reading comprehension, vocabulary* and *grammar* of the students in each subgroup are illustrated in Table 4.56.

Table 4.56 Frequencies of students' participation on WBFA

Subgroup	No. of Students	Freq. of CD	Freq. of RC	Freq. of VO	Freq. of GR	Total of Freq.
AR	3	13	4	3	7	27
ED	7	65	41	38	40	184
EN	12	62	28	23	26	139
HS	25	107	66	62	69	304
MT	3	14	2	4	0	20
NU	4	28	12	9	9	58
PH	6	28	16	15	11	70
SC	15	98	80	72	96	346
SS	1	4	3	0	1	8
Total	76	419	252	226	259	1156

AR = Faculty of Fine and Applied Arts
ED = Faculty of Education
EN = Faculty of Engineering
HS = Faculty of Humanities and Social Sciences
MT = Maritime College
NU = Faculty of Nursing
PH = Faculty of Public Health
SC = Faculty of Science
CD = Completing dialogues
RC = Reading comprehension
VO = Vocabulary
GR = Grammar

4.4 Summary

According to the data analysis from the post-test-only design in the first section (4.1), there was no significant difference between the performance on English skills of the students in CPFA and WBFA groups. Hence, the hypothesis stating that the overall mean of the WBFA group would be higher than that of the CPFA group at alpha level .05 was rejected. In relation to the comparisons of the means between the CPFA and WBFA groups in each discipline, the overall means of the WBFA groups in the Faculty of Education and the Faculty of Public Heath were significantly higher than the means of their CPFA groups. Regarding the comparisons of the four language features, the WBFA overall means in the *reading comprehension* and *vocabulary* sections were significantly higher than the CPFA means. The WBFA group also outperformed the CPFA group in *completing dialogues* and *grammar* sections; however, there were no significant differences between those overall means.

In the second section (4.2), related to the results of the pre-test and post-test design to detect any improvement in the CPFA and WBFA groups during the study, there was no statistical improvement between overall means in the WBFA group. However, the overall mean decreased significantly in the CPFA group. In relation to the overall means of the four language features, there was significant improvement in *reading*

comprehension and *vocabulary* sections in both CPFA and WBFA groups. In the CPFA group, there was significant improvement in *vocabulary* section in the Faculty of Engineering, and the *reading comprehension* and *vocabulary* sections in the Faculty of Humanities and Social Sciences. In the WBFA group, on the other hand, there was significant improvement in *reading comprehension* and *vocabulary* sections in four disciplines: the Faculty of Engineering, the Faculty of Humanities and Social Sciences, the Faculty of Nursing, and the Faculty of Public Health. As a result, the hypotheses stating that, in the WBFA group, the overall mean and means of the four language features in the post-test would be higher than those means in the pre-test at alpha level .05 were rejected.

Also in the third section (4.3), there was no statistical difference between the overall means of ABA and BLA. In addition, there was no significant difference between both groups in the means of the four language features. In each discipline, there were significant differences in the Faculty of Engineering and the Faculty of Science where BLA means were higher than those of ABA. In relation to the four language features, there were also significant differences in *vocabulary* section of the Faculty of Engineering and the *completing dialogues* section of the Faculty of Humanities and Social Sciences where the BLA means were higher than the ABA's. Therefore, the hypothesis stating that the ABA group would have better learning outcomes than those in the BLA group at alpha level .05 was rejected.

In the next chapter, Chapter 5, the data analyses and results from the pre- and post-WBFA attitude questionnaires and the pre- and post-WBFA semi-structured interviews that were administered to determine the students' attitudes toward using the WBFA program in their course are reported.

Chapter 5 Data analyses and results: Students' attitudes

Chapter 4 has reported the results of the analysis of quantitative data obtained from the pre- and post-test scores to compare students' performances on English skills and language features (*completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*) between the CPFA and WBFA groups. In Chapter 5, the results of students' attitudes toward the use of WBFA based on the analysis of quantitative and qualitative data acquired from closed-end items and open-ended responses in the pre- and post-WBFA questionnaires are reported. In addition, the qualitative data from the pre- and post-WBFA semi-structured interviews are analysed and reported. This chapter consists of three main sections. The first section (5.1) is related to the mixed methods data analysis and results with reference to the pre- and post-WBFA questionnaires. The second section (5.2) relates to relevant to the qualitative data analyses and results from the pre- and post-WBFA semi-structured interviews. The findings in this chapter are presented to further explore and illuminate the results presented in the previous chapter. The last section (5.3) provides a summary of the chapter.

5.1 The questionnaire

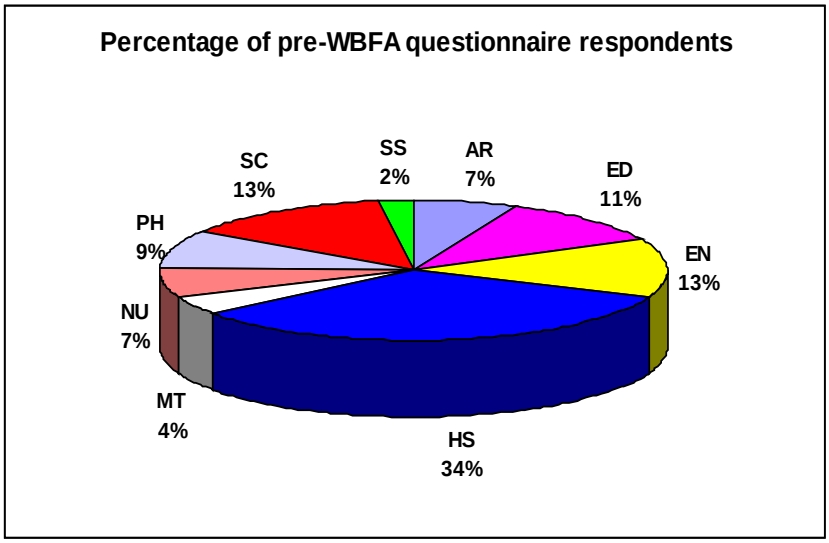
The respondents to the pre- and post-WBFA questionnaires were the stratified random sample of the WBFA group (*n = 76*). They were all Thai first year undergraduate students, who enrolled to study English II (212102) in the second semester of 2004-2005 academic year at the Department of Western Languages, Faculty of Humanities and Social Sciences, Burapha University, Bang Saen Campus. It was anticipated that the WBFA group's attitudes toward the use of the WBFA would inspire positively increase over the duration of the study. Consequently, the information obtained from both questionnaires was used to identify and compare the participants' attitudes prior to and after the program.

5.1.1 The return rate of the pre-WBFA questionnaire

The pre-WBFA questionnaire was delivered to students in the WBFA group at the end of the WBFA tryout session. The researcher had explained the aims of using the

questionnaire to all participants in Thai, their first language. Then they were requested to respond to the questionnaire after they had completed some examples of the WBFA program. Some students completed their questionnaires and returned them to the researcher at the end of the session while other students returned their responses later to their English lecturers because they could not complete their questionnaires in the session. The questionnaires were then delivered by the lecturers to the researcher for the data analysis.

The return rate for the pre-WBFA questionnaire was 59.21% (*n = 45*). As shown in Figure 5.1, percentages of participants who responded to the pre-WBFA questionnaire were categorised in relation to students' nine disciplines: approximately 34% (*n = 15*) of the participants from the Faculty of Humanities and Social Sciences, 13% (*n* = 6) from the Faculty of Engineering, 13% (*n = 6*) from the Faculty of Science, 11% (*n = 5*) from the Faculty of Education, 9% (*n = 4*) from the Faculty of Public Health, 7% (*n = 3*) from the Faculty of Fine and Applied Arts, 7% (*n = 3*) from the Faculty of Nursing, 4% (*n = 2*) from the Maritime College, and 2% (*n = 1*) from the College of Sport Science.



AR = Faculty of Fine and Applied Arts
EN = Faculty of Engineering
MT = Maritime College
PH = Faculty of Public Health
SS = College of Sport Science
ED = Faculty of Education
HS = Faculty of Humanities and Social Sciences
NU = Faculty of Nursing
SC = Faculty of Science

Figure 5.1 Percentage of pre-questionnaire respondents

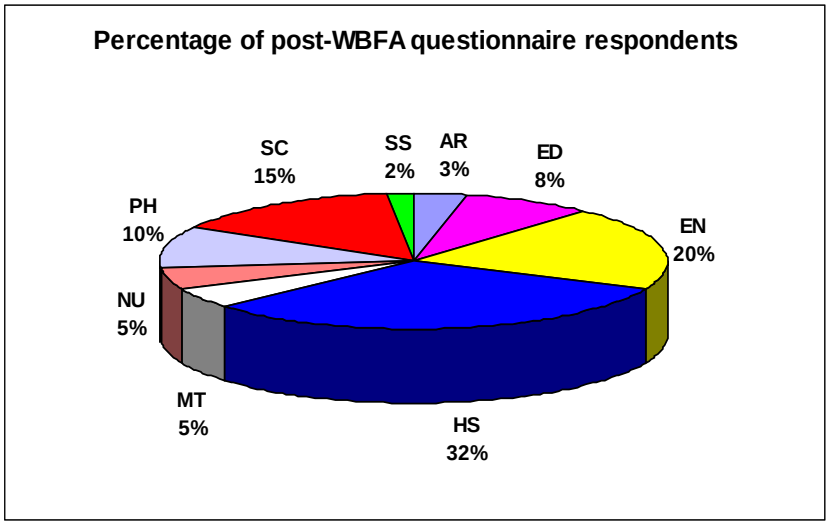
The majority of the respondents in the pre-WBFA questionnaire (*n = 32*, 71%) were female and 29% (*n = 13*) of them were male. The percentage of respondents who stated that they had basic computer skills prior to the study was 56% (*n = 25*) in which 22 respondents had 1-5 year experience in using the WWW and three of the respondents mentioned that they had 6-10 year experience in using the WWW. However, 44% (*n = 20*) of the entire cohort stated that they did not have any experiences in using the WWW prior to the study.

5.1.2 The return rate of the post-WBFA questionnaire

The post-questionnaire was conveyed to WBFA students at the end of the WBFA program. The students were requested to complete the post-WBFA questionnaire, and their lecturers collected the questionnaires and then delivered them to the researcher. After the students' final examinations, this questionnaire was sent one more time by the researcher to these participants through e-mail; however, no more questionnaires were returned to the researcher.

For the post-WBFA questionnaire, the return rate was 78.95% (*n = 60*). As shown in Figure 5.2, percentages of participants who responded to the post-WBFA questionnaire were: approximately 32% (*n = 19*) of respondents from the Faculty of Humanities and Social Sciences, 20% (*n = 12*) from the Faculty of Engineering, 15% (*n = 9*) from the Faculty of Science, 10% (*n = 6*) from the Faculty of Public Health, 8% (*n = 5*) from the Faculty of Education, 5% (*n = 3*) from the Faculty of Nursing, 3% (*n = 2*) from the Faculty of Fine and Applied Arts, 5% (*n = 3*) from the Maritime College, and 2% (*n = 1*) from the College of Sport Science.

In the post-WBFA questionnaire, 72% (*n = 43*) of the respondents were female while 28% (*n = 17*) of them were male. There were 45% (*n = 27*) of the respondents who had basic computer skills prior to the study. Fourteen of these respondents stated that they had 1-5 year experience in using the WWW while 13 of them mentioned that they had 6-10 year experience in using the WWW. However, there were 55% (*n = 33*) of the students in the post-WBFA questionnaire group who did not have any experience in using the WWW before the study.



AR = Faculty of Fine and Applied Arts
EN = Faculty of Engineering
MT = Maritime College
PH = Faculty of Public Health
SS = College of Sport Science
ED = Faculty of Education
HS = Faculty of Humanities and Social Sciences
NU = Faculty of Nursing
SC = Faculty of Science

Figure 5.2 **Percentage of post-questionnaire respondents**

5.1.3 Analyses of quantitative data and results of both questionnares

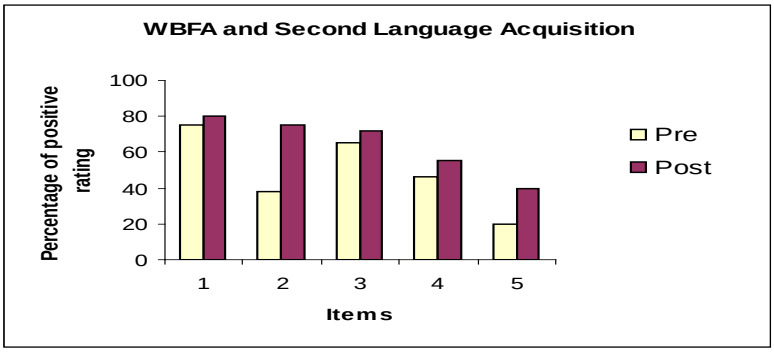
This subsection is related to data analyses and results obtained from the closed-ended items in the pre- and post-WBFA questionnaires which comprised of two parallel forms containing identical items (see Appendix G). The data analyses and results are presented in relation to focused theoretical perspectives in the study as noted below. The results are reported for the entire cohort, regardless of students' gender and discipline.

This part of the questionnaire was in the form of five-point Likert scales with the range of "*Strongly Agree (5), Agree (4), Neutral (3), Disagree (2),*" to "*Strongly Disagree (1).*" The responses were coded and placed in four categories (coded 1, 2, 3 and 4) related to the four theoretical perspectives of second language acquisition, language pedagogy, language assessment, and technology integration, which guided and underpinned the WBFA program. Six of the 25 closed-ended items (i.e., question number 2, 3, 8, 12, 14 and 16) in both questionnaires were reversely coded. This was a deliberate strategy as part of the design to help ensure the variety of the questionnaire items. The Cronbach's alpha of the questionnaire was .80 (see Appendix I). Results on

students' attitudes toward the use of WBFA are presented in the following subsections (5.1.3.1-5.1.3.4) in accordance with the four theoretical perspectives on which WBFA is based.

5.1.3.1 Students' attitudes toward WBFA and second language acquisition (1)

Figure 5.3 shows the results of students' attitudes toward the use of the WBFA program which is based on second language acquisition (SLA). This aimed to increase their achievement by providing a communicative approach to encourage an emphasis on self-study and the constructivist approach to learning, according to Krashen (1985), stated that students can be encouraged to participate in real-world environments to create knowledge for themselves as they study. When asked whether WBFA would *help improve learning language in the course*, most students in both questionnaires stated that using WBFA would help them gain knowledge in learning English in the course. This was confirmed by the increasing percentage from 75.56% at the beginning to 80% (a difference of 4.44%) by the end of the program. The percentage of students who were not certain about this item lowered from 24.44% in the pre-WBFA questionnaire to 18.33% (a reduction of 6.11%) in the post-WBFA questionnaire whereas the percentage of students who did not agree with the item slightly increased to 1.67% at the end from 0% at the beginning of the program. It is evident that the majority of these students thought that WBFA would be great assistance for them in learning English in the course.



1. Help improve learning language in the course (Q5)
2. Encourage self-study (Q6)
3. Enjoy using WBFA in the language course (Q1, Q3)
4. Connect to real-life materials (Q19)
5. Be appropriate for the language course (Q 16)

Figure 5.3 Positive rating of students' attitudes on WBFA and second language acquisition

By reason of encouragement regarding the learner-centred approach which is another pedagogical emphasis of SLA underpinning the WBFA program to stimulate students to enhance self-study to improve their achievement, students were asked to express whether WBFA would *encourage self-study*. From the closed-items in the pre-WBFA questionnaire, most students seemed to be uncertain that WBFA could help them to motivate self-study until they experienced using WBFA. As a result, the percentage increased from 37.77% in the pre-WBFA questionnaire to 76.66% (a difference of 38.89%) in the post-WBFA questionnaire. There were no students who had negative attitudes about this after they used WBFA as the percentage on this item reduced from 20% of students either disagreeing or strongly disagreeing whether WBFA would encourage self-study at the beginning to 0% by the end of the program. Also by the end of the WBFA program students' attitude to whether WBFA would encourage self-study improved by almost 20% with regards to students viewing the program's possible effectiveness with uncertainty at the start to becoming more favourably, since the percentage reduced from 42.22% in the pre-WBFA questionnaire to 23.33% (a reduction of 18.19%) in the post-WBFA questionnaire. Therefore, it can be concluded that WBFA had an encouraging effect on a substantial proportion of students. Thus, this supports the capacity of WBFA to assist students to work independently to construct knowledge about English and to constructively work mostly at their own pace because, in the WBFA program, the students had to do everything by themselves from login to logout.

Furthermore, for both questionnaires, students responded positively that they *enjoyed using WBFA in the language course*. Their attitudes as shown by their positive ratings became increasingly more favourable toward using WBFA. The proportion of positive response ratings increased from 65.56% in the pre-WBFA questionnaire to 71.67% (a difference of 6.11%) in the post-WBFA questionnaire. The percentage of students who did not enjoy using WBFA at the beginning increased slightly from 12.22% to 14.17% (a difference of 1.95%) while the percentage of the students who were not sure whether they enjoyed using the WBFA program decreased from 22.23% at the beginning to 14.17% (a reduction of 8.06%) by the end of the program. It is quite clear that the majority of students had positive experiences and favourable attitudes toward using the WBFA program both before and by the time the program ended.

In addition, the percentage of students who found that WBFA offered them to connect to more authentic materials on the Internet than the conventional paper-and-pencil assessment increased from 46.67% in the pre-WBFA questionnaire to 55% (a difference of 8.36%) after the completion of their WBFA program. They responded to the item that WBFA helped them to *connect to real-life materials*. As a result, more than half of the students in the WBFA group agreed that using WBFA helped them to link to more real-life language features in accordance with the focus of the language communicative approach. This was another evidence to be noted that learning experiences in the WBFA program did actually promote the use of real-life communicative activities and use of language as with the communicative approach. Prior to the introduction of the program students might not be interacting through language to communicate with anyone but were simply doing CPFA in their workbooks. The number of students who neither agreed nor disagreed reduced from 46.67% in the pre-WBFA questionnaire to 41.67% (a reduction of 5%) in the post-WBFA questionnaire as well as the number of students who did not agree reduced 3% from 6.66% at the beginning to 3.33% by the end of the program.

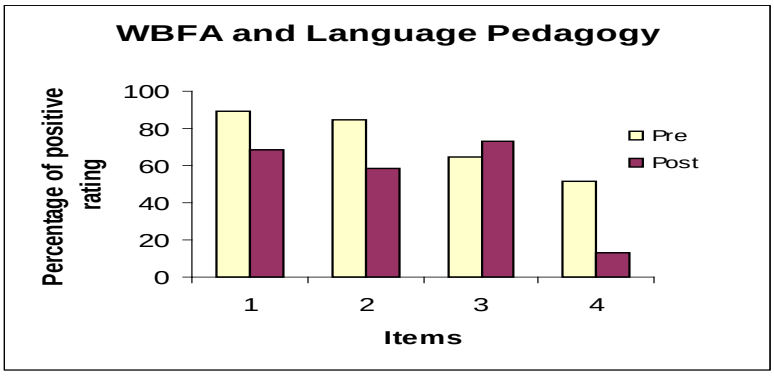
When asked whether WBFA was appropriate to be used in the course they were studying as well as CPFA, the number of students who thought WBFA could *be appropriate for the language course* as suitable as CPFA increased from 20% in the pre-WBFA questionnaire to 40% in the post-WBFA questionnaire. At the same time, the percentage of students who were not satisfied with using WBFA in their course reduced from 42.22% in the pre-WBFA questionnaire to 25% (a reduction of 17.22%) after they used the WBFA program. The number of students who were not sure in relation to this issue decreased only marginally from 37.78% in the pre-WBFA questionnaire to 35% in the post-WBFA questionnaire.

On the basis of this comparison between the pre- and post-WBFA questionnaires, it can be said that the majority of respondents were more positive after than before they used WBFA to support their EFL learning. The WBFA students indicated more favourable acceptances that WBFA did help them improve their learning English in the course, encouraged their self-study, and assisted them to connect to real-life materials. The

students also felt that they enjoyed using WBFA in the course and agreed that the program should be appropriate to be integrated in the course in the future.

5.1.3.2 Students' attitudes toward WBFA and language pedagogy (1)

Figure 5.4 shows the results from the pre- and post-WBFA questionnaires related to WBFA and language pedagogical issues. Relevant to repetition which is quite important in language learning (in keeping with the learning theory underpinning this research), it revealed that the number of students who said they wanted to *repeat doing WBFA to get a good score* decreased from 88.89% at the beginning to 68.33% (a reduction of 20.56%) by the end of the program. The reduction in the percentage of students who did not like reattempting WBFA was negligible, reducing from 2.22% to 1.67%. However, the percentage of students who were uncertain whether they wanted to repeat doing WBFA or not increased from 8.89% to 30% (a difference of 21.11%) by the end of the program. It should be noted that at the time of the WBFA tryout, these students had not yet studied the content in the WBFA program. Their scores may not have been as high as they wished, and this may have made them desire to reattempt to get a higher score. In contrast, the respondents to the post-WBFA questionnaire had studied the lessons before performing the WBFA program. Therefore, they had learned and understood the content to some degree. After certain attempts, they might receive a score they were satisfied with. Thus they did not want to repeat doing WBFA anymore.



1. Repeat doing WBFA to get a good score (Q9)
2. Provide immediate feedback (Q18)
3. Motivate to study the course (Q23)
4. Want to use WBFA as frequently as possible (Q8)

Figure 5.4 Positive rating of students' attitudes on WBFA and language pedagogy

In addition, another possibility might be that correct answers were provided to post-WBFA students once they completed each task while the pre-WBFA students did not have that opportunity.

To *provide immediate feedback* (i.e., a numerical score and correct answers) is another important focus relevant to the language pedagogy in offering the WBFA program to students. Nearly sixty per cent (58.34%) of the students in the post-WBFA questionnaire showed their preference to the immediate score reporting. However, the percentage declined from 84.45% (a difference of 26.11%) at the beginning of the program. The question is what happened to their attitudes during the time of the study. There might be some difficulties they had encountered during that period. The information from the qualitative data analysis and results could possibly provide detailed explanations. On the contrary, the percentage of students who did not want to mention this automatic score reporting increased from 13.33% at the beginning to 40% (a difference of 26.67%) by the end of the WBFA program. It seems that this increasing percentage came from the respondents who had responded positively in the pre-WBFA questionnaire. The percentage between those who did not like this score reporting system reduced from 2.22% at the beginning of the WBFA program to 1.67% by the end of the WBFA program.

The next item pinpoints whether the provision of a non-threatening environment, which is a difference between formative and summative assessment to students, would increase their motivation to learn English in the course. The results show that the percentage of students who believed that WBFA helped *motivate to study the course* increased from 64.45% at the beginning to 73.34% (a difference of 8.89%) by the end of the WBFA program. The percentage of students who did not think that WBFA would motivate them to study the course decreased from 4.44% to 1.67% (a reduction of 2.77%). Similarly, the number of students who were not certain about this item decreased from 31.11% at the beginning to 25% (a reduction of 6.11%) by the end of the WBFA program.

The results of the final area for analysis is shown in Figure 5.4 indicates that 48.89% of the students in the pre-WBFA questionnaire viwed that they would *want to use WBFA as frequently as possible* reduced by 13.33% in the post-WBFA questionnaire. The

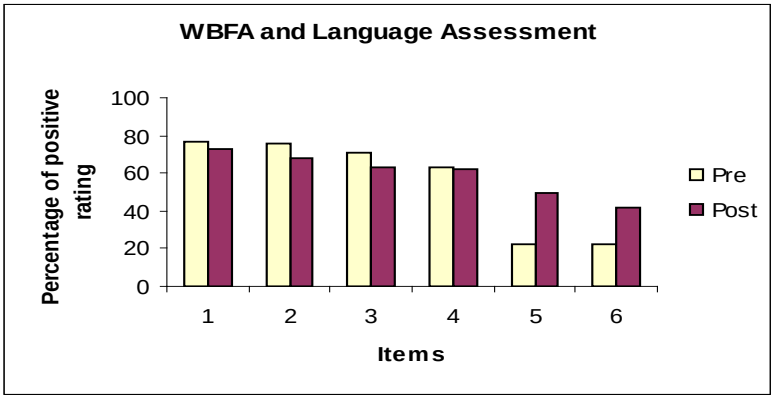
reduction in the percentage of students who accepted that they did not use WBFA as often as they should have, however, was negligible with response rates for the pre-WBFA Questionnaire compared with the post-WBFA ratings being 22.23% and 21.67%, respectively. The percentage of students who were uncertain whether they used WBFA as often as they could or not increased dramatically from 28.89% at the beginning to 65% (a difference of 36.11%) by the end of the WBFA program. This is quite a dramatic increase and indicated that at least the WBFA experiences made them think about their frequency of WBFA access and clearly they were open to considering that there was value in doing it more often given this response.

According to the comparison of the percentage positive ratings of respondents' pre-post questionnaire results, there was a substantial increase in their view of WBFA as an effective language learning pedagogy. Most respondents in the post-WBFA questionnaire had more positive attitudes toward WBFA and the language pedagogy than those in the pre-WBFA questionnaire regarding motivation to study the course. The results support the argument that the students using WBFA tended to increase their incentive to study English in their course. There was evidence in both questionnaires and interviews that students wanted to use WBFA in their course. They also accepted that WBFA motivated them to study in the course.

5.1.3.3 Students' attitudes toward WBFA and language assessment (1)

Figure 5.5 shows the results of students' attitudes toward WBFA and language assessment. The students responded in favour of using WBFA to *assist them to prepare for later summative assessment* as an appropriate tool for them to learn and gain a course grade. The percentages of student responses to this item on both occasions (pre- and post-WBFA questionnaires) were relatively high. The results revealed that 76.67% of the students in the pre-WBFA questionnaire and 72.50% (a reduction of 4.17%) in the post-WBFA questionnaire agreed at the issue. The percentage of students who disagreed or strongly disagreed with this idea that the use of WBFA would yield beneficial effects on their final examination slightly increased from 2.22% at the beginning to 3.34% by the end of the program. The percentage of students who were uncertain that WBFA would help them in the summative assessment slightly increased

from 21.11% at the beginning to 24.17% by the end of the program (a difference of 3.06%).



1. Assist for preparation of later summative assessment (Q7, Q10)
2. Help support self-assessment (Q17)
3. Like MCQ format on WBFA (Q21)
4. Relate to the course content (Q4, Q11)
5. Need alternate question formats (Q22)
6. Should have access time limit (Q24)

Figure 5.5 Positive rating of students' attitudes on WBFA and language assessment

Another important thing in providing the WBFA program is to support students to be independent in self-assessment. In responding whether WBFA could *help support self-assessment*, the percentage of students who stated that WBFA assisted them to promote self-assessment was 75.55% at the beginning and 68.33% by the end of the program (a reduction of 7.22%). The reduction in the number of students who did not think WBFA would help support their self-assessment was negligible, reducing from 2.22% at the beginning to 1.67% by the end of the program. In addition, there was an increase of 7.78% of students who were uncertain that WBFA would support self-assessment by the end of the WBFA program moving from 22.22% at the start to 30% by the end of the program.

With reference to students' attitudes toward preference of MCQ format used in WBFA, the results from both questionnaires showed that there was a decrease in the proportion of students who *liked MCQ format on WBFA.* While 71.11% of students gave positive response ratings at the start of the program, by the end of the program positive response ratings had reduced by 7.78% to 63.33%. However, it needs to be considered that some of this reduction may be explained by students possibly starting with too high an expectation for what WBFA might be able to do to enhance their learning of the

English language. This may also be explained by the fact that the number of students who were not certain about this item rose from 24.44% in pre-WBFA questionnaire to 31.67% in post-WBFA questionnaire (a difference of 7.23%). Interestingly, the percentage or proportion of students whose did not like the MCQ format remained relatively the same 4.44% to 5% thus supporting the idea that at least 7% of students changed from being positive toward the MCQ format to being uncertain about it as a result of their experience in the WBFA program.

In relation to students' attitudes toward whether WBFA content would *relate to the course content,* there was negligible difference in the student's positive ratings from start to finish of the program with over 60% of students in both pre- and post-WBFA questionnaires expressing agreement that there was a correlation between the course content and that of the WBFA. The percentage positive ratings for the pre- and post questionnaire were 63.33% at the beginning of the program and 61.67% (a reduction of 1.66%) by the end. The percentage of students who did not think that the content of WBFA was related to the course content reduced marginally from 4.45% to 0.84% (a difference of 3.61%), and the percentage of students who were uncertain about this issue increased by 5.28% from 32.22% at the beginning to 37.50% by the end of the WBFA program.

To clarify whether they wanted to have different question formats on WBFA, students were asked if there was a *need for alternative question formats*. While 22.22% of students in the pre-WBFA questionnaire preferred to have other question formats, by the end of the program the responses of 50 % (a difference of 27.78%) of the students in the post-WBFA questionnaire showed that they would like to have more question patterns than merely multiple-choice, fill-in and true-false formats. The number of students who did not want other types of questions on WBFA besides MCQ dramatically declined from 48.89% at the beginning to only 8.34% (a reduction of 40.55%) by the end of the program. However, the percentage of students who were uncertain toward this item increased from 28.89% at the beginning to 41.67% (a difference of 12.78%) by the end of the program.

When asked whether there *should be an access time limit* to enter the WBFA program, instead of allowing them to log in the program at any time, the students' opinion

changed between the start and the finish of the program. The limited interval for access to each of the 12 sets of WBFA questions encouraged students to actively take part in performing WBFA after they finished their face-to-face lessons. This was supported by the increasing percentage of students from 22.23% at the beginning to 41.67% (a difference of 19.44%) by the end of the WBFA program. The number of students who did not want to have a limited time frame for using WBFA decreased from 53.33% at the beginning to 10% (a reduction of 43.33%) by the end of the program. Students who were not sure about this issue increased from 24.44% at the beginning to 48.33% (a difference of 23.89%) by the end of the program.

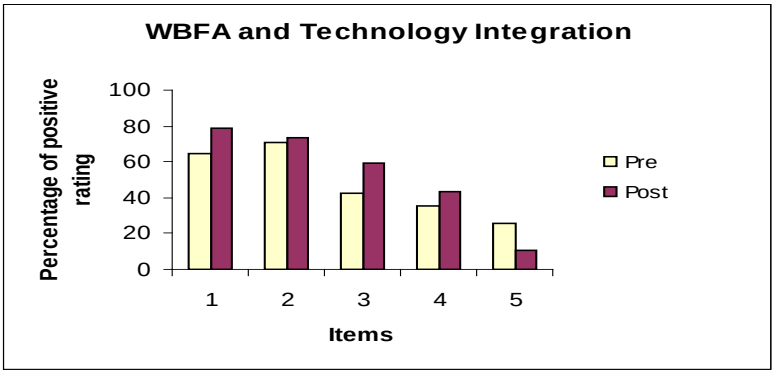
In brief, over 60% of respondents in both questionnaires had optimistic attitudes concerning WBFA and language assessment. Their positive perceptions included the question format and content of the assessment. They also felt that WBFA helped support self-assessment. In addition, they agreed that WBFA helped with their summative test preparation. The results illustrated that post-WBFA respondents had more positive attitudes than those in the pre-WBFA questionnaire relating to alternative questions formats and control of access interval to the program.

5.1.3.4 Students' attitudes toward WBFA and technology integration (1)

Figure 5.6 shows the results from the pre- and post-WBFA questionnaires related to students' attitudes toward WBFA and technology integration. The percentages of students who were satisfied with Web sites connected to WBFA increased from 64.45% in the pre-WBFA questionnaire to 78.34% (a difference of 13.38%) in the post-WBFA questionnaire. There were no students who were against the idea that WBFA helped *provide useful Web sites*. However, the number of students who neither agreed nor disagreed with this issue reduced from 35.56% in the pre-WBFA questionnaire to 21.67% (a reduction of 13.89%) in the post-WBFA questionnaire.

Additionally, the high percentage of students supporting the idea to *integrate WBFA in the course* increased from 71.11% in the pre-WBFA questionnaire to 73.34% (a difference of 2.23%) in the post-WBFA questionnaire. This shows students' acceptance of technology integration in their course, which is a very good sign for stimulating these students to be more autonomous to construct their knowledge in real-world contexts

with extended access to resources/information. It also underpins the idea of learner-centred approach in EFL. The number of students who did not think that WBFA should have been integrated in their course decreased slightly from 8.89% to 6.67% (a reduction of 2.22%). The number of students who were not certain about this was still the same at 20% before and after the WBFA program.



1. Provide useful Web sites (Q25)
2. Integrate WBFA in the course (Q15)
3. Use online interactions on WBFA (Q13, Q14)
4. Offer multimedia (Q20)
5. Feel encouraged and relaxed when using (Q2, Q12)

Figure 5.6 Positive rating of students' attitudes on WBFA and technology integration

The proportion of students who rated positively the *use of online interactions on WBFA* as a collaborative facility arose from 42.23% at the beginning to 59.17% (a difference of 16.94%) by the end of the WBFA program. This might be an indicator to pinpoint that, when students were familiar with using WBFA, they felt more familiar with employing this online facility by themselves. In addition, students felt confident about the e-contacts accommodated to replace in-class personal contacts as the percentage of students who were uncertain about this reduced from 34.44% to 25.83% (a difference of 8.61%). Interestingly, the number of students who disliked these online contacts and expressed that they never used those e-contacts declined from 23.33% in the pre-WBFA questionnaire to 15% (a reduction of 8.33%) in the post-WBFA questionnaire. The WBFA students used the e-contacts to ask questions to the researcher and also exchanged information among themselves via Web board and chat room provided in the WBFA program during the time of the study (see Appendix A).

Students also responded that they preferred WBFA to their previous experience in performing CPFA because WBFA helped *offer multimedia*. The percentage of students

who were satisfied with this technological capability increased from 35.56% in the pre-WBFA questionnaire to 43.34% (a difference of 7.78%) in the post-WBFA questionnaire. The percentage of students who disagreed with this item reduced from 26.67% to 21.67% (a difference of 5%) while the number of those who were uncertain about this item also declined slightly from 37.78% at the beginning to 35% by the end of the program.

With reference to the last item shown in Figure 5.6, almost half of the students (45.56%) in the pre-WBFA questionnaire and more than half of students (55.83%) in the post-WBFA questionnaire were uncertain that WBFA made them *feel encouraged and relaxed when using WBFA*. However, 28.89% of students in the pre-WBFA questionnaire accepted that they encountered some problems while using WBFA. The percentage increased to 33.33% when they responded to this at the end of the program. It should be noted that 25.55% of students who stated that they felt encouraged and relaxed when using WBFA in the pre-WBFA questionnaire reduced to 10.84% in the post-WBFA questionnaire.

In summary, most of the respondents in the post-WBFA questionnaire had more positive attitudes than those in the pre-questionnaire in four different matters. They preferred the links to useful Web sites and they liked the idea of integrating WBFA in their English course. The others referred to online contacts and favoured the multimedia offered in the WBFA program, which they could link to real-world sources and activities related to the content in their textbook for further information. This was different from simply learning in their classroom.

In the following subsections, the analyses of the qualitative data gathered from eight open-ended items in the pre- and post-WBFA questionnaires are reported.

5.1.4 The qualitative data analyses and results of both questionnares

In the final part of the pre- and post-WBFA questionnaires, there were eight open-ended questions for students to respond freely in their own words with respect to convey their attitudes toward the use of WBFA in their study course. The open-ended questions in both questionnaires were identical and asked about the factual information regarding

the students' perceptions of the WBFA program. The questions were based on the four theoretical perspectives of: second language acquisition (5.1.4.1), language pedagogy (5.1.4.2), language assessment (5.1.4.3), and technology integration (5.1.4.4).

The open-ended responses in the pre-WBFA questionnaire were collected at the same time as the closed-ended ones when students completed some examples of the assessment during the tryout of the WBFA program. In addition, the open-ended information in the post-WBFA questionnaire was gathered at the same period of the closed-ended responses at the end of the WBFA program. In both questionnaires, the students answered mostly in the Thai language. Those Thai words were translated into plain English by the researcher for the data analysis. However, some of their answers were written in short English words. The responses to the open-ended questions in both questionnaires were coded into distinct categories with reference to the theoretical perspectives as noted earlier. The findings are reported due to the returned responses of the students in the WBFA cohort, regardless of the students' gender and discipline. The results of the comparison of students' responses to the pre- and post-WBFA parallel questionnaires are reported below.

5.1.4.1 Students' attitudes toward WBFA and second language acquisition (2)

In the pre-WBFA questionnaire, all respondents (*n = 45*) provided mixed responses reflecting the usefulness of WBFA in relation to the ideas in second language acquisition (SLA) relating to the communicative approach and constructivism that emphasises learner-centred approach as evidenced in their responses. Some of their answers to the question, *"What are the advantages of WBFA?"* with reference to WBFA and SLA are as follows:

> *I can study outside my classroom. I can do it anytime. It gives me more time to practice because I can do it anytime I want. I can choose what to study. I want to do it again and again. It is good for self-study. It helps me practice English and makes me feel more confident and I will not feel shy anymore. It makes me learn and understand more English. It can increase my language learning skills. It helps me write English more. I like the link to the Web with 50 conversation lessons with exercises online because it helps me practise more speaking and listening skills using basic everyday English.*

In the post-WBFA questionnaire, 93.33% (*n = 56*) of the respondents wrote the benefits of using the WBFA program which are summarised as follows:

> *It is good for self-study and it needs self-responsibility. The WBFA program is useful for being used in real life. It gives me more knowledge, concerning useful everyday English. I can learn more about English vocabulary. It offers me more opportunities to practice English grammar, vocabulary, etc. It is challenging and interesting. I want to use WBFA more and more often. It helps improve my English language skills. It makes my free time more valuable. I can do it anytime, so I do not have to wait to attend the class. It helps save my study time. It provides me more time to study. It is not boring. It makes me feel more enthusiastic.*

However, in the post-WBFA questionnaire, there were four students (6.67%), two from the Faculty of Public Health, one from the Faculty of Engineering and the other one from the Maritime College, who did not provide any responses to the question.

When asked, "*What are the disadvantages of WBFA*?" 31.11% (*n = 14*) of the students responded to the pre-WBFA questionnaire that there were no disadvantages while one HS student (2.22%) from the Faculty of Humanities and Social Sciences simply wrote, "*I don't like it*" without providing any specific problems that he or she experienced when using WBFA. An EN student from the Faculty of Engineering wrote an interesting response that:

> *I want to have the answer key with an explanation why that those answers are correct or incorrect, at least, to help me learn a bit more by myself. I think many students may not try their best to do this assessment because there is little or no effect to their course grade, providing additional information possibly helps make this program more informative for them.*

In the post-WBFA questionnaire, 55% (*n = 33*) of the students responded that there were no disadvantages while there was one student (1.67%) from the Maritime College who wrote, "*It needs high self-discipline and responsibility. In addition, I do not have a personal computer.*" The rest of the students in both questionnaires stated their problems in using WBFA in relation to the language pedagogy (15.56% and 11.67%), language assessment (15.56% and 11.67%), and technology integration (35.56% and 20%), respectively. The results owing to these responses are reported in the relevant subsections.

With regard to students' attitudes toward WBFA and SLA, it is evident from the responses to both questionnaires that over 90% of them had favourable perspectives.

5.1.4.2 Students' attitudes toward WBFA and language pedagogy (2)

The students' responses to the open-ended questions in the pre-WBFA questionnaire included the benefits of language pedagogy due to immediate and non-threatening feedback which is obviously a great motivation as well. In the questionnaire, when asked "*What do you think about the feedback provided?*" 66.67% (*n = 30*) of the students responded positively that:

> *Good, so good. It is good to get instant feedback to learn what I am good or weak at. I can have the results of the assessment immediately. I think it is very good. It is good because I can know how much I learn once I complete the test. It is exciting to know the results immediately, so it is good to have prompt results. It stimulates my enthusiasm. It increases my interest to study English. When I do not get a good score, I can reattempt. I can repeat doing WBFA as much as I want. By doing this, it helps me understand English more.*

However, there were six students (13.33%) who had different ideas. One of them, a student from the Faculty of Education, responded that, *"It is interesting, but it should have had something more attractive"* while another student from the Faculty of Fine and Applied Arts wrote, *"I gain how to learn the language but it is not as good as learning in the traditional class."* Unfortunately, there were no responses to the question from nine students (20%). They were three HS students, two AR students, two ED students, one MT students and one EN student.

In the post-WBFA questionnaire, the percentage of students who had positive attitudes toward the prompt feedback on WBFA was 63.33% (*n = 38*). Their responses were as follows:

> *It is very good. It is like doing the real test. It is fun; I want to reattempt till I get better scores. I am satisfied with the feedback even I did not get good scores, sometimes. It is good to have the answer key with score reporting. The feedback is very quick. It is modern. Good, it makes me feel more encouraged and interested in studying English in the course. The feedback helps me know how to improve my grade in the course. This feedback helps me prepare to get more test scores in the final exam. The feedback makes me understand my lessons more. I think WBFA provides a good feedback for students; I like it. I can learn from my mistakes. The feedback provided helps me realise my own strength and weakness.*

However, three students (5%) from the Faculty of Public Health, the Faculty of Engineering and the Faculty of Humanities and Social Sciences, had a similar

comment: "*It should have had more explanations for each item to help students understand why it is right or wrong.*" One student (1.67%) from the Faculty of Humanities and Social Sciences rated the feedback as "*Fair*". Unfortunately, there were 18 students (30%), who did not provide any responses to the question asked. They were eight EN students, four HS students, four SC students and two PH students.

When asked about the drawbacks concerning pedagogical issues related to using WBFA, seven students (15.56%) in the pre-WBFA questionnaire responded that:

> *No answer keys or correct answers provided. It should have had correct answers provided with explanation why they are correct in order to help students understand more. When finishing the test, I cannot ask the lecturer why I did it wrong. When I have some questions to ask, there is no lecturer to answer my questions. It should have had translations in the Thai language for difficult English words.*

After the pre-WBFA questionnaire, the correct answers were provided to the students when they completed each test set on WBFA. In addition, the time for performing each set test was expanded to ten, instead of five minutes. The number of students who had problems with using WBFA relevant to the language pedagogy reduced from 15.56% *(n = 7)* in the pre-WBFA questionnaire to 11.67% (*n = 7*) in the post-WBFA questionnaire. These five SC students and two HS students wrote:

> *The computer cannot talk to me. When I have problems, I cannot ask or discuss face-to-face with my lecturer. When using WBFA, it is hard to ask or contact my lecturer. When I have a problem, there is nobody to help, but in the class I can ask directly to my lecturer.*

It should be noticed that even though the e-contacts (i.e., e-mail, Web board and chat room) were provided to the students in the WBFA group, some students might not know how to use or might not try to use them to communicate during the time of the study.

In brief, it can be concluded that over 60% of the students responding in the both questionnaires had positive attitudes toward WBFA and the language pedagogy regarding the instant feedback provided in the program.

5.1.4.3 Students' attitudes toward WBFA and language assessment (2)

The majority of students in the pre-WBFA questionnaire, 71.11% (*n = 32*), provided positive attitudes toward the content of WBFA, when asked, "*What do you think about WBFA content?*" This may be evidence to confirm that the WBFA is comprehensible for most students in the group. Their responses are presented below.

> *It is suitable for learning English in the course. It is appropriate. It is interesting and useful. It is related to the course content. It provides useful clues to learn the English language in the course. It helps me get prepared for the final exam. It will help me have more final exam scores. I can use it to prepare for my final exam. It helps revise my lessons. It helps me revise some exercises. I can practise many different kinds of tests. It makes me know how well I can do it. It is clear. It is very easy. Everything is good. It is good, so good.*

However, six students (13.33%) had different ideas. These students, four from the Faculty of Humanities and Social Sciences and two from the Faculty of Education wrote similarly, "*It is rather difficult. It is very difficult*." One student from the Faculty of Public Health (2.22%) complained, "*The sentences are too long.*" Another student (2.22%) from the Faculty of Fine and Applied Arts simply answered, "*I have no idea,*" while five of them (11.11%) provided no responses to this question. These were AR, ED, EN, HS and SC students.

In the post-WBFA questionnaire, the percentage of students who liked the WBFA content slightly increased to 73.33% (*n = 44*), indicating their positive attitudes as follows:

> *Good. It helps me to prepare for the final exam. It is interesting with a great variety of question formats. It is connected with the lesson content in my class. It is appropriate for the students. It is challenging. The content is related to everyday life. It is relevant to the lessons in the face-to-face class. It is a little bit difficult but it is suitable for the students' level. The content covers what I learnt in my lecture class.*

However, three HS students (5%) who wrote, "*It is difficult,*" recommended, "*There are only texts. It should have been more colourful*." There were 13 students (21.67%) who did not respond to the question. They were eight EN, one MT, one PH, and three SC students.

In relation to the question format in WBFA, over half of the students, 55.56% (*n = 25*), responding in the pre-WBFA questionnaire, showed positive attitudes to the MCQ question format. They wrote:

> *This question type is good. It is appropriate. It is not difficult. I really like it. Please do not change. I do not need any other question formats. I have no problems with MCQ. I am familiar with this question format. I think it is good.*

There were three, ED, HS and NU, students (6.67%) who did not provide any responses to the question. However, 17 students (37.78%) preferred other question formats. These two AR, one ED, three EN, seven HS, two NU, and two SC students wrote similarly that:

> *I want a real conversation test. It should have had various question types for self-practice, something like language games with answer keys. I also want writing and listening tests. I want to listen to a conversation and then answer questions about it. I like a test that I can talk back. I like to test my listening-speaking skills. I want a vocabulary test that looks like a language game.*

In the post-WBFA questionnaire, the percentage of students who liked the MCQ format in the WBFA program increased to 61.67% (*n = 37*). They stated:

> *This one is good. I like MCQ format in the WBFA. I like it because this question pattern is similar to those in the final exam. This type is good enough. I do not want other question formats.*

Other question formats that 11 students (18.33%) recommended in the post-WBFA questionnaire were somewhat similar to those reported in the pre-WBFA questionnaire. However, seven of them asked for real conversation format tests. Other responses to the question were:

> *I want questions about tourism. I like subjective tests. I really want pronunciation tests. I need more about listening tests. I want to have cloze tests concerning English grammar. I need to practise conversational skills from different situations. I want matching and writing tests.*

Unfortunately, there were 12 students (20%) in the post-WBFA questionnaire, who did not respond to the question. They were six EN, five SC, and one SS students.

In relation to the drawbacks associated with WBFA and language assessment, seven students (15.56%) in the pre-WBFA questionnaire wrote:

> *It does not provide enough time to do. The tests are difficult. It is rather difficult. No item randomisation, so I can cheat. It is too easy. It does not provide enough content. It should have had more varieties of content. It should have answer key to explain why it is right or wrong. There is no answer key.*

In the post-WBFA questionnaire, the percentage of the students who were concerned about the language assessment issues in the WBFA program slightly reduced to 11.67% *(n = 7)*. Their problems were:

> *I do not understand the questions. I need more content. It should have 15-20 questions in each test set, instead of ten. I need more explanations for each set of the assessment, not only the correct answer.*

To sum up, over 70% of the respondents in both questionnaires had positive attitudes toward WBFA and the language assessment regarding the WBFA content. In addition, over 55% of those in both questionnaires also had favourable attitudes concerning the MCQ format used in the assessment. Their responses show more acceptances after than before they performed the WBFA program.

5.1.4.4 Students' attitudes toward WBFA and technology integration (2)

When asked, *"What comments do you have on logging into WBFA?"* to determine students' attitudes toward technology integration in the WBFA program, the majority of students, 86.67% (*n = 39*) in the pre-WBFA questionnaire, responded positively to the login process to the program. Their answers were as follows:

> *It is easy to do. It is easy, so easy. I have no problems at all. At first, it is quite complicated, but it is easy when I know how to log in. It is not very difficult. It is convenient. It is quick. It is modern. Only students enrolling to the course can log in, so it is well secured. Therefore, the information is confidential. I can practise my computer skills, revise my lessons after class, as well as do my tests and exercises. I can study through the Internet at the same time. It helps improve my computer skills.*

However, two students (4.44%) wrote, "*It is very difficult,*" while another one answered, "*Sometimes, it is good. Sometimes, it wastes my time*." Three students

(6.67%) stated, "*A few problems, just a few*," but did not mention the specific problems while there was one student (2.22%) who provided no responses to the question asked.

In the post-WBFA questionnaire, over half of the students, 53.33% (*n = 32*) stated positive comments to the login process to the program. Their answers are as follows:

> *It is good, very good. It is enjoyable. I am happy with it. It is convenient. There are not any problems. It is easy to use and suitable for novice computer users. It is sophisticated. To log in is not so complicated. It saves me a lot of time.*

Nevertheless, eight students, (13.33%) specified their problems when logging in to the WBFA program as, "*Registration before getting approved to log in to use the program is rather complicated*," and complained about the speed of the computer when downloading as," *It is slow."* It is quite disappointing that there were 20 students (33.33%) who provided no comments on the question. They were nine EN, five HS, two ED, two PH, one MT, and one NU students.

With reference to the students' responses to the question "*What are the advantages of WBFA?"* in the pre-WBFA questionnaire, all respondents (*n = 45*) wrote their positive answers concerning the technology integration as follows:

> *I can do it anytime I want. I can increase my computer skills. It is convenient. I can test my knowledge anytime I wish and I can reattempt. I can realise my ability. It is classy. I can choose my time to perform the assessment. If I do not understand, I can repeat doing it. It is quick and easy to use. I like it. I can do it anywhere.*

With regard to the to the students' responses to the same question in the post-WBFA questionnaire, 93.33% (*n = 56*) of the respondents had their positive answers relating to the technology integration as follows:

> *It helps me learn more English outside my classroom. It is convenient. It is interesting. It is good for students who want to study by themselves. It is comfortable if you only have the Internet access. I can log in to do the assessment as many times as I want. I can do it on my own because there is nobody to force or control me. It can link to good sources of information on the Internet. It is free of charge. I can gain my computer skills as well. I can contact my lecturer anytime I want. It saves my time. It provides me more time for self-study. It is a new way of leaning the language which focuses on learner-centred approach. I can contact my lecturer through the Internet. I feel more independent to learn the English*

language. This is a very good idea. It is for a modern way of language learning that I have never done before. It is really good indeed.

According to the disadvantages of using WBFA, when asked, "*What are the disadvantages of the WBFA?"* 35.56% (*n = 16*) of the students in the pre-WBFA questionnaire wrote that they had some problems concerning technology integration. They reported the challenges of using WBFA below:

Login process occasionally causes trouble, such as login failure. Students who are not familiar with using computers may experience this login problem. It is slow and there are too many steps to register to the program. Registration before using the program is confusing. There are not enough computers provided in the faculty computer lab, so it is not convenient. Students who do not have computers cannot use the WBFA, so they have to pay for using it in the Internet shops or cafes.

In addition, in the pre-WBFA questionnaire, there was one student from the College of Sport Science who simply wrote "*I do not like to use the Internet.*"

However, the percentage of students who reported the disadvantages of using WBFA reduced in the post-WBFA questionnaire. There were 12 students (20%) who responded the problems of using WBFA as follows:

There are too many steps in a registration process. Maybe, there are not enough computers provided for the students. Sometimes, the Internet causes the problem; it is slow. I prefer using the keyboard, instead of only clicking the mouse. It is rather hard for novice computer users.

According to the technology integration, students were provided flexibility to perform WBFA regardless of time and venue. In response to the open-ended question, "*What time do you like to do assessment on WBFA?"* 42.22% (*n = 19*) of the students in the pre-WBFA questionnaire preferred to perform WBFA in the evening and at night when they did not have classes. Fourteen students (31.11%) liked to do it in the afternoon while nine students (20%) stated that they would perform the assessment anytime when they were available without specific time frame. Two HS students (4.44%) wanted to do it at the same time when the students in CPFA group were in class. Another HS student (2.22%) had no ideas for this issue. Relating to the same question, 18 students (30%) in the post-WBFA questionnaire mentioned that they wanted to perform WBFA anytime when they were free from attending conventional lecture classes while 17

students (28.33%) preferred to perform the WBFA both in the morning and in the afternoon. The number of students who wanted to do the assessment in the evening and at night reduced from over 40% in the pre-WBFA questionnaire to 20% (*n = 12*) in the post-WBFA questionnaire. Two students, HS and ED, (4.44%) liked to do it on weekends while a PH student (2.22%) wrote that he or she performed WBFA before the final examination. Unfortunately, there were ten students (16.67%) who did not provide responses to the question.

In line with the last open-ended question, *"Where do you prefer to perform WBFA?"*23 students (51.11%) in the pre-WBFA questionnaire stated that they would do it off campus while 21 students (46.67%) would do in on campus. There was one student (2.22%) who wanted to do the assessment both at home and at the university. In the post-WBFA questionnaire, according to the similar question, students who liked to do WBFA on campus slightly increased to 48.33% (*n = 29*) while those who did it off campus reduced from over 51% in the pre-WBFA questionnaire to 25% (*n = 15*) in the post-WBFA questionnaire. There were five students (8.33%) who performed WBFA both on and off campus. However, 11 students (18.33%) did not respond to this issue.

In a nutshell, over 90% of the respondents in both questionnaires responded positively to the technology integration in the WBFA program. They liked the convenience provided because they could perform the online formative assessment any time and anywhere they desired. They were also able to reattempt the assessment as many times as they wanted. In addition, they stated that at the same time they could improve their computer skills while using the WBFA. However, there were some difficulties concerning the registration process prior to the use of the program.

In the following section, the analyses of the qualitative data gathered from the pre- and post-WBFA semi-structured interviews are reported.

5.2 The semi-structured interveiew

In relation to the mixed methods research design applied in the study, this section is relevant to the second phase of data collection related to follow-up qualitative data analyses and results collected from students in the WBFA group conducted through the

pre- and post-WBFA semi-structured interviews. This aims to corroborate the findings from the first phase of data collection regarding the quantitative data reported in Chapter 4 and, especially, to strengthen results from the pre- and post-WBFA questionnaires to scrutinise students' attitudes toward the use of WBFA. The following subsections are reported in four parts with reference to the theoretical perspectives underpinning the WBFA program described in Chapter 2. The first one (5.2.1.) is relevant to the students' attitudes toward the WBFA and second language acquisition. The second part (5.2.2) is related to the students' attitudes toward the WBFA and language pedagogy. In the third part (5.2.3), the students' attitudes toward the WBFA and language assessment are detailed. The final part (5.2.4) pertains to the data analyses and results with reference to the students' attitudes toward the WBFA and technology integration.

The participants in the pre- and post-WBFA semi-structured interviews were recruited through the opportunistic sampling technique, which the sample was selected simply because they became available. The advantage of such a sample is that only a few cases may be needed to confirm that a particular performance exists. The researcher selected some specific students whom it was he thought could provide the most valuable data concerning their attitudes toward using the WBFA program in the course. The sample for the pre-WBFA semi-structured interview was six students (7.89%) in the WBFA group. They were interviewed by the researcher at the end of the WBFA tryout session after they completed some test sets of the WBFA program. In addition, ten students (13.16%) in the WBFA group were selected for the post-WBFA interview. They were also interviewed by the researcher after they completed their performance on the WBFA program. The participants in both interviews were in the WBFA group; however, they were not the same students. The selection of participants in both semi-structured interviews did not have to consider students' gender and discipline.

The questions asked in the pre- and post-WBFA semi-structured interviews were identical to the eight open-ended questions in the questionnaire. To collect the follow-up qualitative data, both pre- and post-WBFA semi-structured interviews were conducted as a small group interview. The intent of using the group interview was to encourage the participants to interact with each other so that the quality of the group was developed. Consequently, the interviewees could encourage each other to produce

useful data needed for the study. Furthermore, the aim of the semi-structured interviews was to gather useful information related to the perspectives of the participants as representatives of the entire WBFA group. Therefore, the researcher emphasised on gathering the in-depth information from the interviews as much as possible. The semi-structured interviews did not focus on the characteristics and confidentiality of the individual interviewees as in the case study.

Moreover, the importance of the cultural issues was also considered. In general, Thai students are rather passive and shy, especially when talking face-to-face with a person senior to them. This is because seniority is unique in Thai culture. Hence, Thai students would feel uncomfortable to speak with somebody they were not familiar with. This would also make them feel reluctant and not be confident to speak openly. This is another reason that they might not want to express their attitudes explicitly when they had to talk with a person senior to them in one-on-one interviews. Conversely, it was anticipated that they would feel more secured expressing their attitudes with that person in a group along with their friends. As a result, being interviewed in a small group would make the participants feel more comfortable and relaxed rather than in one-on-one interviews. Therefore, the group interviews would yield more information when the interviewees felt more confident to express those ideas. It should also be noted that it is not appropriate for a man and a woman who are not members of the same family to talk to each other alone behind closed door in the Thai culture. As a result, the researcher decided to organise the small group interview.

The pre-WBFA interviews were taken place in an air-conditioned room on the eighth floor of Queen Sirikit 2 (QS 2) Building at the Department of Western Languages in the Faculty of Humanities and Social Sciences, Burapha University. The room is the lectures' office in which the researcher shared with other six lecturers in the department. The pre-WBFA semi-structured interview was conducted with two small groups consisting of three students each. Both group interviews occurred on the same weekday and in a similar environment. The first group interview was in late morning and the other group interview was in the early afternoon. Each group interview took around 40 minutes.

The post-WBFA interviews were taken place in an air-conditioned computer laboratory on the fourth floor of Queen Sirikit 2 (QS2) Building in the Faculty of Humanities and Social Sciences on two different weekdays. The post-WBFA semi-structured interviews were also in the form of group interviews. There were three small groups consisting of three to four students each. The first two group interviews, with four and three students, respectively, were conducted in the afternoon on the same day. The last group interview, with three students, happened in the afternoon on the other weekday. Each group interview lasted about 30 minutes.

Both pre- and post-WBFA semi-structured interviews took place in a friendly atmosphere between students, their friends, and the researcher. Each individual student in the group interview did not show any sign of suppressing or modifying their true feeling when in the presence of others. Actually, they felt secured to be with the group of friends while being interviewed. They also felt safe to express their attitudes when they were with their friends. The pre- and post-WBFA interviews were in the Thai language. They were also recorded into the cassette tape by the researcher. The students were informed of the objectives of the study and the importance of the interview prior to both interviews. The data were recorded after the students were satisfied to sign the student consent forms to allow the data to be collected by the researcher. After that, the data from both interviews were translated, coded and transcribed into plain English by the researcher.

In the following subsections, the students' attitudes toward the use of the WBFA program from both pre- and post-WBFA semi-structured interviews are reported to determine and compare the students' perceptions prior to and after their experience with the WBFA program.

5.2.1 Students' attitudes toward WBFA and second language acquisition (3)

According to the pre-WBFA semi-structured interviews, students in these small groups identified both the benefits and challenges of the WBFA program they experienced. In accordance with the first theoretical perspective, SLA, students in both group interviews agreed that they preferred the WBFA to the CPFA. One of these students had

mentioned the advantages of the WBFA when asked to compare the situation between both types of formative assessments. The student said:

> *I think one good thing about the WBFA is that it gives chances for us to study by ourselves. It is good because students can search for information and learn independently. There is nobody to force or control us like in the traditional classroom. Another good thing is that we can study outside our classroom any time we want. And the students who pay more attention and perform the WBFA regularly will be the ones who get the most benefits from the program. The students who do not pay much attention or do not perform the program seriously will not think this program is beneficial for them.*

Another student in the pre-WBFA semi-structured interviews agreed with the first student. This student said:

> *Positively, this program is very useful for Thai students if they realise and actually understand the importance of self-study or self-assessment offered through the WBFA program. However, if they know that but are not really interested in participating in the program, or they just want to log in and pretend to perform the assessment by clicking in vain to complete the program without any attention to learn anything from the assessment, or they just want to complete the program to get the 5% of the course grade provided, they will not learn anything. I think they will deserve what they do at the end when they perform their final examination.*

The third student included another favourable reason for choosing the WBFA program that:

> *If the size of the class is small, I think it is not necessary to use the WBFA program. I really think it is not. But for large classes where all students sitting and listening to the lecturer, they will not get anything much from that situation in the face-to-face teaching. I think I am very lucky to participate in this program. It is one of the good ways to test how much you learn and how much you understand the lesson after you study it in the conventional lecture class.*

The rest of students in the pre-WBFA semi-structured interviews also expressed their positive attitudes toward the WBFA in relation with the SLA in the similar way. However, the students in both group interviews also stated some interesting comments about the challenges of this program. One student in the second group said:

> *One thing I am worried is that Thai students may wait until the last minute to perform the WBFA because there is nobody to force them to. As a result, there may be no need for them to hurry to perform or to complete the WBFA program. Allowing students to do the WBFA program anytime they desire may lead them to*

their procrastination. This is different from doing CPFA because the lecturer in the class usually has time limit or deadline for the students to submit their assignment. Another thing to worry is that Thai students are not familiar with this kind of independence, performing the assessment on their own. They probably cannot do it effectively by themselves. Giving them a lot of freedom will make them less enthusiastic to learn or to do their assignment. This situation may become worse when they are allowed to do their own study or assessment out of sight of their lecturers. This would end up with cheating. Personally, I think, Thai students prefer everything that they can complete easily, quickly, and comfortably. They may think that this program is good because it emphasises on self-study, self-assessment, self-discipline, and self-improvement, but they may not want to do because they are too lazy. They may like this program because they do not have to attend the usual class, It means they have more free time available to do anything they want while their friends in the CPFA group do not have this opportunity.

The researcher probed because he wondered whether Thai students liked to be forced to do their assignment. Students in both groups refused and said:

It seems that they do not like to be controlled or forced to do something, but they find it hard to control or encourage themselves. They need someone to encourage them rather than to motivate themselves. It seems that they lack self-discipline, I think. Most Thai students tend to behave that way. I do not know why. It is hard to explain. Maybe, they just try to get along with their friends or it is probably due to their environment, everything surrounding the students, that leads them to do that way.

When asked whether they liked to perform the WBFA on their own or with their friends, all of the students in the pre-WBFA semi-structured interviews said that they preferred to perform the WBFA program alone.

In the post-WBFA semi-structured interviews, most students were satisfied with using the WBFA program. Most students in the three group interviews had similar ideas toward the benefits of the WBFA program. Their answers can be concluded as follows:

I think the program helped stimulate me to learn by myself, to control myself. It was quite different from studying in the conventional face-to-face class where there was always the lecturer to help me. And I could perform the WBFA anytime and anywhere I wanted to. That also made it very convenient for me to test my progress after studying in the lecture class. There was no need for me to attend the class to perform the formative assessment. That was very good. It also helped me learn more about the English vocabulary. This program supported self-study that meant if we did not perform the WBFA, we would not know about it and we would never learn anything from the program. Everything was dependent on me.

When the researcher asked them to compare the CPFA they had experienced before with the WBFA they completed, the students said:

> *I hardly attended the one-period class because the class was scheduled in the late afternoon, which was not convenient for me to attend. I was usually late for the class. I myself liked the WBFA program because there were greater varieties of information than the black and white in the CPFA. There were more things to learn than in the face-to-face class. Most importantly, I could select my time to perform the WBFA due to my convenience. That gave me more time to study and practice English by myself.*

There was one student in the post-WBFA semi-structured interviews who liked both CPFA and WBFA. The student mentioned that the CPFA and the WBFA were both good. The only difference was that the CPFA was offered in the face-to-face class but the WBFA was done with the computers.

It can be said that students in both pre- and post-WBFA interviews were satisfied with the learner-centred approach, the constructivism and the communicative approach offered through the WBFA program. At the same time, there was some concern in the pre-WBFA semi-structured interviews that this kind of assessment might not be successful with some students due to lack of self-discipline.

5.2.2 Students' attitudes toward WBFA and language pedagogy (3)

Owing to the students' attitudes toward the WBFA and the second theoretical perspective, the language pedagogy, the students in the pre-WBFA interviews were satisfied with the instant feedback they received once they completed each test set of the WBFA program. Nonetheless, they were concerned about the interaction between students and lecturers. Although the Web board and chat room were offered in the program for synchronous and asynchronous communication, respectively, the students might not have been familiar with the facilities provided. They said:

> *In the large face-to-face class I could hardly get anything, especially in a one-period class for doing conventional paper-and-pencil assessment. There were a few students attended the class. Those who attended that class accepted that they needed the scores for attendance not what was taught in the class. The WBFA program provided the immediate feedback and it was a good assessment after I studied the chapter in my face-to-face class. However, one drawback of the WBFA was that there was no teacher to answer my questions. In the face-to-face class, I*

could ask what I did not understand at once and got promptly personal answers from my teacher. I could not ask my teacher in person or immediately while doing the WBFA. I could not contact the teacher at once when having a question to ask for help. In the conventional class, I could do that immediately.

When asked about the feedback they received from the CPFA, all students in the pre-WBFA interviews stated that it would take more time for them to get feedback from their teachers in the conventional paper-and-pencil assessment.

In the post-WBFA semi-structured interviews, most students said that they liked the prompt feedback of the WBFA. One reason might be that after the tryout of the WBFA program, the correct answers were also offered to students after the prompt score reporting. The students were able to click to check with the answer keys in case they wanted to see them. For this, some students said that seeing the correct answers was the last resort for them because they wanted to learn from their mistakes and tried to fix their problems themselves first. Some of them said that the scores obtained from their WBFA performance were not their focus. They wanted to learn how much they understood the lesson after they studied in the class. They agreed that it was appropriate to provide correct answers for students to check after the score reporting. They said:

In the face-to-face class, my teacher asked me to do the conventional paper-and-pencil assessment as an assignment and provided the correct answer of the assessment in the following week. But in the WBFA, it was good to know my scores instantly after finishing each test set. It was really good to know the results at once.

In addition, there were two students in the post-WBFA semi-structured interviews who provided interesting answer when talking about the feedback provided.

Getting more or less scores was not important. I wanted to test myself. I did not want to look at the correct answers. It was nice, but I did not pay much attention about correct answers provided because I wanted to test myself first. I would have a look at the answer key when I thought it was necessary. I also liked to repeat doing the test until getting higher scores without looking at the answer key. I tried to reach 80-90 per cent of total scores of each set before moving on to improve myself. I did not expect that I would get full score every time I did the test. That was not important. I just wanted to know how much I understood the lesson. That was what I did.

However, they also stated that in their face-to-face class they could ask their teacher immediately to help them when they had problems while they were unable to do that when performing the WBFA program.

Due to the evidence automatically recorded on the WBFA program, there were not many students in the WBFA group who used the Web board and chat room to communicate with other students and with the researcher. There were 12 messages posted on the Web board. Five students (6.58%) in the WBFA group used the Web board during the time of the WBFA program. The information posted on the board was in both English and Thai. The first student posted that:

> *It is a lot of fun. There is a timer when I am performing the test. It makes me excited because I have to do my test quickly to finish it in time. Using the Internet is very quick and looks modern. The content is relevant to the lesson in the class. The test is not so difficult or so easy. It is another good way to practice my language skills.*

The second and the third students posted to greet the researcher while the last two students posted to ask that:

> *Is it OK, if I log in only once and perform all the 12 test sets? Can you check the login times due to the number of the test sets I did? How can I know whether the computer saves what I did on the WBFA? How can I know if the teacher realises what I have done on the WBFA? How can the teacher check the results of my tests on the WBFA?*

There were seven students (9.21%) in the WBFA group who utilised the chat room for synchronous communication during the WBFA program. There were18 messages posted in the chat room. The first student joined the chat room saying that he or she finished doing the test sets of the first chapter and waited to perform the new assessment of the next chapter. The second student popped up to greet his or her friends. The third one posted to ask how to login to the WBFA program. The fourth and the fifth also entered the chat room to greet their friends online. The sixth student also had the login problem. The last one asked some questions. The questions were related to what she could do after completing all sets of the WBFA program. The researcher posted to answers all her questions in the Web-board for all students could view.

In brief, there was evidence in the pre- and post-WBFA semi-structured interviews that students felt happy with the feedback provided in the WBFA program. However, they were worried about the student-teacher interaction while using the program. One reason might be that they were not familiar with the applications of the e-contacts facilitated.

5.2.3 Students' attitudes toward WBFA and language assessment (3)

In line with the third theoretical perspective, the language assessment, especially the language formative assessment, students in the pre-WBFA semi-structured interviews suggested that there should have been more time provided for doing each test set of WBFA. Five minutes for each test set of ten multiple-choice questions was not sufficient for them to complete each test set in time. They were also critical about the appearance of the WBFA program.

> *I cannot finish a test set of WBFA in five minutes. It is too short because many students are not familiar with the WBFA program, I think, Thai students are not very good at using computers. This is not sufficient time to provide for doing each test set. Most students cannot finish in time, I reckon. Moreover, this test has only text. It should have been more attractive, more eye-catching to arouse the students to perform the assessment. It would be much better if it is user friendly and does not look like a serious test. Its look makes me feel nervous. It makes me feel like I am sitting for the final examination.*

However, they were satisfied with the content of the WBFA, which was relevant to the content in the course they were studying and was related to that in their textbook. Some students said that the WBFA was good for summative test preparation. Two students agreed that the WBFA participation should comprise five per cent of the overall course grade in the similar manners as the students in the CPFA receiving the same score when attending all CPFA class throughout the semester. They said:

> *The content is OK. And I think it is a good idea to provide some scores to encourage students to perform WBFA. The problem is whether those students would really perform the assessment by themselves. Anyway, I think it should be weighted on performing the WBFA because it would stimulate students to engage with the program. The score is the good motivation for them to perform this assessment. Without the score, the students would feel less motivated and less enthusiastic. For me, if I could get scores in performing WBFA, I would do that for sure.*

With reference to the score providing for the WBFA participation, one student was concerned that participants in the WBFA group might perform the assessment because of the scores provided, not because they wanted to improve themselves. Therefore, this should be made clear among students in the WBFA group about the objectives of the program. Another student recommended that the scores for the WBFA participation should not be too high. It should be employed to encourage them and make the participants feel at least they would get something in return.

When asked about the format of the question on the WBFA, the students in the pre-WBFA interviews stated that they liked and were familiar with the MCQ format. However, two students complained that:

> *I cannot finish the reading comprehension section in time, especially the fill-in test because it takes more time to read and to type my answers in the box provided, but I had no problems with multiple-choice questions.*

In the post-WBFA semi-structured interviews, students stated that the WBFA helped them to revise their lessons after they studied them in their face-to-face class. Some of them complained that the content was more difficult than that in their textbook. There were more difficult sentences and words in the WBFA than in their textbook. Their comments can be concluded as follows:

> *The content in the WBFA was rather difficult when compared with what my teacher taught me in the conventional paper-and-pencil formative assessment. There were more difficult words than the CPFA. I think there should have been Thai translation for some difficult English words or sentences in the WBFA. It would be great if the link to the online English-Thai dictionary was provided while we were doing the test in WBFA because we needed explanations in the Thai language. And I wanted to have pictures, comics in conversation tests because they could guide the answers, and make the tests more attractive like playing language games.*

Conversely, there were two students who said that the questions in the WBFA should have been more difficult. Most of the students in the post-WBFA interviews mentioned that they preferred to try three attempts per each test set of the WBFA.

The students said that it was good to have a timer on each test set of the WBFA because it forced them to complete the assessment in time and made them feel as if they had been doing the summative tests. One of them said it was like playing computer games.

> *I liked the timer showing while I was performing the WBFA. It made me excited and a bit nervous. It was like I was playing computer games. I think that the time, ten minutes provided for ten multiple-choice questions, was sufficient for me. It was not more or less.*

With regard to the time and the number of items provided for each test set of the WBFA, when asked, some students in the post-WBFA semi-structured interview had several different ideas. They said:

> *I needed more time for the reading section because I could not finish the reading tests in time. Ten minutes for each set of ten reading items was not enough. It should have been 15-20 minutes instead of 10 for each set of the reading test. But for the conversation tests, ten minutes was adequate. And I think there were not so many questions in the WBFA. Ten questions for each test set were not enough for me to test myself. It should have been 15 to 20 questions for each test set. I think 20 questions for each set is the most appropriate.*

There were two students in the post-WBFA interviews who preferred to have far more items in each test set of the WBFA. They recommended that:

> *There should have been more items, something like an item bank, for instance. Or 100 items for each test set would be great for me, so I could have performed greater varieties of test items, instead of doing only ten items each time.*

However, most students in the post-WBFA interviews were satisfied that the time limit was set at 10 minutes for each test set of 10 multiple-choice questions. They said it was sufficient and much better than five minutes set at the time of the pre-WBFA interviews.

When asked about the question format, most students revealed that they liked the MCQ format provided on the WBFA. Most of then wanted to try at least three attempts per each WBFA test set. One student recommended that the written test should have been included in the WBFA.

> *I think MCQ was OK. It was good. It was appropriate. But I needed some other types like something to test my writing. Even though, I thought it was hard because I did not have a good background in writing English. Most Thai students are familiar with only MCQ. I wish to try written tests although I did not like writing.*

This might be because there were not many writing tests when I was in my primary or even high school. MCQ was not bad but it could lead to guessing. I was also worried about cheating when correct answers were provided to each test set.

At the end of one group in the post-WBFA interview, one student was concerned about the scores obtained when completing each test set of the WBFA. The student wondered whether the scores would be recorded or not and how he or she was able to know that their scores were accurately recorded. The researcher explained that the objectives of the WBFA were to encourage self-assessment after the students learnt their lessons in the face-to-face class and described that the high scores were not as important as the gaining knowledge from his or her WBFA performance. The answer to the question had also been placed on the Web board since the beginning of the WBFA program.

In general, students in both semi-structured interviews provided positive comments about the content and format of the WBFA. There were some interesting issues raised during the interviews that the number of question items for each set of the WBFA would be increased to promote the effectiveness of the formative assessment. In addition, appropriate procedures need to be considered to reduce guessing owing to the format of the MCQ questions.

5.2.4 Students' attitudes toward WBFA and technolgy integration (3)

In the pre-WBFA semi-structured interviews, students stated that there were so many steps in the process of registration. There was a lot of information to fill in the registration page. One student wondered whether it was necessary for the students to fill in their date of birth because it was personal information. Another student wondered whether the registration page had to be filled in every time when he or she logged in to the program. It may be that the student did not read and follow the login instructions written in the Thai language on one page of an A4 sheet which was provided to all students in the WBFA group by the researcher. Another possibility was that he or she was not familiar with using this technology. Their comments on the login to the WBFA program are as follows:

I wish the process of login were easier because some students are not very good at using this technology. They may get confused and do not know where to start. I think even though there is a manual or instructions for login provided, some Thai

students may be unable to follow the instructions because they never get used to it. I do not understand why the registration to the program has so many steps. It is quite confusing. Do I have to login, put my user name, and my password every time I use the WBFA? Why do I have to put in my date of birth in the registration form? I just wonder. Why can't I use only my user name and my password to log in? Why do I have to register? Why do I have to put a lot of details before I can use the program? Do I have to do this every time I want to perform the WBFA? The login box should be placed where students can log in easily, instead of scrolling down the page to log in. Thai students seem to know well and learn a lot from the computer but they are still very slow in using it. They need more guides more information to lead them. It is possible that they may not know how to log in even after reading the instructions sheet even they were written in Thai. It would be great if the manual instructing how to use the program was placed online, instead of a hard copy.

The researcher explained that the main reason for filling in the New Student Registration page was for the system security and for the confidentiality of the students' information. Another issue raised by another student was that the WBFA should have been more attractive to capture students' attention. They said that the look of the WBFA did not arouse the attention of most Thai students who thought the program should have had less academic appearance. They felt like that they were doing a high-stakes test, rather than the ongoing test to help motivate them. The look of the WBFA made them feel serious. One of them worried about the security of the WBFA because it could be easily duplicated by some students. Two students said that they were quite certain that this kind of assessment provided a great chance for the students to cheat because there was no one to invigilate them.

However, all of the students in the pre-WBFA interviews were satisfied with the links to relevant Web sites provided on the WBFA program so that they could learn more about that lesson before clicking the "Exam" button to perform the WBFA. They also agreed that the WBFA was suitable for a large class.

The good thing about the WBFA program is that the students can increase their computer skills while learning the English language.

Most of them accepted that they wanted to perform the WBFA assessment during their free time when they did not have class, sometimes in the morning, sometimes in the afternoon or in the evening. The venue they preferred to perform the assessment was at the university computer centre and at their dormitory in which some computers with

Internet access were provided. Some performed the assessment at a computer room in the university library.

In the post-WBFA semi-structured interviews, one student stated that it was difficult to type in a lot of information in the registration stage. Another student also complained about waiting for the approval to perform the WBFA program after the registration.

> *Why did I have to wait for the lecturer's approval? It was not difficult to log in but it was quite annoying about registration and waiting for the approval. I only had a problem after I logged in for the first time because I did not know where to access the test questions. Even though there was an instruction sheet provided for the login process I did not think I understood it thoroughly. When I got familiar with that login step, there were no more problems. And I never felt nervous when used the WBFA.*

The rest of the students in the post-WBFA interviews said that it was not difficult to register and log in when they were asked what comments they had on logging into the WBFA program. They said:

> *No problem. No difficulty at all. It was only the first time that I had a minor problem, probably, because I had never performed this kind of assessment before. There was not a problem when I logged in the second time. When I used this program more often and when I got used to it, no problem.*

It should be noted that some students did not want to express their attitudes when asked about the challenges of WBFA. Maybe, they might think that would hurt the researcher's feeling. They kept silent and smiled. There is another cultural issue to be considered because Thai people care much about the feeling of the listener when they are talking about something not very good in relation to the listener. They are relatively sensitive about the issue.

When asked about the time and place they preferred to perform the WBFA program, most of them answered:

> *I did it in my free time. There was no specific time for that, just anytime when I was free from attending classes. The place that I liked to perform WBFA most was at the computer corner in my dormitory and at the university computer centre which opened until midnight every day.*

To sum up, students in both interviews were pleased with the technology integration including in the WBFA program. Nevertheless, there were some students who were not familiar with working via the computer.

5.3 Summary

Results from the pre- and post-WBFA questionnaires confirm that, after they performed WBFA, students had more positive attitudes than before in several items asked. They felt that WBFA really helped improve learning English in the course. The assessment also encouraged self-study and connected them to real-life activities which are very important for learning EFL. As a result, they accepted that they enjoyed using WBFA and suggested that the program be appropriate for their course. With reference to the language pedagogy, the students strongly agreed that WBFA motivated them to study the course. In addition, they preferred to have alternate question formats in WBFA and supported that it would be suitable to limit access time to the program. The students also showed their preference toward the technology integration afforded by the use of WBFA. They liked useful Web sites linked with the program and thought that online interactions on WBFA were useful. Finally, they liked links to the multimedia offered in the program.

Findings from the pre- and post-WBFA semi-structured interviews reveal that the majority of the students had favourable attitudes toward the use of the WBFA program in their course. There were satisfied with the great opportunity to learn and test by themselves. Additionally, they liked the instant feedback and content of the assessment. It is interesting to discover that some students requested more question items in the WBFA to serve their formative purposes.

The next chapter provides a discussion of the research results and draws conclusions about the possible future applications of WBFA for such large enrolment EFL classes as those at the centre of this research and also makes recommendations for future research.

Chapter 6: Discussion and conclusions

Chapter 4 and Chapter 5 have reported the data analyses and results of this investigation into the impact of WBFA on large enrolment English classes in tertiary education. Chapter 6 is the final chapter and provides the discussion and conclusions of the study. This chapter has 10 sections. The first section (6.1) pertains to discussion on the effect of the Web-based formative assessment (WBFA) program as a tool to elevate students' overall English skills achievement that was targeted. The second section (6.2) discusses whether students who used WBFA received higher achievement scores on the particular language features of *completing dialogues*, *reading comprehension*, *vocabulary*, and *grammar* than students who used the conventional paper-and-pencil formative assessment (CPFA), which was the standard practice with large enrolment classes at the time of the research. The third section (6.3) discusses the issue of frequency of students' participation in WBFA in relation to their achievement over the duration of the program. The fourth section (6.4) discusses students' opinion of WBFA and their perceptions of how WBFA compared with the university's standard approach that involved pencil-and-paper tests with much longer feedback time. In the fifth section (6.5) the importance of this research to the teaching and pedagogy for enhancing learning in large EFL classes at the tertiary level is discussed. The sixth section (6.6) discusses the value of the research findings for those involved in teaching EFL and the potential of WBFA. The seventh section (6.7) explores the potential of universities with large enrolments adopting the WBFA approach as developed in this study. While the limitations of the present research are addressed in the eighth section (6.8), the ninth section (6.9) deals with recommendations for EFL pedagogy and assessment for large tertiary level enrolment classes and future research, and the final section (6.10) presents the conclusions.

6.1 WBFA as a tool to elevate students' achievement

This section discusses the research findings in relation to whether the WBFA program served as a worthwhile tool for students in large EFL classes to elevate their levels of achievement compared with those who undertook CPFA. This discussion relates to the first research question: *Does WBFA serve as a tool for students to elevate their levels of*

achievement compared with CPFA when measured by objective testing? It focuses on the four theoretical perspectives—second language acquisition, language pedagogy, language assessment, and technology integration—that underpinned and guided the research and development of the WBFA program.

6.1.1 WBFA and second langauge acquisition

The second language acquisition (SLA) theory, which underpinned and directed the WBFA program, emphasises that students are able to improve their linguistic competence by connecting new information with what they have already learned. According to Krashen's (1985) input hypotheses, which are essential to SLA, students can obtain new linguistic knowledge through comprehending messages or obtaining understandable input. As a result, in keeping with the theory, the intervention of WBFA as the comprehensible input was utilised in the study. WBFA was also implemented as a pilot study to investigate its impact on student's achievement and attitudes towards this form of pedagogy and assessment as a strategy that is specifically relevant to large EFL courses in universities such as the one at the focus of this study.

In addition, SLA used in the WBFA program was aimed to help promote a more learner-centered approach where students could choose their own study place and time that was convenient to them outside their conventional EFL classes. The WBFA program was created to support EFL students to construct their own learning and understanding through flexible delivery and self-assessment. It focused on this approach because it is consistent with the constructivist's pedagogical theory (Vygotsky, 1962), as noted by Wilson and Lowry (2000), which highlights the students' responsibility in building their own knowledge as opposed to simply receiving the information by means of transmission from their EFL lecturers. WBFA was aimed to encourage students to construct their understanding through interaction with the milieu provided. By using the WBFA, students could learn mostly by themselves how to find their answers and correct their own mistakes. Furthermore, this is related to the implementation of a communicative approach to encourage students to learn from more real-life activities than what they did inside their large, traditional EFL classes.

Prior to the implementation of the WBFA program, students were to some extent apprehensive about it as a form of learning and assessment. Some students mentioned in the pre-WBFA semi-structured interviews that it might be difficult for them to learn on their own although they all agreed with the idea of the learner-centered approach. They were uncertain because most of them were familiar with learning the language only in the way it was taught in the classroom by their EFL lecturers. It was likely that learning in that way had made them passive and shy students who typically, did not initiate their own learning but depended on the teacher's directions. In addition, they also mentioned that Thai students preferred to be dependent because they were not used to receiving or seeking knowledge by themselves. In this context, it seemed to them that performing the WBFA would increase their learning burdens since they had to study mostly by themselves and they had no experience with this kind of assessment. Besides this, it might be possible that they would not try their best in their WBFA performance, especially when they were allowed to do the WBFA independently without the presence of their lecturers whom they placed so much dependence on. With these somewhat pessimistic attitudes, it seemed that they were concerned with both self-discipline and self-confidence to connect themselves to new linguistic knowledge.

Consequently, this is a challenge for Thai EFL lecturers because they have to try rigorously to encourage these students to be more autonomous, and more independent in seeking knowledge by themselves. It is the lecturers who have to clarify this misunderstanding to every student in order to improve the achievement and negative attitudes of their students in the future. However, this research shows that there are potential benefits in providing more independence to students in large EFL classes through the use of the WBFA program. For instance, the outcomes of the research showed that students' involvement in WBFA caused them to realise the importance of independent learning that is so crucial for lifelong learning which is the focus of the 21st educational policy and curriculum. Similarly, the research showed that the provision of the opportunity for students to self-assess and be involved in using English for communicative and functional purposes along with the integration of technology was an experience that they viewed as enhancing their EFL learning.

The findings from the mixed methods design revealed that the application of SLA through the learner-centered approach, constructivism, and communicative approach by

the use of the WBFA yielded beneficial impacts on both students' achievement and their attitudes. With reference to the increased overall mean scores of the students in the WBFA group and their positive attitudes toward the WBFA program, thus confirming WBFA could help change the idea that Thai students might not do well in self-assessment, especially in the large EFL classes of passive, low-proficiency students. The research findings can also verify that Thai students are able to increase their achievement when they are encouraged and provided with independent opportunities to learn by themselves. Hence, this research will serve as a way of encouraging them to try strategies to be independent learners and have the potential to increase their enthusiasm and ambition to be successful in their EFL learning. Similarly, it provides a message for their teachers that the students are capable of learning in this way and that their pedagogical approach can be thriving if they incorporate WBFA.

In summary, the study provides sound evidence to support the argument that the use of the WBFA program underpinned and guided by the second language acquisition theory was a valuable and worthwhile experience for students in the large EFL classes at the focus of the research. Besides helping to elevate their levels of achievement it introduced them to a new form of pedagogy and assessment that helped them gain confidence to learn more independently and understand independent learning better, thus, providing strong evidence that WBFA should be supported for further implementation in the future.

6.1.2 WBFA and language pedagogy

With regard to the language pedagogy theory, providing students with a non-threatening setting and instant feedback through the applications of the WBFA program would stimulate students' motivation to learn, particularly for passive students with low-proficiency in the English language. As noted in Lin's (2003) research, students were enthusiastic and their motivation became heightened when there was less tension in the learning environment. As a result, the learning occurred more effectively.

In their previous conventional EFL classes, prior to the study, the students in the WBFA group studied English as a foreign language by listening and taking notes from attending the two-period lectures which tended to engage with their receptive language

skills rather than the productive skills they needed to enhance their communicative competence. Then, after the lectures, they were assigned to perform CPFA in their workbook which they received feedback from their lecturers a week or longer later. Also, there were very few opportunities to encourage them to use their English language skills to communicate with other people outside the classroom. This kind of EFL learning and teaching does not support the communicative language teaching approach that involves and stimulates students to interact in their use of English for functional purposes and also be proactive and enthusiastic in their learning. In addition, this kind of EFL teaching focuses mainly on a teacher-centered approach rather than a learner-centered approach. As a result, students are likely to lack motivation to learn when English is mainly taught by means of translation into Thai, the students' native language, simply through their lecture classes.

Furthermore, in this standard practice, the large size of the class makes it very difficult for teachers and students to implement effective in-class interactive/communicative language learning experiences. According to the students responding to the questionnaires and semi-structured interviews, there were few students who wanted to attend the one-period class scheduled for the CPFA. They thought they would not gain any worthwhile learning experience from doing it. They said they were less motivated to attend. According to Krashen (1985), language students will learn more when their affective filter or mental block is down. That is when they are well motivated while learning in a non-threatening environment, or when they can control they own pace of study. Therefore, good motivation is very important and should be promoted in language learning, specifically in large EFL classes with somewhat passive and low-proficient learners.

In relation to pedagogical issues, the WBFA students in the study complained that it was not convenient for them to ask their lecturers when they had some questions while performing the WBFA program compared with their experience in their previous CPFA context. For this, it is essential to reconsider about the teacher-student interaction online, as noted by Heinemann (2003) that the interaction is an important factor in cognitive and affective learning. However, the students stated that a major benefit of WBFA was that it provided them with faster feedback than when they performed CPFA. This is supported by Dalziel (2001) who discovered that instant feedback could

supply students with an opportunity to test their emergent understanding and immediately identify their problem areas during the learning process. By this means, the students had the potential to improve the cognitive skill of self-assessment to correct specific mistakes. As a result, regular self-assessment and prompt feedback during learning by the use of WBFA was shown to be considerably helpful in increasing students' learning outcomes.

With respect to the results regarding the language pedagogy in the WBFA program, there was evidence that the students in the WBFA group showed a favourable attitude to the non-threatening nature of the online learning environment and they valued highly the instant feedback provided by the WBFA program. These positive attitudes could clearly be observed in their increased achievement scores when compared with those of the CPFA group.

6.1.3 WBFA and language assessment

In relation to language assessment theory (Bachman & Palmer, 1997), the formative assessment applied through the implementation of the WBFA program involved students in a process that allowed them to know their strengths and to identify and address their weaknesses. It was anticipated that employing WBFA would help students to focus more on self-improvement through independent flexible learning and self-assessment. Black and Wiliam (1998b) suggested that frequent, short, formative assessments are better than one infrequent long one. With regard to the formative assessment of the progress of language learning, WBFA was employed to offer frequent, short, objective and ongoing tests to assist individual students to elevate their achievement in learning EFL in the challenging learning context of large classes. This was supported by Aiken (1996) who argued that for evaluating student performance in classrooms, an objective, teacher-made test, which consists of true/false, multiple-choice, fill-in, and short answer questions, is one of the most common and effective testing tools.

There was evidence in both the pre-WBFA questionnaire and the pre-WBFA semi-structured interview that the WBFA participants were worried when responding to the multiple-choice question (MCQ) format used in the program. Some of them were

concerned about the possibility that this approach would encourage students' to be dishonest when performing the WBFA independently. There was also apprehension that students might remember the correct answers which they could look at, if they wanted to, after receiving instant feedback on completion of each test set. WBFA students raised their concern that this would offer some students an opportunity to get higher scores when reattempting instead of stimulating them to really seek new knowledge from their mistakes. To ease the problem, it was thoroughly explained to the students that the objectives of the WBFA was not to necessarily get full scores when performing each set of the WBFA, but to gain knowledge and to learn from their mistakes to improve themselves. With regard to this issue, the mixed methods results of the study, namely the results of the pre- and post-tests, pre- and post-WBFA questionnaires and pre- and post-WBFA semi-structured interviews, verified that the majority of the students using the WBFA understood this objective. According to Aggarwal and Bento (2003), a Web-based assessment program could not be a success without responsible, motivated students whose aims are to learn and not to simply get a passing grade.

In brief, the provision of Web-based frequent, objective-formative assessment substantiated that the WBFA could yield an effective influence on students' achievement and attitudes in large EFL classes such as those involved in this study. Moreover, the research findings suggested that it would be more appropriate if more objective questions were used in the WBFA as students' responses in the post-WBFA questionnaire and post-WBFA interview indicated that they needed an item bank to be enlarged so that they could practise even more. This was supported by Dalziel (2000) who found that using question banks also offered greater flexibility for teachers in constructing and editing questions and designing feedback that provided positive washback and pedagogical relevance.

6.1.4 WBFA and technology integration

With regard to the integration of technology into the WBFA program, there was concern from some students about the security of the assessment in that they thought that it could be easily duplicated. However, it needs to be emphasised that WBFA is a

low-stakes assessment created to help low-proficient students in large EFL classes to practise and test themselves outside their classrooms, thus supporting them make better use for their time outside the lectures. A further issue is that when students are allowed to perform on their own as with WBFA, it is relatively difficult to know whether they actually performed the assessment by themselves. Nevertheless, this situation is no different from students in the standard formative assessment CPFA routine when students engage with take-home assignments. Similarly, it was difficult to know whether they did the assignments by themselves. However, it is worth noting here that the WBFA program was a kind of teacher-made test or classroom assessment. It was used in combination with other evaluations before the course grade was finalised. Furthermore, these abovementioned issues and problems were better understood by students and so were easily resolved after students were actually made aware of the objectives of WBFA and they experienced its non-threatening quality.

Thus, this study provided strong evidence from the overall qualitative analyses of WBFA usage and students' opinion that there were several benefits that emerged from the integration of technology into learning and assessment in this way. Firstly, employment of the technology caused students in the WBFA group to construct knowledge about the English language at their own rate. Secondly, students found it more motivating because it brought students into contact with more authentic materials and more real-life learning activities. Thirdly, it was a more interactive approach to EFL pedagogy compared with the standard approach that students had experienced before when performing their CPFA. Since WBFA could be performed anywhere and at anytime via the Internet, it could provide students with the skills' practice and knowledge regardless of time and venue. Hence, it was suitable for them to obtain knowledge by themselves and come to grips with the idea of independent learning and what this means for language learning. Fourthly, with swift automatic marking on the Web, WBFA was able to give students instant feedback on completion of each WBFA test set. This was a tremendous benefit for reducing lecturers' marking loads in the large enrolment EFL classes involved in this research. Finally, the WBFA program was found to have increased students' computer skills, which they needed to learn to use in their future workplaces, since the acquisition of the English language is valued highly in the Thai education system and so it is tested accordingly. For these reasons the research findings provided strong support for the integration of technology into this EFL

learning program and, therefore, corroborates the usefulness of such a WBFA program that encourages students to be proactive EFL learners who, according to Wilson and Lowry (2000), benefit from the Web as an influential source of information, of great importance for developing and monitoring independent learning. In addition, because of its use of hypertext, Web users found they had more control over the learning experience than students who depended on listening to lecturers or reading books.

Importantly, this WBFA program would not have been successful without technology integration. It was impossible to solve the problems of students' low motivation and achievement, and the slow feedback cycle with regard to the lecturers' extremely high, time consuming, marking loads for the large EFL classes involved in the study.

In summary, WBFA underpinned and guided by the four theoretical perspectives provides strong evidence that such a program can serve as a tool for the majority of students to elevate their level of achievement compared with CPFA when measured by objective testing, particularly in a large enrolment tertiary English course. This is supported by the quantitative results that indicated WBFA students in two subgroups (Faculty of Education and Faculty of Public Health) outperformed the students in their CPFA groups on the post-test-only design. The overall means of WBFA students in the two subgroups were significantly higher than the overall means of those in the CPFA groups (p <.05). Owing to the four language features tested, the WBFA students also surpassed those of the CPFA group in two sections of the four language features. The overall means of the WBFA group in the sections of *reading comprehension* and *vocabulary* were significantly higher than those of the CPFA group (p <.05) on the post-test of English skills.

The following section discusses the overall CPFA group's performance compared with the overall WBFA group's performance in relation to their improvement on the test of English skills. Then the comparisons of the CPFA and WBFA means related to each of the four language features (*completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*) in each subgroup are discussed. This is with reference to the second research question: *Do students who use WBFA get higher achievement scores on completing dialogues, reading comprehension, vocabulary and grammar for English than students who use CPFA?*

6.2 Comparison between WBFA and CPFA

When the results of the WBFA group were compared with those of the CPFA group with respect to whether WBFA facilitated greater success with English language learning, students who used WBFA obtained generally higher achievement scores on the sub-components of *completing dialogues, reading comprehension, vocabulary,* and *grammar.*

The quantitative result of the overall post-test of English skills mean (27.68) in the WBFA group was almost identical to its pre-test mean (27.64). On the contrary, the overall post-mean (26.13) in the CPFA group was significantly lower than its pre-test mean (27.09). The quantitative results shown here provide evidence though not conclusive that students using WBFA improved at least as well as and better in some English skills compared with the CPFA group. In addition, it can be argued on the basis of these results that using CPFA in large enrolment EFL classes is quite problematic and possibly detrimental to students' English language learning and because they do not receive the systematic, regular, timely feedback on their responses like the WBFA group do. Worth noting here is that the present research raises the question as to whether it is time to change the mode of EFL learning in large classes from a teacher-centered approach to a learner-centered approach such as this one and whether it is time to encourage dependent and less-motivated students to be independent and better-motivated through WBFA. From the results, it is obvious that the overall WBFA students' performance yielded more satisfactory effects on the students' achievement scores than the CPFA's.

Regarding the four language features tested, *completing dialogues, reading comprehension, vocabulary,* and *grammar,* students in both CPFA and WBFA groups gained statistically significant improvement in the *reading comprehension* and *vocabulary* sections of the English skills test (p <.05). Perusal of descriptive statistics also revealed that the mean of the CPFA's section of *reading comprehension* increased 6.60% over the duration of the program, while the WBFA's increased 10.20%. In the *vocabulary* section, the CPFA's mean gained 10.10% while the WBFA's raised 16.30%. Conversely, students' performance over the period of the program as measured by the pre- and post-English skills test surprisingly lowered for both WBFA and CPFA

groups. The mean scores in the sections of *completing dialogues* and *grammar* reduced significantly (p <.05) for both CPFA and WBFA groups. The CPFA's mean for c*ompleting dialogues* reduced 10.20% and the WBFA's decreased 9.10%. In addition, the CPFA's mean for *grammar* section lowered 10.93% while the WBFA's decreased 11.20%. What to be reconsidered is why WBFA students were successful in two sections of *reading comprehension* and *vocabulary* but unsuccessful in the other two sections of *completing dialogues* and *grammar*. One possible reason might be that the question formats in the first two sections were different from the last two. They were multiple-choice, fill-in, and true/false question formats for *reading comprehension* and *vocabulary* sections while the questions in *completing dialogues* and *grammar* were simply in multiple-choice format with four alternatives (a, b, c, and d). This is the single visible evidence found in the study. Should this be the real cause of the failure, it is crucial to reconsider to provide more alternative question formats in those sections in the future research.

In the following subsections, the students' performance on *completing dialogues*, *reading comprehension*, *vocabulary* and *grammar* in the CPFA and WBFA groups regarding each subgroup is discussed.

6.2.1 WBFA and CPFA in Faculty of Fine and Applied Arts

Regarding the four language features tested in the CPFA group, there was a little improvement in *reading comprehension* and *vocabulary* sections, yet the group failed to increase its achievement scores in the sections of *completing dialogues* and *grammar* shown by the comparisons of pre- and post-test means. There were no significant differences in all four sections. In the WBFA group, there was a significant difference in the section of *reading comprehension* where the pre-test mean was statistically larger than that of the post-test. The WBFA group increased its means in the last two sections of *vocabulary* and *grammar* although there was no statistical significance. However, the WBFA group failed to gain its achievement score in the *completing dialogues* section.

It is noted that the Arts students in the CPFA group achieved more satisfactorily than those in the WBFA group. One AR student wrote in responding to the pre-WBFA questionnaire that, "*I gain how to learn the language [from the WBFA program], but it*

is not as good as learning in the traditional class." Related to this response, which is opposed to the majority of the students in the WBFA group, it can imply that WBFA's aim to encourage its participants to learn EFL independently in a non-threatening situation via formative assessment on the Web did not have positive effects on AR students' achievement and attitudes toward the program. This attitude could explain why 50% of students in this subgroup did not perform any WBFA set test. Their frequency of WBFA participation was under the average and was the second lowest of all subgroups in the study. In addition, there was no one in the subgroup who completed all 12 test sets in the WBFA program.

6.2.2 WBFA and CPFA in Faculty of Education

In this subgroup, there were no significant differences between the pre- and post-test means in the CPFA group pertaining to the *reading comprehension* and *vocabulary* sections. However, there was a significant difference in the *completing dialogues* section where the post-test mean was smaller than that of the pre-test ($p < .05$). In the WBFA group, there were no significant differences between the pre- and post-test means concerning the *completing dialogues*, *reading comprehension* and *vocabulary* sections. However, WBFA students failed to gain any improvement in the *grammar* section where the post-test mean was significantly smaller than the pre-test mean ($p < .05$).

Owing to the overall means within this subgroup, the CPFA's pre- and post-test means were surprisingly the same whilst the WBFA's post-test mean was slightly lower with no significance than its pre-test mean. One possible cause of the WBFA group was obviously from the low achievement score in its *grammar* section. Due to the qualitative results, there was evidence that could possibly relate to this quantitative result when two ED students wrote to respond to the pre-WBFA questionnaire, "*It is interesting but it should have had something more attractive,"* and "*It is very difficult*." These attitudes might influence their participation in the WBFA program since there were seven from 15 students (46.67%) who actually performed the online assessment during the time of the study; however, the rest of them logged in to the program but did not perform the assessment, according to the automatic records on the Web.

Although the majority of WBFA students were satisfied with e-contacts provided on the WBFA program, there were not many students who used them to communicate with their lecturers or the researcher when they had problems. It was possible that they were not familiar with using those e-facilities.

6.2.3 WBFA and CPFA in Faculty of Engineering

There were not any significant differences between overall means of the pre- and post-tests in the CPFA group. In relation to the four language features tested, there were significant differences in two sections of *vocabulary* and *grammar*. The CPFA mean increased statistically in the *vocabulary* section whereas its mean in the *grammar* section decreased significantly.

In the WBFA group, there were no significant differences between overall means of the pre- and post-tests. In relation to the four language features, there were significant differences between means of the pre- and post-tests in all four sections. The WBFA means of *reading comprehension* and *vocabulary* sections increased significantly while its means in the *completing dialogues* and *grammar* sections decreased significantly.

It should be noted that EN students were one of the three subgroups where all participants (100%) in the WBFA group did perform the WBFA program. However, the qualitative results revealed that many of them were uncertain how to respond to the open-ended items in both pre- and post-WBFA questionnaires, which were written by the researcher in plain English with Thai translation. Though the WBFA program could positively encourage all EN participants in the WBFA group to participate and increased achievement scores in the *reading comprehension* and *vocabulary* sections, it failed to encourage them to gain their improvement in the *completing dialogues* and *grammar* sections. The last two language features seemed to be the major problems for EN WBFA students.

6.2.4 WBFA and CPFA in Faculty of Humanities and Social Sciences

There were no significant differences between overall means of the pre- and post-tests in this subgroup. Regarding the four language features tested the average points in the *reading comprehension* and *vocabulary* sections increased with significance in both CPFA and WBFA groups. However, there were no significant differences in the sections of *completing dialogues* and *grammar* in both groups.

From the quantitative results, it was obvious that the WBFA program had positive effects on the achievement of this subgroup in the two sections of *reading comprehension* and *vocabulary* but failed to enhance their improvement in the other two sections of *completing dialogues* and *grammar*. Some evidence that probably helped explain their problems was found in the qualitative results. It revealed as one HS student wrote, "*I do not like it*." In the same questionnaire, other four HS students complained, "*It is rather difficult. It is very difficult*." In addition, in the post-WBFA questionnaire, three HS students wrote, "*There are only texts,*" and "*It should have been more colourful*," and another HS student suggested, "*It should have had more explanations for each item to help students understand why it is right or wrong*." This useful information is worth considering for the future applications of the WBFA program as well as for improving learning and teaching approaches in those two language features in the course.

6.2.5 WBFA and CPFA in Maritime College

In relation to this subgroup, there were no significant differences between overall means of the pre- and post-tests in the CPFA group. In addition, owing to the four language features tested, there was no improvement through mean comparisons of the pre- and post-tests in all four sections of *completing dialogues, reading comprehension, vocabulary,* and specifically in the section of *grammar,* where the CPFA students in this subgroup had their achievement score reduced significantly ($p < .05$).

In the WBFA group, there were no significant differences between overall means of the pre- and post-tests. The MT students failed to gain their achievement scores in the sections of *completing dialogues,* and *grammar*. However, they increased their means

in the sections of *reading comprehension* and *vocabulary*. In the post-WBFA questionnaire, one MT student wrote, "*It needs high self-discipline and responsibility [to perform WBFA]. In addition, I do not have a personal computer*." Thus, the importance of familiarity/readiness and access to a computer is identified here.

6.2.6 WBFA and CPFA in Faculty of Nursing

The results revealed that NU students in the CPFA group did not achieve in the sections of *completing dialogues* and *grammar* where post-test means were lower than their pre-test means. In the section of *reading comprehension*, the CPFA group had slightly improvement while pre-post means of the *vocabulary* section were identical. However, there were no statistical differences in all sections of this group.

In the WBFA group, NU students had positive achievement results in three sections of *completing dialogues*, *reading comprehension*, and *vocabulary*. Particularly, there was a significant increase in the *vocabulary* section whereas the pre-post means of the *grammar* section were similar. It cannot be discounted, therefore, that students' language learning style may be a consideration of future research in this area.

6.2.7 WBFA and CPFA in Faculty of Public Health

The overall average score of PH students in the CPFA group indicated no improvement at the end of the course. With reference to the four language features tested, their post-test means in *completing dialogues, vocabulary* and *grammar* sections were smaller than the relevant pre-test means while there was a little increase in the section of *reading comprehension*. However, there were no significant differences. On the contrary, students in the WBFA group gained their achievement scores in all four language feathers tested. Specifically, there was significant improvement in the two sections of *reading comprehension* and *vocabulary*.

Furthermore, PH students were the single subgroup that its WBFA group had the overall post-test mean significantly higher than the pre-test mean. Surprisingly, their frequencies in the WBFA participation were lower than the average frequency. In relation to this, some students responded to qualitative instruments that it would be

much greater if there were more items in each set test of the WBFA program. They commented that ten items for each WBFA test set were not adequate for self-assessment. The small number of items in each test set might have failed to attract students to reattempt because there were not many challenges. This would lead to the consideration of constructing item banks for the future WBFA program.

The majority of PH students had positive attitudes toward the WBFA program. However, one PH student wrote in the pre-WBFA questionnaire pertaining to content of the online assessment, "*The sentences are too long*." Another PH student wrote in the post-WBFA questionnaire, "*It should have had more explanations for each item to help students understand why it is right or wrong*," in order to recommend for further WBFA modification. It seems obvious from the quantitative results that the WBFA program had positive impacts on PH students in every language feature tested.

6.2.8 WBFA and CPFA in Faculty of Science

There was no improvement in terms of overall average scores within the CPFA group from the Faculty of Science. In addition, the post-test means of the four language features tested (*completing dialogues*, *reading comprehension*, *vocabulary* and *grammar*) were all lower than their pre-test means. However, they were not statistically different.

In the WBFA group, there was no improvement in terms of overall average scores. There was a little improvement in the *reading comprehension* and *vocabulary* sections in the WBFA group; however, there were no significant differences. In contrast, their post-test means in the two sections of *completing dialogues* and *grammar* were smaller than their pre-test means. It appeared that the mean in the section of *grammar* of the WBFA group decreased significantly (p <.05). SC students had the overall highest frequencies in performing WBFA. They had the highest frequencies in three sections of *reading comprehension*, *vocabulary*, and *grammar*. They were supposed to be the most active and most responsible in terms of their enthusiasm in doing WBFA. According to the qualitative results, SC students showed positive attitudes toward the program. The quantitative results of this subgroup revealed the most surprising effect of WBFA on their achievement in the *grammar* section. Had all 20 SC students (100%) performed

the WBFA program, instead of 15, it is anticipated that the quantitative results may have been different.

6.2.9 WBFA and CPFA in College of Sport Science

For SS students, there were no significant differences between overall means of the pre- and post-tests within the CPFA group in three sections (*reading comprehension*, *vocabulary*, and *grammar*) of the four language features. However, the CPFA achievement reduced significantly in the section of *completing dialogues* where the pre-test mean was higher than the post-test mean.

In the WBFA group, SS students had the lowest frequencies in performing the WBFA program in all four sections when compared with other subgroups. There was one student (33.33%) who performed WBFA but did not complete all the WBFA test sets. With one student left in the WBFA group, the SPSS software program refused to compare means of the CPFA and WBFA groups due to insufficient students in the WBFA group.

In summary, with reference to the results of the pre-test and post-test design, this entire section reports that there was no significant difference between overall pre and post-means in the WBFA group. Neither was there a statistical difference in the CPFA group's. In relation to the four language features, there was significant improvement in the *reading comprehension* and *vocabulary* sections in both groups whereas students' overall means in the *completing dialogues* and *grammar* sections reduced significantly in both groups.

Regarding the results of the pre- and post-tests in the CPFA and WBFA groups in each subgroup, there was significant improvement in the WBFA group of the students from the Faculty of Public Health (PH). In relation to the four language features tested in each subgroup. For the CPFA group, there was significant improvement in the *vocabulary* section of the students in the Faculty of Engineering (EN), and in the *reading comprehension* and *vocabulary* sections in the Faculty of Humanities and Social Sciences (HS). Concerning the WBFA group, there was significant improvement

in the two sections of *reading comprehension* and *vocabulary* in four subgroups: the Faculty of Engineering, the Faculty of Humanities and Social Sciences, the Faculty of Nursing, and the Faculty of Public Health.

In the next section, whether greater numbers of attempts to perform WBFA would affect WBFA students' achievement in each subgroup is discussed.

6.3 Frequency of WBFA participation and students' achievement

In keeping with the learning theory underpinning this research, repetition is seen as crucial to language learning (Inoue & Bell, 2006). In performing WBFA, the students were allowed to reattempt any test set as frequent as they desired after they had studies related lessons in their lecture classes. Therefore, it is interesting to investigate how the frequency of the students' participation in the WBFA program affected their learning outcomes. The relevant overall results reported in Chapter 4 revealed that there was no statistical evidence to indicate that students who performed more than the average frequency would achieve more than those who did not. In relation to respond to the third research question (*How does frequency of participation in WBFA affect students' learning outcomes?*), the results related to the WBFA students' frequency of WBFA participation and their achievement in each subgroup are discussed in the following subsections.

6.3.1 WBFA frequency and achievement in Faculty of Fine and Applied Arts

In this subgroup, there were six students in the WBFA group with regard to the stratified random sampling; however, there were only three students (50%) who actually performed the WBFA during the time of the study. It was uncertain why the other three students in the group simply logged in but did not perform WBFA. The single evidence obtained was from the pre-WBFA questionnaire in which one student from the subgroup responded to one of the open-ended questions that the WBFA program was not as good as the traditional CPFA class, with no further reasons. This might be possible to verify that the attitude indicated in the questionnaire would discourage his or her intent to participate the program. However, these six participants

signed the student consent form to join the group, and they showed no sign of withdrawal from the WBFA group during the time of the study.

Among those three students who performed WBFA, there were no students who completed all 12 sets of WBFA. The most frequency that one student (16.67%) performed was 14 times while the least frequency performed by another student was once. The total frequency of WBFA students in this subgroup was the second lowest rate of attempts when compared to the students in other subgroups. The average frequency of the AR students was nine which was lower than the overall average (15.20). This might be another problematic factor that affected the students' attainment.

According to the post-test only design, the WBFA participants in this subgroup were the ones who had their overall mean significantly lower than those in the CPFA group. Nonetheless, their achievement scores with reference to the four language features tested were not significantly different from those in the CPFA group. This can be concluded that the intervention of the WBFA program did not necessarily have any positive effect on their achievement. In addition, in the pre-test-post-test design, their attainment in the section of *reading comprehension* indicated the disappointing result when their post-test mean was statistically lower than that of their pre-test.

In relation to the results of the students in this subgroup, it is essential to reconsider for further and effective actions to assist them. It was not certain that their low achievement was due to their lack of interest in learning the language or their low interest in participating WBFA. More time for the training session prior to the commencement of the program should be reconsidered as well. However, it was evident that the average scores of the AR students were the lowest in all WBFA subgroups.

6.3.2 WBFA frequency and achievement in Faculty of Education

In this subgroup, there were 12 students in the WBFA group; however, there were seven students (56.33%) who actually performed WBFA. It should be noted that one ED student wrote in the pre-WBFA questionnaire that the WBFA program was not so attractive. Another two ED students complained in the same questionnaire that the

WBFA was difficult for them. These should be considered in relation to the small number of ED students who performed WBFA.

There were four students (57.14%) of those participating in the program who completed all 12 sets of the WBFA program. The most frequency performed was 57 while the least was twice. The average of the ED students' attempts was 26.29, which was much higher than the overall average. This high frequency pinpointed the positive effect to the WBFA students' achievement because their overall mean (29.29) was statistically higher than that of the CPFA group (23.75) when compared in the post-test only design. In addition, the means of the four language features tested in the WBFA group were larger than those of the CPFA group, especially in the *reading comprehension* section where the WBFA mean was significantly larger than that of the CPFA group. This indicated that the WBFA program may have helpful impacts on these students' achievement.

In summary, the WBFA students in this subgroup were among the six subgroups that outperformed the students in the CPFA group on the basis of the post-test-only design. Additionally, they were the one in two subgroups which had their overall means statistically higher than those of the CPFA group. This can be confirmed that there is evidence-based to indicate that the frequency of WBFA participation did have beneficial effect to students' outcomes in this subgroup.

6.3.3 WBFA frequency and achievement in Faculty of Engineering

There were 12 EN students in the WBFA group, and all of these 12 students (100%) participated in the WBFA program. The students were among the three subgroups in which all students in the WBFA group did perform the WBFA program. They did not simply log in to the WBFA program and then log out without any performance. In addition, there were four students (33.33%) who completed all 12 sets of WBFA. The most frequency was 33 while the least was once. Their average frequency was 11.58, which was lower than the overall average.

With reference to the quantitative results from the post-test-only design, the achievement score of the EN students' overall mean in the WBFA group (23.08)

slightly surpassed that of the CPFA group (22.92). In addition, the WBFA means in the sections of *completing dialogues* and the *reading comprehension* were higher than the CPFAs' but there were no significant differences. With regard to the pre- and post-test scores, EN students in the WBFA group did better than the CPFA group in the sections of *reading comprehension* and *vocabulary*. It should be noted that their post-test means were statistically higher than those of the pre-test. The findings demonstrated that the WBFA had a favourable effect on the achievement of the EN students.

6.3.4 WBFA frequency and acievement in Faculty of Humanities and Social Sciences

The 27 WBFA students in this subgroup were the largest in the study. However, there were 25 students (92.59%) who actually performed WBFA. There was evidence in the pre-WBFA questionnaire that one HS student responded that he or she did not like the WBFA program while another HS student complained that the assessment was rather difficult without providing any reasons.

Among those who performed WBFA, ten students (40%) completed all 12 sets of the WBFA program. It should be noted that this was the highest number of students who completed the WBFA program in the study. The most frequency was 32 while the least was twice. However, the average number of attempts was 12.16, which was lower than the overall average. According to the record, HS students had the highest number of attempts in the section of *completing dialogues*.

According to the post-test-only design, the overall mean of the students in the WBFA group (28.76) was not significantly different from the overall mean of the CPFA group (27.59). Additionally, the WBFA group also had larger means than the CPFA group in the sections of *reading comprehension*, *vocabulary*, and *grammar*. The WBFA group had smaller mean than the CPFA group in the section of *completing dialogues*. This is not what was expected for the reason that the HS students had more attempts in performing the *completing dialogues* section than any other subgroup. This also raises an interesting issue that the greater number of attempts to perform the WBFA program did not necessarily mean that the students' would obtain higher achievement scores.

In general, the participants of this subgroup were expected to have the more outstanding achievement scores than other subgroups because there were some students who chose to study English as their major subject in the faculty. Nevertheless, it should be noted that there are fifteen departments in the faculty (HUSO, 2008). It is impossible to select specific major subjects of the participants in the WBFA group recruited through the stratified random sampling. In addition, the stratified random sample in the study was randomly chosen from the same group of students who received the lowest English scores from the national university entrance examination, which was previously described in Chapter 3. It was possible that most students whose major was English might not be represented in the stratified sample as they needed to obtain higher English scores in the entrance examination to be qualified by the relevant department to study in that major subject.

6.3.5 WBFA frequency and achievement in Maritime College

There were three MT students in the WBFA group, and all of them (100%) actually performed the WBFA program. However, there were no MT students who completed all 12 sets of WBFA. The most frequency was 10 while the least was five. The average of their attempts was 6.67, which was much lower than the overall average. Owing to the low number of MT students' attempts in performing the WBFA program, there was evidence in the post-WBFA questionnaire where one MT student responded that he or she needed to have high self-discipline and responsibility in performing the WBFA program. The student also stated that not having a personal computer was one problem he or she experienced in using WBFA.

According to the overall mean comparison, the WBFA group (29.67) outperformed the CPFA group (28.33) without a statistical difference. MT students were the only WBFA participants in the study who did not have a single record on performing the *grammar* section. Surprisingly, their average score in the *grammar* section was identical to that in the CPFA group when the means were compared in the post-test-only design. It was anticipated that this result might be positively different if they did perform WBFA. Additionally, the WBFA group did better than the CPFA group in *reading comprehension* and *vocabulary* sections, but they failed to surpass the CPFA group in

the section of *completing dialogues*. However, there were no significant differences between means in each of these three sections.

6.3.6 WBFA frequency and achievement in Faculty of Nursing

In this subgroup, there were four students in the WBFA group, and all of these students (100%) performed WBFA. However, there was one student (25%) who completed all 12 sets of the WBFA program. The most frequency was 17 while the least was 11. Their average frequency was 14.5, which was slightly lower than the overall average.

The overall mean of NU students in the WBFA group (30.50) exceeded that of the CPFA group (28.25) when compared in the post-test-only design but there was not a statistical significance. In the four language features tested, there was a significant difference in the section of *vocabulary* where the mean of the WBFA group was higher than that of the CPFA group. There were no significant differences between CPFA and WBFA groups in other language features tested.

Generally, it can be concluded that the WBFA program had positive effect on the NU WBFA students. Their average scores indicated in both the post-test-only design and the pre-test-post-test design could confirm this. Interestingly, in the pre-test prior to the implementation of the WBFA, the mean of the CPFA group was 30.00 and that of the WBFA group was 26.50. After the WBFA program, the mean of the WBFA group increased from 26.50 to 30.50 whereas that of the CPFA group decreased from 30.00 to 28.25. These findings may be the observable corroboration regarding the effectiveness of the WBFA program when compared to the results in the CPFA group.

6.3.7 WBFA frequency and achievement in Faculty of Public Health

In this subgroup, there were seven students in the WBFA group. Nonetheless, there were six students (87.71%) who actually performed WBFA. Two students (28.57%) completed all 12 sets of WBFA. The most frequency was 21 while the least was four. The average frequency was 11.67, which was lower than the overall average. It should be noted that the PH students had the highest means in both CPFA (31.14) and WBFA groups (36.33) when compared in the post-test-only design. This subgroup was one of

the two subgroups whose WBFA overall means were significantly higher than those of the CPFA. In addition, it seemed that the intrusion of the WBFA program had the beneficial impacts on the PH students. The findings illustrated that the WBFA means surpassed those of the CPFA in all sections of the language features tested. Particularly, there were significant differences in the *reading comprehension* and *vocabulary* sections.

It is interesting to consider why and how the mean of the WBFA group in the subgroup was significantly higher than the students in the CPFA group although they did not have high frequency rate in WBFA participation. The evidence indicated that PH students in both group had high pre-test means. Their pre-test means were higher than those in any other subgroup. The pre-test mean of the CPFA group was 33.14 while that of the WBFA group was 29.00. As a result, it might be possible that they were better at English than the students in other subgroups. However, it is more interesting to discover that, after the WBFA program, the mean of the WBFA group increased from 29.00 to 36.33 while that of the CPFA group decreased from 33.14 to 31.14. Similar to the findings in the Faculty of Nursing, this may be another noticeable verification of the usefulness of the WBFA program on PH students in the WBFA group when compared to those in the CPFA group.

6.3.8 WBFA frequency and achievement in Faculty of Science

In the Faculty of Science, there were 20 students in the WBFA group; however, there were 15 students (75%) who actually performed WBFA. There were eight students (53.33%) who completed all 12 sets of WBFA. The most frequency was 63, which was the highest, while the least was five. The average attempt was 23.07, which was much higher than the overall average.

However, the overall mean of the WBFA group (27.87) was smaller than that of the CPFA group (28.30) with no significance when compared by the post-test-only design. Although the WBFA means of the *completing dialogues* and *reading comprehension* sections were better than those in the CPFA group, they had lower means than those in the CPFA group in the *vocabulary* and *grammar* sections. Nonetheless, there were no significant differences.

It is interesting to consider that what happened to the SC WBFA students. They seemed to be the most enthusiastic students in the study in terms of their high rate of attempts in performing the WBFA program. There was evidence in the pre-WBFA questionnaire that five students from this subgroup responded in one of the open-ended questions that they were unable to contact their lecturers while performing the WBFA. Although there were facilities such as the Web board, chat room and e-mail for both asynchronous and synchronous communication in the WBFA program, some students might not be familiar with them. However, it was evident that there was no significant correlation between SC students' achievement and the numbers of their attempts in performing the WBFA program.

6.3.9 WBFA frequency and achievement in College of Sport Science

In this subgroup, there were three students in the WBFA group; however, there was one student (33.33%) who really performed the WBFA. This student did not complete all 12 sets of the WBFA. This subgroup had the lowest rate of attempts in performing the WBFA program.

According to the insufficient number of students in the WBFA group, it is impossible to compare their CPFA and WBFA means. There was single evidence which might be relevant to this issue in the pre-WBFA questionnaire where one SS student wrote to respond one of the open-ended questions that he or she did not like to use the Internet. As a result, this might possibly be the cause of their non-performance in the WBFA program.

In summary, at the end of the WBFA program, the total number of attempts for *completing dialogues* was 419, *reading comprehension* 252, *vocabulary* 226 and *grammar* 259. There were 1,156 attempts in total during the time of the study. It should be noted that students logged in to perform the section of *completing dialogues* more than any other section. This may be because the section was the first section of the WBFA program. It may not be accurate to conclude that students liked to perform the *completing dialogues* section more than the others. It should also be noted that the number of attempts to perform the *vocabulary* section was the smallest.

In the post-test-only design, the means of the WBFA group were significantly higher than those in the CPFA group in the sections of *reading comprehension* and *vocabulary*. In addition, in the pre-test-post-test design, there were statistical differences in the sections of *reading comprehension* and *vocabulary* where the post-WBFA means were higher than the pre-WBFA means. It is interesting to perceive that all questions in both *completing dialogues* and *grammar* were in the multiple-choice format whereas those in the *reading comprehension* and *vocabulary* sections were the combination of multiple-choice, fill-in and true/false question formats. This evidence may be useful for further considerations. For example, it is interesting to investigate whether using different question formats would influence students to have different achievement scores, or whether utilising multiple-choice questions to assess conversational skills as in the *completing dialogues* section is appropriate or not. Some possible changes in teaching and testing for formative and summative purposes need to be reconsidered for the benefits of the students to use the language to serve their communicative purposes effectively.

To answer the third research question, the findings indicate that even though the students who completed the WBFA program did have higher mean than those who did not, there was no significant correlation between the number of students' attempts to perform the WBFA and their achievement scores.

According to automatic records on WBFA during students' login time, it showed that some students in the WBFA group used less time but obtained more scores when they reattempted than their first WBFA performance. This might be another evidence to indicate that they might remember the answer key for their later WBFA performance. However, it was not very surprising when the correct answers were provided. It was uncertain to specify how they performed the WBFA and what reason they had in mind when they did the WBFA. The evidence from the questionnaires and the semi-structured interviews indicated that there were two main groups of students who performed the WBFA. Students in the first group wanted to use the assessment to test themselves while those in the second group wanted to get high scores and reattempted as many times as they could to reach their target when performing the WBFA.

Nonetheless, whatever motivations they used to perform the WBFA, it was certain that they were exposed to the WBFA. It was anticipated that the students in the WBFA group had learned something from engaging in the WBFA program. It is relatively difficult to find supporting evidence regarding how and why they performed the assessment. What is more important is that they should have a clear understanding with reference to the objectives of the WBFA, which was utilised to help them to increase their achievement scores. The high numbers of their attempts might not be crucial in that case. It might be possible that some students repeated so many times to ensure that they would get 5% of the course grade for participating in the WBFA program. In this study, WBFA students had to complete 12 WBFA test sets to get the 5% of the course grade, whereas CPFA students had to attend and complete CPFA to receive the 5% of the course grade. The high frequency of attempts might indicate the students' enthusiasm; however, it might have no relation with their achievement scores, according to the results demonstrated in the study. There was the potential that some students might memorise the correct answers after their first attempts to take advantage when they reattempted later. For instance, there were records showing that some students spent almost ten minutes to complete their first test and got five out of ten points, but in the second attempt or more attempts later they got ten out of ten in less than one minute. In this case, it seemed that the students spent time to learn something only in their first attempt while doing WBFA. Nevertheless, some students got some scores in their first attempts within less than a minute. They simply wanted to obtain scores through guessing. It was difficult to conclude that they learned anything from their WBFA performance.

The following section discusses issues regarding the last research question, *what are students' attitudes toward the use of WBFA and learning English in the course?*

6.4 Students' attitudes toward WBFA

6.4.1 Students' attitudes toward WBFA and SLA

With reference to their attitudes toward SLA underpinning the WBFA program, the students in the WBFA group responded positively in both questionnaires and semi-structured interviews. Over 90% of the students in both pre- and post-WBFA

questionnaires indicated their favourable acceptances toward WBFA. Firstly, the WBFA program did help them improve their learning English in the course. Secondly, the program encouraged their self-study. Thirdly, WBFA assisted them to connect to real-life materials via the Internet. Fourthly, the students felt that they enjoyed using WBFA in the course. Finally, they agreed that the WBFA program was appropriate to be used in the course. In both pre- and post-WBFA semi-structured interviews, the students expressed that they were satisfied with the learner-centered approach, the constructivism and the communicative approach offered through the WBFA program.

According to the second language acquisition theory, WBFA was used as a comprehensible input to enhance students to increase their achievement scores by means of self-assessment to solve the problems related to large-sized EFL classes. This was conceptualized in relation to a learner-centred approach aiming to encourage students to construct and obtain new knowledge by themselves. If Thai EFL students wanted change their role from dependent to independent students, they would revolutionize their role as passive to active learners. Findings illustrated here are evident that Thai EFL students in WBFA groups desired to be more proactive students. Meanwhile, Thai EFL teachers had to change their position from knowledge arbitrators to facilitators to assist their students as well. Results from both questionnaires and interviews indicated that their students were satisfied with the WBFA program because it provided more communicative activities to students than in their traditional classrooms. According to SLA, students will gain their achievement when they feel content to learn. It is anticipated that this means will lead students to be better motivated than learning in the conventional class and performing simply CPFA.

6.4.2 Students' attitudes toward WBFA and language pedagogy

According to the percentage of positive ratings shown in Chapter 5, it can be concluded that most respondents in the post-WBFA questionnaire had more positive attitudes toward WBFA and the language pedagogy than those in the pre-WBFA questionnaire regarding motivation to study the course. They were certain that the use of WBFA did increase their incentive to study English in their course.

In relation to the WBFA students' responses in the open-ended items in their questionnaires, over 60% of the students responding in the both pre- and post-WBFA questionnaires had positive attitudes toward the language pedagogy underpinning the WBFA program regarding the non-threatening setting and the instant feedback provided in WBFA.

With reference to the pre- and post-WBFA semi-structured interviews, there was evidence that the students felt happy with the feedback provided in the WBFA program. However, they were concerned about the student-teacher interaction while using the program.

Regarding the provision of feedback to students, Boston (2002) stated that the most helpful type of feedback on tests should provide specific comments about errors and specific suggestions for improvement and encourage students to focus their attention on the task rather than on simply getting the right answer. This type of feedback may be particularly helpful to lower achieving students because it emphasises that students can improve as a result of effort.

In this study, prompt score reporting and correct answers were provided to let students in the WBFA group know their level and try to improve themselves through the revision in their own textbooks and attempt to get higher scores in their later attempts. Many students learnt this way, so they paid more attention at the first time they logged in to perform the WBFA program. However, some students might take advantage by memorising the answers for their later performance. Some stated that they would not even want to look at the correct answers but kept trying to repeat doing the WBFA until they were satisfied with their scores. There was evidence of this in the semi-structured interview at the end of the WBFA program. Students in the CPFA group usually received their feedback from their lecturers about a week or longer after they had performed their traditional assessment. The feedback was mostly spoken to provide correct answers from teacher's manual to CPFA students.

It was hoped that WBFA would better motivate students to have more productive learning outcomes. One reason was that WBFA offered students with a non-threatening atmosphere and prompt feedback, which was not possible in the CPFA learning

environment. The WBFA allowed students to communicate via Web-board and chat room. These communicative language use strategies were intended to promote communication for real-life purposes and lead to positive attitudes and more fruitful learning outcomes for the students.

Although any assessment can be effective when immediate feedback is presented to students (Alagumalai, Toh, and Wong, 2003), providing feedback with correct answers to students in the WBFA program might have caused some challenges. Firstly, according to Ellis and Ratcliffe (2004), it is difficult for teachers to use the questions again after the correct answers were provided to students without the risk of students recognising the answers although those items were randomised. Secondly, students might not try their best to perform the formative assessment when it does not have the crucial influence of being counted for final grading. Thirdly, if students did not receive a teacher's personal explanations why their answers were wrong, it might be difficult for them to understand by themselves. WBFA students also mentioned these three problems in their questionnaires and the interviews. Finally, teachers might not be convinced how well their students were learning the materials without supervised summative results.

However, according to Emberger (2002), giving students correct answers has a moderate positive effect. Explaining with what is correct and what is incorrect has the greater effect. Allowing students to continue working until successful has the greatest effect. In this study, the real objective of providing students with correct answers was offering them to have a look when they needed. Some students said in one of the semi-structured interview that they did not want to look at the correct answers provided because they wanted to pursue the correct answers by themselves. However, some students took advantage of this offer and memorised correct answers to use in their later attempts. This could be confirmed, regarding to the automatic record on the WBFA program, that some students performed each WBFA test set less than one minute to get the full score (10 points) in their second attempts after taking more than 10 minutes and simply got a few points (lower than 5) in their first attempts. It is possible that they might misunderstand the objective of the formative nature of assessment. As noted by Conradson and Pedro (2004), the explanation of the objective of any assessment was essential to make students feel certain that it was introduced to assist them to gain their

achievement scores in the course rather than to make decision on their course grade. For teachers, the motivating challenge should not be figuring out how to "catch" students who plagiarise, but how to frame the assignment in such a way that students will not see the need to plagiarise and cheat.

6.4.3 Students' attitudes toward WBFA and language assessment

The majority of respondents in both pre- and post-WBFA questionnaires had positive attitudes toward WBFA. Their positive perceptions included liking of the question formats used in WBFA (i.e., multiple-choice, fill-in, and true/false) and the content. They also felt that WBFA helped support self-assessment. In addition, they agreed that WBFA helped for their summative assessment preparation. The results illustrated that post-WBFA respondents indicated preference of alternative question formats and strict access time to each test set of WBFA to those in the pre-WBFA questionnaire.

There were two interesting issues raised during the post-WBFA semi-structured interview. The number of question items for each set of the WBFA would be increased to promote the effectiveness of the formative assessment. In addition, appropriate procedures need to be reconsidered to reduce students' guessing in performing WBFA.

With formative purposes, the WBFA program would help students and teachers to realize their weaknesses and strengths. It would assist students to have more achievement scores because the assessment offered students with appropriate level of difficulty with relevant content to the lessons in the course. This assessment aimed to support students instead of judging them.

WBFA can be applied to assess several kinds of language skills, for instance, reading comprehension, vocabulary, and grammar (Morrison, 2002). However, there might be some problems when using WBFA. It is unable to identify whether the students performed the assessment by themselves when they are not invigilated, as noted by (Luecht, 2001). One possible solution was using a login system with a user name, and a password, several short tests, large pools for randomization of questions (Olt, 2002; University of Illinois, 2007). Roever (2001) also suggested that cheating is less applicable when WBFA was a low-stakes assessment such as here. This is because that

the low-stakes assessment was generated to give students feedback on their performance and progress in reaching their learning target. It also helped students to prepare for summative tests.

6.4.4 Students' attitudes toward WBFA and technology integration

There were some students who were not familiar with utilising the computers. They experienced frustration with technological breakdowns, difficulties associated with their limited computer skills and withdrawal symptoms from the lack of face-to-face contacts with teachers. There was evidence illustrated in the questionnaires and the semi-structured interviews that some students complained that they could not contact face-to-face with their lecturers when they were performing the WBFA program. However, there were students who completed all WBFA test sets and gained their achievement scores after using WBFA without any contacts with the researcher.

On the other hand, with regard to the to the students' responses to the post-WBFA questionnaire, over 93% of the respondents had their positive attitudes relating to the technology integration. Most of the respondents in the post-WBFA questionnaire had more positive attitudes than those in the pre-WBFA questionnaire in four different matters: (1) they preferred the links to useful Web sites; (2) they liked the idea of integrating WBFA in their English course; (3) they preferred online contacts; (4) they liked the multimedia offered via the WBFA program.

Hence, technology integration engaged students in active, constructive, intentional and cooperative learning. It helped shift from an old to a new educational paradigm. They needed to change their roles. Technology integration would link students to modern information sources and learning tools. It also provided students with rich sources of information. Analysis reported significant cognitive and affective learning gains. These results support the hypothesis that teacher-student interaction is an important factor in cognitive and affective learning (Heinemann, 2003).

6.5 WBFA and a large enrolment EFL course

Problems of teaching and learning in large EFL classes are: (1) less time and inflexible, physical environment for teaching and for practising communicative language skills, and difficulty to arrange effective in-class activities to develop language competence, including student to student interactions; (2) less time for teacher-student interaction; (3) slow and virtually unachievable quality feedback from the lecturers to students; (4) more marking workloads; and (5) less motivation which leads to low achievement. It seemed that WBFA could help solve every problem mentioned above because: (1) with the capacities of the Web, students could perform the assessment any time and anywhere they wished outside the classroom with links to real-life activities on the Internet; (2) in the WBFA program, students were provided with online facilities to communicate with their lecturers both synchronously and a synchronously; (3) the WBFA provided students with instant feedback once they submitted their answers to the system; (4) with automatic scoring, lecturers would not have to spend much time in marking; and (5) according to the students' attitude responses in both questionnaires and semi-structured interviews, there was confirmation that the program did help motivate students to learn English in the course. From the findings, it was obvious that overall students' achievement increased when they fully used WBFA; specifically, the reading comprehension and vocabulary skill also increased significantly when compared with students who used the CPFA.

One thing from the study suggested that there should have been more question items in each test set of WBFA since it seemed that ten questions for each test set was inadequate for students to practise to increase their achievement. Another reason for construction of the item bank was to reduce the students' cheating when performing WBFA.

This in accordance with Dalziel (2000) who stated that one solution to the problem of creating large number of multiple-choice questions is provided by textbook publishers, who sometimes provide question banks to teachers as a supplementary material when textbooks are prescribed for students in the teacher's course. This is one of easy ways to make an item bank. The other way is the cooperation between the lecturers in the same EFL course to help each other to construct the bank of MCQ items.

In relation to the students' attitudes toward the use of the WBFA, there was evidence that the majority of the students in large EFL classes who used WBFA showed positive perspectives toward it. This could be a good sign to make lecturers who want to use this kind of assessment in the future feel more confident to use it in the future. The instant feedback seemed to a major favourite matter for them. The other issue might be the flexibility that the students were able to use the assessment regardless of time and venue. They were able to control their study instead of being controlled like studying in the large face-to-face classes where they could not do anything rather than being passive students listening and taking notes in the lecture class. This would encourage the students to have more opportunities to connect to more appealing information linked with the WBFA.

McKenna (2001) stated that to reduce high marking loads is the main reason for adopting online assessment. One aim of constructing WBFA is to reduce the marking time for lecturers in large EFL classes. Although the lecturers had to spend a lot time to construct the question items and provided correct answers or explanation to offer effective feedback to students once they deliver their answers to the system, their effort should yield fruitful results in the long run. It was also more cost-effective than constructing the paper-and-pencil assessment because the lecturers were able to disseminate the question or make any change more easily than with the paper-based assessment.

6.6 WBFA and teaching staff

The WBFA program may provide certain benefits to lecturers in large EFL classes. Firstly, they could get the feedback from the students' performance in the program instantly, which was a good source for fixing things up both in their teaching and the students' learning the course. This would be a great benefit for improving their teaching and assisting students to be on the right track.

Secondly, the lecturers would increase their confidence and knowledge of how to construct online assessment. In addition, they would learn how to make their classroom assessment more reliable and valid. This is also to ensure the reliability and the validity

of their assessment. It would lead to creation of item banks when all lecturers teaching the same EFL foundation course will corporate to construct the questions for the assessment together. It would increase the effectiveness of the assessment in the long run. This would be the best opportunity for them to increase their computer skills because, according to Chapelle and Doulas (2006), lecturers in second language teaching will have to deal with computer-based assessment some day in the future. It is to confirm that computer technology plays a role in language assessment, and particularly in its future.

In language assessment, exploration of technology for testing has increased to the point that today no matter where second language learners live; they will sooner or later take a computer-assisted language test. Language teachers need a solid understanding of this online assessment because it will help learners to develop self-assessment strategies, test learners in the classroom, select or develop test for language programs and prepare learners to take other tests beyond the classroom and language program. Computers can be used to display and process large amounts of data rapidly allowing for the input the examinee receives on a language test to include rich contextual information consisting of images, sounds, and full-motion video, potentially enhancing authenticity in both the input and response.

One issue to be considered is that lecturers do not want to change their roles as they used to have because they may lose their authority to control the class when students are encouraged to study on their own. The main roles will be on students. Lecturers will change their roles to be simply facilitators. However when these lectures are well trained and understand clearly about using the WBFA program in large class, WBFA will be of great help for them in the long run. Additionally, the university and the administrative board should support them by providing facilities and budget to support the WBFA program.

Web-based technologies in teaching and learning require fundamental changes in many areas of an institution as noted by Inoue and Bell (2006). Faculty members' knowledge and skills in using technology are very important. Their reluctance is often related to actual fear of technology. In addition, as noted by Eustace (2003), the major transition from traditional paper-based questions to Web-based format did increase their burden

of work from learning how to use this technology prior to placing the questions online. This takes time and effort; consequently, most lecturers ignore this. In addition, there is a need to mention they have to learn how to use the technology and of course in this context their monthly salary and considerations of promotion at the end of the financial year are further elements that influence lecturers' perspectives and likelihood of wanting to change to WBFA.

To encourage the use of WBFA in the faculty and the entire university, all facilities should be ready to promote the assessment system. Astin, et al. (2003) argued that introducing this online assessment to students is a campus-wide responsibility. It may start small, but assessment's questions cannot be fully addressed without participation by student-affairs educators, librarians, administrators, and students. How to integrate computerised assessment with traditional study methods is considered as the biggest challenge for lecturers, according to Walker and Delius (2004). Therefore, teachers have been reluctant to administer regular tests because they consider testing too time-consuming, taking away valuable class time.

Writing appropriate questions is time consuming initially because it often involves a shift to new question formats as agreed by McKenna (2001). Organisational and pedagogical difficulties may be raised, including reluctance by individuals to share questions; a reluctance to use questions written elsewhere; variance in standards; operational concerns; and a possible narrowing of the curriculum. Teachers still appear unwilling to lose their control over assessment and many learners also continue to adhere to traditional power roles (Bachman & Palmer, 1997). These are still some problems among the lecturers who do not want to use technology such as WBFA in their course because they may think that this will increase their teaching load which is rather heavy when compared to lectures loads in the other departments. One potential solution is to train them and provide more facilities to assist them to realise the benefits of the technology integration.

Additionally, the benefits of implementation of technology in assessment yield prosperous outcomes. It provides teachers more than student scores. It also provides much more convenience for them to compile a report for a conference (Juliana, Miller, & Edwards, 2004). This is another good reason to provide to the lecturers who are

hesitant to use the technology with assistance. This shows further value-adding to the initiative. Another major difficulty with the project was getting students to register on the WBFA program. In addition, students should understand clearly about the objectives of the WBFA, which emphasised the formative purposes.

It was concluded by Russell (2006) that teachers must see how the technology supports the curriculum and must be trained to use the technology, both in a technical manner and in an instructional manner. Teachers must have or must develop positive beliefs about the educational value of the technology.

6.7 WBFA and the Univesity policy

With reference to Morris and Milam (2004) and Northcote (2002), online assessment provided authentic, relevant tasks in order to assess student-learning outcomes. It was suitable for managing large groups of students because of its potential to reduce the time and cost associated with administering, distributing and marking high numbers of student assignments and examinations. Furthermore, administrative supports with strong media specialists and lecturers collaboration while providing greater access to technology could help increase students' achievement scores.

The results of this study can reaffirm that technology integration is useful for learning and teaching large enrolment EFL classes. This is in accordance with Burapha University's policy to issue a five-year plan (2007-2011) for the implementation of technology to support teaching staff. It is planned to enhance technology for the benefits of academic outcomes of the entire community. The results can be employed as evidence for further distance learning or e-learning which is the possible for the future direction of the university to support the lifelong learning approach. The results also provide reasons for the university to prepare relevant facilities to support the increasing number of students. It is also important to specify EFL lecturers for brainstorming the ideas of creating huge item banks to serve large classes and assessments. It is obvious that using technology integration in the large EFL classes encourages students to have more communicative activities than providing them with simply traditional face-to-face teaching. It is a better way to motivate enthusiastic students to yield better academic outcomes and meanwhile motivate other students to focus more on their study by the

provision of self-study and self-assessment. The ideas of formative assessment are suitable for stimulating students to be interested in their summative assessment. Hence, their learning outcomes will be increased. Students will have more purposeful ways of their EFL learning, instead of being passive students and become very passive citizen of the whole society which is never the nationwide direction.

Another issue to be considered is that using WBFA proves to be affordable for all students since they did not have to buy more workbooks. They can have their self-assessment through performing on WBFA with no additional cost. This not only saves their money but increases their motivation to learn the course as well. Lecturers in large EFL classes have to corroborate to assist their students. This means additional working loads at first, but a lot less loads later. This is a very good idea because at the moment the university provides some facilities for them. They just use those facilities in the more purposive way for their students. However, the University must have a very intensive training for these teaching staff to make them more confident to integrate this technology into their classrooms. By this, the assessment system will be improved. More emphasis will be on item analysis, test reliability and test validity. Far more will be on the usefulness of the assessment. This will increase the test quality. It is anticipated that this would be more preferable than creating a test to evaluate their students and end up at that stage, and no further test improvement is continued.

With more institutions integrating technology component in their courses, the need for faculty technical support would increase, as noted by Keengwe (2004). In this regard, support must include the desktops, classrooms, and the development of course materials, as well as the technical infrastructure that supports everything. With the increase in technology resources and access, it makes more sense for faculty to invest their time and effort into use technology for their teaching, especially when they realise that doing so will make teaching easier in due course.

6.8 Limitations

The results revealed in the research are limited by the following circumstances which occurred during the time of the study. Firstly, the numbers of stratified WBFA sample in some subgroups were rather low. This is because there were non-performing students

as mentioned above. They did not perform their WBFA actions despite frequent notification and encouragement from their teachers and stimulating emails from the researcher. In addition, they did not desire to resign from the group, so the other stratified samples were not selected to replace these students. This caused many difficulties when student numbers were not adequate for tests of statistical significance to be able to be calculated to compare their mean scores with students in the CPFA group.

Secondly, there was concern about the WBFA security. WBFA was a formative assessment; therefore, it was considered as low-stakes assessment, as noted by Crisp (2007) and Roever (2001). As a result, there was not a serious problem about the WBFA security. The students could even copy the assessment and printed it out for their study if they wanted to. Their performance on WBFA was weighted 5% of the course grade. However, in the post-WBFA questionnaire, no students mentioned about cheating on WBFA. The researcher had explained the objectives of delivery this assessment to all students in the experimental group to clarify that there was no need to cheat because this assessment was created to help them learn from their mistakes. It was used as a tool to help them. It was not employed to judge them. Some students said in the post-WBFA semi-structured interview that they did not want to look at the correct answers provided since they wanted to seek the right solutions for the questions by themselves. That was the major aim of the formative assessment. The WBFA was just one teacher-made test, a classroom test, which was used in combination with other assessments to provide the overall course grade to the students. It was constructed by the researcher. If there were more lecturers involved to construct this kind of assessment, it was anticipated that that it would increase its effectiveness.

Thirdly, students in the WBFA group might perform in WBFA and then attend CPFA. This was beyond the control of the study. For this, Creswell (2005) stated that in practice, personal factors that participants brought to an experiment could never be totally controlled. Finally, there is one occasion during the program that students could not log in to perform WBFA due to technical problems. That was between 1 January and 2 February 2005. This event might be one reason to discourage some students to continue doing the assessment later.

6.9 Recommendations

As noted by Wilson and Lowry (2000), the Web is indeed a place where constructivist learning can happen all the time. There are three crucial principles for effective use of the Web. First, it provides access to rich sources of information. Second, it encourages meaningful interactions with content. Finally, the Web brings people to challenge, support, or respond to one another. Consequently, EFL lecturers should contribute themselves to the usefulness of the WBFA to encourage their students with this huge source of information and communicative interactions outside their classroom.

The question which remains is why the WBFA is beneficial to students in large EFL classes. While these results indicate that the intervention had an effect, they tell us little about how or why. To answer these questions, more work is required, at the very least, an investigation of how students actually used the system and how they used the feedback they were given, according to Buchanan (2000). It is related to the ideas of finding the answers of how and why the students used the WBFA. The only evidence might be from the questionnaires and the semi-structured interviews. There were some students stating about cheating whereas the others saying they did not pay much attention to their WBFA scores. They wanted to focus on testing their knowledge after they studied the lesson in the lecture class. As a result, there were at least two groups of students who had different ideas about their use of the WBFA. This does not necessarily mean that WBFA use leads to higher marks. Students who are dedicated, conscientious, or highly motivated to succeed might be expected to use the exercises extensively. They might also be expected to work hard in other ways—and it could be this hard work in other areas which influences their exam performance, not their WBFA use. This may be unpacked by looking at the role a third variable which should be strongly influenced by the same factors, attendance at classes. Should the findings outlined above simply reflect hard work in other areas, then if the effect of class attendance is controlled for in an analysis, WBFA use may not emerge as a significant predictor of performance. In conclusion, it seems to have demonstrated that the use of WWW-based formative assessment, when an integral part of a course syllabus, seems to have be associated with superior performance.

In the future research related to the WBFA, it should be noted that the number of students in each sample should be sufficient for statistical analyses to be conducted and backup plan incorporated. This would prevent the problems experienced in the study where there were some non-performing students in the WBFA group and this ended up with additional analyses concerning the unequal sample sizes. That took more time to calculate the homogeneity of the variance by using *F* –test.

The next issue to be considered is the researcher can use the TOETL or IELTS tests in stead of pre- and post tests. However, there will be some issues to think about regarding costs and objectives of the test whether those test aims are in accordance with the course objectives or not. The good thing about using those standard tests is that the students will have better opportunities to practice all four language skills, instead of the items in relation to the student final examination.

In future WBFA research can emphasise computer adaptive testing (CAT), or computer adaptive formative testing (CAFT), where students are able to perform the assessment due to their actual abilities. For example, if the students answer correctly to the first question, the next question will be more difficult. Conversely, if they submit the wrong answer, the next question will be easier than the previous question. This can save test time because the students will not have to do all questions that are irrelevant to their abilities.

Therefore, if EFL lecturers do not try to learn how to use the technology as soon as possible, they will encounter several problems regarding technology development because the growth of the Internet moves so fast. They should start to be interested in this WBFA to avoid to be left behind in this technology age.

6.10 Conclusions

The evidence-based findings in this mixed method research design revealed that WBFA has positive effects on both students' achievement and their attitudes. The quantitative results indicated the mean scores of the sample in the WBFA group surpassed those of the CPFA group. Statistics showed that the WBFA groups in two subgroups (Faculty of Education and Faculty of Public Health) outperformed their CPFA groups with

significance (p <.05). Students' EFL learning was also enhanced by the WBFA program, particularly in the *reading comprehension* and *vocabulary* sections where the WBFA overall achievement scores increased with statistical differences (p <.05).

In addition, qualitative results disclosed that students in the WBFA group had more positive attitudes toward the use of WBFA after they completed the program than before they used it. Findings in this study indicated that they enjoyed using WBFA in the language course. The students felt that WBFA helped improve their language learning in the course. It also encouraged them to study by themselves. WBFA could connect them to real-life activities. WBFA was appropriate for the language course. WBFA motivated them to study in the course. WBFA provided useful Web sites. WBFA should be integrated in the course. They could use online interactions on WBFA.

The conclusions demonstrate a positive correlation between using WBFA and students' achievement when the WBFA program was integrated in the large enrolment foundation English course in a tertiary level. The students' overall academic achievement in the experimental group increased when compared with that in the control group. Although there was no significant difference between the overall means of both groups. However, it is evident that WBFA was, at least, as good as the CPFA. That is, using the WBFA made no differences from using the conventional paper-and-pencil assessment in the course, regarding students' achievement. Furthermore, the results from the questionnaires and semi-structures interviews indicated that the students had positive attitudes toward the WBFA program. Findings indicated that students' achievement was enhanced when they were provided with well-planned formative assessment to allow them to study on their own pace while they still received the similar opportunity as they did when they used CPFA in their previous course. This certainly encouraged them to be more autonomous to stimulate life long learning. If they are provided sufficient facilities from the university, they will be better motivated to emphasise on studying by themselves at their own convenience regardless of time and venue.

There is nothing unique about the WBFA program. It is probably the fact that it provides timely feedback which meets the requirements for formative use. Its value is

probably also enhanced by the fact that it does not supply additional explanation with correct answers to encourage students to do further work on their own. It would be interesting to test experimentally by comparing a system such as this with one which provides explanations with correct answers. In this study even the correct answers were provided, it did not mean that students never learnt anything from their performance in WBFA. The score reported immediately after the completion of each set test would lead to disclose something to students. They eventually knew what they strengths or weaknesses were and at the same time, they learnt how much they knew from the class what needed to be revised or studied later to got improvement. This system is an example of the meaningful interaction between student and instructional materials, which is an essential component of successful pedagogy — a component that can be provided through technology.

In particular term, it seems that the WBFA actually yielded productive impacts in terms of students' achievement and attitudes. This is reassuring to the number of EFL lecturers to accept such testing system. The system may not be the vital teaching tool since there are language features which are difficult to assess through simply the multiple-choice questions in the WBFA program. In addition, there are types and sources of feedback which may be more valuable than simply providing the correct answers to students. However, the WBFA program did offer the opportunity to supply some kind of individualised in a flexible and cost-effective manner. As part of a balanced curriculum utilising an appropriate selection of teaching, learning, and assessment methods, it offers another form of learning opportunity, and these findings suggest that students who engaged in that opportunity are likely to benefit from the experience.

By using WBFA, not only the students' achievement scores in the WBFA group are better than those in the CPFA group, even though there was no significant difference, but also the students' attitudes are positive from the results reported in Chapter 4 and Chapter 5. It is obvious that WBFA is suitable for large enrolment EFL classes. It supports life-long learning, self-study which is in accordance with the constructivism. All lecturers in the large classes should be supported to use WBFA in their course. They should be supported to be trained to integrate this technology in their large EFL classes. All relevant facilities should be offered to them. The roles of the teachers and the

students should be revolutionised. The students should be explained how to use the e-contacts to ensure they will never lack interaction with their lecturers. The university should focus the integrity of technology by stating it in the university policy. It would be a great idea to refer to the BUU document about this issue. The result of this study can be adapted seriously to further issue like distance e-learning because the university has already had these facilities on campus. Lecturers should be encouraged to focus on the quality of their assessment by using item analysis and creating item banks for students' benefits. This would be better than creating the tests then do nothing after grading their students at the end of the semester. The assessment in the large enrolment English courses should be emphasised to communicative approach to increase the students' language skills.

References

AEN. (2003). Country report Thailand. Retrieved 30 June, 2004, from www.asia-elearning.net/content/aen_conference_2003/files/q_results/q_thailand.pdf

Aggarwal, A., & Bento, R. (2003). Web-based education. In A. Aggarwal (Ed.), *Web-based learning and teaching technologies: Opportunities and challenges* (pp. 2-14). Hershey, PA, USA: Idea Group.

Aiken, L. R. (1996). *Rating scales and checklists evaluating behavior, personality, and attitudes*. New York: John Wiley & Sons.

Alagumalai, S., Toh, K., & Wong, J. (2000). Web-based assessment: Techniques and issues. In A. Aggarwal (Ed.), *Web-based learning and teaching technologies: Opportunities and challenges* (pp. 246-256). Hershey, PA, USA: Idea Group.

Al-Amri, S. (2007). Computer-based vs. paper-based testing: Are they the same? In M. Danson (Ed.), *11th International Computer Assisted Assessment Conference, Loughborough University*. Leicestershire, UK: Loughborough University.

Aluisio, S. M., & Oliveira Jr, O. N. (1999, 16-17 June). *An innovative computer assisted proficiency test of English for academic purposes.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Arkkelin, D. (2001, 20-24 July). *Trials and triumphs of testing online: Use of online tests in relation to students' attitudes towards an experience with computer.* Paper presented at the Syllabus 2001, Santa Clara, CA, USA.

Astin, A., Banta, T., Cross, K., El-Khawas, E., Ewell, P., Hutchings, P., Marchese, T., McClenney, K., Mentkowski, M., Margaret, A., Moran, E., & Wright, B. (2003). *9 principles of good practice for assessing student learning*. Retrieved 30 March, 2004, from http://ultibase.rmit.edu.au/Articles/june97/ameri1.htm#9

Atkinson, T., & Davies, G. (2000). *Computer aided assessment (CAA) and language learning*. Retrieved 12 September, 2003, from http://www.ict4lt.org/en/en mod4-1.htm

AUTC. (2002). *Teaching large classes*. Retrieved 22 September, 2004, from http://www.tedi.uq.edu.au/largeclasses/

Babbie, E. (1986). *The practice of social research*. Belmont, CA: Wadsworth.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.

Bachman, L. F., & Palmer, A. S. (1997). *Language testing in practice: Designing and developing useful language tests* (2nd ed.). Oxford: Oxford University Press.

Baggot, G., & Rayne, R. (2004, 6-7 July). *Student perceptions of computer-based formative assessment in semi-distance module.* Paper presented at the 8th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Baker, D. E. (2003). *Online technology teaching model: A pilot test by nursing faculty.* Unpublished doctoral thesis, University of Delaware, Newark.

Balnaves, M., & Caputi, P. (2001). *Introduction to quantitative research methods: An investigative approach*. Thousand Oaks, CA: SAGE Publications.

Barnett, V. (2002). *Sample survey: Principles & methods* (3rd ed.). New York: Oxford University Press.

Basanta, C. (1995). Coming to grips with progress testing. *English Teaching Forum Online, 33*(3), 55.

Basham, J. D. (2002, 22 February). *Technology integration: What this means for our students and teachers.* Paper presented at the Midwest Symposium for Leadership in Behavior Disorders, University of Illinois, USA.

Beale, J. (2002). Is communicative language teaching a thing of the past? *Babel, 37*(1), 12-16.

Beattie, K. K. (2000). *The effects of intensive computer-based language intervention on language functioning and reading achievement in language impaired adolescents.* Unpublished doctoral thesis, George Mason University, Virginia.

Bennett, R. E. (2001). How the Internet will help large-scale assessment reinvent itself. *Educational Policy Analysis Archives, 9*(5).

Bennett, R. E. (2002a). Using electronic assessment to measure student performance. *The State Education Standard, 3*(3), 22-29.

Bennett, R. E. (2002b). Inexorable and inevitable: The continuing story of technology and assessment. *The Journal of Technology, Learning, and Assessment, 1*(1).

Bennett, S., Oliver, A., & Pinn, A. (1999, 16-17 June). *The uptake of CAA: Some theory! Some practice!* Paper presented at the 3rd International Computer Assisted assessment Conference, Loughborough University, Leicestershire, UK.

Benson, M. R. (2003). Assessing participants learning in online environments. *New Directions for Adult and Continuing Education, 100*, 69-78.

Benson, S. J. (2000). *Technology integration: An in-depth, reflective position paper*. Retrieved 10 March, 2007, from http://mathstar.nmsu.edu/educ621/ sharon4.html

Berg, K. E., Latin, R.W. (2008). *Essentials of research methods in health, physical education, exercise science, and recreation* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

Beverly, J., Beverly, E., Clarke, M., & White, S. (2001, 2-3 July). *Changing perceptions in language learning and testing.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Black, P., & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-149.

Black, P., & Wiliam, D. (1998b). *Assessment and classroom learning: A review of 25 years of world wide research on formative assessment*. Retrieved 23 July, 2004, from www.jh.retsd.mb.ca/rbl/appendices/formativeres.ppt

Blayney, P., & Freeman, M. (2003, 8-9 July). *Automated marking of individualised spreadsheet assignments: The impact of different formative self-assessment options.* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Bonham, S., Beichner, R., & Deardorff, D. (2001). Online homework: Does it make a difference? *The Physics Teacher, 39,* 293-296.

Bonham, S., Titus, A., Beichner, R., & Martin, L. (2000). Education research using Web-based assessment system. *Journal of Research on Computing in Education, 33*(1).

Born, A. D. (2003). Web-based student assessment. In A. K. Aggarwal (Ed.), *Web-based education: Learning from experience* (pp. 165-188). Hershey, USA: Idea Group Inc.

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation, 8*(9).

Bouma, G. D., & Ling, R. (2004). *The research process* (5th ed.). Oxford: Oxford University Press.

Boyle, T. (2000, 23-25 August). *Constructivism: A suitable pedagogy for information and computing sciences?* Paper presented at the 1st LTSN-ICS Annual Conference, 1st LTSN-ICS Annual Conference, Heriot-Watt, Edinburgh, UK.

Bringsjord, E. L. (2001). *Computerized-adaptive versus paper-and-pencil testing environments: An experimental analysis of examinee experience.* Unpublished doctoral thesis, State University of New York, Albany.

Britten, J. S., & Cassady, J. C. (2005). The technology integration assessment instrument: Understanding planned use of technology by classroom teachers. In C. D. Maddux, & D. LaMont Johnson (Ed.), *Classroom integration of type II uses of technology in education* (Vol. 22, pp. 49-61). New York: Haworth Press.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing students learning in higher education*. London: Routledge.

Brown, H. D. (1994). *Principles of language learning and teaching* (3rd ed.). New Jersey: Prentice Hall Regents.

Brown, J. D. (1995). *Understanding research in second language learning* (1st ed.). Cambridge: Cambridge University Press.

Brown, J. D. (1997). Computer in language testing: Present research and some future directions. *Language Learning & Technology, 1*(1), 44-59.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press.

Brown, J. D., & Rodgers, T.S. (2002). *Doing second language research*. Oxford: Oxford University Press.

BUCC. (2007). *History of Burapha University Computer Center*. Retrieved 22 January, 2006, from http://ict.buu.ac.th/index.php?option=com_content&task=view&id=25&Itemid=37

Buchanan, T. (2000). The efficacy of a World-Wide Web mediated formative assessment. *Journal of Computer Assisted Learning, 6,* 193-200.

Bugbee, A. C. J. (1996). The equivalence of paper-and-pencil and computer-based testing. *Journal of Research on Computing in Education, 28*(3), 282-299.

Bull, J. (1999). Computer-assisted assessment: Impact on higher education institutions. *Educational Technology & Society, 2*(3).

Bull, J. (2003). *Online assessment workshop*. Retrieved 4 March, 2004, from http://www.eduology.com/cats/Bhambooklet.pdf

Burns, R. B. (1999). *Introduction to research methods* (4th ed.). Frenchs Forest, NSW: Pearson Education.

BUU. (2003). *Numbers of first year students in 2003*. Retrieved 13 June, 2004, from http://service.buu.ac.th/WebAademic/news/stat/b.pdf

BUU. (2004). *Numbers of students in 212102 English II course*. Retrieved 13 July, 2004, from http://reg.buu.ac.th/registrar/class_info_2.asp?backto=home&option =0&courseid=21210145&acadyear=2547&semester=1&avs419197262=12

Calhoon, M. B., Fuchs, L. S., & Hamlett C. L. (2000). Effects of computer-based test accommodation performance assessments for secondary students with learning disabilities. *Learning Disability Quarterly, 23*(4), 271-282.

Carbone, A., Schendzielorz, P. & Zakis, J. (2000). Electronic assessment and self-paced learning on the Web using a multiple-choice quiz generator. *International Journal of Electrical Engineering Education, 37*(2), 119-125.

Carneson, J., Delpierre, G., & Masters, K. (2002). Designing and managing MCQ's. *The Computer ASsisted Teaching & LEarning (CASTLE).*

Caygill, R., & Eley, L. (2001, 13-15 September). *Evidence about the effects of assessment task format on student achievement.* Paper presented at the Annual Conference of the British Educational Research Association, University of Leeds, England.

CEC. (2006). *Continuing Education Center Chulalongkorn University: History*. Retrieved 1 August, 2006, from http://www.cec.chula.ac.th/webcec/profile.html

Chalhoub-Deville, M. (2001). Language testing and the technology: Past and future. *Language Learning & Technology, 15*(2), 95-98.

Chambers, E. A. (2002). *Efficacy of educational technology in elementary and secondary classrooms: A meta-analysis of the research literature from 1992-2002.* Unpublished doctoral thesis, Southern Illinois University, Carbondale.

Chang, E. A. (2002). *The efficacy of asynchronous online learning in the promotion of critical thinking in graduate education.* Unpublished doctoral thesis, Columbia University Teachers College, New York.

Chapelle, C., & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge: Cambridge University Press.

Charman, D., & Elmes, A. (1998). A computer-based formative assessment strategy for a basic statistics module in geography. *Journal of Geography in Higher Education, 22*(3), 381-385.

Charnitski, C., Molinaro, J., Corabi, J., & Nolan, K. (2003). Comparing student achievement in a graduate level research methods course using face-to-face and Web-based instruction: Results of a pilot study. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003* (pp. 174-177). Chesapeake, VA: AACE.

Charnsripinyo, C. (2006). Research and education networks in Thailand and international collaborations.

Chen, H., Sullivan, H., & Savenye, W. (2002). Perspectives on the future of computer use in China. *Educational Technology Research and Development, 50*(1), 92-101.

Cheung, H. (2004). Preparing tomorrow's teachers to use media technologies in language education. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 2039-2046.

Chin, K. (2001). *Attitudes of Taiwanese nontraditional commercial institute students toward computers.* Unpublished doctoral thesis, University of South Dakota, South Dakota.

Choi, I., Kim, K., & Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. *Language Testing, 20*(3), 295-320.

Chuang, T., & Chen, W. (2004). *E-learning in ESL classrooms: A literature review.* Paper presented at the Society for Information Technology and Teacher Education International Conference (SITE), Orlando, Florida, USA.

Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology, 33*(5), 593-602.

Clarke, C., & Dawson, R. (1999). *Evaluation research: An introduction to principles, methods and practice*. London: SAGE Publications.

Conole, G., & Sclater, N. (2005, 5-6 July). *Using evaluation to inform the development of a user-focused assessment engine.* Paper presented at the 9th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Conradson, S., & Pedro, H. (2004). Computers, the Internet, and cheating among secondary school students: Some implications for educators. *Practical Assessment, Research & Evaluation, 9*(9).

Cook, L. J. (2000). *The effects of Web-based instruction on preservice teachers' attitude, confidence, and perceived value regarding Internet-based learning resources.* Unpublished doctoral thesis, Georgia State University, Georgia.

Coombes, H. (2001). *Research using IT*. London: Palgrave.

Creswell, J., Plano C., Gutmann, M., & Hanson, W. (2003). Advanced mixed methods research designs. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 209-240). Thousand Oaks, CA: SAGE Publications.

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks: SAGE Publications, Inc.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: SAGE Publications.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Merrill Prentice Hall.

Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education, Inc.

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education. Inc.

Crisp, G. (2007). *The e-assessment handbook.* New York: Continuum International Publishing Group.

Crofts, A. (1999, 16-17 June). *Enabling reuse of CAA by design.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Crooks, T. (2001, 13-15 September). *The validity of formative assessments.* Paper presented at the Annual Conference of the British Educational Research Association, University of Leeds, England.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Cross, J. (2001, 2-3 July). *Computerised formative assessment in very large laboratory classes.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Cucchiarelli, A., Panti, M., & Vanlenti, S. (2000). Web-based assessment in student learning. In A. Aggarwal (Ed.), *Web-based learning and teaching technologies: Opportunities and challenges* (pp. 175-197). Hershey, PA, USA: Idea Group.

Cummings, C., Kalkman, D., & Underwood, J. (2004). Integration of technology in higher education: Projects that rock. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 2072-2074.

Curtin, J. (2002). WebCT and online tutorials: New possibilities for student interaction. *Australian Journal of Educational Technology, 18*(1), 110-126.

Dabbagh, N. H. (2000). The challenges of interfacing between face-to-face and online instruction. *Tech Trends, 44*(6), 37-42.

Dalziel, J. (2000, 21-22 June). *Integrating computer assisted assessment with textbooks and question banks: Options for enhancing learning.* Paper presented at the 4th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Dalziel, J. (2001, 2-3 July). *Enhancing Web-based learning with computer assisted assessment: Pedagogical and technical considerations.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Davies, P. (1999, 16-17 June). *Learning through assessment: OLAL .. on-line assessment and learning.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Davies, P. (2001, 2-3 July). *Computer aided assessment must be more than multiple-choice tests for it to be academically credible?* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Davies, P. (2002). "There's no Confidence in Multiple-Choice Testing, ……". In M. Danson (Ed.), *6th International Computer Assisted Assessment Conference, Loughborough University*. Leicestershire, UK: Loughborough University.

Davies, M. B. (2007). *Doing a successful research project: Using qualitative or quantitative methods*. New York: Palgrave Macmillan.

Dempsey, M. S., PytlikZillig, L. M., & Bruning, R. (2005). Building writing assessment skills using Web-based cognitive support features. In L. M. Pyllikzillig,

Bodvarson, M, & Bruning, R. (Eds.), *Technology-based education: Bringing researchers and practitioners together* (pp. 83-105). Greenwich, Connecticut: Information Age Publishing, Inc.

Denscombe, M. (2003). *The good research guide for small-scale social research projects* (2nd ed.). Maidenhead: Open University Press.

Di, X., Dunn, D., & Lee. S. J. (2000). An integrated approach to instructional technology. *Action in Teacher Education, 22*(2A), 1-13.

Dornyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dow, R. (2003). The introduction of Web-based formative assessment in medical science: An action research project. *CAL-laborate, 10*.

Dufresne, R., Mestre, J., Hart, D. M., & Rath, K. A. (2002). The effect of Web-based homework on test performance in large enrolment introduction physics courses. *Journal of Computer in Mathematics and Science Teaching, 21*(3), 229-251.

Dunkel, P. (1999). Considerations in developing and using compute adaptive tests to assess second language proficiency. *Language Learning & Technology, 2*(2), 77-93.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs: Prentice Hall.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Ellis, W., & Ratcliffe, M. (2004, 6-7 July). *Improving results with positive directed feedback in summative assessments.* Paper presented at the 8th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Emberger, M. (2002). Focused feedback. *Maryland Classroom, 7*(3).

Ervin-Tripp, S. M. (1974). Is second language learning like the first? *TESOL Quarterly, 8*(2), 124.

Eustace, D. (2003). *On-line assessment for large groups*. Retrieved 25 February, 2004, from http://www.edu.salford.ac.uk/scd/tlqis/reports/eustace.pdf

Ewing, J. P. (2000). *Teaching language towards computer-based technology*. Unpublished doctoral thesis, University of Alberta, Alberta.

FairTest. (2007). *The value of formative assessment*. Retrieved 3 March, 2008, from http://www.fairtest.org/examarts/winter99/k-forma3.html

Farrell, G. (2006). A comparison of an innovative Web-based assessment tool utilizing confidence measurement to the traditional multiple choice, short answer and problem solving questions. In M. Danson (Ed.), *10th International Computer Assisted Assessment Conference, Loughborough University*. Leicestershire, UK: Loughborough University.

Farrell, G., & Leung, Y. (2004, 6-7 July). *A comparison of two student cohorts utilizing Blackboard CAA with different assessment content: A lesson to be learnt.* Paper presented at the 8th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Farrell, G., Farrell, V., & Leung, Y. (2005, 5-6 July). *A comparison of Blackboard CAA and an innovative self-assessment tool for formative assessment.* Paper presented at the 9th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Farthing, D., & McPhee, D. (1999, 16-17 June). *Multiple choice for honours-level students? A statistical evaluation.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Felzer, J. S. (2002). *The Effects of Internet-based testing versus traditional testing on elementary algebra students' achievement and attitudes.* Unpublished doctoral thesis, North Carolina State University, North Carolina.

Fenno, C. (2000). *Assessment in higher education*. Retrieved 9 March, 2004, from http://www.glencoe.com/ps/teachingtoday/educationupclose.phtml/9

Ferero, M. (2002). *Effect of classroom use of the Web on the self-efficacy of minority students is at-risk learning situations.* Unpublished doctoral thesis, Stanford University, California.

Ferguson, G. A., & Takane, Y. (1989) Statistical analysis in psychology and education. New York: McGraw-Hill.

Fernandez-Garcia, M. (2001). Review of issues in computer-adaptive testing of reading proficiency. *Language Learning and Technology, 5*(2), 19-22.

Finch, A. (2000). *A formative evaluation of a task-based EFL programme for Korean university students.* Unpublished doctoral thesis, Manchester University, Manchester.

Finnis, J. A. (2004). Learning technology: The myths and facts. *International Journal of Instructional Technology & Distance Learning, 1*(5), 53-62.

Fisher, R. A., & Yates, F. (1963). *Statistical tables for biological, agricultural and medical research* (6th ed.). Edinburgh: Oliver and Boyd.

Foltos, L. (2003). Technology and academic achievement. *New Horizons for Learning Online Journal, 9*(1).

Fox, R., Justus, M., & Bachman, L. (2004). Technology changes how we teach and learn: A close look at the implementation of technology in a foreign language methods course. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 3303-3311.

Frank, M., & Barzilai. (2004). Designing course web sites for supporting lecture-based courses in higher education--some pedagogical aspects. *International Journal of Instructional Technology & Distance Learning, 1*(12), 37-50.

Frankel, J. T. (2002). *An evaluation of a Web-based model of assessment for the New Jersey state core curriculum content standards in music.* Unpublished doctoral thesis, Columbia University Teachers College, New York.

Frary, R. (1995). More multiple-choice item writing do's and don'ts. *Practical Assessment, Research & Evaluation, 4*(11).

Frary, R. (2002). *A brief guide to questionnaire development*. Retrieved 19 May, 2004, from http://www.testscoring.vt.edu/fraryquest.html

Frechtling, J., & Sharp, L. (1997). *User-friendly handbook for mixed method evaluations*. Retrieved 7 July, 2004, from http://www.nsf.gov/pubs/1997/nsf97153/start.htm

Frew, P. (2001). *Evaluation and assessment strategies for Web-based education*. Retrieved 4 August, 2004, from http://naweb.unb.ca/02/Frew.ppt

Fulcher, G. (2000). Computers in language testing. In P. Brett, and G. Motteram (Eds.), *A special interest in computers: Learning and teaching with information and communications technologies* (pp. 93-107). Manchester, UK: IATEFL publications.

Gardner-Medwin, A., & Gahan, M. (2003, 8-9 July). *Formative and summative confidence-based assessment.* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Ghani, R., & Daud, N. (2003). Computer-mediated communication: Its pedagogical aspects and considerations. *Teaching English with Technology A Journal for Teachers of English, 3*(2).

Gillham, B. (2000). *Developing a questionnaire*. London: Continuum.

Gillham, B. (2000). *The research interview*. London: Continuum.

Ginn, W. (1995). *Jean Piaget--Intelligence development*. Retrieved 21 July, 2004, from http://www.sk.com.br/sk-piage.html

Glass, G. V., & Hopkins, K.D. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston: Allyn & Bacon.

Godwin-Jones, B. (2001). Emerging technologies: Language testing tools and technologies. *Language Learning & Technology, 5*(2), 8-12.

Goldberg, A. L. (2000). *Test-level, item-level, and experiential differences on computer and paper-and-pencil versions of a practice Graduate Record Exam (GRE).* Boston College.

Gomez, S. (2002). *Student attitudes to learning with new technologies*. Retrieved 3 August, 2003, from http://www.uwe.ac.uk/library/nlsf/about/Stephen_Gomez.pdf

Gorard, S. (2001). *Quantitative methods in educational research*. London: Continuum.

Greene, J., & Caracelli, V. (1997). Defining and describing the paradigm issue in mixed-method evaluation. In J. C. Greene, & Caracelli, V.J. (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 95-17). San Francisco, CA: Jossey-Bass.

Greene, J., & Caracelli, V. (2003). Making paradigmatic sense of mixed methods practice. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 91-110). Thousand Oaks, CA: SAGE Publications.

Gretes, J., & Green, M. (2000). Improving undergraduate learning with computer-assisted assessment. *Journal of Research of Computing in Education, 33*(1), 46-54.

Gur, B., & Van Schaack, A. (2004, 6-7 July). *Approaches to assessment of online learning: Conceptual challenges.* Paper presented at the 8th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Hall, R., Butler, L., Kestner, N., & Limbach, P. (1999). *Combining feedback and assessment via Web-based homework*. Retrieved 26 September, 2003, from http://taddeo.emeraldinsight.com/vl=5230438/cl=21/nw=1/fm=html/rpsv/cw/mcb/10650741/v16n1/s3/p24

Hamilton, L., Klein, S., & Lorie, W. (2000). *Using Web-based testing for large-scale assessment*. Retrieved 21 August, 2003, from http://www.rand.org/publications/IP/IP196/IP196.pdf

Hanson, J., Millington, C., & Freewood, M. (2001, 2-3 July). *Developing a methodology for online feedback and assessment.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Harper, R. (2001, 2-3 July). *Web-based delivery of courseware, tutorials and assessment with WebOL*. Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Haycock, B. A. (2001). *Developing and using Web-based management systems for large-enrollment courses: Homework, testing, and delivery*. Unpublished doctoral thesis, The University of Nebraska, Lincoln.

Hayden, J. (2002). *The development and formative evaluation of an assessment instrument based on the National Educational Technology Standards for Teachers.* Unpublished doctoral thesis, Georgia State University, Georgia.

Hein, G. (1991, 15-22 October). *Constructivist learning theory*. Paper presented at the CECA (International Committee of Museum Educators) Conference, Lesley College, Massachusetts, USA.

Heinemann, M. (2003). *Teacher-student interaction online and learning in Web-based graduate theological education.* Unpublished doctoral thesis, Trinity Evangelical Divinity School, Illinois, USA.

Henly, D. (2003). Use of Web-based formative assessment to support students learning in a metabolism/nutrition unit. *European Journal of Dental Education, 7*(3), 116-122.

Higgins, E., & Tatham, L. (2003). Exploring the potential of multiple-choice questions in assessment. *Learning & Teaching in Action, 2*(1).

Higgins, J., Russell, M., & Hoffman, T. (2004). *Examining the effect of computer-based passage presentation on reading test performance: Part of the New England compact enhanced assessment project*. Retrieved 17 March, 2005, from http://escholarship.bc.edu/cgi/viewcontent.cgi?article=1028&context=intasc

Hricko, M., & Howell, S. L. (2006). *Online assessment and measurement: Foundations and Challenges*. Hershey: Information Science Publishing.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.

Hunt, N., Hughes, J., & Rowe, G. (2002). Formative automated computer testing (FACT). *British Journal of Educational Technology, 33*(5), 525-535.

HUSO. (2008). *Departments in the Faculty of Humanities and Social Sciences*. Retrieved 1 January, 2008, from http://www.huso.buu.ac.th/work.asp

Ingram, D. (2003, October). *Methodology in the new millennium: Towards more authenticity in language learning and assessment.* Paper presented at the First International Conference on Pedagogies and Learning: New meaning for a new millennium, University of Southern Queensland, Australia.

Inoue, Y., & Bell, S. T. (2006). *Teaching with educational technology in the 21st century: The case of the Asia-Pacific region*. London: Information Science Publishing.

Irving, A., Read, M., Hunt, A., & Knight, S. (2000, 21-22 June). *Use of information technology in exam revision.* Paper presented at the 4th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Jacobs, L. C. (1991). *Test reliability*. Retrieved 19 May, 2004, from http://www.indiana.edu/~best/test_reliability.shtml

Jao, F. (2001). *An investigation of preservice teachers' attitudes and confidence levels toward educational technology standards and selected instructional software applications.* Unpublished doctoral thesis, The University of Toledo, Ohio.

Jackson, S. L. (2009). *Research methods and statistics: A critical thinking approach*. Belmont, CA: Wadsworth.

Jefferies, P., Constable, I., Kiely, B., Richardson, D., & Abraham, A. (2000, 21-22 June). *Computer aided assessment using WebCT.* Paper presented at the 4th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Jia, J. (2004). The study of the application of a Web-based chatbot system on the teaching of foreign languages. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1201-1207.

Johnson, B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative and mixed approaches* (2nd ed.). Boston, MA: Allyn & Bacon.

Johnson, B., & Turner. L. (2003). Data collection strategies in mixed methods research. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 297-319). Thousand Oaks, CA: SAGE Publications.

Johnson, L., Liu, L., & Cheney, C. (2004). An approach to improving learning through authentic learning projects. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1737-1742.

Jonassen, D. (1994). Thinking technology. *Educational Technology, 34*(4), 34-37.

Jonassen, D., Howland, J., Moore, J., & Marra, R. (2003). *Learning to solve problem with technology: A constructivist perspective* (2nd ed.). Columbus: Merrill Prentcie Hall.

Juliana, M., Miller, D., & Edwards, M. (2004). Linking technology with assessment: Are we assessing what we think we are? *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 127-129.

Keengwe, J. (2004). Faculty development, support and training: Effectively integrating educational technology in the classroom. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 2316-2318.

Kemper, E., Stringfield, S., & Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 273-296). Thousand Oaks, CA: Sage Publications.

Kendall, M., & Prowse, A. (2005). Supporting first year undergraduates through blended learning. *Learning & Teaching in Action, 4*(1).

Kenyon, D., & Malabonga, V. (2001). Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments. *Language Learning & Technology., 5*(2), 60-83.

Khan, K., Davies, D., & Gupta, J. (2001). Formative-self assessment using multiple true-false questions on the Internet: Feedback according to confidence about correct knowledge. *Medical Teacher, 23*(2), 158-163.

Kiattananan, P., Koanantakool, T., Chairatanayut, T., Kiatisevi, P., & Beck, R. (1999, 22-25 June). *Network design and resource management scheme in SchoolNet Thailand Project.* Paper presented at the Internet Global Summit, San Jose, CA, USA.

Kilickaya, F. (2007). The effect of computer assisted language learning on Turkish learners' achievement on the TOEFL exam. *IATEFL Poland Computer Special Interest Group Teaching English with Technology A Journal for Teachers of English, 7*(1).

Kim, M., Rich, P., & Hannafin, M. (2004). The challenges of teaching with computer tools: Teachers as facilitators of student learning. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 2329-2334.

Kjollerstrom, B., & Martensson, M. (1999, 16-17 June). *Success in CAA will change the way we learn.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Klein, S., & Hamilton, L. (1999). *Large-scale testing: Current practices and new directions*. Retrieved September 12, 2003, from http://www.rand.org/publications/IP/IP182/

Klingner, B. G. (2003). *The relationship between learning styles of adult learners enrolled in online courses at Pace University and success and satisfaction with online learning.* Unpublished doctoral thesis, Walden University, New York.

Knäuper, B. (2002). *Item analysis*. Retrieved 25 May, 2005

Koanantakool, T. (1995). *A perpetual chronicle on the Internet in Thailand: Our milestones*. Retrieved 5 March, 2007, from http://www.nectec.or.th/users/htk/milestones.html

Koanantakool, T., and Thuvasethakul, C. (2002). *National ICT Policy in Thailand*. Retrieved 21 April, 2007, from http://www.nectec.or.th/users/htk/publish/ 20020302-National-ICT-Policy-11- ppt.pdf

Kobrin, J. L., & Young, J. W. (2003). The cognitive equivalence of reading comprehension test items via computerized and paper-and-pencil administration. *Applied Measurement in Education, 16*(2), 115-140.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Torrance, CA: Laredo Publishing Company, Inc.

Krashen, S. D. (1987). *Principles and practice in second language acquisition*. New York: Prentice-Hall International.

Krashen, S. D. (1988). *Second language acquisition and second language learning*. New York: Prentice Hall.

Kurubacak, G. (2000). *Online learning: A study of students' attitudes towards Web-based instructions*. Retrieved 10 October, 2003, from http://wwwlib.umi.com/dissertations/preview_all/9973125

Lao, T. M. (2000). *A position paper on technology integration in the classroom*. Retrieved 11 March, 2007, from http://pt3.nmsu.edu/educ621/teresa4.html

Lee, H., & Lenzo, S. (2004). *Web-based testing with classroom instruction: A first step toward blended instruction.* Paper presented at the Society for Information Technology and Teacher Education International Conference 2004, Atlanta, GA, USA.

Lee, Y. (2003). *Effects of learning styles and attitudes on achievement within a Web-based course in Korea*. Retrieved 20 February, 2004, from http://web.syr.edu/~ylee30/portfolio/ide843/Final%20Prospectus.pdf

LeLoup, J., & Ponterio, R. (2000). *Enhancing authentic language learning experiences through Internet technology*. Retrieved 12 January, 2004, from http://www.cal.org/resources/digest/0002enhancing.html

LeLoup, J., & Ponterio, R. (2003). *Second language acquisition and technology: A review of the research*. Retrieved 1 January, 2006, from http://www.cal.org/resources/digest/0311leloup.html

Lewis, B. A. (2002). *Learning effectiveness: Efficacy of quizzes vs. discussions in on-line learning.* Unpublished doctoral thesis, Syracuse University, New York.

Liang, X., & Creasy, K. (2004). Classroom assessment in Web-based instructional environment: Instructors' experience. *Practical Assessment, Research & Evaluation, 9*(7).

Lilley, M., Barker, T., & Britton, C. (2005, 5-6 July). *Automated Feedback for a Computer-Adaptive Test: A Case Study.* Paper presented at the 9th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Lin, A. (2003). An initial study on EFL learners' attitude towards multimedia application in language learning. *Teaching English with Technology: A Journal for Teachers of English, 3*(2).

Lin, S., Liu, E., & Yuan, S. (2001). Web-based peer assessment: Attitudes and achievement. *IEEE Transactions on Education, 44*(2).

Liu, Y. (2004). Effects of online instruction vs. traditional instruction on students' learning. *International Journal of Instructional Technology & Distance Learning, 2*(3), 57-64.

Liu, Y. (2006). A comparison study of online versus traditional student evaluation of instruction. *International Journal of Instructional Technology & Distance Learning, 3*(4), 15-29.

Luanganggoon, N. (2001). *Improving English language teaching in Thailand.* Unpublished doctoral thesis, Deakin University, Melbourne.

Luecht, R. (2001, 11-13 April). *Challenges of Web-based assessment.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Seattle, Washington, USA.

Maccini, P., Gagnon, J., & Hughes, C. (2002). Technology-based practice for secondary students with learning disabilities. *Learning Disability Quarterly, 25*(4), 247-261.

Macdonald, J. (2002). *Developing competent e-learners: The role of assessment*. Retrieved 9 March, 2004, from http://www.leeds.ac.uk/educol/documents/00002251.htm

MacKay, B., & Emerson, R. (2000, 21-22 June). *Web-based assessment of writing skills.* Paper presented at the 4th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Mackenzie, D. (2003, 8-9 July). *Assessment for E-Learning: What are the features of an Ideal E-assessment system?* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Maddux, C. D., & LaMont Johnson, D. (2005). Information technology, type II: Classroom integration, and the limited infrastructure in schools. In C. D. Maddux, & D. LaMont Johnson (Eds.), *Classroom integration of type II uses of technology in education* (Vol. 22, pp. 1-5). New York: Haworth Press.

Malaiwong, K. (2000). *IT and schools in the future*. Retrieved 5 May, 2004, from http://www.drkanchit.com/ict_education/articles/it_school02.pdf

Martinez, L. (2002). *Internet-based vs. paper-pencil test administration: An equivalency study using the Ability Explorer.* Unpublished doctoral thesis, Texas A&M University, Texas.

Mascuilli, A. B. (2004). Comparing an online course to its classroom counterpart. *International Journal of Instructional Technology & Distance Learning, 1*(5), 29-31.

Mason, B. (2001). *Relative effectiveness of human and computer-based feedback in a mastery context.* Unpublished doctoral thesis, The University of Nebraska, Lincoln.

Mason, B., Patry, M., & Bernstein, D. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research, 24*(1), 29-39.

Maughan, S., Peet, D., & Willmott, A. (2001, 2-3 July). *On-line formative assessment item banking and learning support.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Maunder, P. (2002, 12-14 September). *In support of multiple choice questions: some evidence from curriculum 2000.* Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England.

Maxwell, J., & Loomis, D. (2003). Mixed methods design: An alternative approach. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 241-271). Thousand Oaks, CA: SAGE Publications.

McAlpine, M., Clark, R., & Hesketh, I. (2003, 8-9 July). *Multiple response questions-allowing for chance in authentic assessments.* Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

McCormack, A. (2002). *Interactive computer-based technology in EFL.* Paper presented at the Society for Information Technology and Teacher Education International Conference (SITE), Orlando, Florida, USA.

McKenna, C. (2001, 2-3 July). *Academic approaches and attitudes towards CAA: A qualitative study.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

McKenna, C., & Bull, J. (1999). *Designing effective objective test questions: An introductory workshop*. Retrieved 6 March, 2004, from http://caacentre.lboro.ac.uk/dldocs/otghdout.pdf

McKenna, C., & Hesketh, I. (2000). A review of online resources for computer-assisted assessment. *Educational Developments, 1*(2), 16-18.

McMillan, J. H. (2004). *Educational research: Fundamentals for the consumer* (4th ed.). Boston, MA: Pearson Education, Inc.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry* (6th ed.). Boston, MA: Pearson Education, Inc.

McMurty, K. (2001). *E-cheating: Combating a 21st century challenge*. Retrieved 23 March, 2004, from http://www.thejournal.com/articles/15675

McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.

Mehrens, W. A., & Lehmann. I. J. (1973). *Standardized tests in education* (2nd ed.). New York: Holt Rinehart and Winston, Inc.

Millsap, C. M. (2000). *Comparison of computer testing versus traditional paper-and-pencil testing.* Unpublished doctoral thesis, University of North Texas, Texas.

Ministry of Education Thailand. (1999). *National Education Acts of B.E. 2542 (1999)*. Retrieved 22 January, 2005, from http://www.moe.go.th/English/edu-act.htm#Chapter%209

Ministry of Education Wellington New Zealand. (1998). *Exploring formative assessment – Assessment for learning*. Retrieved 6 May, 2004, from http://www.tki.org.nz/r/assessment/atol_online/ppt/online_workshop_1.ppt#257,1,Slide1

Ministry of Education Wellington New Zealand. (1998). *Planning an English unit assessment against the curriculum*. Retrieved 3 August, 2003, from http://english.unitecnology.ac.nz/resources/units/assessment.html

Moore, D. (2004). Using technology to collect and assess non-discrete multiple choice data. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1803-1809.

Moore, N. B. (2001). *An evaluation of a computer-based assessment simulation.* Unpublished doctoral thesis, Northern Illinois University, Illinois.

Morote, E., & Pritchard, D. (2004). Technology closes the gap between students' individual skills and background differences. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 826-831.

Morris, B., & Milam, P. (2004). Instructional technology improves test scores: Implications from research. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1289-1294.

Morrison, S. (2002). *Interactive language learning on the Web*. Retrieved 4 October, 2003, from http://www.cal.org/resources/digest/0212morrison.html

Mory, E. H. (1992). The use of informational feedback in instruction: Implications for future research. *Educational Technology Research and Development, 40*(3), 5-20.

Mueller, J. (2006). *Authentic assessment toolbox*. Retrieved 19 May, 2007, from http://jonathan.mueller.faculty.noctrl.edu/toolbox/whatisit.htm

Mulligan, B. (1999, 16-17 June). *Pilot study on the impact of frequent computerized assessment on student work rates.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Murphy, E. (1997). *Characteristics of constructivist learning & teaching*. Retrieved 26 June, 2005, from http://www.cdli.ca/~elmurphy/emurphy/cle3.html

NECTEC. (1994). *History of NECTEC*. Retrieved 23 May, 2004, from http://www.nectec.or.th

NECTEC. (2006). *NECTEC, National Electronics and Computer Technology Center: Background*. Retrieved 12 March 2007, 2007, from http://www.nectec.or.th/2007/nectec-info.html

NETS. (2000). *Connecting curriculum and technology: Essential conditions to make it happen*. Retrieved 16 April, 2007, from http://cnets.iste.org/students/pdf/ess_cond.pdf

Nguyen, D. M. (2002). *Developing and evaluating the effects of web-based mathematics instruction and assessment on student achievement and attitude.* Unpublished doctoral thesis, Texas A&M University, Texas.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Northcote, M. (2002). Online assessment: friend, foe or fix? *British Journal of Educational Technology, 33*(5), 623-625.

NSTDA. (2007). *NSTDA: History*. Retrieved 28 March 2007, 2007, from http://www.nstda.or.th/en/index.php?option=com_content&task=view&id=211&Itemid=43

Nunley, L. (2001). *A national on-line survey of education faculties' use of technology in preservice teacher education courses.* Unpublished doctoral thesis, United States International University, California.

O Murchu, D. (2005). New teacher and student roles in the technology-supported, language classroom. *International Journal of Instructional Technology & Distance Learning, 2*(2), 3-10.

OECD. (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD, Organisation for Economic Co-operation and Development.

O'Hare, D. (2001, 2-3 July). *Student views of formative and summative CAA.* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Olsen, J. (2000). *Guidelines for Computer-Based Testing*. Retrieved 3 April, 2004, from http://www.isoc.org/oti/articles/0500/olsen.html

Olson, H. (1995). *Quantitative "versus" qualitative research: The wrong question*. Retrieved 23 October, 2003, from http://www.ualberta.ca/dept/slis/cais/olson.htm

Olt, M. R. (2002). *Ethics and distance education: Strategies for minimizing academic dishonesty in online assessment*. Retrieved 6 July, 2004, from http://www.westga.edu/~distance/ojdla/fall53/olt53.html

O'Malley, J. M., & Valdez Pierce, L. (1996). *Authentic assessment for English language learners: Practical approach for teachers*. Virginia, USA: Longman.

O'Neill, S., & Gish, A. (2008). *Teaching English as a second language.* South Melbourne: Oxford University Press.

Oppenheim, A. (1999). *Questionnaire design, interviewing and attitude measurement* (New ed.). London: Pinter.

Peat, M., & Franklin, S. (2002). Supporting student learning: The use of computer-based formative assessment modules. *British Journal of Educational Technology, 33*(5), 515-523.

Polyson, S., Saltzberg, S., & Godwin-Jones, R. (2002). A practical guide to teaching with the World Wide Web. *ERIC (Educational Resources Information Center) Digest*.

Pope, C., & Golub, J. (2000). Preparing tomorrow's English language arts teachers today: Principles and practices for infusing technology. *Contemporary Issues in Technology and Teacher Education [Online serial], 1*(1).

Poulton, J. (2001, 28 - 30 August). *Computer aided assessment: Web based testing with Active Server Page.* Paper presented at the LTSC-ICS 2nd Annual Conference, University of North London, London, UK.

Powers, D. E. (2001). Test anxiety and test performance: Comparing paper-based and computer-based adaptive versions of the Graduate Record Examination (GRE) general test. *Journal of Educational Computing Research, 24*(3), 249-273.

Quesada Pacheco, A. (2000). *Using the Web to practice and learn grammar: ESL student perspectives.* University of Kansas, Kansas, USA.

Richards, J. C., & Rodgers, T.S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.

Richardson, J., & Schaffer, S. (2004). Supporting technology integration across the teacher education system. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1383-1388.

Ricketts, C., & Wilks, S. (2001, 2-3 July). *Is computer-based assessment good for students?* Paper presented at the 5th International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Robles, M., & Braathen, S. (2002). Online assessment techniques. *Delta Pi Epsilon Journal, 44*(1), 39-49.

Roblyer, M. D. (2003). *Integrating educational technology into teaching* (3rd ed.). Columbus, Ohio: Merrill Prentice Hall.

Rocco, T., Bliss, L., Gallagher, S., Perez-Prado, A., Alacaci, C., Dwyer, E., Fine. J., & Pappamihiel, N. (2003). The pragmatic and dialectical lenses: Two views of mixed methods use in education. In A. T. Tashakkori, C. (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 595-615). Thousand Oaks, CA: Sage Publications.

Rodgers, T. (2001). *Language teaching methodology*. Retrieved 14 September, 2004, from http://www.cal.org/resources/digest/rodgers.html

Roever, C. (2001). Web-based language testing. *Language Learning & Technology., 5*(2), 84-94.

Roos, B. (2001, 10-13 June). *From distance education to online learning: Formative assessment in higher education.* Paper presented at the 2001 - EDEN 10th Anniversary Conference - Learning Without Limits: Developing the Next Generation of Education, Stockholm, Sweden.

Rovai, A. P. (2000). Online and traditional assessments: What is the difference? *Internet and Higher Education, 3*(3), 141-151.

Rudestam, K., & Newton, R. (2001). *Surviving your dissertation: A comprehensive guide to content and process* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Rudner, L. M., & Schafer, W. D. (2001). *Reliability. ERIC Digest.* Retrieved 18 May, 2004, from http://www.ericdigests.org/2002-2/reliability.htm

Russell, M. (2006). *Technology and assessment: The tale of two interpretations*. Greenwich: Information Age Publishing.

Russell, M. (1999). Testing on computer: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20).

Russell, M., & Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives, 5*(3).

Russo, A. (2002). Mixing technology and testing. *School Administrator, 59*(4), 6-12.

Ryan, S., Scott, B., Freeman, H., & Patel, D. (2000). *The virtual university: The Internet and resource-based learning* (1st ed.). London: British Library Cataloguing in Publication Data.

Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori, & C .Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 321-350). Thousand Oaks, CA: Sage Publications.

Satoshi, K., Shinichi, F., Seinosuke, N., & Yin, M. (2002). *A computer-based English exercise system through pictures using TTS.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA).

Saunders, G., & Pincas, A. (2004). Student attitudes towards information and communication technologies in teaching and learning in the UK. *International Journal of Instructional Technology & Distance Learning, 1*(8), 3-11.

Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. *Language Learning & Technology, 5*(2), 38-59.

Scarino, A., & Liddicoat, A. J. (2009) *Teaching and Learning Languages: A Guide*. Melbourne: Curriculum Corporation.

Schulz, R. (2006). *Language acquisition - language learning*. Retrieved 26 June, 2007, from http://www.sk.com.br/sk-laxll.html

Schulze, S. (2003). Views on the combination of quantitative and qualitative research approaches. *Progressio, 25*(2), 8-20.

Schutz, R. (2004). *Vygotsky & language acquisition*. Retrieved 24 November, 2006, from http://www.sk.com.br/sk-vygot.html

Schutz, R. (2007). *Stephen Krashen's theory of second language acquisition*. Retrieved 5 April, 2008, from http://www.sk.com.br/sk-krash.html

Shelton, K., & Saltsman, G. (2004). Tips and tricks for teaching online: How to teach like a pro! *International Journal of Instructional Technology & Distance Learning, 1*(10), 53-63.

Shuey, S. (2002). Assessing online learning in higher education. *Journal of Instruction Delivery System, 16*(2), 13-18.

Sly, L., & Rennie, L. (1999, 16-17 June). *Computer managed learning: Its use in formative as well as summative assessment.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Smith, G., Kromrey, J., Barron, A., Carey, L., Hogarty, K., & Hess, M. (2004). Assessing the pedagogical and technological quality of online courses. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1047-1054.

Sydenstricker-Neto, J. (1997). *Research design and mixed-method approach: A hands-on experience*. Retrieved 17 November, 2005, from http://www.socialresearchmethods.net/tutorial/Sydenstricker/bolsa.html

Tantasanawong, P. (2007). *Thailand Research and Education Networks (ThaiREN).* Retrieved 25 March, 2007, from http://ewh.ieee.org/r10/w_australia/main/activities/seminars/2007/IEEE-CS-Seminar-070207-1.pdf

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches* (Vol. 46). Thousand Oaks, CA: SAGE Publications.

Tashakkori, A., & Teddlie, C. (2003a). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research* (pp. 3-50). Thousand Oaks, CA: SAGE Publications.

Tashakkori, A., & Teddlie, C. (2003b). The past and future of mixed methods research: From data triangulation to mixed model designs. In A. Tashakkori, & C. Teddlie (Ed.), *Handbook of mixed methods in social & behavioral research.* (pp. 671-702). Thousand Oaks, CA: SAGE Publications.

ThaiSarn. (2005). *Thai Social/Scientific Academic and Research Network*. Retrieved 16 March, 2007, from http://thaisarn.nectec.or.th/htmlweb/index.php

The University of Western Ontario. (2002). *Teaching large classes*. Retrieved 16 October, 2004, from http://www.uwo.ca/tsc/tlc/index.html

Thongdhamachart, N. (2005). *e-Learning in Thailand*. Retrieved 12 August, 2006, from http://www.asia-elearning.net/content/conference/2005/file/Thailand_learning_aen.pdf

Toral, P. (2004). A comparative study of Hispanic students' achievement in traditional in-class courses and Internet-based courses. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1528-1535.

Toriskie, J. (1999). *The effect of Internet usage on student achievement and student attitudes.* Unpublished doctoral thesis, Loyola University of Chicago, Chicago.

Tuckman, B. W. (1988). *Testing for teachers* (2nd ed.). San Diego: Harcourt Brace Jovanovich.

Tunc, Y., & Armstead, M. (2000). *Computer based testing: The Ball State experience.* Paper presented at the 29th annual ACM SIGUCCS conference, Portland, Oregon, USA.

University of Illinois. (2007). *Strategies to minimize cheating online*. Retrieved 4 July, 2004, from http://www.ion.uillinois.edu/resources/tutorials/assessment/cheating.asp

Van Belle, G. (1997). *How cheating helps drive better instruction*. Retrieved 23 January, 2004, from http://www.plagiarized.com/vanb.html

Vendlinski, T. (2001). *Affecting United States education through assessment: New tools to discover student understanding.* Unpublished doctoral thesis, Massachusetts Institute of Technology, Massachusetts.

Vendlinski, T., & Stevens, R. (2002). Assessing student problem-solving skills with complex computer-based tasks. *The Journal of Technology, Learning, and Assessment, 1*(3).

Vonderwell, S. (2003). Student and Instructor experiences in an online technology in education course: A case study. In C. Crawford (Ed.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2003* (pp. 543-552): Chesapeake, VA: AACE.

Vygotsky, L. S. (1985). *Thought and language*. Cambridge, MA: The M.I.T. Press.

Wahlstrom, D. (2001). *Technology-mediated assessment in a secondary manufacturing technology program: A study of the relationship between participants' goal-orientation and the calibration of performance expectations to performance feedback.* Unpublished doctoral thesis, University of Michigan, Michigan.

Walker, D., Adamson, M., & Parsons, R. (2004, 6-7 July). *Staff education - learning about online assessment, online.* Paper presented at the 8th International CAA Conference, Loughborough University, Leicestershire, UK.

Walker, R., & Delius, G. (2004, 6-7 July). *Integrating on-line assessment with class-based teaching and learning: A preliminary study of the AIM marking system.* Paper presented at the 8th International CAA Conference, Loughborough University, Leicestershire, UK.

Wall, J. (2000). *Technology-delivered assessment: Guidelines for educators traveling the technology highway*. Retrieved 23 September, 2003, from http://www.ericdigests.org/2001-3/highway.htm

Wang, C. (2004). Use of technology in second-language education teaching. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 2728-2731.

Wang, C., & Wang, Y. (2004). Integrating technology into ESL learning environment: Use process writing approach as an example. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1442-1447.

Wang, F., Zhang, K., & Sun, H. (2004). Barriers and concerns of teachers using technology in their instruction: A pilot study. *Society for Information Technology and Teacher Education International Conference (SITE), 2004*(1), 1070-1074.

Wang, K. H., Wang, T.H., Wang, W. L., & Huang, S. C. (2006). Learning styles and formative assessment strategy: enhancing student achievement in Web-based learning. *Journal of Computer Assisted Learning, 22*(2), 207-217.

Wang, T. H. (2007). What strategies are effective for formative assessment in an e-learning environment? *Journal of Computer Assisted Learning, 23*(3), 171–186.

Weinberg, A. (2001). Comparison of two versions of a placement test: Paper-pencil version and computer-based version. *Canadian Modern Language Review, 57*(4), 607-627.

Whigham, D., & Houston, J. (1999, 16-17 June). *On-line delivery and marking of Excel based assessment.* Paper presented at the 3rd International Computer Assisted Assessment Conference, Loughborough University, Leicestershire, UK.

Whittington, D., Bull, J., & Danson, M. (2000, 12-17 June). *Web-based assessment: Two UK initiatives.* Paper presented at the Sixth Australian World Wide Web Conference, Rihga Colonial Club Resort, Cairns, Australia.

Whitworth, C. K. (2001). *Equivalency of paper-pencil tests and computer-administered tests.* Unpublished doctoral thesis, University of North Texas, Denton.

Widmayer, S., & Gray, H. (2000). *On-line course management systems and language teachers.* Paper presented at the CALL in the 21st Century Conference, Barcelona, Spain.

Willet, K., & Bouldin, A. (2004). Instructional design and assessment: Development and assessment of an online elective toxicology course. *American Journal of Pharmaceutical Education, 68*(3), 1-9.

Williams, J., & Cambiano, R. (2005). Online vs. face to face: Does delivery method make a difference in achievement? In *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005* (pp. 3230-3233). Vancouver, Canada.

Wilson, B., & Lowry, M. (2000). Constructivist learning on the Web. *New Directions for Adult and Continuing Education, 88*, 79-88.

Wong, C., Wong, W., & Yeung, C. (2001). Student behaviour and performance in using a Web-based assessment system. *Innovations in Education and Teaching International, 38*(4), 339-346.

Wong, T. (2003, 1-4 October). *Incorporating technology in pedagogy.* Paper presented at the First International Conference on Pedagogies and Learning: New meanings for a new millennium, Toowoomba, Queensland, Australia.

Wright, V. L. (2002). *Technology and high-stakes tests in reading: The effect of a computer-based test preparation program.* Unpublished doctoral thesis, University of South Florida, Florida.

Yanes, M. J., & Curts, J. B. (2004). A technology-based process for teaching learner-centered classroom assessment. In C. Crawford (Ed.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2004* (pp. 1536-1538). Chesapeake, VA: AACE.

Yang, S. (2004). *A case study of technology-enhanced language learning.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA) 2004, Lugano, Switzerland.

Yao, Y., Weissinger, E., & Grady, M. (2003). Faculty use of student evaluation feedback. *Practical Assessment, Research & Evaluation, 8*(21).

Yoshida, M. (2004, 30 August - 6 September). *Selecting suitable and feasible strategies for e-learning enterprises of educational institutions.* Paper presented at the 3rd Seminar-Workshop of the 7th Programme Cycle of Asia and the Pacific Programme of Educational Innovation for Development (APEID) Activities: Seminar-Workshop on e-Learning, Tokyo and Kyoto, Japan.

Zhang, W. (2003). *Doing English digital: An assessment model for a new college English curriculum in China.* Unpublished doctoral thesis, Columbia University, New York.

Zhao, J., & Yu, X. (2003). Evaluation of an on-line EFL/ESL course in China. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003*. Chesapeake, VA: AACE.

Zin, A., Darus, S., Nordin, M., & Yusoff, A. (2003). Using a coursework management system in language teaching. *Teaching English with Technology: A Journal for Teachers of English, 3*(1).

Zurawski, R. (1998). Making the most of exams: Procedures for item analysis. *The National Teaching & Learning Forum (NTLF), 7*(6).