ASSESSING AND INTERPRETING STUDENTS’ ENGLISH ORAL PROFICIENCY USING D-VOCI IN AN EFL CONTEXT

DISSERTATION

Presented in Partial Fulfillment of the Requirement for
The Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Tae-Young Jeong, M.A.

The Ohio State University
2003

Dissertation Committee:
Dr. Charles R. Hancock, Advisor
Dr. Keiko K. Samimi
Dr. Suzanne K. Damarin

Approved by

Advisor
College of Education
ABSTRACT

This study investigates the nature of multimedia-enhanced English oral proficiency interview with an emphasis on the relationship between students’ electronic literacy and their multimedia-based oral proficiency interview (OPI) scores in an EFL (English as a Foreign Language) context. A sample of 144 Korean college students participated in this study by completing The Electronic Literacy Questionnaire (ELQ) and taking both the face-to-face English interview and the multimedia-enhanced English oral proficiency interview during the 2001 and 2002 sessions of the Summer Intensive English Course.

Information about subjects’ electronic literacy and their attitudes towards the computer-enhanced OPI was gathered in the ELQ; validity and reliability were established by a panel of experts and a number of statistical analyses.

The multimedia-enhanced English Oral proficiency interview program (d-VOCI: digital-Video Oral Communication Instrument) was a collaboration of the Language Acquisition Resource Center at the San Diego State University with the English faculty of the Korea Military Academy. Following the ACTFL (American Council on the Teaching of Foreign Languages) oral proficiency guidelines, the d-VOCI assessed students’ language proficiency in terms of their ability to use the language effectively and appropriately in real-life situations (ACTFL, 1999).
The findings of the research support the argument for the multimedia-enhanced English oral proficiency test (e.g., d-VOCI) as a promising instrument for teachers and students in the EFL context to meaningfully incorporate the English speaking tests into their classrooms.

In addition, the results of the study indicated a positive moderate relationship between the electronic literacy and the English oral proficiency of the students in the study. Further research across a wider range of EFL contexts could determine the extent to which these two constructs may be related in different EFL settings.
Dedicated to my parents
ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere appreciation to many individuals who have stayed with me along the ‘long and winding’ road and made possible the completion of this dissertation.

First of all, I must say that it has been my privilege to have the opportunity to study and work with such a wonderful person as Dr. Charles Hancock. He has not only been my academic advisor, but also an outstanding counselor and boss as well as sincere friend throughout my studies at the Ohio State University. I wonder how many graduate students could have studied, worked, and even praised the Lord together with their advisors as I have.

My sincere thanks go also to the other committee members, Dr. Keiko Samimy and Dr. Suzanne Damarin. Dr. Samimy has been an excellent role model for many international graduate students in demonstrating how well they could serve and contribute to the field of Second Language Acquisition as non-native English speaking researchers. Dr. Damarin helped me tremendously with synthesizing ideas and clarifying concepts. Her professional expertise and trust motivated me to pursue my professional goal from the very beginning of my study here at OSU.

I also deeply appreciate the never-ending support from Dr. Chan Park and Dr. Galal Walker from the Department of East Asian Languages and Literatures,
from whom I have learned so much about the teaching and learning of foreign languages and cultures.

   Special thanks goes to the ROK Army and the Korea Military Academy for giving me this exciting opportunity to study abroad for this lengthy hiatus from my routine duties. I am fortunate to have received various supports — financial as well as spiritual — from these institutions throughout my stay in the States.

   Finally, my deepest appreciation goes to my family, especially my wife, Min-Hee who has been my best friend and colleague ever since we first met. I cannot sufficiently express in words my gratitude for her constant support, help, and love.
VITA

Aug 17, 1967 ....................... Born — Seoul, Korea

1990 .............................. B.A. German Language,
                        Korea Military Academy, Seoul, Korea

1994 ......................... M.A. English Language & Literature
                        Yonsei University, Seoul, Korea

1994 – 1998 ................. Full-time Instructor, English
                        Korea Military Academy, Seoul, Korea

1998 – 2001 ................ Graduate Teaching and Research
                        Associate, The Ohio State University

2001 – 2003 ................ Graduate Administrative Assistant
                        The Ohio State University

FIELDS OF STUDY

Major Field:  Education

Area of Focus: Teaching English to the Speakers of Other Languages (TESOL)

Foreign Language Testing

Minor Field:  Educational Technology

Quantitative Research, Measurement, and Evaluation
TABLE OF CONTENTS

Abstract .................................................................................................................... ii
Acknowledgments .................................................................................................... v
VITA ...................................................................................................................... vii
Table of Contents ................................................................................................ viii
List of Tables .......................................................................................................... xi
List of Figures ........................................................................................................ xii

CHAPTER 1: OVERVIEW OF THE STUDY ........................................................ 1
1.1 Introduction........................................................................................................ 1
1.2 Significance of the Study .................................................................................. 2
1.3 Purpose of the Study ......................................................................................... 4
1.4 Korean Setting of the Study ............................................................................ 4
1.5 Research Questions ......................................................................................... 6
1.6 Basic Assumptions .......................................................................................... 7
1.7 Definition of Terms ........................................................................................ 8
1.8 Organization of the Study .............................................................................. 13
1.9 Summary ......................................................................................................... 14

CHAPTER 2: REVIEW OF LITERATURE ......................................................... 15
2.1 Introduction...................................................................................................... 15
2.2 Second Language Acquisition and CALL ....................................................... 15
  2.2.1 Programmed Instruction ......................................................................... 16
  2.2.2 Information Processing Approach .......................................................... 17
  2.2.3 Socio-Constructivist Perspective ............................................................ 18
  2.2.4 Discourses of CALL, SLA, and Pedagogy: Appearance and nature .... 22
2.3 Second Language Testing from Communicative Perspectives .................... 23
  2.3.1 Communicative Language Teaching (CLT) and Alternative Assessment ............................................................................. 24
  2.3.2 Oral Proficiency Interview as Performance Assessment ................. 26
  2.3.3 Portfolio Assessment and d-VOCI ...................................................... 31
2.4 Considerations of an Electronic literacy framework in Korean context ....... 35
  2.4.1 Technology and Non-native English teachers in EFL context ......... 35
  2.4.2 Critical Pedagogy and Power Issues in Language Classrooms .......... 37
  2.4.3 Ownership, Authenticity and Autonomy in Language Pedagogy .... 40
  2.4.4 NNESTs Empowered by Technology in EFL context .................... 44
  2.4.5 Refining Literacy (Multimedia and electronic) in Korean Context ... 45
2.5 Summary ......................................................................................................... 48
CHAPTER 3: METHODOLOGY ..........................................................................................50
3.1 Introduction ...........................................................................................................50
3.2 Location and Participants ..................................................................................50
3.3 Procedures .........................................................................................................52
3.4 Materials ............................................................................................................54
3.5 Data Analysis .....................................................................................................58
3.6 Summary ............................................................................................................66

CHAPTER 4: DATA ANALYSIS AND DISCUSSION ..................................................67
4.1 Introduction ........................................................................................................67
4.2 Descriptive Statistics .........................................................................................68
    4.2.1 Demographic Characteristics of the Research Sample .......................68
    4.2.2 Electronic Literacy ..............................................................................69
    4.2.3 English Test Scores ..........................................................................71
    4.2.4 Students’ Attitudes towards the d-VOCI and the face-to-face interview ..73
4.3 Multiple Regression Analysis ..........................................................................76
    4.3.1 Explaining the Variance of d-VOCI ...............................................76
    4.3.2 Multicollinearity Analysis ..................................................................83
    4.3.3 Residual Analysis .............................................................................85
4.4 Structural Equation Modeling ..........................................................................87
    4.4.1 Fit of Measurement proposed Model .............................................87
    4.4.2 Parameter Estimates and Hypothesis Testing ..................................89
4.5 Summary ..........................................................................................................94

CHAPTER 5: FINDINGS, IMPLICATIONS, RECOMMENDATIONS FOR
FUTURE RESEARCH, AND CONCLUSION .........................................................96
5.1 Introduction ....................................................................................................96
5.2 Review of the Conceptual Framework of the Study .......................................97
5.3 Research Questions and Discussion ..............................................................100
5.4 Recommendations for Further Research .....................................................106
5.5 Implications of the Study ...............................................................................108
5.6 The Limitations of the Study .......................................................................111
5.7 Conclusion .....................................................................................................112

REFERENCES ......................................................................................................115
Appendix A: d-VOCI Evaluation Sheet (a sample page) ........................................ 127
Appendix B: Students Survey Permission Form .................................................. 128
Appendix C: Electronic literacy survey (Pilot survey) ......................................... 129
Appendix D: Electronic literacy survey (Final version) ...................................... 136
Appendix E: Script for the English d-VOCI ...................................................... 143
Appendix F: Main Screen of d-VOCI ............................................................... 145
Appendix G: d-VOCI log-in screen ................................................................. 146
Appendix H: d-VOCI interview screen ............................................................. 147
Appendix I: d-VOCI video screen .................................................................. 148
Appendix J: Summary of Reliability Estimates for Electronic Literacy scales ... 149
LIST OF TABLES

Table 1: Pedagogical Foci in Structural, Cognitive, and Sociocognitive Frameworks .....21
Table 2: Traditional Assessment vs. Alternative Assessment .................................................25
Table 3: The Hierarchy of Global Tasks ...............................................................................28
Table 4: Assessment Criteria ...............................................................................................29
Table 5: Earlier Approaches vs. Electronic Literacy Approach ..........................................47
Table 6: Questions for Face-to-face interview ....................................................................53
Table 7: Demographic Characteristics of the Research Sample .........................................68
Table 8: Descriptive Statistics and Correlations of Electronic Literacy ...............................70
Table 9: Descriptive Statistics of English Tests .....................................................................71
Table 10: Descriptive Statistics of Students’ Attitudes toward d-VOCI .................................74
Table 11: Independent Groups t-Test between students who preferred d-VOCI and those
who preferred Face-to-Face interview ............................................................................75
Table 12: Descriptive and correlations output from multiple regression .............................76
Table 13: Correlations between variables in the Multiple Regression Analysis ....................79
Table 14: Model Summary and ANOVA Results-Simultaneous Regression .......................81
Table 15: Regression Coefficients ......................................................................................84
Table 16: Fit analysis for the measurement model ..............................................................88
Table 17: Estimates of Standardized Regression Weights ....................................................91
Table 18: Model Summary (d-VOCI and TEPS sub-skills) ...............................................104
LIST OF FIGURES

Figure 1: The Changing perspective in OPI .........................................................34
Figure 2: The Conceptual Framework: Technology in EFL context ......................49
Figure 3: Hypothesized model between students’ electronic literacy and their English
proficiency with their measured predictors .......................................................62
Figure 4: Histogram of the Standardized Residuals .............................................86
Figure 5: Normal P-P Plot of Regression Standardized Residual ............................87
Figure 6: Hypothesized structural relationship between electronic literacy and English
Proficiency (Model A) ......................................................................................90
Figure 7: Hypothesized structural relationship between electronic literacy and English
Proficiency (Model B) ......................................................................................93
Figure 8: Hypothesized Conceptual Framework: d-VOCI in EFL context ...............98
CHAPTER 1: OVERVIEW OF THE STUDY

1.1 Introduction

Over the years researchers of SLA (Second Language Acquisition) and CALL (Computer-Assisted Language Learning) have sought to understand the impact of the technological revolution in society and its potential as a tool for language teaching and learning. In fact, beginning with language labs of the 1950s and 1960s, new technologies such as the microcomputers of the 1970s and 1980s and today’s digital technology-based language labs have been considered useful tools to help teachers teach and students learn. Thus, not surprisingly, the speed and intensity of implementing technology into the language classroom as well as the amount of research conducted to discuss these phenomena have continuously gained momentum in use within the second language profession.

Research has shown that the use of computer technology in foreign language assessment at different levels and for various purposes has also been steadily increased along with the development of new technology (Laurel, 1993; Murison-Bowie, 1993; Levy, 1997; Kern & Warschauer, 2000; Chapelle, 2001). For example, the use of CAT (Computer Adaptive Testing) such as TOEFL, GRE, and GMAT for proficiency/achievement testing purpose has become an increasingly appealing method of assessment in the field of education both in general and in second/foreign language
education. Among many other types of computerized language assessments, multimedia enhanced oral proficiency tests such as the COPI (computerized oral proficiency instrument) developed by CAI (Center for Applied Linguistics) and d-VOCI (digital Video Oral Communication Interview) by the Language Acquisition Resource Center at the San Diego State University, employ authentic tasks to elicit speech ratable on the ACTFL (American Council on the Teaching of Foreign Languages) proficiency guidelines.

The purpose of the COPI/d-VOCI is to use the advantage of multi-media computer technology to improve the previous face-to-face OPI (Oral Proficiency Interview) by giving interviewees more control over various aspects of the testing situation and increasing the rater’s efficiency in scoring the language samples.

Recently, several researchers have asserted that a computerized language testing development project should inspire more teachers and researchers to begin thinking about using and developing computerized testing for their own institution’s assessment purposes, where the participants are concerned about developing valid, reliable, and useful language assessment instruments (Chalhoub-Deville, 1999; Dunkel, 1999; Chapelle, 2001).

1.2 Significance of the Study

The administration and scoring programs of the COPI or d-VOCI are relatively new concepts, lacking sufficient research on their reliability, validity, and practicality, in comparison with other CALT (Computer-Assisted Language Testing) instruments.
designed to measure reading and listening skills, and the traditional face-to-face interview method. Nevertheless, these multimedia-enhanced language testing tools continue to draw much interest from language professionals as well as language students, especially in the EFL context where the population of qualified or certified OPI raters is quite small, whereas the students’ interest in knowing their own oral proficiency is tremendously high. Another issue with regard to English speaking tests in the EFL context is lack of confidence of teachers who are not native speakers of English in public school settings. The results of this research show that multimedia-enhanced English oral proficiency test using d-VOCI can build a promising system where the collaboration between the teachers from both sides, native and non-native speakers of English, can be meaningfully employed.

In addition, few studies have been conducted to investigate various aspects of the technology-enhanced oral proficiency test; for instance, the relationship between students’ electronic/media literacy and their oral proficiency, or the reliability between the technology-enhanced oral proficiency test and the traditional face-to-face interview. Therefore, this research provides useful information not only to those who are interested in implementing multimedia enhanced oral proficiency interviews into their school curriculum or teacher education program, but also to those concerned with the estimation and interpretation of computerized oral proficiency scores in a foreign language education setting.
1.3 Purpose of the Study

The purpose of this study was to investigate the nature of the multimedia enhanced English oral proficiency interview with an emphasis on the relationship between students’ electronic literacy and their multimedia based OPI scores in an EFL context.

More specifically, this study is focused on assessing the applicability of the multimedia-based English oral proficiency interview in a Korean college classroom environment. Since the multimedia-enhanced English oral proficiency test in Korean context is a new method for language testing, and has been hardly studied, the characteristic of this research is rather exploratory.

To accomplish its purpose, this study mainly investigated 1) the structural relationship between two latent constructs (electronic literacy and English proficiency) using the statistical analysis with seven manifest variables (electronic communication, cyberspace construction, academic research, d-VOCI, TEPS, face-to-face interview, and self-assessment) designed to measure students’ electronic literacy and English proficiency, and 2) the results of the survey designed to measure students’ electronic literacy as well as their attitudes towards the computerized format in English oral proficiency interviews. Finally, the research discussed pedagogical and other implications.

1.4 Korean Setting of the Study
This research was conducted in a Korean college setting, and readers need to understand the context of the Korean Information Technology industry as well as the status of English in current Korean society.

For the past 20 years, the Korean government has been pursuing a policy of building an information-based society, in the belief that doing so would be essential for the advancement of the nation’s competitive edge and the improvement of the standard of living and education. As of 2001, for example, Internet access is being provided for all 10,400 grade schools in the nation, free of charge. According to the Korean government (2001), Korean citizens tend to have an enthusiasm for education and enjoy one of the world’s highest literacy ratios, and seem to have enhanced capabilities to command good computer skills, which have contributed to narrowing the digital divide between the people from different socio-economic backgrounds and among different regions in Korea.

Another key aspect of the Korean context is the status of English as a foreign language. As Crystal pointed out (1997), English is the medium of a great deal of the world’s knowledge, especially in such areas as business, science, and technology. So, it is understandable that several nations, including Korea in particular, have in recent years made English an official language or chosen it as their chief foreign language in schools. As an example, the Ministry of Education in Korea recently mandated that all English classes in secondary schools in Korea should be conducted in English by the end of 2002 (Korean Government, 2001). According to a survey conducted by a congressman in Korea (Lee, 2000), however, the majority of active English teachers in secondary schools (95%) answered that they have either little or no confidence in conducting class in English while most of them acknowledged the importance of L2 use in their English
classroom. In later chapters, this issue is further elaborated and several suggestions are provided in Chapter 5.

1.5 Research Questions

This study investigated the research questions listed below; they are categorized as primary and secondary questions. Primary questions address the estimation and interpretation of students’ d-VOCI scores. Secondary questions focus on the attitudes of students toward d-VOCI. The last research question is to examine the implications of this study for EFL context.

Primary research questions:

1. What is the structural relationship between students’ electronic literacy and English oral proficiency?

2. Which variable(s) of this study best explain(s) the variance of the students’ English oral proficiency?

3. Is there a statistically significant difference in results between the computerized format and the conventional face-to-face format in English oral proficiency interviews?

4. What is the relationship between the four constructs (READING, LISTENING, GRAMMAR, VOCABULARY) measured by the TEPS and d-VOCI scores of the students?
Secondary questions:

5. What patterns of student views about the electronic oral assessment format were reported by participants?

6. To what extent do electronic literacy and English proficiency influence students’ attitudes toward d-VOCI and conventional methods?

7. What magnitude of difference characterizes student’ attitudes toward d-VOCI versus face-to-face English interviews?

8. What is the implication of this study in EFL context?

1.6 Basic Assumptions

For readers to better understand the nature of this study, the following assumptions were developed and are presented below:

1. The use of technology in testing in second/foreign language instruction is increasing. The L2 researchers, practitioners, and experts in testing organizations need an empirical base for designing, developing and using computerized placement and achievement tests. Thus, research into the use of d-VOCI is needed.

2. In d-VOCI format, various questions can be delivered in a way that is impossible via a face-to-face format, thus requiring new concepts about English oral proficiency interview administration and its implication.

3. The subjects’ English oral proficiency scores rated by qualified interviewers and the results of their TEPS were assumed to be appropriate measures to identify the
English proficiency level of the subjects. It is further assumed that a comparison of these scores was appropriate despite their inherent differences.

4. Finally, students’ responses on questionnaires and in interviews were assumed to be honest representations of their views.

1.7 Definition of Terms

The following terms are defined in two ways: a constitutive definition and an operational definition of selected terms as used in the present study.

**Attitude scale**

Constitutive definition

A measure of the degree of favorableness or unfavorableness a subject has toward a group, institution, construct, or object.

Operational definition

A 60-item questionnaire designed by the researcher based on the literature to measure subjects’ attitudes toward d-VOCI and face-to-face formats of oral proficiency interview.

**COPI (Computer Oral Proficiency Interview) and d-VOCI (digital Video Oral Communication Interview)**

Constitutive definition
Both d-VOCI and COPI use digital video delivery on the computer as the means for testing oral proficiency. The test-taker’s oral response is also recorded on the computer digitally, and may then be sent via e-mail to an evaluator at some other remote location.

Operational definition
A multimedia-enhanced oral proficiency interview program developed by a team of both the Language Acquisition Resource Center at the San Diego State University and the faculty of the Korea Military Academy. First, the Macromedia Director template was designed by LARC staffs, and the context-bound scenario, actual scripts, as well as digital video/audio files for the questionnaire were created by the faculty of the KMA.

**Electronic literacy**

Constitutive definition
The knowledge and skill required to make marks in an electronic age with electronic devices. Such knowledge and skill generally includes alphabetic literacy as well as at least the rudimentary grasp of a computer’s interface and some specialized knowledge for communicating, constructing, and conducting research electronically (Shetzer & Warschauer, 2000).

Operational definition
Quantified survey results from a questionnaire with 45 items developed to measure subjects’ electronic literacy in terms of three factors: communication, cyberspace construction, and academic problem-solving research.
Face-to-face OPI (Oral Proficiency Interview)

Constitutive definition
A standardized procedure for the global assessment of functional speaking ability. OPI measures language production “holistically” by determining patterns of strengths and weaknesses. It also establishes a speaker’s level of consistent functional ability as well as the clear upper limitations of that ability (ACTFL, 1999).

Operational definition
A type of language face-to-face interview assessment employed by the faculty of the Korea Military Academy to evaluate students’ oral fluency at the end of the Summer Intensive English Course. Two instructors decide the level of students’ English oral proficiency via a ten-minute face-to-face interview. Two students take the interview at the same time, and each pair is interviewed for 10 minutes. The average of the two grades given by each instructor will serve as the final grade.

Language Proficiency

Constitutive definition
Language proficiency is a coherent orchestration of discrete elements, such as vocabulary, discourse structure and gestures, to communicate meaning in a specific context (Canales, 1994).

Operational definition
An average of students’ scores from four different types of language assessments: 1) TEPS, 2) face-to-face interview, 3) d-VOCI, and 4) self-assessment.

**Reliability**

**Constitutive definition**

The extent to which a score or measure is free of measurement error. Theoretically, reliability is the ratio of true score variance to observed score variance. The ratio can be estimated using a variety of correlational methods, including coefficient alpha, KR20, test-retest, and parallel forms. Types of reliability includes inter-rater reliability, test-retest reliability, and internal consistency reliability.

**Operational definition**

1) Index of the consistency between one or more ratings made by separate raters. It is indexed by the correlation between the ratings of two or more raters. In this study, two inter-rater correlations were calculated to establish the inter-rater reliabilities of the d-VOCI and the conventional face-to-face interview at KMA.

2) Index of the consistency of students’ answers in their attitudes survey.

**TEPS (Test of English Proficiency by Seoul National University)**

**Constitutional definition**

The TEPS is the first English proficiency test that was developed in Korea by a Korean institution (Seoul National University). It is argued that the TEPS should fit the needs of Korean learners because the test developers know the problems of Korean students better.
than anyone else from different context. The four areas measured by TEPS are the
listening comprehension, the grammar, the vocabulary, and the reading comprehension.
The TEPS is designed to test proficiency in "real-life" English in a refreshingly
innovative way.

Operational definition

Students’ sores from three sets of the TEPS. Each set has the minimum score of 10 and
maximum of 990. There are five variables for TEPS score: TEPS_R (reading
comprehension), TEPS_L (listening comprehension), TEPS_G (grammar), TEPS_V
(vocabulary), and finally TEPS_T (total score of the four sub tests).

Validity

Constitutive definition

The extent to which a test measures the quality it purports to measure. Types of validity
include content validity, criterion validity, and construct validity.

Operational definition

1) The validity of the d-VOCl is established both by the panel of experts from the
   English Department of the KMA (content/criterion validity) and by the
   comparison with the students’ scores from TEPS test (construct validity).
2) The validity of the survey instrument is established by the panel of experts from
   the English Department of the KMA.
1.8 Organization of the Study

Chapters 1 and 2 introduce the relevant literature of the theoretical background of second language testing and CALL. Chapter 3 describes the methodological procedures used to collect and analyze data to answer the research questions presented in Chapter 1, including location/participants, procedure employed for data collection, materials, and statistical methods for data analyses. Chapter 4 begins with descriptive statistics of the variables in the study including demographic characteristics of the sample, mean score, standard deviations, and normality of the data. Then, the results of the multiple regression analysis are introduced with emphasis on the variance of the students’ English oral proficiency score measured by d-VOCI. Finally, the results from the structural equation modeling are presented. Chapter 5 concludes this study with the discussion of the findings related to research questions, implications, and further research to be studied. This chapter provides a discussion of the study based upon the data reported in Chapter 4. First, the conceptual framework in this study is briefly reviewed, and then the findings of this study are explained and summarized based on the research questions of the study. Then, different types of implications of the study to the field of English education in the EFL context are identified and described. The chapter continues with suggestions for future research projects, followed by major conclusions drawn from the study.
1.9 Summary

Chapter 1 introduced the reader to the content and context of this study. Beginning by the significance of the study, the chapter addressed the purpose, and the Korean setting of the study. As introduced earlier, the main focus of this study is to examine the applicability of the multimedia-based English oral proficiency interview in a Korean college classroom environment. The second section of this chapter provided research questions, basic assumptions of the study followed by definition of key terms and the organization of the dissertation.
CHAPTER 2: REVIEW OF LITERATURE

2.1 Introduction

This chapter provides a review of the relevant literature of the theoretical background of second language testing and CALL (computer assisted language learning), including the descriptions of the following areas in language teaching: 1) second language acquisition and CALL, 2) second language testing from communicative perspectives, 3) electronic literacy and English teaching in EFL context, and 4) issues in technology-enhanced language testing.

2.2 Second language acquisition and CALL

Traditionally, language teaching emphasized the formal analysis of the system of structures that constitute a given language. For example, the grammar-translation method trained students to memorize verb paradigms, apply prescriptive rules, parse sentences, and translate texts. American structural linguist’s work from the 1920s through the 1950s influenced the development of various methods of language instruction, (e.g., Bloomfield, 1933), culminating in the audiolingual method of the 1940s and 1950s. These historical and interdisciplinary perspectives will be provided in more detail to help us provide a
context for both SLA and CALL, thus allowing the identification of topics and issues as well as meaningful linkages between those two discourses.

**2.2.1 Programmed instruction**

In the 1950s and early 1960s empiricist theory was predominant in language teaching, a theory described by Stern (1983) as “pedagogically audiolingualism, psychologically behaviorism, linguistically structuralism” (p.169, italic added). The principles emanating from these three schools of thought were mutually supportive when applied to language teaching and learning (Ellis, 1997; Levy, 1997). The audiolingual approach to language teaching and learning emphasized use of the target language in spoken form and students were expected to learn the language through a process of habit-formation, that is, endless practice. The teacher presented new vocabulary and structures through model dialogues which students learned through imitation and repetition.

At the time, the key concept for both CALL and SLA was programmed instruction (Levy, 1997). Skinner (1954), for example, advocated the use teaching and testing machines for individualized instruction which would be responsive to the preferred pace of the learner. Underlying principles included the use of instructional steps or “frames,” and active responses from the learner which would be followed by immediate feedback. In other words, the emphasis in speaking, reading, and writing was on the achieved linguistic product, not on cognitive or social processes (Kern and Warschauer, 2000). To understand the relationship between SLA and CALL, it is important to begin with this era because the programmed instruction explained above directly influenced the grammatical sequencing that was very much in evidence in early CALL. Language teaching and
testing software developers realized that drill and practice exercises advocated in the audiolingual approach were readily programmable on the computer because of their “systematic and routine character” and “their lack of open-endedness” (Kenning and Kenning 1983, p. 53). In contrast to behaviorism, learning in cognitive theories is assumed to result from the mental activity in the human mind rather than from the learner’s reaction to the external environment.

2.2.2 Information processing approach

According to McLaughlin (1990), cognitive approaches focus on the mental processes in acquisition and use of knowledge. In addition, in these views, learners are said to voluntarily act, construct, and plan for building their knowledge beyond the simple reception of stimuli from the external environment.

One of the most widely accepted variables influencing learning in cognitive approaches has been prior knowledge. Prior knowledge refers to all the knowledge that an individual has previously acquired. The information processing approach is the common theoretical approach used by cognitive psychologists; it is not a single unified theory, but rather a broad approach to the problems of understanding human knowledge and action (Best, 1989). McLaughlin (1990), in arguing for a cognitive psychological approach to second language teaching, emphasizes “automaticity and restructuring” in SLA. He maintained that practice improves performance because lower-level skills become automated; however, he also says that increased practice can also cause a temporary reduction in performance as learners reorganize their schema.
With their “dynamic” rather than “fixed” views of intelligence information processing or cognitive perspectives have been considered by a number of authors in relation to CALL. Cook (1985) discussed the information processing approach as a suitable model of language teaching and SLA appropriate for CALL activities, and Warschauer (2000) also argues that this cognitive, constructive generation of CALL was a significant advance over earlier programmed instruction (tutorial and drill programs).

Along with the changing perspective of educational theory, there were significant changes in the field of SLA and CALL, too; the birth of communicative language teaching (CLT) and the invention of the microcomputer. CLT is an approach rather than a method (Richards and Lockhart, 1994), which aims to make communicative competence the goal of language teaching and to develop procedures for the teaching of the four language skills that acknowledge the interdependence of language and communication. With the widespread availability of cheap microcomputers and humanistic methods in language teaching and learning such as CLT and Total Physical Response, as Levy (1997) argues, “the interest in CALL grew dramatically and much software was produced but without a unified theory supporting its structure and content” (p. 22).

Literature regarding CLT and language testing are elaborated later in Chapter 2.3.

### 2.2.3 Socio-constructivist perspective

Currently, the most productive direction in SLA is toward the socio-constructivist approach (McGroarty, 1998). As some constructivists have also attended to the role of the social systemic environment in which learners build their knowledge (Kramsch, 1992), socio-constructivist theories have been distinguished from the simple constructivist view.
Socio-constructivist educators agree that objectivism has become inadequate to meet educational needs. Jonassen (1991) strongly argued that behaviorist and most cognitive theories have been nurtured by the Western cultural beliefs and limited by a restrictive set of Western traditional philosophical assumptions. Objectivism holds that “the world is completely and correctly structured in terms of entities, properties, and relations” (Jonassen, 1991, p.8), and that knowledge is stable, staying independent of the individual because the essential properties of objects are knowable and relatively unchanging.

Constructivism, however, is generally considered as a viewpoint in learning theory which holds that a person actively and subjectively constructs his/her own ways of thinking as a result of innate capacities interesting with his/her own experience (Molenda, 1991). It grew from cognitive theories, but constructivists reject the objectivist tradition (Jonassen, 1991).

According to the socio-constructivist view, knowledge is a “social product.” This view includes the premise that knowledge building is significantly influenced by the process of negotiation within the communities of learners. Consequently, knowledge is created by a socially shared experience rather than an individually formed experience (Prawat & Floden, 1994). Vygotsky (1978), the primary figure in this school of thought, proposed that the human beings are products not only of biology, but also of their human cultures. According to him, intellectual functioning is the product of our social history, and language is the key mode by which we learn our cultures and through which we organize our verbal thinking and regulate our actions.

His ZPD (Zone of Proximal Development) made meaningful implications for the cooperative CALL situation. According to Vygotsky (1978), ZPD is “the distance
between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (p. 86). In ZPD framework, language users should have opportunities to communicate with other users by negotiating meaning. To meet this Vygotskian condition, the cooperative CALL can introduce the authentic communication situations for negotiation and interactivity. This authentic context motivates computer users “to learn language, learn about language, and through language” (Warschauer, 1997, p. 471). In his later article (2000), Warschauer still maintains his assumption that the principal role of CALL in current socio-constructivist perspective is to provide alternative contexts for social interaction, and to facilitate access to existing discourse communities and the creation of new ones.

Table 1 (Kern & Warschauer, 2000, pp. 6-7) insightfully summarizes the instructional foci associated with structural (programmed instruction), cognitive, and sociocognitive (socioconstructivist) approaches to language teaching.
<table>
<thead>
<tr>
<th>Who are some key scholars?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leonard Bloomfield, Charles Fries, Robert Lado</td>
<td>Noam Chomsky, Stephen Krashen</td>
<td>Dell Hymes, M.A.K. Halliday</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How is language viewed?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>As autonomous structural system.</td>
<td>As a mentally constructed system.</td>
<td>As a social and cognitive phenomenon.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How is language understood to develop?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Through transmission from competent users. Internalization of structures and habits through repetition and corrective feedback.</td>
<td>Through the operation of innate cognitive heuristics on language input.</td>
<td>Through social interaction and assimilation of others’ speech.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What should be fostered in students?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mastery of a prescriptive norm, imitation of modeled discourse, with minimal errors.</td>
<td>Ongoing development of their interlanguage. Ability to realize their individual communicative purposes.</td>
<td>Attention to form (including genre, register, and style variation) in contexts of real language use.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How is instruction oriented?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Toward well-formed language products (spoken or written). Focus on mastery of discrete skills.</td>
<td>Toward cognitive processes involved in the learning and use of language. Focus on development of strategies for communication and learning.</td>
<td>Toward negotiation of meaning through collaborative interaction with others. Creating a discourse community with authentic communicative tasks.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What is the primary unit of analysis?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Isolated sentences.</td>
<td>Sentences as well as connected discourse.</td>
<td>Stretches of connected discourse.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How are language texts (spoken or written) primarily treated?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>As displays of vocabulary and grammar structures to be emulated.</td>
<td>Either as “input” for unconscious processing or as objects of problem-solving and hypothesis testing.</td>
<td>As communicative acts (“doing things with words”).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Where is meaning located?</th>
<th>Structural</th>
<th>Cognitive</th>
<th>Sociocognitive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In utterances and texts (to be extracted by listener or reader).</td>
<td>In the mind of the learner (through activation of existing knowledge).</td>
<td>In the interaction between interlocutors, writers and readers; constrained by interpretive rules of the relevant discourse community.</td>
</tr>
</tbody>
</table>

Table 1: Pedagogical Foci in Structural, Cognitive, and Sociocognitive Frameworks
(Kern & Warschauer, 2000, pp. 6-7)
Literature regarding CALL and SLA also showed some contradictions between the appearance and the nature of the field both in terms of where teachers and students are situated and to what extent their voice influence the linking theories from SLA and CALL. In the next chapter, this contradiction is elaborated.

2.2.4 Discourses of CALL, SLA, and Pedagogy: Appearance and nature

Although the apparent goal of many methodologists is to prepare students to communicate in the foreign language, the means to this alleged goal is highly variable. The notion of communicative competence for language and the communicative approach is open to individual interpretation and variation (Levy, 1997; Warschauer 2000). Practitioners differ in the way in which this philosophy or approach translates into teacher and learner roles, materials, classroom activities and techniques. As a result, instead of the set of optimal prescriptions for language teaching that was used in the 1970s and 1980s, contemporary views are more circumspect and they reflect the complexity of language teaching and learning when a number of interrelated competencies need to be acquired. For CALL, this means that an author with a CLT approach may in fact develop and use numerous kinds of CALL activities which all fit the general orientation, but which may in themselves vary significantly. Thus, a set of CALL materials labeled “communicative” has little meaning at the implementation or procedural level, and is of limited assistance in the precise definition of the interactions taking place or the distinctions between CALL programs. In the field of foreign language education, in particular, this phenomenon has been frequently observed: for instance, a communicative CALL material developed in ESL context has less meaningful value in EFL context, and
CLT oriented textbooks designed to convey a specific target culture might lack authentic feature that students in EFL context intend to learn. Therefore, it is quite imperative that a CALL material meets the needs from both students and teachers in a specific context, where the role of the technology is mainly to bridge the gap between what it is and what it should be in terms of empowering both parties meaningfully. It is also important that, when a new technology is introduced into language classroom, there should be a close examination as to how much freedom and opportunity it provides teachers and students for them to better use for learning purpose in their own context, not just follow the fixed direction.

The instrument employed for this research (d-VOCI), in that sense, is believed to effectively support teachers since it is a well-designed template with much flexibility. It not only offers an opportunity for EFL teachers to create an authentic language assessment environment, but also presents a room for their own research and reflection.

2.3 Second language testing from communicative perspectives

For a long time, tests have served the interests of psychometricians and policy makers and not those of students and teachers (Wiggins, 1994). According to Hancock (1994), these traditional tests are justified on the basis of accountability to answer whether students are learning what they are supposed to be learning. As a result, it was not the teacher-student community who benefited from the tests but decision-makers from the school administration, business, and politics, whose interests were gathering objective evidence to make judgments as gatekeepers of various communities. Since tests have been used to make such decisions about students, and also been playing powerful role
within the educational system, many teachers are forced to teach to the tests, not to mention that students’ own goals have been getting a good grade/score/result from the tests. In the field of language testing, particularly, the traditional view of language ability was one derived largely from a structuralist linguistics view that saw language as being composed of discrete components and skills (Bachman & Cohen, 1998).

The view of language ability as communicative competence currently accepted by many SLA and Language Testing (LT) professionals, however, represents a major paradigm shift from the traditional “skills and components” trait view toward more “multi-dimensional” (Canal & Swain, 1980) and “performance-based” testing (Hancock, 1994; Valette, 1994). In line with the changing view of language ability, many researchers and practitioners also began to emphasize on the process of language learning rather than the product. To sum up, the role of tests as mere “decision making tools” has been changed to more complex and interactive process reflecting, thus requiring a new approach reflecting this change.

2.3.1 Communicative Language Teaching (CLT) and Alternative Assessment

Influenced by the changing perspectives of language learning and testing, alternative assessment basically assumes that the goal of testing should be to support the improvement of learning and teaching. Tests designed to improve learning and teaching can be effective if they are connected to the whole learning system and not used in isolation (Shohamy, 1990). Connecting testing and learning in the classroom can occur in three main areas: 1) integrating assessment with teaching; 2) involving students and teachers in the assessment process, and 3) using multiple assessment sources, not just
tests (Shohamy, 1990). Also, an important component of this approach (integrating assessment with teaching) is that the process of language learning is continuously recorded.

As noted before, to connect language testing and teaching in a more meaningful way that reflects changing perspectives of language ability in the communicative approach and to focus on the process, we need to understand the essence of alternative assessment in contrast to traditional assessments. Bailey’s table (1998) below provides concise yet primary ideas as to how the alternative assessments differ from traditional formats.

<table>
<thead>
<tr>
<th>Traditional</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-shot tests</td>
<td>Continuous, longitudinal assessment</td>
</tr>
<tr>
<td>Indirect tests</td>
<td>Direct tests</td>
</tr>
<tr>
<td>Inauthentic tests</td>
<td>Authentic assessment</td>
</tr>
<tr>
<td>Individual projects</td>
<td>Group projects</td>
</tr>
<tr>
<td>No feedback provided to learners</td>
<td>Feedback provided to learners</td>
</tr>
<tr>
<td>Speeded exams</td>
<td>Untimed exams</td>
</tr>
<tr>
<td>Decontextualized test takes</td>
<td>Contextualized test tasks</td>
</tr>
<tr>
<td>Norm-referenced score interpretation</td>
<td>Criterion-referenced score interpretation</td>
</tr>
<tr>
<td>Standardization</td>
<td>Classroom-based tests</td>
</tr>
</tbody>
</table>

Table 2: Traditional Assessment vs. Alternative Assessment (Bailey, 1998, p. 45)

Since CLT focuses on “what students can actually do?” rather than “what students know?” teachers should introduce this framework when assessing students’ language competence.
Hancock (1994) explained that assessment is an on-going evaluation or check by a teacher to determine how a learner is progressing, often for the purpose of making decisions about what instruction is needed by the learners. However, since classroom tests should be used as an interactive process between the testers and the students, the teacher should provide the students with feedback on their knowledge and problems so they can improve, and also the students are required to provide the teacher with feedback on the quality of the teaching as well as of the tests. An example of alternative assessment that reflects the changing perspectives in SLA and LT might be the oral proficiency interview as a portfolio assessment.

2.3.2 Oral Proficiency Interview as performance assessment

It is obvious that speaking samples of students that have been carefully gathered and analyzed over the years can be better than any other records to show students’ progress and achievement in their language acquisition process. The question, then, is how to gather the speech samples that are valid, reliable as well as recognized by other parties outside the classroom. In line with this, many researchers suggested that the ACTFL OPI is an interactive and learner-centered assessment that is also valid and reliable. While some scholars including Bachman (1998) and Halleck (1992) criticized the validity and reliability of the ACTFL OPI, the test in general has been considered a useful tool for the efficient elicitation of a ratable sample of the student (Clark & Clifford, 1988; Dandonoli & Henning, 1990; Kuo & Jiang, 1997).

As briefly introduced earlier, the ACTFL Oral Proficiency Interview (OPI) is a standardized procedure for the global assessment of functional speaking ability. The OPI
assesses language proficiency in terms of the ability to use the language effectively and appropriately in real-life situations (ACTFL, 1999, p.1).

The ACTFL OPI rating scale is based on the hierarchy of global tasks comprising four major levels of language performance. Table 3 and 4 show the hierarchy of global tasks and assessment criteria of ACTFL OPI, respectively. The hierarchy of global tasks in Table 3 can be summarized in a rating scale spanning a wide range of performance profiles; for instance, 1) novice learners can communicate minimally with formulaic and rote utterance, lists and phrases, 2) intermediate learners can create with language, ask and answer simple questions on familiar topics, and handle a simple situation or transaction, 3) advanced learners can narrate and describe in all major time frames and handle a situation with a complication, and finally 4) superior learner can support opinion, hypothesize, discuss topics concretely and abstractly, and handle a linguistically unfamiliar situation (ACTFL, 1999).
<table>
<thead>
<tr>
<th>Level</th>
<th>Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speakers at the Superior level</strong></td>
<td>participant fully and effectively in conversations in formal and informal settings on topics related to practical needs and areas of professional and/or scholarly interests; provide a structured argument to explain and defend opinions and develop effective hypotheses within extended discourse; discuss topics concretely and abstractly; deal with a linguistically unfamiliar situation; maintain a high degree of linguistic accuracy (no patterns of errors); satisfy the linguistic demands of professional and/or scholarly life.</td>
</tr>
<tr>
<td><strong>Advanced level</strong></td>
<td>participate actively in conversations in most informal and some formal settings on topics of personal and public interest dealt with concretely; narrate and describe in major time frames with good control of aspect; deal effectively with unanticipated complications through a variety of communicative devices; sustain communication by using with suitable accuracy and confidence, connected discourse of paragraph length and substance; satisfy the demands of work and/or school situations.</td>
</tr>
<tr>
<td><strong>Intermediate level</strong></td>
<td>participate in simple, direct conversations on generally predictable topics related to daily activities and personal environment; obtain and gibe information by asking and answering questions; initiate, sustain, and bring to a close a number of basic, uncomplicated communicative exchanges, often in a reactive mode; create with the language and communicate personal meaning to sympathetic interlocutors by combining language elements in discrete sentences and strings of sentences; satisfy simple personal needs and social demands to survive in the target language culture.</td>
</tr>
<tr>
<td><strong>Novice level</strong></td>
<td>respond to simple questions on the most common features of daily life; convey minimal meaning to interlocutors experienced with dealing with foreigners by using isolated words, lists of words, memorized phrases and some personalized recombinations of words and phrases; satisfy a very limited number of immediate needs.</td>
</tr>
</tbody>
</table>

Table 3: The Hierarchy of Global Tasks (ACTFL, 1999, p. 10)
<table>
<thead>
<tr>
<th>Proficiency Level</th>
<th>Global tasks and functions</th>
<th>Context/Content</th>
<th>Accuracy</th>
<th>Text Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior</td>
<td>Discuss topics extensively, support opinions and hypothesize. Deal with a linguistically unfamiliar situation.</td>
<td>Most formal and informal settings/ Wide range of general interest topics and some special fields of interest and expertise.</td>
<td>No pattern of errors in basic structures. Errors virtually never interfere with communication or distract the native speaker from the message.</td>
<td>Extended discourse</td>
</tr>
<tr>
<td>Advanced</td>
<td>Narrate and describe in major time frames and deal effectively with an unanticipated complication.</td>
<td>Most informal and some formal settings/Topics of personal and general interest.</td>
<td>Understood without difficulty by speakers unaccustomed to dealing with non-native speakers.</td>
<td>Paragraphs</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Create with language initiate, maintain, and bring to close simple conversations by asking and responding to simple questions.</td>
<td>Some informal settings and a limited number of transactional situations/Predictable, familiar, topics related to daily activities.</td>
<td>Understood, with some repetition, by speakers accustomed to dealing with non-native speakers.</td>
<td>Discrete sentences</td>
</tr>
<tr>
<td>Novice</td>
<td>Communicate minimally with formulaic and rote utterance, lists and phrases.</td>
<td>Most common informal settings/Most common aspects of daily life.</td>
<td>May be difficult to understand, even for speakers accustomed to dealing with non-native speakers.</td>
<td>Individual words and phrases</td>
</tr>
</tbody>
</table>

Table 4: Assessment Criteria (ACTFL, 1999, p. 31)

While the ACTFL OPI guidelines are specific about the conversation format/content used to obtain ratable speech samples as shown in Table 3 and 4, there are several research criticizing this procedure. For example, Salaberry (2000) argued that the type of linguistic dialogue portrayed in a typical OPI interview is not typical of a natural conversation. He also pointed out that one of the clearest weaknesses of the OPI test is the fact that a limited range of tasks is assessed even though there exist other communicative
discourses including small-group discussion, playing a game, or conducting a survey (Salaberry, 2000). Another critique came from Lowe (1986) who emphasized that the ACTFL guidelines address performance in general rather than job-specific language.

At the end of this chapter, the general and standardized nature of the ACTFL OPI is contrasted with the contextualized and task-oriented characteristic of the multimedia-enhanced English oral proficiency test.

SOPI (Simulated Oral Proficiency Interview)

Due to the limited availability of certified/qualified ACTFL OPI raters, and other constraints such as time, cost, and location, it is common that universities and school systems administer SOPI, rather than traditional face-to-face OPI for their placement tests as well as other language requirements. The simulated oral proficiency interview, which is another type of performance-based speaking test, is basically tape-mediated, and it relies on audiotaped instructions and a test booklet to elicit language samples from the examinee. Several research projects focus on work done by different testers and languages. However, in general, they support the view that SOPI can be a valid and reliable surrogate to the OPI. Clark and LI’s study in 1986, for example, showed that there is a strong positive relationship between SOPI and OPI ($r=0.93$). Another study that supported the strong correlation between SOPI and OPI was conducted by Stansfield et al. (1990), where they developed three forms of a SOPI in Portuguese. Interestingly, this study shows that not only is there a positive strong correlation between SOPI and OPI ($r=0.93$), but also that the SOPI was slightly more reliable and easier to rate than the OPI. Other studies conducted to investigate the relationship between SOPI and OPI are
research by Shohamy, Gordon, Kenyon, and Stansfield (1989) Stansfield (1992), and Kenyon (1993), all of which support a strong correlation between the two tests.

Subsequently, Stansfield (1996) argued that the SOPI format has many practical benefits because any teacher, language lab technician or aide can administer the SOPI and the SOPI has proved to be an advantage in locations where a trained interviewer is not available or in languages that lack ACTFL-certified testers. In addition, the SOPI can be administered simultaneously to a group of examinees by one administrator, whereas a live interview can only be administered individually.

In the next chapter, the latest form of OPI (d-VOCI) in language classroom is introduced in conjunction with its role as a portfolio of language students.

2.3.3 Portfolio assessment and d-VOCI

A portfolio is a collection of documents which provide evidence of the language level of the learners (Shohamy, 1990; Valette, 1994), and portfolio assessment is an ongoing process involving the student and teacher in selecting samples of student work for inclusion in a collection, the main purpose of which is to show the student’s progress (Hancock, 1994; Hammadou, 1996). According to Hancock (1994), the greatest overall benefit of using portfolio assessment is that the students are taught by example to become independent thinkers, and the development of their autonomy as learners is facilitated. Regarding the key decisions for portfolio assessment, Hammadou (1996) suggested that teachers need to answer what will be the main purpose of the portfolio? Who owns the portfolio? How will the portfolio be used? And who is the intended audience of the portfolio? While there could be many other issues than what is raised by Hammadou,
these issues, in general, can be grouped into three main categories: validity, reliability, and practicality of portfolio assessment.

The validity of a test is related how well the test does what it is supposed to do, namely, to inform us about the examinee’s progress toward some goal in a curriculum or course of study, or to differentiate levels of ability among various examinees on some task. The reliability of a test is a matter of how consistently it produces similar results on different occasions under similar circumstances. Questions of reliability have to do with how consistently a test does what it is supposed to do, and thus cannot be strictly separated from validity questions. Moreover, a test cannot be any more valid than it is reliable. A test’s practicality must be determined in relation to the cost in terms of materials, time, and effort that it requires. This must include the preparation, administration, scoring, and interpretation of the test.

Bailey (1998) emphasized that wide variation in the learners’ response is inherent in portfolios, due to the fact that students self-select much of what will be included. This intentional uniqueness is thought to contribute to the high validity and washback potential of the procedure. To establish high validity, however, teachers need to gather opinions from other teachers and expertise as well as students. Not surprisingly, the diversity of the products to be evaluated can create problems in terms of reliability and practicality. In addition, the whole process of portfolio assessment can be very time-consuming. One way to establish high reliability of portfolio assessment would be a combined assessment that includes various sources such as teacher-, peer-, self-, and even parent assessment.
Arguably, what teachers need to know is the comprehensive assessment of any language skills of students as complex and dynamic as communicative language teaching would require such an effort.

As noted before, language testing in CLT approach should be multi-dimensional and performance-based testing. Given that traditional language testing in EFL School context (e.g., Korean or Japanese) has been heavily dependent upon grammar-translation method, in part due to the national college entrance examination (Yu, 1999), and also in part due to the lack of qualified and/or certified oral proficiency interviewers (Kim, 1997), there exists an urgent need for teachers from this context to combine the new language teaching approach and the assessment that is supposed to foster students’ communicative competence as well as linguistic/cultural competence.

Recently, some researchers have begun to investigate the effectiveness and the practicality of computer enhanced oral proficiency interview instrument (Kenyon & Malabonga, 2001, 1996; Norris, 2001, 1999). For instance, the d-VOCI is introduced to be a solution for teachers of the EFL context. d-VOCI uses digital video delivery on the computer as the means for testing oral proficiency. The test-taker’s oral response is also recorded on the computer digitally, and may then be sent to via e-mail to an evaluator at some other remote location.

It is argued that d-VOCI could be an ideal instrument for OPI test, especially, in EFL context where the students hardly have the opportunities to have their oral proficiency assessed by a qualified evaluator. Once teachers learned how to design and develop d-VOCI procedure—from script writing to digitizing the media sources—they can use it not only for assessment purposes but also for providing the students with
meaningful opportunities where they can deeply situate themselves with real-life tasks as well as people of other cultures.

Yet, while the number of new ways of assessing students’ oral proficiency using computer technology is steadily increasing, the literature on the field has established few theoretical grounds to guide the selection and administration of meaningful ways in terms of how new language teaching approach (CLT) can be meaningfully connected to the way we teach, test, and assess with an emphasis on EFL context. In addition, there is little research conducted to link students’ new electronic literacy to classroom language teaching/testing environment, especially in EFL context.

Figure 1 shows the changing perspectives in oral proficiency interviews in foreign and second language context.

<table>
<thead>
<tr>
<th>OPI</th>
<th>COPI</th>
<th>d-VOCI, COPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face-to-face</td>
<td>Tape-based</td>
<td>Computer-based</td>
</tr>
<tr>
<td>Phone-based</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater’s notes</td>
<td>Audio cassettes</td>
<td>Audio files</td>
</tr>
</tbody>
</table>

Figure 1: The Changing perspectives in OPI

So far, several test approaches and methods in English oral proficiency test have been discussed. As the figure above indicates, it is apparent that along with other types of computer mediated language tests such as GRE, TOEFL, and GMAT, computer assisted
oral proficiency interview is going to be administered at more institutions for various purposes. Yet, only few researches are available to discuss the characteristics of the new proficiency test to answer such research questions as “What is the construct of the d-VOCI?” “Does d-VOCI measure the same proficiency that OPI and COPI measure?” “How the computer technology should be integrated into speaking test context?’

When discussing the pedagogical implications of information technology in language classroom, Warschauer (1999) pointed out that the use of computer is part of the environment of language learning, rather than an optional instrument. He further claimed that for success in the new millennium, students also need to learn language in electronic environments as they do in print environments. His electronic literacy approach framework along with the status of English and technology in Korean context is critical to understanding the impact of technology on Korean society and its implication in and integration into the English curriculum. In the remaining part of Chapter 2, this electronic literacy framework to improve teachers and students in English classrooms in Korean context is elaborated.

2.4 Considerations of an Electronic Literacy Framework in Korean Context

2.4.1 Technology and Non-native English Teachers in EFL Context

As briefly introduced in earlier chapter, the consequence of English teachers’ powerlessness in Korean classrooms is quite important in this research. In line with that, recently, the issue of Non-native English Speaking Teachers (NNEST) in the English classroom has been frequently addressed in major ESL/EFL organizations such as
TESOL (Brutt-Griffler & Samimy, 1999; Cook, 1999; Lung, 1999; Milambuling, 2000) and other relevant communities (Braine, 1999; Medgyes, 1992). Milambiling (2000), however, says, “The number of articles and research studies on this subject [Nonnative English professionals] is growing, but awareness and systematic study are still in the beginning stages of development” (p.327). Braine also argued (1999) that the notion of the native speaker (NS) and non-native speaker (NNS), and all the accompanying linguistic, social, and economic connotations are still open to controversy. It is, in part, due to the hegemonic perspective and linguistic caste system (Kachru & Nelson, 1996), and in part, due to the ‘world Englishes’ issue (Crystal, 1997; Widdowson, 1994).

While some researchers have agreed that the search for “ideal teacher” should not involve the native-non-native dichotomy, but “how qualified an individual is as an EFL teacher” (Brutt-Griffler & Samimy, 1999) and/or “teacher’s language competence” (Liu, 1999; Medgyes, 1994), few studies have been conducted to answer further questions with regard to how to better empower NNS teachers in EFL contexts or what is the different pedagogical and methodological implementation for NNS professionals in EFL contexts compared with NS in ESL context, not to mention what “qualification” and “competence” of NNEST should matter.

Given the fact that the goal of today’s English education in EFL contexts is more broadly situated with global and cross-cultural communication than ever, and that integration of technology into a classroom is one of the major issues in the field of global education, it is assumed that technology in language classroom in EFL context can have the greatest affect on non-native English professionals’ teaching and learning when applied through the critical lens of creating generative understanding.
2.4.2 Critical Pedagogy and Power Issues in Language Classrooms

English has become a world language so rapidly because of the political, economical, technological and military progress by the USA in the past 50 years since the end of the second World War. According to Crystal (1997), most of current knowledge of human being is written in or translated into English. Regarding the impact of current technology on social aspects, Castells (1998) said, “Technology, and the ability to use it and adapt it, is the critical factor in generating and accessing wealth, power, and knowledge in our time” (p. 92). Without deep understanding of the essence of “power” issues referred to English and technology of today, it seems impossible to answer as to how we better empower NNESTs in EFL context who are being challenged by the powerful trend toward technology-enhanced language environment and also by the “communicative language teaching” approach that requires linguistic as well as cultural competence of teachers.

Critical theory addresses the relationship among schooling, education, culture, society, economy, and governance. According to Popkewitz and Fendler (1999), “Critical theory is concerned with the workings of power in and through pedagogical discourses” (p. xiii), and “The study of the effects of power enables us to focus on the ways that individuals construct boundaries and possibilities” (p. 6). Drawing on the principles of Freire’ s critical pedagogy and Weedon’ s theory of subjectivity that critical pedagogy is to be lead to change and empowerment, Brutt-Griffler and Samimy (1999) suggested, “the process of empowerment of NNESTs is neither linear nor simple but can nevertheless be generated within and by teachers engaged in critical praxis” (p. 413). Why is this issue so complicated and critical to researchers as well as practitioners in the English education
community? What are other issues intertwined with native versus nonnative dichotomy from critical perspectives? One of the reasons making this issue so complicated can be found from Freire’s notion that education is politics (Shor, 1993).

Shor (1993) interpreted Freire’s critical pedagogy as follows.

The fundamental effort of education is to help with the liberation of people, never their domestication. Teaching and learning are human experience with profound social consequences. Thus, education is not reducible to a mechanical method of instruction. The whole activity of education is political in nature. All forms of education are political, whether or not teachers and students acknowledge the politics in their work (p. 25)

Freire posited that society is controlled by an elite which imposes its culture and values as the standard (Shor, 1993). As a result, this imposed standardization is everywhere in educational environment from the students need assessment, curriculum development, syllabus design, to testing and assessment. As introduced at the beginning, in EFL context, the white Native English speaking teacher, another standard version among many English teachers, has been considered the only and qualified language teacher, especially in communicative language teaching environment where students are supposed to learn how to interact and communicate with native speakers of English.

In line with the English language, the impact of technology on our society in general and on EFL context in particular has been discussed quite often recently. As for the power of technology, Bromley (1998) pointed out, “Computers, like other technologies, are involved in many ways in the construction and use of power: in the way they are designed and built, in how they are sold and to whom, and in how they are used” (p. 2). And Stone (1998) further argues that the kind of education and skills needed in our society today includes the ability to use technology—to use it as opposed to being
manipulated by it or by those few who do have the knowledge to use it (p.187). Here, we can find a common catch phrase from both of English and technology driving forces: namely, “You must master English (or Technology) so as not to be mastered by those who have the skills.” Not surprisingly, NNESTs of EFL context face these unquestionable challenges without any proper support and time to prepare, nor meaningful pedagogical implications from research field resulting in a good general feeling of powerlessness.

The following four qualities of the “critical consciousness,” the goal of Freirean education (Shor, 1993), thus, gave an insightful perspective as to how to problematize generative themes from everyday life as well as topical issues from society and academic subject matter from specific disciplines.

1) power awareness: Knowing who exercises dominant power in society for what ends and how power is currently organized and used in society.

2) critical literacy: Analytic habits of thinking, reading, writing, speaking, or discussing which go beneath surface impressions, traditional myths, mere opinions, and routine clichés.

3) desocialization: Recognizing and challenging the myths, values, behaviors, and language learned in mass culture; critically examining the regressive values operating in society, which are internalized into consciousness.

4) self-organization/self-education: Taking the initiative to transform school and society away from authoritarian relations and undemocratic, unequal distribution of power (Shor, 1993, pp. 32-33).
One of several reasons explaining NNESTs’ powerlessness in the classroom is the notion of acquiring “standard English” as a goal of language teaching. Then, it seems crucial to ask following questions such as “Who owns English?” and “What is meant by authenticity and autonomy?” In the following, the ownership of English with regard to language teaching/testing context is discussed.

2.4.3 Ownership, authenticity and autonomy in Language pedagogy

As regard to the ownership of English and its relationship with authenticity and autonomy, Widdowson (1994) said, “The very fact that English is an international language means that no nation can have custody over it. To grant such custody of language is necessarily to arrest its development and so undermine its international status” (p. 385). In line with the standardness of English, he wisely predicted, “English will naturally stabilize into standard form to the extent required to meet the needs of the communities concerned” (p. 385). Important here is that this natural tendency towards standardization will be reinforced by the extending of networks of interaction through developments in telecommunications and information technology, not by native speakers of English nor the imperialistic nature of English.

Even though his clear concepts of the ownership of English and standard English help and motivate the periphery communities for their English education, one of his arguments in the same article is still problematic, which is the relationship between authenticity and autonomy. He said, “a pedagogy which combines authenticity of use with autonomy of learning is a contradiction. because the notion of authenticity privileges native-speaker use and native-speaker teachers” (p. 387). According to him,
only native speakers can be the arbiters of what is authentic since authenticity can be
determined by insiders alone. Nunan (1999), however, takes a different approach when
defining authenticity. According to him, “authentic data are samples of spoken and
written language that have not been specifically written for the purposes of teaching
language” (p. 27). For example, a textbook written by an experienced ESL teacher who is
a native speaker of English, and especially published in America or in England, would be
a perfectly authentic material according to Widdowson’s definition. On the contrary, a
newspaper article written in English yet by a Korean correspondent, or films and videos
that are produced not for educational purposes can be authentic materials, according to
Nunan. After all, authenticity is a construct that has a variety of meanings depending on
contexts. What sounds rather convincing, however, is that after the advent of Internet in
language classroom authentic materials, both Widdowson’s and Nunan’s, are easily
accessed and instantly retrievable on the Web. Thus, it is hard to agree to the notion that
authenticity only privileges native-speaker teachers. The point here is that any materials
designed by the native or the non-native, for the purpose of language teaching or not,
could work as long they help students build their own linguistic as well as cultural
knowledge to be eventually autonomous learners of English.

Autonomy is the uttermost important goal of language education according to
Nunan (1997, p.192). Like authenticity, autonomy is not an absolute concept but has
several degrees that can be increased through different contexts and strategies. Denying
context-free language materials, he asserts, “most commercial materials can be modified
and adapted in ways which are likely to enhance learner autonomy” (p. 202). Again, it is
not a matter of which authentic material you select, but how you provide students with
meaningful situations where they can actually be emerged in authentic context and use
the target language. As examples of those authentic contexts raising students’ degree of
autonomy, Nunan (1997) emphasized the use of information technology and the Internet
introducing his own EFL students in Hong Kong who, once discovered and began to tap
the potential of the Internet and the World Wide Web, found it both liberating and
empowering.

Pennycook (1997) suggested that it can never be possible to achieve more than
partial cultural or ideological autonomy. He further argues that we can never step
completely outside the cultural and ideological worlds around us, but we can learn to
question their boundaries, to become aware of those worlds, to search out other
possibilities and to pursue cultural alternatives (p. 46). In language classrooms operating
within a transmission and banking mode, learners practice patterns provided by teachers,
textbooks, and tapes. Maybe this is the setting where a native English teachers can play
an important role because students in this transmission model are just cast into a relatively
passive role: namely, to observe and mimic. They might learn how to communicate in
model and predictable situations, but they don’t know how to respond appropriately in
novel and authentic communicative situations. In no way do they become autonomous
language learners.

Thus, autonomy is not something students automatically acquire as their learning
progresses. According to Pennycook (1997), “to become the author of one’s world, to
become an autonomous language learner and user, is not so much a question of learning
how to learn as it is a question of learning how to struggle for cultural alternatives” (p.
35). It is impossible to encourage students’ autonomy without making them aware of the
social, cultural and political context in which they are living and working. For an instance, “PC Bang” is a Korean local colloquialism for public-access computer rooms. According to Magnier (2000), at least 15,000 of these service centers blanket Korea, up from just 100 in 1998, bringing high-speed access to users for a dollar for an hour. Most of regular customers of PC Bang are teenagers, whose ultimate goal of life is becoming a professional computer gamer. Here lies their intrinsic and extrinsic motivation of English learning: they need to learn English to read and study the computer game manuals, to chat and discuss various strategies with speakers of other languages on the cyberspace, to even travel for the participation in an international competition. Their goal of English learning may or may not be the same as that of Somalia ESL children in the United States or German students intending to study abroad in England, for example.

Since English needs to be understood not only in terms of its global position, to help students to develop their “degree of autonomy,” again, teachers should know the cultural, social, educational, and last but not least individual context in which their students are situated. Here, lies an imperative importance of NNESTs in EFL countries, who, knowing students’ individual as well as collective social needs, already experienced how hard yet achievable it is to learn how to struggle those cultural alternatives. Thus, we need to focus more on the comprehensiveness of research on teacher learning and acquisition of professional knowledge as well as linguistic/communicative knowledge in the target language. Once it clear as to how NNESTs increased their own levels of autonomy, and what facilitated their L2/C2 acquisition in conjunction with their sociocultural context, it should be transferred to the language classroom in such a way that more authentic actions and activities are implemented to raise students’ levels of
autonomy, thus making it possible that students can claim the ownership of English in
their everyday lives.

2.4.4 NNESTs empowered by Technology in EFL context

Previously noted, teachers’ role in today’s language classrooms is not to play a
model that students are asked to mimic. Besides, when it is only native speakers of
English who are projected as qualified teachers and speakers of English in students’
textbook as well as classrooms, students’ linguistic/cultural boundaries and possibilities
will be seriously limited since being a “native speaker” is not an option the Korean
students nor having a “native-like” competence cannot be a realistic goal, for them to
acquire English.

Due to the widespread multimedia technology, there are relatively enough
authentic sources in and outside of classrooms that can provide students with much more
flexible and meaningful opportunities thus help them to improve their cultural and
linguistic competence. As an experienced language learner who has already been there
where students are situated now, who knows students’ local context, NNEST can and
should be in a position where s/he can help students to keep asking questions such as
“What is the goal of (my) English learning?” “Why did English become so powerful a
language and what are the consequences?” “What does it mean to be an autonomous
language learner?” Certainly with the proliferation of personal computers, Internet, and
CD-Roms, there is less need for teacher to spend great amounts of time just telling
students the content knowledge. Instead, the teacher could play the role of learning
specialist and guide students to seek and apply information themselves. In order to use
the technology as an informational resource, many researcher argue that a teacher should have considerable expertise gained through experience (Levy, 1997; Warschauer, 1997).

The suggestion above could be more meaningful both teachers and students only after teachers have a clear picture of what it means by being a NNEST in EFL context, and what kind of challenges and missions they have, what is their own power with new technology-enhanced inside and outside classroom environment, and furthermore, what is meant by the new electronic literacy that students need to acquire for them to function appropriately in their future global society.

2.4.5 Refining Literacy (Multimedia and electronic) in Korean Context

As Shetzer and Warschauer (2000) asserted, “the overlap between English language and the development of electronic literacy is especially pronounced” (p. 171). In the past, technology was introduced to and integrated with English curriculum based upon theories and assumptions from CALL that it would augment students’ learning. Today, however, it is English that helps students to build and develop their electronic literacy. According to Lemke (1996), the generic literacies of the Information Age will certainly include multimedia authoring skills, multimedia critical analysis, cyberspace exploration strategies, and cyberspace navigation skills. If the goal of education is geared toward helping students to better function in their future society, and if teachers can agree on the notion that in their future English will play more a important role as an international language yet with different varieties than ever, they need an approach connecting English and electronic literacy with critical perspectives, which will empower our students as well as teachers.
The electronic literacy cannot be explained without understanding the relations between literacies and the societies in which they operate. Here, it is imperative to look at the unique feature of “literacy” in Korean context. In fact, there is no proper word in Korean equivalent to the word “literacy.” If someone looks up the word literacy in any English-Korean dictionary, he/she will find just a sentence (not a word!) saying “the ability to read and write,” whereas there is a word, 문맹 (文盲: unmeng), meaning “illiteracy” in English. The reason why Korean vocabulary misses literacy in spite of its vast amount of vocabularies is that Koreans have considered literacy as a natural and normal status every human being could reach without any specific obstacles.

On the contrary, munmeng (illiteracy) is considered a special and abnormal occasion that needs prompt treatment. Officially, the illiteracy rate among Korean population is less than 1%, which is quite understandable given the 5,000 year-long monolingual and monocultural Korean history. It is not surprising that people easily come to believe that Koreans have almost no social conflict regarding literacy. At the same time, the very low illiteracy rate in the modern Korean society is one of those pressures driving Korean people to be literate in English and technology, which was noted at the beginning of this paper. Since becoming literate is not just a matter of learning how to read and write, but rather a matter of mastering processes that are deemed valuable in particular societies, cultures, and contexts, and a matter of acquiring how to better function in students’ future communities, the burden of development and implementation of an electronic literacy-in-English approach rests on the Korean NNEST.
When introducing an electronic literacy framework, Shetzer and Warschauer (2000) divided electronic literacy skills into three broad, overlapping areas: communication, construction, and research.

<table>
<thead>
<tr>
<th>Earlier Approaches</th>
<th>Electronic Literacy Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication</strong></td>
<td></td>
</tr>
<tr>
<td>Based on speaking and listening</td>
<td>Also includes computer-mediated communication</td>
</tr>
<tr>
<td>Based on linear texts</td>
<td>Also includes hypertexts</td>
</tr>
<tr>
<td>Excludes non-print media</td>
<td>Combines texts and other media</td>
</tr>
<tr>
<td>Tends to focus on individual writing</td>
<td>Strong focus on collaboration</td>
</tr>
<tr>
<td><strong>Construction</strong></td>
<td></td>
</tr>
<tr>
<td>Based on linear texts</td>
<td></td>
</tr>
<tr>
<td>Excludes non-print media</td>
<td></td>
</tr>
<tr>
<td><strong>Reading &amp; Research</strong></td>
<td></td>
</tr>
<tr>
<td>Restricted to print sources</td>
<td>Includes online sources</td>
</tr>
<tr>
<td>Focuses on linear texts</td>
<td>Also includes hypertexts</td>
</tr>
<tr>
<td>Excludes non-print media</td>
<td>Combines texts and other media</td>
</tr>
<tr>
<td>Tends to separate reading skills from critical evaluation skills</td>
<td>Views critical evaluation as central to reading</td>
</tr>
<tr>
<td>Focuses on library search skills</td>
<td>Includes searching and navigating online sources</td>
</tr>
<tr>
<td><strong>Learning paradigm</strong></td>
<td></td>
</tr>
<tr>
<td>Often based on curricular learning paradigm</td>
<td>Based on interactive learning paradigm on autonomous learning</td>
</tr>
</tbody>
</table>

Table 5: Earlier Approaches vs. Electronic Literacy Approach
(Shetzer & Markschauer, 2000)
According to them, computer-mediated communication is so powerful that it transfers how we interact in business, education, and personal life. So, learning how to communicate effectively via computer involves more than just translating from one communication medium to another; it involves new ways of interacting and collaborating (p.174). As for construction, they also emphasize the importance of collaboration, creativity, and interaction. While traditional writing appears in a linear form, the hypertext requires more collaboration and creativity since it represents a radically different way of presenting information. When producing a hypertext, students need to know not only how to write but also how to link and combine other media such as graphics, audio, and video to better convey the meaning of the text.

While the traditional resource for students’ research has been the library, students of the electronic age heavily resort to on-line resources. It is not just because of physical convenience, but also its function to provide multiple resources in response to students’ requests depending on students’ different searching and evaluation skills.

The electronic literacy survey of this study was designed to measure these three sub factors (communication, construction, and research) of the participants (n=144).

2.5 Summary

Many researchers have argued that computer technology in today’s language classroom has the potential to help teachers better understand their target culture as well as their students different needs and motivations. They also advocate the need for a new educational environment that requires teachers to adopt a different role which requires
more interaction and communication with students, and to have a deep understanding of electronic literacy.

However, there is little information available on the impact of the new electronic literacy on students’ language acquisition as well as their attitudes toward technology enhanced language learning and testing. Also, few studies have provided solutions to help NNEST in the EFL context when their status may be marginalized due to the so-called “English Only” policy when it is adopted without appropriate and meaningful guidelines.

Thus, the results of this study are believed to be useful and informative not only to teachers responsible for teaching and testing foreign/second language in classrooms, but also to those who are planning course work in language education in general.

The Figure 2 below presents the conceptual framework of this study based upon the review of the relevant literature in this chapter.

**Figure 2: The Conceptual Framework: Technology in EFL context**
CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter describes the methodological procedures used to collect and analyze data to answer the research questions presented in Chapter 1. More specifically, the methodology is described in relation to the following aspects of the study: 1) location and participants, 2) procedures employed for data collection, 3) materials, and 4) data analysis for describing statistical techniques used in the study. The chapter ends with a summary statement.

3.2 Location and Participants

The data for this study were gathered at the Department of English of the Korea Military Academy located in Seoul, Korea during 2001 and 2002. The target population of this study consisted of the cadets studying at the Korea Military Academy (N=1,050). Among that population, 150 freshman students enrolled in the Summer Intensive English Course were selected to participate. Based on previous knowledge of the population and the specific purposes of the research, the researcher chose only freshman students, since it is believed that the samples selected not only have all the test score data required for this study, but they also were accessible for further data collection, test administration and
surveying. Of 150 students who finished both the d-VOCI and the electronic literacy questionnaire, six cases were dropped because of their lack of TEPS scores (n=144). Another reason for selecting only freshman cadets is that the researcher believes that they possess the necessary information about high school students in Korea. All students in this study were graduates of Korean high schools from across the country.

The Korea Military Academy’s policy of selecting new cadets in proportion to the population size of each province indicates that the participants in this study provided a more accurate representation of the high school student population than any other college academic institution.

During the two-month long Summer Intensive English Course period, the 150 students were divided into 10 classes of approximately equal size. The survey and the face-to-face OPI were conducted in the students’ own classrooms, and the d-VOCI was administered in two multimedia labs with 150 IBM personal computers.

For a statistical analysis of multiple regression, Cohen (1977) suggested that the number of cases per variable should be 14. To acquire generalizability of results, and accuracy and stability of estimates, SEM researchers generally suggest large sample sizes, though a few of them have come up with actual minimum sample sizes.

Ding, Velicer and Harlow (1995) indicated 100 to 150 subjects and Molennar and Boomsma (1987) recommended 400. Another method of determining sample size involves the consideration of sample size in relation to the number of variables: Bentler and Chou (1987) suggested a ratio as low as five subjects per variable for normal and elliptical distributions and 10 subjects per variable for other distributions. Therefore, the sample size of 144 in the present study is considered to meet the above requirements.
3.3 Procedures

The procedures employed in this study are as follows. At the end of the Summer Intensive English Course, all participants participated in a face-to-face interview with two experienced KMA English instructors. Students were interviewed in pairs for approximately ten-minute long sessions. The reason for pairing students was twofold: first, the interviewers were able to ask the pair to perform a role-play by giving them an arbitrary situation, and it also allowed the interviewers to conduct as many interview sessions as possible to include all the students in their classes.

At the end of the interview, the average of the two grades given by two instructors served as the final grade. It was assumed that two grades from different interviewers would help increase test credibility.

The students were given two or three interview questions out of three different types of question pools. The types of questions were: 1) stating facts (e.g., everyday life, life as a cadet, living in Korea) (Easy), 2) description and storytelling (moderate), and 3) describing personal opinions (difficult). There were six categories used for interviews to evaluate students’ performance: 1) understanding of the questions, 2) appropriate response, 3) pronunciation, 4) grammar, 5) fluency, and 6) use of conversational expressions.

Table 6 shows interview questions asked during the face-to-face interview session.
<table>
<thead>
<tr>
<th>POOL</th>
<th>LEVEL</th>
<th>QUESTIONS</th>
</tr>
</thead>
</table>
|      |       | - What did you like and dislike the most about your high school life?  
|      |       | - What is your favorite section of the newspaper? Why?  
|      |       | - Discuss the best news and the worst news of your life. |
| 1    | E     | - What was your dream when you were in elementary school?  
|      |       | - Describe your ideal house. What does it look like?  
|      |       | - How would you like to change by the time you graduate from KMA?  
|      |       | - Talk about a recent news story that you found interesting. Why?  |
|      | M     | - A shy friend needs help going out on a blind date. How would you help your shy friend get ready for the big date?  
|      |       | - What do you think is a value that a cadet must treasure?  
|      |       | - Which older people have had an influence on your life? Why? What kind of people are they? |
| 2    | E     | - What kind of conversational style do you have? (Bragger, complimenter, wandering eye, bore, interrupter, conversation starter..)  
|      |       | - What stresses you out? Is there something you would like to learn in the future (i.e. a new hobby) to help you cope with stress? What would you have to do to learn it?  
|      |       | - What do you usually talk about with your parents, friends, and girlfriend/boyfriend? |
|      | M     | - What was your worst nightmare? What do you think it means?  
|      |       | - What is your hobby? Let’s say that your friend wants to learn it, how would you help him/her?  
|      |       | - What is your biggest regret? Do you have any stories that you would like to tell?  
|      |       | - Talk about your summer military training. Do you have any stories to tell?  |
|      | D     | - Imagine you are making a speech at the age of 80. What do you think could have been the biggest influence on you?  
|      |       | - If you become a parent, what values would you teach your children?  
|      |       | - What do teenagers need to know in order to make the most of their lives? Why? |
| 3    | E     | - What would you like to major in?  
|      |       | - What are your usual topics for small talk?  
|      |       | - Describe a typical cadet. |
|      | M     | - What are some of the rules in your family that are different from others’?  
|      |       | - What was the best dream that you’ve ever had? What do you think it means?  
|      |       | - What things are unique about you? How are you different from other cadets?  
|      |       | - Is there anything in particular that was interesting about your summer military? |
|      | D     | - When you have a family of your own, how would you like your family to be?  
|      |       | - You have to entertain some preschool children after school. What will you do?  
|      |       | - Should we require high school students to wear uniforms? |
| 4    | E     | - What subject would you like to learn the most? How will you learn it? Why do you want to learn it?  
|      |       | - What would be the top three news of KMA in 2001?  
|      |       | - Which city would you like to visit the most? Why? |
|      | M     | - Let’s say that you are planning a backpacking trip to Europe. What would your itinerary be? Explain your itinerary.  
|      |       | - What keeps you awake? What are your methods of getting to sleep that you know of?  
|      |       | - Discuss 2-3 things that Koreans are most concerned about.  
|      |       | - What was the biggest trouble you’ve had with your parents and why? |
|      | D     | - Say that you’ve become a sophomore cadet and your new freshman cadet needs some advice. It’s early in the spring semester and the freshman cadet is having trouble getting used to his new life as a cadet.  
|      |       | - It’s your friend’s birthday, and you want to plan a surprise party he or she will never forget. What can you come up with?  
|      |       | - Should we learn a foreign language at an early age? |

Table 6: Questions for Face-to-face interview
The d-VOCI was administered after the face-to-face OPI. Each student individually answered the questions provided by the d-VOCI computer program of which specific features are elaborated in the material part later, and his/her answers were audio recorded on the computer’s hard drive. The audio files were then sent to the three qualified OPI raters via e-mail or post mail for evaluation on OPI rating.

After the two English OPI tests were administered, a survey of students’ electronic literacy and attitudes towards d-VOCI was conducted. Detailed information regarding the instrument is described in the next section.

3.4 Materials

English tests

Multimedia-enhanced English oral proficiency testing was carried out on IBM personal computers situated in the Multimedia Lab at the Korea Military Academy. The software used for this study is called ‘d-VOCI’ (digital Video Oral Communication Instrument) designed by the staff at Language Acquisition Resource Center (LARC) at San Diego State University.

After logging in using their personal information, students first read the directions and then answered two practice questions. Then, upon the instructor’s confirmation that students understood the procedures of the test and were ready to continue through their tests without any technical questions, they went to the test page and answered 10 video questions, each of which corresponded to an ACTFL OPI level.
When logging in, students were required to type in such information as their names, student numbers (instead of social security number), and their e-mail addresses (see Appendix G). The information given by each student was automatically saved in a text file (*.txt) in the computer, and was sent to the three evaluators along with the audio files (*.wav). In terms of the content of the video questions, it is noteworthy that the interviewer of the video test is one of the instructors in English at KMA (see Appendix H) and is a native Korean who learned English as a second language. Other personnel characters in the video segments also were played by two teaching assistants from the same department; they possess at least a BA degree from an English speaking country (see Appendix I). The instructors at the Korea Military Academy designed and developed all the digital video- audio clips for the d-VOCI so that the content and construct of the d-VOCI would reflect students’ own educational/curricular context. As noted before, the levels of the ten video questions were aligned with the ACTFL OPI guidelines. Appendix A shows the script for the d-VOCI used in this research.

After answering a couple of simple questions, the examinee was asked to give directions to someone using a map on the screen, to describe a particular place based on a picture, and to pose some questions to a virtual friend who had just returned from a trip. In later parts, d-VOCI requires the examinee to tailor his or her discourse strategies to selected topics and real-life situations.

Students’ Test for English Proficiency (TEPS) scores were acquired from the English department of the KMA: each cadet at the KMA is required to take the TEPS test at least once per semester to demonstrate English skills that are a minimum graduation requirement level.
Additional test scores (e.g., GPA, TOEFL, and English presentation) were also secured for further analysis but not used in this study because of the constraints of variable quantity required for statistical analysis (SEM).

ELQ (electronic literacy questionnaire)

Information about subjects’ electronic literacy and their attitudes towards computer-enhanced OPI was gathered in a non-anonymous survey. Although the lack of anonymity may have compromised some students’ willingness to answer honestly, this was a necessary decision to allow correlations to be calculated between the students’ answers on the survey and their English test scores.

The final version of the instrument is found in Appendix D, electronic literacy survey, whereas the version used for the pilot test is presented in Appendix C. Given the paucity of comprehensive and psychometrically sound measures of students’ electronic literacy in Korea, the researcher developed a new instrument to adequately measure this construct. As Churchill (1979) suggested, the researcher specified the constructs (i.e., electronic communication, cyberspace construction, and academic problem solving research) and developed an initial sample of items for each construct through an extensive review of electronic literacy literature (Douglas, 1998; Dunkel, 1999; Shetzer & Warachauer, 2000; Chapelle, 2001: Swan, 2001). The proposed research model for electronic literacy and English proficiency was then developed based upon the literature review. When selecting and developing items, the researcher used criteria such as the length of the items, level of difficulty, relevance to the scale’s purpose, and clarity (DeVellis, 1991).
The initial item pool needed to be purified to identify and eliminate those items which are either unreliable or that cross-load on other constructs (Churchill, 1979). To establish face validity, the researcher simply asked the experts to indicate whether the instrument looks like it is measuring what it claims to be measuring (Gliem, 1996).

Content validity refers to “a qualitative means of ensuring that indicators tap the meaning of a concept as defined by the analyst” (Bollen, 1989, p.186). In this study, to establish content validity, five experts were selected because of their experience, expertise, and knowledge in the content area, and research methodology. They were asked to rate each item as well as the instrument as a whole for content, clarity, wording, format, thoroughness, ease of use, focus, and appropriateness. In addition, they were asked to identify which items fit under each of the sub-dimensions.

After the revisions were made based upon the feedback from the panel members, a field test was conducted by administering the revised version of the instrument to a representative group similar to the target population. This group included 20 sophomore students at the Korea Military Academy who participated in the Summer Intensive English Courses. Appropriate revisions were made to the instrument based on the feedback from the field test members. For example, some general terms categorizing a group of software (e.g., graphic tools, web development tools, and authoring tools) were replaced by the title of more specific software (e.g., “Adobe Photoshop,” “Dreamweaver/FrontPage,” and “Director”/“Flash”) to avoid ambiguity.

Before the survey began, the researcher explained the option to withdraw from the study, but there were no dropouts during the whole process of the study.
The survey for electronic literacy and student attitude towards d-VOCI included ten personal information questions and 70 additional questions answered on a six-point Likert scale. Of these 70 questions, 45 questions addressed the subjects’ electronic literacy in terms of electronic communication (15 items), cyberspace construction (15 items), and academic problem solving research (15 items). The other 30 questions were designed to measure students’ attitudes towards d-VOCI in terms of personal preference, difficulty (both content and procedure), and usefulness in comparison to face-to-face OPI. Coefficient alpha provided an estimate of reliability based on inter-item correlations and yields an estimate comparable to the average of all possible split-halves. Item-total statistics were examined to detect any items that did not contribute significantly to the internal consistency of the scale. Appendix J summarizes the reliability analysis statistics for electronic literacy by reporting item-total statistics, analysis of variance, and reliability coefficients for electronic communication ($\alpha = .92$), cyberspace construction ($\alpha = .93$) and academic problem solving research ($\alpha = .91$).

3.5 Data Analysis

The data gathered from the survey as well as the English tests were analyzed using the software Statistical Package for Social Studies 9.0 (SPSS) and Analysis of Moment Structure 4.0 (AMOS). To show an overall of students’ electronic literacy and their attitudes towards d-VOCI, descriptive statistics of means, frequencies, skewness and standard deviation were calculated for each factor.

The results from the d-VOCI test were then evaluated by three native speakers of English at The Ohio State University in the United States. A CD containing 144 folders,
each containing 10 audio samples, was given to each instructor, and the three instructors evaluated the audio samples separately to establish inter-rater reliability. All raters were instructors in the Spoken English Program with sufficient background and experience in the Speak Test. Each rater evaluated 10 audio files from each student and assigned a level from novice to superior using ACTFL OPI rating guidelines. The inter-rater correlation was then analyzed to establish an inter-rater reliability. Both face-to-face and d-VOCI inter-rater correlations are reported on Table 9 in Chapter 4.

To analyze the variance of students’ d-VOCI scores, Multiple Regression Analysis was used to gauge the contribution of each independent variable (TEPS, Face-to-face Interview, electronic literacy, and self-assessment) to explain the variance of the dependent variable, d-VOCI. Multiple Regression Analysis is the prediction of a criterion using a combination of two or more predictor variables. Each predictor is weighted in proportion to its contribution to prediction accuracy. The equation showing the weights assigned to each predictor is the multiple regression equation.

The procedure for Multiple Regression Analysis is as follows. First, correlations are computed between each independent variable and the dependent variable. Then, correlations among all the independent variables are computed. Finally, regression parameters are estimated and a test of significance (e.g., F, t) is applied. All these calculations can increase error and limit the applicability of the results to the overall population, which requires a larger sample size. Shannon and Davenport (2001), for example, recommended for Multiple Regression that the sample contain approximately ten subjects for each independent variable. Since this study has four independent variables, the sample size of 144 was appropriate for this multiple regression.
In this study the simultaneous model was employed where all four independent variables are treated simultaneously and entered into the regression equation on a single step. This strategy was used since there was no logical or theoretical basis for considering any independent variable prior to any other independent variable.

To investigate the structural relationship between students’ electronic literacy and their English oral proficiency, and to test the efficacy of the proposed model, the researcher employed structural equation modeling (SEM). A model in this context is an a priori hypothesis about a pattern of linear relationships between the two latent variables above.

SEM includes investigations of both structural and measurement models. The structural model is the path model, which relates the independent to the dependent variables. Structural model analysis is an essential tool for the identification of the causal relationship between several constructs in which separate multiple regression equations are estimated simultaneously. The measurement model allows the researcher to use several variables for a single independent or dependent variables, and assesses the contribution of each scale item as well as incorporate how well the scale measures the concept into the estimation of the relationship between the dependent and independent variables (Fassinger, 1986; Hair et al., 1998). As Browne (2002) argued, the Confirmatory Factor Analysis (CFA) allows the user to specify relationships of multiple latent variables to their indicators, but allows only correlational (non-directional) influences of LVs to each other. In a structural equation model, s/he can test models of directional influences among latent variables.
A measurement model “specifies a structural model connecting latent variables to one or more measures or observed variables” (Bollen, 1989, p.182). As stated earlier, the measurement model describes the relationship between the measured variables or indicators and latent variables. Based on the data from the electronic survey and test scores of students, the researcher constructed the following hypothesis:

Hypothesis 1:
The three manifest variables (electronic communication, cyberspace construction, and academic problem solving research) adequately capture the latent variable, electronic literacy. And the four manifest variables (TEPS, face-to-face interview, d-VOCI, and self-assessment) adequately capture the latent variable, English proficiency. The two latent variables are correlated.

Figure 3 depicts the hypothesized structural relationship between those two latent variables with their measured predictors.
Figure 3: Hypothesized model between students’ electronic literacy and their English proficiency with their measured predictors
The model in Figure 3 supposes the left three sub-domains (i.e., electronic communication, cyberspace construction, and academic problem-solving research) depend on an unobserved variable called “electronic literacy.” Likewise, the right four variables are supposed to depend upon “English proficiency.” In this model, the two common factors, electronic literacy and English proficiency, are allowed to be correlated based upon the data from the multiple regression analysis described in the previous section. In addition, the error variables (i.e., e1-e3, z1-z4) represent any and all influence on measurement in the measured variables that are not explained by the common factors, electronic literacy and English proficiency.

The hypothesis was tested by SEM, and the AMOS output provided a chi square statistic and a number of goodness-of-fit statistics, which are helpful in interpreting the fit between the data and the hypothesized model. As for the use of a chi square statistic, however, Browne (2002) warned that it is necessary to employ fit indices other than a chi square test because it is highly sensitive to sample size. Therefore, to compensate for the limitations of the chi square statistic, the researcher also employed several fit indices obtained from the output of the AMOS 4.0 analysis. The purpose of the criterion is to determine the degree to which the hypothesized relationships are identical to the observed data (Kline, 1998; Maruyama, 1998). The method of maximum likelihood was specified for estimating the structural equation model. There are three different types of indexes (i.e., absolute indexes, relative indexes, and adjusted indexes). Specific indexes used for the study include: 1) absolute index (e.g., Chi-square/df, Root Mean Residual (RMR), Goodness of Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA), 2) relative or incremental index (e.g., Tucker-Lewis Index (TLI), Comparative Fit Index
(CFI), and Adjusted Goodness-of-Fit Index (AGFI), and 3) adjusted index (e.g., Parsimonious Goodnees-of-Fit Index (PGFI) and Parsimonious Normed Fit Index (PNFI) (Hair, et al., 1998; Kline, 1998, Maruyama, 1998).

Although there is no clear-cut guideline, the ratio of values of Chi-square and degree of freedom less than 3 was considered to be good fit or accepted fit (Carmines and McIver, 1981: Kline, 1998). A Chi-square/df ratio less than 2.0 indicates an excellent model fit (Hayduk, 1996). The use of chi-square is appropriate for sample sizes between 100 and 200, which is the case of this study. As sample size increase (over 200), a significant difference is found for the proposed model. In contrast, when sample size decreases (below 100), Chi-square test shows acceptable fit, even when the model relationship is not significant (Hair, et al., 1998). The sample-size-dependent fit indices are affected by sample size, either directly or indirectly (Tanaka, 1993).

The RMR is the average of the residuals between the observed and predicted matrices (Hair et al., 1998). As the RMR of zero indicates a perfect fit, the smaller the RMR is, the better (Arbuckle and Wothke, 1998).

GFI indicates the degree to which the overall model predicts the observed correlation matrix. It is analogous to a squared multiple correlation in that it represents the proportion of the observed covariances explained by the model-implied covariances. The GFI has an absolute interpretation because it concerns only the proposed research model (Kline, 1998). The value of this index ranges from 0 (poor fit) to 1 (perfect fit). Higher values indicates better fit, but greater than .90 is desirable (Hair, et al., 1998; Kline, 1998).
The RMSEA is the discrepancy (the average of residuals between the observed and estimated matrices) per degree of freedom. RMSEA is recommended with relatively larger sample as the discrepancy is measured in terms of the population, not just the sample. According to Browne and Cudeck (1993), values of .08 or less for the RMSEA provide evidence for reasonable fits and values less than or equal to about .01 may be satisfactory for exploratory research.

As for the relative index, the TLI combines a measure of parsimony into a comparative index between the proposed and null models (Hair, et al., 1998). First proposed as a means of evaluating factor analysis, the TLI has been extended to SEM. The range of TLI lies between 0 and 1. A value of .90 or greater is recommend (Hair, et al., 1998).

The CFI is a relative comparison of the researcher’s proposed model to the null model (i.e., independence model) which is a measure ranging from 0 (not fit at all) to 1.0 (perfect fit). The value of .80 indicates that relative overall fit of the researcher’s model is 80 % better than of the null model. A recommended value is .90 or greater (Bentler & Bonnet, 1980; Kline, 1998)

For a more thorough description of the goodness-of-fit indices, refer to the article written by Mulaik and his colleagues (1989). When the proposed model satisfied the above criteria, descriptions of “good” and “adequate” were used (Kline, 1998).
3.6 Summary

To answer the primary and secondary research questions of this study, the following statistical procedures were used: 1) Descriptive statistics including means, standard deviations, and frequencies were used to determine and examine the pattern shown in the results of the survey and tests, 2) Reliability of the questionnaire and inter-rater reliability for d-VOCI were established, 3) Correlations between d-VOCI scores and each of TEPS scores, electronic literacy, and attitudes towards d-VOCI were calculated, 4) Multiple regression analysis was used to investigate the variable that best explains the variance in students’ d-VOCI scores, and finally, 5) Structural Equation Modeling (a two-construct model) was carried out to investigate the structural relationship between all the variables employed in this study with an emphasis on the structural relationships between students’ electronic literacy and their English proficiency as measured in both standardized tests (e.g., TEPS) and oral interviews.
CHAPTER 4: DATA ANALYSIS AND DISCUSSION

4.1 Introduction

The purpose of this study was to examine the nature of a multimedia-enhanced English oral proficiency interview with emphasis on the relationship between students' electronic literacy and their multimedia-based OPI scores in an EFL context. As explained in Chapter 3, the multimedia enhanced English oral proficiency scores were acquired via d-VOCI, and quantitative survey data were collected using the Electronic Literacy Questionnaire with three sub-domains—electronic communication, electronic construction, and cyberspace construction. This chapter begins with descriptive statistics of the variables in the study including demographic characteristics of the sample, mean score, standard deviations, and normality of the data. Then, the results of the multiple regression analysis are introduced with emphasis on the variance of the students' English oral proficiency score measured by d-VOCI. Finally, the results from the structural equation modeling using AMOS are presented to analyze the structural relationship between the two latent variables: electronic literacy and English proficiency.
4.2 Descriptive Statistics

4.2.1 Demographic Characteristics of the Research Sample

A sample of 150 freshman cadets at the Korea Military Academy (KMA) who were participants in the Summer Intensive English Course were asked to provide data. Due to the missing data, the sample was reduced to a total of 144 cases. Table 1 shows the demographic characteristics of the sample.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Range</th>
<th>Mean</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>18-21</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>M (135) F (9)</td>
<td></td>
<td>Male (94%), Female (6%)</td>
</tr>
<tr>
<td>Years of Studying English</td>
<td>6-9</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Experience with an English Speaking Test</td>
<td>Yes (13) No (121)</td>
<td>Yes (9%), No (91%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Demographic Characteristics of the Research Sample

As shown in Table 7, the students ranged in age from 18 to 21 years old, and there were only 9 female students out of 144 useable samples, meaning the results may not be applicable to groups with a different gender ratio. The unbalanced gender ratio of the sample simply resulted from the fact that 90% of the students enrolled in the KMA are male, and the other 10% are female.

While the average years of English studied for this sample is 7 years, it is noteworthy that only 13 students (9%) had experience in taking one or more English-speaking tests prior to enrollment in college, and none of those 13 students took the test in school, but rather in private language institutes in Seoul. Again, considering the representative characteristics of the cadets at KMA to the population of high school
students in Korea, it is safe to say that there are few, if any, English-speaking tests being given in the Korean high school setting.

4.2.2 Electronic Literacy

Table 8 shows the descriptive statistics of students’ electronic literacy in terms of three sub-domains; electronic communication, cyberspace construction, and academic research. There were fifteen items for each three sub-domains, and the numbers in “mean” column indicated the average scores obtained by the 144 participants from each sub-domain.

Academic Research showed the highest mean score of 4.02, which is followed by Electronic Communication (3.6), and Cyberspace Construction (3.16). In other words, students’ self-rated capability to integrate electronic technology into their own academic research is reported to be higher than the capability to communicate using electronic media including e-mail, cellular phone, and audio/video conferencing. In general, students self-reported higher literacy levels in these two areas (Academic Research and Electronic Communication), yet the mean score of Cyberspace Construction (3.16) is slightly higher than the middle of the scale (1-6), indicating that students are either particularly confident or insecure in their cyberspace construction capability. The mean score of the three sub-domains is 3.6.

The three sub-domains are moderately related ranging from .56 (academic research and cyberspace construction) to .70 (academic research and electronic communication). As mentioned in the previous chapter, the internal reliability of this instrument was moderately high (see Appendix J).
<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>electronic communication</td>
<td>144</td>
<td>1.93</td>
<td>5.27</td>
<td>3.5975</td>
<td>.7447</td>
</tr>
<tr>
<td>cyberspace construction</td>
<td>144</td>
<td>1.27</td>
<td>5.00</td>
<td>3.1689</td>
<td>.8199</td>
</tr>
<tr>
<td>academic research</td>
<td>144</td>
<td>1.67</td>
<td>5.52</td>
<td>4.0234</td>
<td>.7695</td>
</tr>
<tr>
<td>electronic literacy</td>
<td>144</td>
<td>1.67</td>
<td>4.89</td>
<td>3.5966</td>
<td>.6737</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>144</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>electronic communication</th>
<th>cyberspace construction</th>
<th>academic research</th>
</tr>
</thead>
<tbody>
<tr>
<td>electronic communication</td>
<td>Pearson Correlation</td>
<td>.624**</td>
<td>.696**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>cyberspace construction</td>
<td>Pearson Correlation</td>
<td>.624**</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>academic research</td>
<td>Pearson Correlation</td>
<td>.696**</td>
<td>.560**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>144</td>
<td>144</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Table 8: Descriptive statistics and correlations of Electronic literacy
### 4.2.3 English test scores

The table below contains the descriptive statistics of the results from all the English tests administered in the study.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Std. Error</th>
<th>Statistic</th>
<th>Statistical Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TEPS-L</strong></td>
<td>144</td>
<td>83</td>
<td>346</td>
<td>230.63</td>
<td>42.14</td>
<td>1775.60</td>
<td>-.109</td>
<td>.570</td>
<td>.401</td>
<td>.202</td>
<td>.570</td>
</tr>
<tr>
<td><strong>TEPS-R</strong></td>
<td>144</td>
<td>87</td>
<td>366</td>
<td>240.88</td>
<td>47.70</td>
<td>2274.90</td>
<td>-.194</td>
<td>.274</td>
<td>.401</td>
<td>.202</td>
<td>.274</td>
</tr>
<tr>
<td><strong>TEPS-G</strong></td>
<td>144</td>
<td>34</td>
<td>87</td>
<td>57.43</td>
<td>10.10</td>
<td>102.10</td>
<td>.313</td>
<td>.401</td>
<td>.202</td>
<td>.202</td>
<td>.202</td>
</tr>
<tr>
<td><strong>TEPS-V</strong></td>
<td>144</td>
<td>19</td>
<td>84</td>
<td>50.22</td>
<td>9.95</td>
<td>98.995</td>
<td>-.009</td>
<td>.274</td>
<td>.401</td>
<td>.202</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>TEPS-T</strong></td>
<td>144</td>
<td>225</td>
<td>883</td>
<td>579.15</td>
<td>92.44</td>
<td>8545.21</td>
<td>-.020</td>
<td>.401</td>
<td>.202</td>
<td>.202</td>
<td>1.30</td>
</tr>
<tr>
<td><strong>F_RATER1</strong></td>
<td>144</td>
<td>10.00</td>
<td>20.00</td>
<td>16.36</td>
<td>2.00</td>
<td>4.023</td>
<td>-.355</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>F_RATER2</strong></td>
<td>144</td>
<td>11.00</td>
<td>20.00</td>
<td>16.44</td>
<td>2.49</td>
<td>6.192</td>
<td>-.255</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Face-to-face</strong></td>
<td>144</td>
<td>11.0</td>
<td>20.0</td>
<td>16.40</td>
<td>1.94</td>
<td>3.747</td>
<td>-.325</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D_RATER1</strong></td>
<td>144</td>
<td>2</td>
<td>10</td>
<td>4.35</td>
<td>1.25</td>
<td>1.573</td>
<td>.919</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D_RATER2</strong></td>
<td>144</td>
<td>2</td>
<td>10</td>
<td>4.03</td>
<td>1.37</td>
<td>1.880</td>
<td>.911</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D_RATER3</strong></td>
<td>144</td>
<td>1</td>
<td>10</td>
<td>3.97</td>
<td>1.35</td>
<td>1.824</td>
<td>.927</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>d-voci score</strong></td>
<td>144</td>
<td>1.67</td>
<td>10.00</td>
<td>4.12</td>
<td>1.21</td>
<td>1.469</td>
<td>1.21</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S-Assessment</strong></td>
<td>144</td>
<td>1</td>
<td>9</td>
<td>5.24</td>
<td>1.31</td>
<td>1.706</td>
<td>-.22</td>
<td>.202</td>
<td>.401</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Valid N</strong></td>
<td>144</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>144</td>
</tr>
</tbody>
</table>

**TEPS_L**: TEPS Listening part  
**TEPS_R**: TEPS Reading part  
**TEPS_G**: TEPS Grammar part  
**TEPS_V**: TEPS Vocabulary part  
**TEPS_T**: TEPS Total score  
**F_RATER1**: Score from face-to-face interview rater 1  
**F_RATER2**: Score from face-to-face interview rater 2  
**Face-to-face**: face-to-face interview score for each student  
**D_RATER1**: Score from d-VOCI rater 1  
**D_RATER2**: Score from d-VOCI rater 2  
**D_RATER3**: Score from d-VOCI rater 3  
**d-VOCI score**: d-VOCI score for each student  
**S-Assessment**: Self-Assessment score  

face-to-face Interview Inter-rater Reliability Coefficient (alpha) = .64  
d-VOCI Interview Inter-rater Reliability Coefficient (alpha) = .90

Table 9: Descriptive Statistics of English Tests
In Table 9, it is noteworthy to look at the shape of the distribution for two oral proficiency interviews in terms of skewness. While the skewness of the face-to-face interview scores measured by English instructors at KMA is -.33, the skewness of d-VOCI scores measured by three American ESL teachers is 1.208. Since the closer the values are to 0 the more likely it is that the variables follow a normal distribution, the numbers indicate that there are some extreme scores on the positive, upper end of the d-VOCI distribution, and some extreme scores on the negative, upper end of the face-to-face distribution. Another finding from this table is the mean difference between d-VOCI and Self-Assessment. Given that both tests have the same scales from one to ten, with one meaning novice and ten meaning superior, the mean difference can be interpreted to show that the students in this study, in general, overestimate their English oral proficiency. In a practical sense, the mean level of the d-VOCI was ‘Intermediate Low,’ but the results from the students’ self-assessment imply that they locate their English proficiency somewhere between ‘Intermediate Mid’ and ‘Intermediate High’ according to the ACTFL OPI proficiency guidelines.

Finally, the inter-rater reliability between raters of face-to-face interview at KMA was .64 whereas the raters from the ESL program at The Ohio State University had a stronger inter-rater reliability of .90. Shannon and Davenport (2000) emphasized the particular importance of the inter-rater reliability when information is collected through an observational method. Greater amounts of consistency, or agreement, indicate the higher reliability that is essential when the Oral Proficiency Interview is incorporated into the school language curriculum.
4.2.4 Students’ Attitudes towards the d-VOCI and the face-to-face interview

As can be seen in the Table 10, in general, the participants showed strong positive attitudes towards d-VOCI. When asked whether d-VOCI was interesting, none of the participants selected item one or two from the 5 item Lickert-type questionnaire (1: strongly disagree,… 5: strongly agree). Ninety-seven students (67%) rated this test either fun or very fun, and only five students (3.5%) did not agree that this test was fun. Ninety-three students (65%) agreed or strongly agreed that the English oral proficiency test would represent their correct English proficiency as opposed to other tests purported to measure reading, listening, and writing skills, and just seven students (5%) disagreed that their oral English skills were correctly reflected. Students’ responses to the item asking whether the d-VOCI test was difficult (DIFFQUES) had a regular normal distribution.

Only three students reported that they could not understand the interview questions from d-VOCI while the rest of the subjects (98%) reported no problem understanding the questions. It is noteworthy that the majority of the students reported that their English proficiency could be improved through the d-VOCI test (83%), and even more students wanted to know the results of their performance, graded by the qualified OPI raters (90.3). Considering that the multimedia-enhanced English oral proficiency interview was a format relatively new to the students, it is interesting to note that just two students (1.4%) reported difficulty dealing with the computer-mediated testing procedure.

Out of 144 subjects, there were only thirteen students (9%) reporting any previous experience with English Oral proficiency tests. Yet, none of them had taken any English
speaking tests in schools, although some (9%) had taken speaking tests in private institutions off campus.

Finally, when asked which speaking test was preferred, d-VOCI or a traditional face-to-face interview, about 70% of the subjects reported a preference for the face-to-face interview over the d-VOCI. The reason most commonly sited was the interaction between the interviewers and the interviewees. More than half of the participants (63%) reported that they enjoyed both types of English oral proficiency tests, but when asked to choose one format over the other, they opted for the face-to-face interview because d-VOCI cannot react to students’ answers or tailor the interview based upon students’ responses, which is possible in the case of a face-to-face interview. The students seemed to prefer human over technology-mediated interaction during an English oral interview.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERES</td>
<td>144</td>
<td>3.00</td>
<td>5.00</td>
<td>3.9236</td>
<td>.6154</td>
</tr>
<tr>
<td>FUN</td>
<td>144</td>
<td>2.00</td>
<td>5.00</td>
<td>3.7500</td>
<td>.6946</td>
</tr>
<tr>
<td>EXACT</td>
<td>144</td>
<td>2.00</td>
<td>5.00</td>
<td>3.7500</td>
<td>.7710</td>
</tr>
<tr>
<td>DIFFQUES</td>
<td>144</td>
<td>1.00</td>
<td>5.00</td>
<td>3.1875</td>
<td>.8018</td>
</tr>
<tr>
<td>UNDERST</td>
<td>144</td>
<td>1.00</td>
<td>5.00</td>
<td>1.9514</td>
<td>.7417</td>
</tr>
<tr>
<td>MULTIME</td>
<td>144</td>
<td>2.00</td>
<td>5.00</td>
<td>3.6042</td>
<td>.9024</td>
</tr>
<tr>
<td>IMPROVE</td>
<td>144</td>
<td>2.00</td>
<td>5.00</td>
<td>3.9792</td>
<td>.6943</td>
</tr>
<tr>
<td>CURIOUS</td>
<td>144</td>
<td>1.00</td>
<td>5.00</td>
<td>4.3750</td>
<td>.7278</td>
</tr>
<tr>
<td>DIFFHAND</td>
<td>144</td>
<td>1.00</td>
<td>4.00</td>
<td>1.8056</td>
<td>.8217</td>
</tr>
<tr>
<td>PRIOR EXPERIENCEa</td>
<td>144</td>
<td>0</td>
<td>1</td>
<td>9.03E-02</td>
<td>.29</td>
</tr>
<tr>
<td>PREFERENCEb</td>
<td>144</td>
<td>0</td>
<td>1</td>
<td>.32</td>
<td>.47</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>144</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10: Descriptive Statistics of Students’ Attitudes toward d-VOCI

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRIOR EXPERIENCEa</td>
<td>144</td>
<td>0</td>
<td>1</td>
<td>9.03E-02</td>
<td>.29</td>
</tr>
<tr>
<td>PREFERENCEb</td>
<td>144</td>
<td>0</td>
<td>1</td>
<td>.32</td>
<td>.47</td>
</tr>
</tbody>
</table>

Table 10: Descriptive Statistics of Students’ Attitudes toward d-VOCI
Table 11 reports the results from two Independent Groups t-tests between students who preferred the d-VOCI and those who preferred the face-to-face interview in terms of their d-VOCI scores as well as their electronic literacy levels. To test the assumption of homogeneity of variance, Leven’s Test for Equality of Variance was used since unequal variances can impact the results of t-tests, making the probability of committing Type I and Type II errors more likely (Shannon & Davenport, 2000). For the dependent variable d-VOCI score, an F ratio of .623 with a probability (p value) of .431 was produced. This indicates that the difference between the two group variances is likely to occur only by chance as many as 43.1 times out of 100. Likewise, results of the t-test for Equality of means indicate that two groups do not differ significantly in terms of their d-VOCI scores nor electronic literacy levels.

<table>
<thead>
<tr>
<th></th>
<th>Leven's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>d-voci score</td>
<td>.623</td>
<td>.431</td>
</tr>
<tr>
<td></td>
<td>Equal variances assumed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equal variances not assumed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equal variances assumed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equal variances not assumed</td>
<td></td>
</tr>
</tbody>
</table>

Table 11: Independent Groups t-Test between students who preferred d-VOCI and those who preferred face-to-Face interview
4.3 Multiple Regression Analysis

4.3.1 Explaining the variance of d-VOCI

One of the research questions in this study sought to explain the variance of the media-enhanced English Oral Proficiency Interview scores of the students in a Korean college setting. With predictions involving just one variable, the contribution of that variable to the prediction of the dependent variable is easily determined. With multiple predictors, however, the influence of a specific predictor cannot be determined solely by the correlation between the two variables. It is likely that the predictors influence each other, and in doing so, also influence the relationship between the predictor and the dependent variable. In this chapter, after the descriptive and correlation tables were interpreted, the partial and semi-partial correlations were reported and analyzed to determine the extent of each independent variable’s contribution to explaining the probable variance of d-VOCI scores.

Descriptive information, including means, standard deviations, and correlations, is summarized in Table 12 below. Note that this analysis includes 144 students who had a valid responses to each of the five variables for the regression analysis.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>d-voci score</td>
<td>4.12</td>
<td>1.21</td>
<td>144</td>
</tr>
<tr>
<td>TEPS-T</td>
<td>579.15</td>
<td>92.44</td>
<td>144</td>
</tr>
<tr>
<td>face-to-face</td>
<td>16.40</td>
<td>1.94</td>
<td>144</td>
</tr>
<tr>
<td>electronic literacy</td>
<td>3.60</td>
<td>.674</td>
<td>144</td>
</tr>
<tr>
<td>self-assessment</td>
<td>5.24</td>
<td>1.31</td>
<td>144</td>
</tr>
</tbody>
</table>

Table 12: Descriptive and correlations output from multiple regression
The average score of d-VOCI was 4.12, which is slightly above an Intermediate-Low rating according to the ACTFL oral proficiency guidelines, while the mean of students’ self-assessment scores is 5.24, between Intermediate-Mid and Intermediate-High using same ACTFL system. The means for TEPS total scores, the face-to-face interview scores, and students’ electronic literacy were reported as 579.15, 16.40, and 3.60, respectively.

The table of correlations (Table 13) is useful in determining the relationships between the dependent variable (d-VOCI) and each of the predictor variables as well as the relationships among the predictor variables. All of the correlations between the predictors and d-VOCI were positive. The strongest of these was between TEPS scores and d-VOCI scores, reported as .492, implying that students with higher TEPS scores tend to acquire higher d-VOCI scores. The second strongest relationship was with students’ self-assessment and d-VOCI score (.40). Interestingly, the correlation between the d-VOCI and the face-to-face interview was relative low (.364) given that both of these tests are supposed to measure the same construct, English Oral Proficiency. While the correlation coefficient is statistically significant (alpha =.05), it indicates that knowing students face-to-face interview score would be of little help in predicting students’ d-VOCI score (12 %). Finally, the correlation between electronic literacy and d-VOCI was reported as .325, indicating that those students with higher electronic literacy scores tended to have higher d-VOCI scores.

All the correlations among the four predictor variables were positive, ranging from .160 to .405, and they were statistically significant (p<.05). Multicollinearity occurs when some or all of the independent variables are substantially correlated with each other,
or when one or more of the independent variables are almost linear combinations of the other independent variables. Multicollinearity should be suspected when none of the partial regression coefficients is statistically significant when the $R^2$ for the full model is significant. Hair et al. (1998) suggested that, when multicollinearity is found, the researcher should combine two or more highly correlated independent variables into a single variable and use the composite variable in the regression in place of the correlated variables. This substitution is appropriate only when the variables combined are multiple indicators of the same concept. In this case, none of the correlations was higher than .8, and thus the highest correlation was safely assumed to be .405 (between self-assessment and TEPS). Other issues regarding multicollinearity are discussed later in relation to the coefficient table.

In general, Table 13 indicates that students with higher scores from TEPS, fact-to-fact interview, electronic literacy and self-assessment tend to have higher d-VOCI scores.
Table 13: Correlations between variables in the Multiple Regression Analysis

Table 14 summarizes information pertaining to the overall relationship (R) between the predictors and the dependent variable, the amount of variance explained by these predictors ($R^2$), and the results of the significance test used to test the regression model (F change).

The overall relationship between the four predictors and d-VOCI was reported as .619 indicating that 38.3% of the variance in d-VOCI can be explained using these four predictors. The adjusted $R^2$ was .349, which may not be that much different from the sample $R^2$ of .383. The adjusted $R^2$ is an estimate that exists in the population. Since the adjusted $R^2$ was close to the $R^2$ reported for the sample, the fit between the sample and population was considered good.
The results of the F-test revealed a statistically significant F value of 21.553 (p<.001), and the degrees of freedom for this F-test are 4 and 139. The value of the Durbin-Watson statistic was reported as 1.891. This statistic describes the serial correlation among residuals. This test value can range from 0 to 4, with values close to 0 indicating a positive correlation among residuals and those close to 4 indicating a negative relationship. Values between 1.5 and 2.5 are expected and since the value of 1.891 falls in this range, the correlation of residuals was not of concern for this analysis.
Table 14: Model Summary and ANOVA Results—Simultaneous Regression

Table 15 summarizes the data pertaining to the regression coefficients. This information is necessary for making predictions about the dependent variable. Whereas the unstandardized coefficients are dependent upon the scales used to measure each predictor and can rarely be compared directly, the standardized coefficients are based on the same scale. Comparisons can be made to assess the relative contribution of each
predictor to the variance of the dependent variable. In this study, Table 15 was used to
determine the relative importance of each independent variable (TEPS, face-to-face
interview, electronic literacy, and self-assessment) in predicting the variance of students’
d-VOCI scores.

All the unstandardized coefficients were positive, indicating that these predictors
are positively related to the dependent variable, d-VOCI, and all the t-tests were
statistically significant (p<.05). Therefore, it can be assumed that the variables TEPS,
face-to-face, electronic literacy, and self-assessment were contributing significantly to the
prediction of d-VOCI.

Examining standardized coefficients revealed that the TEPS score had a value
of .346, which was followed by the face-to-face interview score (.231), electronic literacy
(.184), and then self-assessment (.170).

The highest partial correlation was also reported for TEPS. This relationship
represents the correlation between d-VOCI and TEPS, with the influence of the remaining
predictors (face-to-face, e-literacy, and self-assessment) removed from both d-VOCI and
TEPS.

The removal of these three influences resulted in the parts of d-VOCI and TEPS that were
not related to these three variables or the residual parts. The resulting partial correlation
between d-VOCI and TEPS was .370. When squared, this coefficient represents the
amount of variation shared between the residuals of both d-VOCI and TEPS. Specifically,
the parts of d-VOCI and TEPS that were not related to face-to-face interview, electronic
literacy, and self-assessment shared approximately fourteen percent of the variance in
common (13.69%).
Next, the highest semi-partial correlation was also reported for TEPS (.313). This coefficient represented the extent to which TEPS relates to d-VOCI, after the influence of the remaining predictors has been removed from TEPS. This represented the unique contribution of TEPS. Based on a semi-partial correlation between TEPS and d-VOCI, 9.8 percent of the variance in d-VOCI can be explained by TEPS and was not influenced by the other predictor variables. Likewise, of all the variance to be explained that could be explained in the dependent variable (100%), face-to-face explained about 5%, electronic literacy 3 %, and finally self-assessment 2.3 %.

4.3.2 Multicollinearity Analysis

As noted before, the independent variables are moderately related. To determine the extent to which the predictors overlap, however, the Tolerance statistic was examined to show the percentage of each variable that is not related to the other predictors.

Tolerance statistic (1-\(R_j^2\): \(R_j^2\) is the squared multiple correlation when regressing the \(j^{th}\) independent variable on all other independent variables) is the proportion of a predictor’s variance that is not explained by the linear combination of the other predictors. Lower tolerance values indicate that there is a significant overlap with other predictors, and therefore little of that predictor’s variance remains available to contribute to the dependent variable. In this case, it is noteworthy that the Tolerance of TEPS (.818) is lower than those for the face-to-face interview (.932) and electronic literacy (.916), but it has the highest semi-partial correlation coefficient. This can be explained by the fact that the zero-order relationship between TEPS and self-assessment was .405, which is higher than any other correlation among the four predictor variables. In other words, students’
self-assessment score might be moderately influenced by the TEPS score since, in a practical sense, students in this study had few criteria other than TEPS scores to which they could refer when they self-assessed their own English language proficiency. In general, high values (near 1.0) indicate that multicollinearity is not a problem and low values (near .00) indicate multicollinearity. In this case, no independent variable is of concern.

Another indicator dealing with the multicollinearity is Variance Inflation Factor (VIF). This statistic is equal to 1 divided by the tolerance. Therefore, if the tolerance were a perfect 1, the VIF would be a 1. Lower tolerances, which indicate overlap among the predictors, result in higher VIFs. Therefore, as the VIF becomes larger, greater overlap exist among the predictors. When there is a significant overlap among predictors, the variability of regression coefficients (B) becomes inflated. In other words, high values for VIF indicate that a particular independent variable is a linear combination of the other independent variables. Since all the VIF values were low, close to 1 in this study, there was no reason for concern that the variability of regression coefficients became inflated.

<table>
<thead>
<tr>
<th>Coefficients(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Model (Constant)</td>
</tr>
<tr>
<td>TEPS-T</td>
</tr>
<tr>
<td>face-to-face</td>
</tr>
<tr>
<td>e-literacy</td>
</tr>
<tr>
<td>s-assessment</td>
</tr>
</tbody>
</table>

\(^a\) Dependent Variable: d-voci score

Table 15: Regression Coefficients
4.3.3 Residual analysis

Finally, analysis of residuals indicates whether assumptions of the regression model appear to be satisfied. For the researcher to conduct a multiple regression analysis, it was assumed that 1) the residuals are normally distributed, 2) they have a mean of zero and a constant variance across levels of the independent variables, 3) they are independent across all subjects, and 4) they are not systematically related to the independent variables. Violation of these assumptions increases the variances of the partial regression coefficients, therefore the standard errors of partial regression coefficients are not unbiased estimators of the true standard deviations. Thus, when assumptions are violated, tests of statistical significance will be inaccurate. As noted before, the Durbin-Watson statistics is a serial correlation among residuals. The value of 1.891 in the study indicated that the residuals were not correlated.

Figure 4 displays a histogram of the standardized residuals with a superimposed normal curve. It visually shows the extent to which the residuals are normally distributed. Although the residuals do not adhere perfectly to a normal distribution, neither do they depart significantly from it.
Figure 4: Histogram of the Standardized Residuals

Figure 5 displays a normal probability plot for the standardized residuals. If the residuals in the analysis followed a normal distribution perfectly, their cumulative probabilities would follow a straight line. Again, the line is not perfectly straight, but no significant departures appear. Therefore, it can be concluded that the assumptions were not violated.
In this section, the results from the Multiple Regression analysis were reported to examine the relative influence of four predictor variables (TEPS, face-to-face interview, electronic literacy, and self-assessment) on a dependent variable (d-VOCI). In the following section, the results from the Structural Equation Modeling (SEM) are introduced.

4.4 Structural Equation Modeling

4.4.1 Fit of Measurement proposed model

The purpose of testing the measurement model is to describe the relation between the measured specific indicators and construct of interest. The results of the measurement
model test determine how well the indicators capture their specific constructs (Bollen, 1989). The results of the analyses supported that the overall fit of the measurement model was adequate (See table 16). Specifically, throughout all measurement tests, the Chi-square value was not significant compared to the degree of freedom (i.e., Chi-square/df = 1.03). The Chi-square/df ratio was lower than the suggested threshold (i.e., less than 3.0) (Carmines & McIver, 1981; Kline, 1998). This indicated that the results did not reject the null hypothesis and suggested that the data represents the proposed model.

The RMSEA value was .01 for the measurement model, which is a reasonable fit (Browne & Cudeck, 1993), and this indicates that measurement model was approaching a close fit and reasonably accepted. In addition, all other indices (i.e., GFI, TLI, and CFI estimates) were greater than the recommended .90 threshold throughout the fit analysis. The results of the analysis is found are Table 16.

<table>
<thead>
<tr>
<th>Measurement Model (n=144)</th>
<th>CMIN</th>
<th>DF</th>
<th>CMIN/DF</th>
<th>RMR</th>
<th>RMSEA</th>
<th>GFI</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13.36</td>
<td>13</td>
<td>1.03</td>
<td>2.97</td>
<td>.01</td>
<td>.98</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 16: Fit analysis for the measurement model

In summary, the analyses of the measurement model represented the data well and suggested that the sub domains used in the study adequately captured the latent variables (i.e., electronic literacy and English proficiency).
4.4.2 Parameter Estimates and Hypothesis Testing

Figure 6 is the path diagram with standardized estimates. In this figure, the squared multiple correlations can be interpreted as follows. For example, 79% of the variance of “electronic communication” was accounted for by the variance in electronic literacy. The remaining 21% of the variance of “electronic communication” cannot be explained by this model, and is thus attributed to the unique factor e1. If e1 represented measurement error only, it could be said that the estimated reliability of “electronic communication” would be .79. As it is, however, the term e1 may comprise systematic unique variance components in addition to random error. Thus, the figure .79 has to be regarded as a lower-bound estimate of the reliability.

An important finding of this analysis was the low squared multiple correlation (.17) between “face-to-face interview” and “English proficiency.” As can be seen in Figure 6, only 17% of the variance of “face-to-face interview” was accounted for by the variance in English proficiency while 63% of “d-VOCI” was explained by English proficiency. Again, this result indicates that the students’ face-to-face interview scores might have been heavily influenced by unknown constructs other than their English proficiency. The present study did not allow for a definitive analysis of this type of interaction.
Figure 6: Hypothesized structural relationship between English proficiency and Electronic Literacy (Model A)
Table 17 below represents the estimates of the standardized regression weights of the model.

<table>
<thead>
<tr>
<th>Standardized Regression Weights</th>
<th>Estimate</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMM &lt;-- Electronic Literacy</td>
<td>0.89</td>
<td>0.00</td>
</tr>
<tr>
<td>CONSTRUC &lt;-- Electronic Literacy</td>
<td>0.71</td>
<td>0.00</td>
</tr>
<tr>
<td>RESEARCH &lt;-- Electronic Literacy</td>
<td>0.78</td>
<td>0.00</td>
</tr>
<tr>
<td>TEPS_T &lt;-- English Proficiency</td>
<td>0.62</td>
<td>0.00</td>
</tr>
<tr>
<td>FATOFA &lt;-- English Proficiency</td>
<td>0.42</td>
<td>0.00</td>
</tr>
<tr>
<td>DVOCI &lt;-- English Proficiency</td>
<td>0.79</td>
<td>0.00</td>
</tr>
<tr>
<td>SELFASSE &lt;-- English Proficiency</td>
<td>0.55</td>
<td>0.00</td>
</tr>
</tbody>
</table>

COMM: electronic communication
CONSTRUC: cyberspace construction
RESEARCH: academic problem solving research

TEPS_T: TEPS score (total)
FATOFA: face-to-face interview
DVOCI: d-VOCI score
SELFASSE: self-assessment

Table 17: Estimates of Standardized Regression Weights

The standardized regression weights in Table 17 can be interpreted as factor loadings on the two latent variables (electronic literacy and English proficiency). The highest factor loading on electronic literacy was .89 (electronic communication), and on English Proficiency, .79 (d-VOCI). The factor loading of face-to-face interview was the lowest (.42), while all values were statistically significant (alpha = .05).

Testing Model B against Model A

Hypothesis 2: Students’ electronic literacy is not correlated with their English Proficiency.
Considering the exploratory nature of the study, the researcher did not place any constraint on any parameter in Model A. Based on the result of Model A, however, the researcher developed another model explaining the relationship between the two latent variables. The rationale for developing the two models was to suggest the more parsimonious and representative of the two by comparing multiple models using statistical analyses.

As Arbucke & Wothke (1999) explained, a direct comparison can be performed whenever one of the models can be obtained by placing additional constraints on the parameters of the other (p.140). In the study, the researcher obtained Model B by imposing an additional constraint on the parameters of Model A to test the hypothesis that the two common factors were mutually exclusive. More specifically, the coefficient parameter between electronic literacy and English proficiency was set as zero to test the second hypothesis.

The results of the re-specified model (Model B) are shown in Figure 7.
Figure 7: Hypothesized structural relationship between electronic literacy and English proficiency with a constraint (Model B)
A test of Model B against Model A can be obtained by subtracting the smaller Chi-square statistic from the larger one (Arbuckle & Wothke, 1999). In the study, the new statistic is 18.447 (i.e., 31.807-13.360). If Model B is correctly specified, this statistic will have an approximate Chi-square distribution with degrees of freedom equal to the difference between the degrees of freedom of the competing models. In the present study, the difference in degrees of freedom was 1 (i.e., 14-13). That is, Model B imposed all of the parameter constraints of Model A, plus an additional one. With one degree of freedom, Chi-square values greater than 3.84 were significant at the .05 level. Therefore, the researcher rejected the null hypothesis that students’ electronic literacy is not correlated with their English Proficiency. In other words, there exists a moderate positive relationship between the electronic literacy and English proficiency of the students who participated in this study.

4.5 Summary

Chapter 4 began with descriptive statistics of the variables in the study including demographic characteristics of the sample, mean score, standard deviations, and normality of the data. There were only thirteen students (9%) reporting any previous experience with English oral proficiency tests, thus, not surprisingly, the majority of the participants in the study showed strongly positive attitudes towards the d-VOCI test.

The results of the multiple regression analyses indicate that the variance of the d-VOCI scores can be explained by four predictor variables (i.e., TEPS, face-to-face interview, electronic literacy, and self-assessment). The overall relationship between the
four predictors and the d-VOCI was reported as .619 meaning that 38.3% of the variance in d-VOCI can be explained using these four variables. Examining standardized coefficients revealed that the TEPS score had a value of .346, which was followed by the face-to-face interview (.231), electronic literacy (.184), and then self-assessment (.170). By employing a series of test statistics, the researcher concluded that the assumptions for the multiple regression analysis were not violated.

After the researcher confirmed that students’ electronic literacy was related with their d-VOCI scores, a SEM analysis was conducted to investigate the relationship between two latent variables (i.e., electronic literacy and English proficiency) with their measured indicators. The results of the two hypotheses tests revealed that the sub-domains used in the study adequately captured the latent variables, and these two factors were correlated with each other.
CHAPTER 5: FINDINGS, IMPLICATIONS, RECOMMENDATIONS FOR FUTURE RESEARCH, AND CONCLUSION

5.1 Introduction

The purpose of this study was to investigate the nature of a multimedia-enhanced English oral proficiency interview with an emphasis on the relationship between students’ electronic literacy and their multimedia-based OPI scores in an EFL context. This research was conducted in a Korean college setting where traditionally there have been few opportunities for students to know their own English oral proficiency levels even though the new status of English in Korean society suggests the need for a higher degree of communicative and linguistic competence in English than ever. The focus of this study, thus, was assessing the applicability of the multimedia-based English oral proficiency interview (d-VOCI) in a Korean college English classroom environment for students as well as teachers. The main goal was to study an instrument to measure and possibly improve students’ English oral proficiency. Since using computer technology for oral proficiency tests is a new method, the characteristic of the research was exploratory.

This chapter provides a discussion of the study based upon the data reported in Chapter 4. First, the conceptual framework in this study is briefly reviewed, and then the findings of this study are explained and summarized based on the research questions of
the study. Then, implications of the study to the field of English education in the EFL context are identified and described. The chapter continues with suggestions for future research projects, followed by major conclusions drawn from the study.

5.2 Review of the conceptual framework of the study

Many researchers have argued that one of the most important goals of foreign and second language education is helping students become autonomous language learners (Nunan 1997; Lemke 1998; Shetzer & Warschauer 2001). While language professionals in the EFL context agree with this “learner autonomy” notion, there have been few theories or examples of instructional practice for English teachers in Korea to employ this notion for meaningful English education that empowers EFL learners. Focusing on learner autonomy, the review of relevant literature of this study identified among the constraints four major concerns in the English education in Korea: 1) lack of teacher confidence (NNEST), 2) lack of collaboration between NEST and NNEST, 3) lack of pedagogical practices based on the CLT approach, and 4) lack of performance-based classroom assessments.

Another major constraint in relation to the four concerns stated above that impedes “learner autonomy” notion in the Korean English education setting is the absence of a use of the standardized speaking test in most K-12 English curricula. These five constraints, correlated with each other in one way or another, form a barrier that may be keeping students from becoming autonomous language learners.
The fundamental purpose of this study was to provide pedagogical implications to bridge this gap for better and more meaningful English education in the Korean English education setting and in other similar EFL contexts elsewhere based upon the hypothesized conceptual framework depicted below.

AUTONOMOUS LANGUAGE LEARNER

TEACHING/LEARNING

Empowered NNEST

Communicative Language Teaching/Learning

TESTING

Collaboration between NEST and NNEST

Alternative Assessment (e.g., on-going, portfolio, self-assessment)

Communicative Competence in English

Figure 8: Hypothesized Conceptual Framework: d-VOCI in EFL context

In Chapter 2, the researcher proposed a hypothesized conceptual framework of the study (Figure 2) based upon a review of relevant literature regarding the use of technology to foster language learning students in an EFL context from three different perspectives: 1) SLA and CALL, 2) SLT from communicative perspectives, and 3) an electronic framework in the Korean context. To provide a meaningful and feasible model
for the use of the technology in the framework, the researcher tested the applicability of
the d-VOCI in an EFL English testing setting in Korea.

The findings of this research support the view that a multimedia-enhanced English
oral proficiency test (d-VOCI) is a promising instrument and a possible system to address
the five constraints stated above in an EFL context. Therefore, the previous hypothesized
conceptual framework was revised with d-VOCI replacing the more generic term
“technology.” (Figure 8).

As indicated earlier, the non-native interviewer reported a gain in confidence
during the course of the English d-VOCI development process and appreciation for the
opportunity given him to assess his own language performance.

It is evident that the d-VOCI encompasses a wide range of collaboration between
NEST and NNEST from the early to final stages: for instance, the d-VOCI template was
developed by a team of NEST, but the specific and contextualized interview questions
were developed and put into the template by NNEST. Again, the final audio files were
sent to another team of NEST specialized in ESL/OPI, and the levels along with written
feedback were sent back to the Korea Military Academy.

The purpose of the d-VOCI is not just to measure student’s linguistic knowledge
but to evaluate their communicative competence in any target language: therefore it is to
note that the teacher can use the instrument as an alternative form of assessment which is
more on-going as well as performance-based.

Since the system is “multimedia- enhanced” (i.e., technology-based), it also must
be taken into account how the construct measured by these multimedia-enhanced
instruments differs from the construct assessed by conventional methods, and whether
there exists any relationship between students’ electronic literacy and their English proficiency (Sheltzer & Warschauer, 2001).

In the next section, the revised hypothesized conceptual framework is examined with respect to each research question posed in the study as well as the relationship between students’ electronic literacy and their English proficiency as measured in the present study.

**5.3 Research Questions and Discussion**

There were four primary research questions dealing with the variance of student’s d-VOCI scores in relation to the other variables, and three secondary research questions about students’ attitudes toward d-VOCI. The following section is a description and discussion of each research question based upon the results of data analysis (Chapter 4) and the hypothesized conceptual framework.

Question 1: What is the structural relationship between students’ electronic literacy and English oral proficiency?

Following Shetzer and Warschauer’s notion of the electronic literacy approach (Shetzer & Warschauer, 2001), the researcher developed an electronic literacy survey with three sub-factors: 1) electronic communication, 2) cyberspace construction, and 3) academic research.

Four predictors (TEPS score, face-to-face interview score, d-VOCI score, and self-assessment rating) were used to describe students’ English oral proficiency (n=144).
The data from SEM analysis revealed that there was a moderate positive relationship between students’ electronic literacy and their English oral proficiency (.44). This finding supports Warschauer’s assertion (1999) of the status of English as an international and global language and its evolving relationship with students’ electronic literacy.

Because of the possibility that the correlation between the students’ electronic literacy and English oral proficiency might have been strongly influenced by the technology-based d-VOCI score that is one of four predictors for English proficiency, the researcher also analyzed a second model excluding the d-VOCI score as one of the predictors. While the coefficient showed a minimal change (.44 → .40), it still implies the notion of a moderate positive relationship between those two latent variables.

Question 2: Which variable(s) of this study best explain(s) the variance of the students’ English oral proficiency?

To explain the variance of the students’ d-VOCI scores, this research employed the multiple regression analysis with four predictor variables (TEPS score, face-to-face interview score, electronic literacy, and students’ self-assessment) and the d-VOCI score as the independent variable. All four predictors contributed significantly to explain the variance of the d-VOCI score, and the variable that contributed most when predicting students’ d-VOCI was the TEPS followed by the face-to-face interview, electronic literacy, and finally self-assessment. Again, students’ electronic literacy levels turned out to be helpful to predict their English oral proficiency as measured by d-VOCI.
Question 3: Is there a statistically significant difference in results between the computerized format and the conventional face-to-face format in English oral proficiency interviews?

This question mainly examined the reliability of the two English oral proficiency tests used in this study. While the construct of both tests should be identical (English oral proficiency), requiring a high positive relationship between these two variables, the statistical result showed only a weak relationship (.30). The correlation value .30 is just high enough to reach statistical significance, yet it is low in the practical sense since it indicates that knowing the results of students’ face-to-face English interview does not seem to contribute significantly in predicting students’ d-VOCI scores (9%). There are several possible explanations for this weak relationship: for instance, the low inter-rater reliability between the face-to-face interview raters (.64) might indicate that different raters measured students’ proficiency from different perspectives and rubrics whereas the high inter-rater reliability established by d-VOCI raters (.90) suggests that the raters in general followed the same procedures and principles when evaluating student’s English proficiency.

It is, however, still questionable whether this low inter-rater reliability was caused by just a format difference or a lack of evaluator training. While it is true that the face-to-face interview gives evaluators more flexibility and freedom in terms of conducting the interview and getting ratable speech samples from the students, the raters usually resorted to their notes taken during the course of the interview or their own memories to assign a rating after the interview is over. This factor needs to receive further study.
In the case of d-VOCI, the drawbacks regarding the lack of the interviewer-interviewee interaction as well as flexibility of topic and format may be compensated for by the fact that the raters have recorded samples (i.e., audio files and CDs) that they can revisit for valid and reliable oral proficiency evaluation. It would, therefore, be an excellent research topic to investigate the different inter-rater reliabilities between the face-to-face interview and d-VOCI in terms of evaluator experience and the format of the tests, which was not the focus of this study.

Question 4: What is the relationship between the four language skills (READING, LISTENING, GRAMMAR, VOCABULARY) in the TEPS and d-VOCI scores of the students?

As previously stated, the TEPS (total score) in general was a good indicator for predicting students’ English oral proficiency score as measured by d-VOCI. However, the statistical analysis of the relationships between the four sub-domains of TEPS and d-VOCI revealed that there existed some unequal coefficients when each sub-skill was individually compared to the d-VOCI. For example, the sub skill most strongly correlated with d-VOCI was VOCABULARY (.51) followed by LISTENING (.48), GRAMMAR (.41), and READING (.33). As a matter of fact, based upon the results of a multiple regression analysis with d-VOCI as the dependent variable and the four skills as independent variables, the “reading” part of TEPS did not seem to contribute at all in predicting students’ d-VOCI score (Table 18).
Table 18: Model Summary (d-VOCI and TEPS sub-skills)

Question 5: What patterns of student views about the electronic oral assessment format were reported by participants?

The subjects in this study showed strong positive attitudes towards the d-VOCI. As represented by Table 10 in Chapter 4, the majority of the students reported that they found the d-VOCI to be fun and interesting. In general, they had few problems understanding the procedures of the computer-assisted test, and the interview questions on the computer screen were seemingly not difficult to understand.

More importantly, the participants indicated that they thought that this test format (d-VOCI) would help them improve their English oral proficiency when it was aligned with their regular English curriculum. The highest interest of the students regarding the
d-VOCI was the final grade they would be assigned by certified raters after the test. This concentration of interest was not surprising, considering that only 13 of 144 participants in this study reported any type of prior English-speaking test experiences.

Question 6: To what extent do electronic literacy and English proficiency influence students’ attitudes toward d-VOCI and conventional methods?

The result of the independent-samples t-test (Table 11) showed that the students who reported a preference for d-VOCI did not get particularly high or low scores from d-VOCI or the electronic literacy survey. In other words, neither English proficiency nor electronic literacy seemed to influence students’ preference between the d-VOCI and the face-to-face interview.

Question 7: What magnitude of difference characterizes student’ attitudes toward d-VOCI versus face-to-face English interviews?

As explained in Chapter 4, approximately 70% of the participants reported that they preferred the face-to-face interview, and 30% reported that they preferred the d-VOCI. Among those whose preference was for the face-to-face interview, more than 80% reported that the face-to-face interview seemed to be the better method since it is “real” and “interactive.” It is noteworthy that while the majority of the students found the d-VOCI helpful to improve their English oral proficiency, they still regarded the interaction between the interviewer and interviewee as more valuable when it comes to an oral proficiency examination.
Among those who preferred the d-VOCI, about 40% reported that the face-to-face interview could not be an objective test due to the interviewers’ subjectivities. During the interview between the researcher and the participants, one student said that he believed that the instructors tended to give higher grades to those students who had been active and performed well in the classroom, not particularly in the face-to-face interview session. Students also complained about the inter-rater reliability, saying that there were rather generous instructors as opposed to those who were at the other end of the spectrum. As has been reported in the previous chapter, the inter-rater reliability test for the face-to-face interview indeed showed a relatively low result (.64). This topic suggests a need for further research.

5.4 Recommendations for Further Research

As suggested before, a quantitative research study needs to be conducted to investigate inter-rater reliability when d-VOCI and the face-to-face interviews are used and ratings are performed by raters with various degrees of experience (e.g., extensive training versus minimal training) as well as with different linguistic and cultural background. This type of research is needed to eliminate raters and their assigned ratings as statistically significant variables in establishing the reliability of d-VOCI as an appropriate format for EFL contexts with a limited availability of competent English instructors to serve as raters. If d-VOCI is established via research to be a viable alternative to the face-to-face English interviews, the d-VOCI can then be used without reservations.
It is also recommended that a longitudinal mixed methodology including both quantitative and qualitative data collection be conducted for intermediate-level college students to investigate the impact of individualized feedback provided by raters utilizing d-VOCI on students’ achievement, particularly their oral usage of English. The researcher envisions a study where individualized feedback are given to students along with their audio data to identify to what extent the feedback affects students’ oral usage of English over the time.

In addition, further quantitative research should be conducted with students of other foreign language commonly taught in Korea, such as Japanese, Chinese, and French. Replications of this study with different languages would provide in particular further evidence for the relationship between students’ electronic literacy and their foreign language proficiency. In this particular study recommended above, the researcher suggests that students from different age groups be selected to examine the interaction between their age, their foreign language proficiency, and their electronic literacy. The researcher’s belief is that younger students who are, in general, more exposed than older students to electronic and network-based environment might show different foreign language learning style and a different level of electronic literacy.

Another qualitative research proposed by the researcher involves reflective and collaborative action research for teachers and students in an EFL context. In this study, teachers and students collaboratively conduct a d-VOCI project including script writing, and videotaping the questions as well as self-, peer-, and teacher evaluation. The researcher believes that such critical reflection can trigger deeper understanding of language learning and teaching. Thus, it would be interesting to examine to what extent
the d-VOCI project and the critical reflection upon it would stimulate and promote students’ language acquisition.

Finally, there are many unresolved questions with regard to the meaningful integration of the d-VOCI test into the English classrooms in the EFL context. For example, how can the d-VOCI be used not just as a testing tool but also as an on-going assessment tool? How can teachers implement d-VOCI in creating a collaborative learning environment? To what extent do teachers’ own electronic literacy as well as their English proficiency impact the uses of the multimedia-enhanced tests? How can teachers create useful yet authentic interview questions that best elicit the ratable speech samples from the students? Further research may provide some insight into these issues and possible solutions.

5.5 Implications of the Study

As noted in Chapter 1, one of the research questions of the study was to provide pedagogical implications and suggestions with regard to the multimedia-enhanced English oral proficiency test in an EFL context. Based upon the data analysis, the researcher concluded that the d-VOCI, as an example of a multimedia-enhanced English oral proficiency tests, might be a promising tool for both teachers and students in English courses in EFL context, particularly is further research conducted to substantiate this usage.

First, the results of the study have implications for language testing. Language performance is recognized by both SLA and LT researchers as being highly complex,
multidimensional, and variable according to a variety of social and contextual factors (Bachman, 1998). Unfortunately, the current language testing system in Korean English classrooms does not reflect this complex and multidimensional nature of language performance, only measuring the fragmented “knowledge” in English, not the “performance,” or “proficiency” of the students. Supporting Hancock’s notion of the ongoing language assessment (1994), d-VOCI can be used to assure that students’ process of language performance is continuously recorded. It also can help teachers to make institutional decisions about what instruction is needed by the learners to improve their language proficiency since it provides the possibility for teachers to gather and analyze students’ speaking samples over time. Being different from the traditional standard tests where the test-takers receive no feedback other than a grade or a score, the d-VOCI, which can be more authentic, classroom-based, and contextualized to the students, allows the teachers and the evaluators to provide individualized, thus potentially more meaningful, feedback to each test-taker based upon his/her own performance in the test. This topic requires further research.

In addition, this study has some implications for English teacher education in EFL context in terms how to empower them. As noted earlier, being a “native speaker” or having a “native-like” competence is not a realistic goal for most EFL students to acquire English. The researcher also pointed out that when it is only native speakers of English who are projected as qualified teachers and speakers of English in students’ learning/testing environments, their linguistic/cultural boundaries and possibilities might be seriously limited. In this regard, the fact that the interviewer in the d-VOCI program of the study was performed by a NNEST who, as an language learner himself, had already
been where most EFL students are situated and knew the student’s local context better than anybody else from the outside classroom, is noteworthy. In fact, it is the researcher’s belief that, through the process of the study (i.e., from the script development and rehearsals, to final video taping), the NNEST interviewer gained confidence in his own language performance, although this was not a focus of the present study. When properly used, the d-VOCI might be a valid and reliable tool for language teachers’ diagnostic tests to closely identify their particular strengths and weaknesses. The d-VOCI, as an diagnostic measure, might also be used in learning centers or self-access study programs to help student and teachers to decide where to focus their attention during EFL study.

Finally, for the multi-media enhanced English oral proficiency to be better incorporated into the school curriculum in the EFL context, it is imperative that qualified and experienced raters evaluate the students’ performances. While the d-VOCI in the study resorted to widespread network technology to deliver the Korean students’ audio samples to the evaluators located in the United States, it was only because of the absence of the qualified/certified raters available in the Korean setting, not necessarily because the raters have to be the native speakers of English. In the future, therefore, this study suggests that more Korean English teachers should be aware of the importance of speaking tests in their classrooms, and given appropriate learning opportunity to grasp the nature of English oral proficiency tests based upon the current ACTFL proficiency guidelines so that eventually they themselves can become qualified and certified raters who are able to provide meaningful feedback to their own EFL students.
5.6 The Limitations of the Study

Several limitations of this study warrant discussion. First, since the data collected in the study were obtained from one single research site (i.e., the Korea Military Academy), if the proposed model is to be generalized to other EFL school contexts (e.g., middle, high schools, and other higher institutions), further research is necessary from different groups of students with different levels of language proficiency. For example, the content of d-VOCI tests included some portion of military terms which were specifically targeted to the students of the Korea Military Academy in terms of context-bound authentic assessment. Thus, the result of this study cannot be generalized to other populations other than similar military institutes.

Another limitation of this study concerns the subject effects. Since this study employed a new way of conducting an English proficiency assessment, subjects might have responded differently than when the newness wears off (novelty effect). Therefore, results obtained in this study may not be generalizable until the treatment is institutionalized.

In addition, as noted in Chapter 4, the sample of this study was heavily male-oriented (94%), thus the result of this study may not apply to other populations with a different gender ratio. At the same time, however, these limitations imply the need for replication of the study. For example, although the results of the analyses supported the survey on students’ electronic literacy as a valid scale, it should be noted that construct validity is not established through a single study, but rather through several additional studies (Kline, 1998). In this regard, further investigations of electronic literacy need to be conducted using different populations within the participant EFL context.
5.7 Conclusion

The professional literature and other research have shown English oral proficiency tests such as the ACTFL OPI, SOPI, and d-VOCI to be an informative and yet neglected tool in English classes in EFL contexts. One of the major concerns in this particular context regarding English education is the absence of standardized speaking tests in the regular EFL school curriculum.

It is obvious that language learners need to develop a certain level of grammatical control to communicate in the language, but with the absence of limited opportunities for teachers and students to assess students’ oral communicative proficiency, the goal of language learning—helping students to be autonomous language learners—remains distant.

Currently, new technologies and changing foreign language education environments provide meaningful and novel ways to incorporate English oral proficiency tests, a real challenge in the past, into EFL curricula. In fact, this phenomenon draws increasing attention with contemporary changing perspectives of the definition of literacy: from the conventional reading and writing focus to the more functional and current electronic literacy for students in a future global environment. In line with these shifting trends, the results of this study highlighted some important issues and posed several important recommendations for future research on English oral proficiency assessment in EFL contexts.
In answer to the research questions of this study, it was found that the multimedia-enhanced English oral proficiency test suited to participants’ particular interests and instructional situations showed promise as an oral proficiency assessment tool. The focus of this multimedia enhanced evaluation system, however, should not be tied to computers or even technology per se, but it needs to be limited human beings. It is important to note that while students found the d-VOCI to be interesting and fun, even helpful to assess and improve their foreign language proficiency, the majority of the participants preferred the face-to-face interview. Most students reported that the purpose of oral proficiency testing should be to measure students’ inter-personal communication skills, but not with machines. The main problems are, of course, the limited opportunities and the dubious legitimacy of the results from raters with little or no training, as the participants also reported.

In this sense, one of the most important implications for this research is that multimedia-enhanced English OPI can be an efficient and effective supplement to the conventional face-to-face interview where the opportunities for the students having human interviews are often quite limited. In fact, the d-VOCI provides not only the opportunity for speaking tests to students, but it can also produce benefits to non-native English speaking teachers to meaningfully incorporate the communicative language teaching/testing approach into their classrooms and to effectively collaborate with native English speaking professionals in setting up a media-enhanced oral English testing system.

Another important finding of the study was investigating the relationship between students’ English proficiency and their electronic literacy. The cases examined in this
study revealed a correlation between students’ English literacy and their electronic literacy skills. These findings confirm the findings of previous studies, and particularly, the notion of an electronic literacy approach to network-based language teaching (Shetzer & Warschauer, 2001).

In conclusion, this study contributed a valid, reliable, and practical language test measuring students’ English oral proficiency based on the ACTFL OPI guidelines. The 10 interview questions of the study were developed to reflect the content and context of the goal of the English department of the Korea Military Academy. With the d-VOCI template, however, teachers from different institutions can create their own interview questions by replacing the existing media files with their own files. Teachers might benefit from using the d-VOCI to conduct a needs analysis, create a collaborative and constructive language learning environment, provide individualized feedback to students, and also evaluate the language curriculum. For researchers interested in SLA and LT, the d-VOCI might be a useful instrument to collect, save, retrieve, and analyze students’ speech samples. Last but not least, students might also benefit from using this type of English oral proficiency test to identify their own proficiency level, improve their own linguistic, communicative, and cultural competence, and eventually become autonomous language learners. Given the importance of developing efficient opportunities for instructors and students to learn English as a foreign language in myriad contexts throughout the world, this study has provided important baseline data for further research.
REFERENCES


APPENDICES

APPENDIX A: d-VOCI Evaluation Sheet (a sample page)

APPENDIX B: Student Survey Permission Form

APPENDIX C: Electronic literacy survey (Pilot survey)

APPENDIX D: Electronic literacy survey (Final version)

APPENDIX E: Script for the English d-VOCI

APPENDIX F: Main Screen of d-VOCI

APPENDIX G: d-VOCI log-in screen

APPENDIX H: d-VOCI interview screen

APPENDIX I: d-VOCI video screen

APPENDIX J: Summary of Reliability Estimates for Electronic Literacy scales
Appendix A: d-VOCI Evaluation Sheet (a sample page)

<table>
<thead>
<tr>
<th>Date:</th>
<th>Evaluator:</th>
<th>Signature:</th>
</tr>
</thead>
</table>

**ACTFL (1-10)**
1: Superior
2: Advanced high
3: Advanced mid
4: Advanced low
5: Intermediate high
6: Intermediate mid
7: Intermediate low
8: Novice high
9: Novice mid
10: Novice low

<table>
<thead>
<tr>
<th>Student Number</th>
<th>ACTFL</th>
<th>OSU SPEAK</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>14674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>677</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>678</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>683</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>686</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>688</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>689</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>693</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>694</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>695</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>696</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>697</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Students Survey Permission Form

Dear Students:

My name is Tae-Young Jeong, and I am a doctorate candidate in the department of Teaching and Learning at the Ohio State University. My academic advisor, Dr. Charles Hancock, and I are conducting research to investigate the effectiveness of a new English Oral Proficiency Test (d-VOCI: digital-Video Oral Communication Instrument), and the relationship between electronic literacy and English proficiency.

We would like permission to ask you to participate in this study as follows.
- You are to fill out a 45-item questionnaire designed to measure your electronic literacy
- You are to take a computer based English Oral Proficiency Test (d-VOCI), of which result will be sent to you in a month directly from the evaluators in USA.

We also request permission to use the test data for our research explained above.

If you give us permission to collect data from you and to use the data for our research, please sign this form and send it back to your English instructor as soon as possible. We strongly believe that this research will help you to accurately measure and effectively improve your English oral proficiency in short time. If you have any questions regarding this request, please contact me via e-mail at jeong.32@osu.edu, and my phone number in Korea is 545-0517.

Thank you very much.

Tae-Young Jeong, Graduate Administrative Assistant

Associate Dean’s Office, The School of Education, The Ohio State University
Tel: 614) 292-9723

Class:        Student Number:
Name:        Dates:
Signature:  

128
Appendix C: Electronic literacy survey (Pilot survey)

- Electronic literacy (electronic literacy) [电子素养] (d-VOCI [电子口试]) [O] [O] [O] [O] [O] [O] [O].
- This survey is designed to measure your electronic literacy as well as your attitudes towards the d-VOCI and the face-to-face English Oral Interview.

PART A: Electronic Literacy

Please answer every question by circling the item which best describes your opinion. Thank you very much.

<table>
<thead>
<tr>
<th>SD</th>
<th>Strongly disagree</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Disagree</td>
<td>2</td>
</tr>
<tr>
<td>SLD</td>
<td>Slightly disagree</td>
<td>3</td>
</tr>
<tr>
<td>SLA</td>
<td>Slightly agree</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td>Agree</td>
<td>5</td>
</tr>
<tr>
<td>SA</td>
<td>Strongly agree</td>
<td>6</td>
</tr>
<tr>
<td>NA</td>
<td>Not applicable</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Example: Learning English is fun.**

<table>
<thead>
<tr>
<th>SD</th>
<th>D</th>
<th>SLD</th>
<th>SLA</th>
<th>A</th>
<th>SA</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

129
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>e-mail is quite convenient for me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2.</td>
<td>My typing speed is fast enough for chatting and e-mail writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3.</td>
<td>From time to time, I use the Internet to communicate with people of various cultural backgrounds.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4.</td>
<td>I send questions to my professors through e-mail.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5.</td>
<td>I understand Netiquette, and observe Netiquette as much as possible.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6.</td>
<td>I enjoy reading and writing articles on various website bulletin boards.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7.</td>
<td>I occasionally use the Internet to chat in English with non-native speakers of Korean.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8.</td>
<td>I am an active member of Internet-based communities.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9.</td>
<td>I exchange information using P-to-P based software (Napster, Soribada, Guruguru, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>10.</td>
<td>I communicate with non-native speakers of Korean via email.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>I frequently communicate using a cellular phone and/or e-mail.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12.</td>
<td>I use e-mail to communicate with my friends and family members.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>13.</td>
<td>Communicating with foreigners through the Internet can improve my English proficiency.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>14.</td>
<td>E-mail and/or other on-line communication methods are less efficient than face-to-face interaction.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>15.</td>
<td>E-mail and/or other on-line communication methods are less convenient than face-to-face interaction.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>16.</td>
<td>I can create web pages.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>17.</td>
<td>I am proficient in/familiar with web editing software.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>18.</td>
<td>I can explain what a ‘Hypertext’ is.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>19.</td>
<td>I understand the copyright issues regarding on-line documents.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>20.</td>
<td>I can use graphics software such as ‘Photo Shop,’ and/or ‘Paint Shop.’</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
21. **Firewall** (Firewall) 

I can explain what an ‘Internet Firewall’ is.

22. I can download electronic documents, pictures, and/or movie/sound files from the Web onto my computer.

23. I can create movie files.

24. I can edit audio/video files.

25. I can create advanced HTML files.

26. Having my own homepage is quite useful.

27. Making my homepage is difficult.

28. I am interested in new web design methods.

29. I am not confident with editing multimedia.

30. I can use my homepage to express myself to others.

31. I use multimedia for my class presentations.

32. Using the Internet search engines is convenient for me.
<table>
<thead>
<tr>
<th></th>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>33.</td>
<td>I tend to use the Internet when creating or planning something.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.</td>
<td>I subscribe to one or more on-line newspaper(s) and/or magazine(s) written in English.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td>I can make appropriate and legal citations of on-line documents.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td>I cannot explain the differences among various Internet search engines.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td>I use the Internet to read the news.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38.</td>
<td>I can find books or documents using on-line libraries.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td>The Internet is my primary source of information on other cultures/cultures other than my own.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td>The Internet is helpful for my acquisition of second/foreign languages.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td>I do not use the Internet for my homework/projects.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td>I am following the current events using the Internet.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I use electronic dictionaries as part of my acquisition of other languages.

The Internet is very helpful for my studies.

The Internet is the most effective way to gather information.

**PART B: d-VOCI [omens] Attitudes towards d-VOCI**

The following questions are designed to measure your attitudes towards d-VOCI. Your honest answers will help for us to develop more effective and efficient English Oral Proficiency Assessment.

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I want to take the d-VOCI test again.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>It was interesting to take the d-VOCI test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>It was fun to take the d-VOCI test.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>I believe that this d-VOCPI test fairly represents my English Proficiency better than other tests such as writing, reading, and Listening tests.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>The interview questions were difficult to answer.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

example) I want to take the d-VOCI test again

( 4 )
50. I couldn’t understand the interview questions.

51. The multimedia in the software helped me to answer the interview questions.

52. I believe that the d-VOCI will help me improve my English oral proficiency.

53. I am curious about the result of my d-VOCI test that will be evaluated by English OPI specialist.

54. The procedure of the test (handling computer and proceeding Test, etc) was difficult.

55. a. Yes   b) No

I’ve had taken an English Speaking Test before (other than d-VOCI). a) Yes  b) No

(10: novice …….5: intermediate ………10: superior) (Only for those who said “YES” to question 55)

Please describe your previous English Oral Proficiency Test in terms of where, by whom, when.

56. Name:
57. Gender:   (male)   (female)
58. Years of English learning:   (years)
59. d-VOCI

On a scale of one through ten, my proficiency in English would rank .

(1:….5:….10: ) (1: novice….5: intermediate….10: superior)

Thank you very much!!
Appendix D: Electronic literacy survey (Final version)

- (electronic literacy) (d-VOCI (O) °ú ·µÎ°¡Áö ÇüÅÂÀÇ ¿µ¾î ¿©·¯ºÐÀÇ ÀǰßÀ» ÁøÁ¤Çϱâ)

- This survey is designed to measure your electronic literacy as well as your attitudes towards the d-VOCI and the face-to-face English Oral Interview.

PART A: Electronic Literacy

This survey is designed to measure your electronic literacy as well as your attitudes towards the d-VOCI and the face-to-face English Oral Interview.

Please answer every question by circling the item which best describes your opinion. Thank you very much.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>Strongly disagree</td>
<td>1</td>
<td>D</td>
<td>Disagree</td>
<td>2</td>
<td>SLD</td>
<td>Slightly disagree</td>
</tr>
<tr>
<td>SLA</td>
<td>Slightly agree</td>
<td>4</td>
<td>A</td>
<td>Agree</td>
<td>5</td>
<td>SA</td>
<td>Strongly agree</td>
</tr>
<tr>
<td>NA</td>
<td>Not applicable</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example) [ ]  [ ]  [ ]  [ ]  [ ]  [ ]  [ ]  [ ]

Learning English is fun.

<table>
<thead>
<tr>
<th>Learning English is fun.</th>
<th>SD</th>
<th>D</th>
<th>SLD</th>
<th>SLA</th>
<th>A</th>
<th>SA</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1.</td>
<td>Using e-mail for communication is quite convenient.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2.</td>
<td>My typing speed is fast enough for chatting and e-mail writing.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3.</td>
<td>From time to time, I use the Internet to communicate with people of various cultural backgrounds.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4.</td>
<td>I send questions to my professors through e-mail.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5.</td>
<td>I understand Netiquette, and observe Netiquette as much as possible.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6.</td>
<td>I enjoy reading and writing articles on various website bulletin boards.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7.</td>
<td>I occasionally use the Internet to chat in English with non-native speakers of Korean.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>8.</td>
<td>I am an active member of Internet-based communities.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>9.</td>
<td>I exchange information using P-to-P based software (Napster, Soribada, Guruguru, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>10.</td>
<td>I communicate with non-native speakers of Korean via email.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
11. I frequently communicate using a cellular phone and/or e-mail.

12. I use e-mail to communicate with my friends and family members.

13. Communicating with foreigners through the Internet can improve my English proficiency.

14. E-mail and/or other on-line communication methods are less efficient than face-to-face interaction.

15. E-mail and/or other on-line communication methods are less convenient than face-to-face interaction.

16. I can create web pages.

17. I am proficient in/familiar with such software as ‘Dreamweaver,’ ‘Front Page’ and/or ‘Namo Web editor.’

18. I can explain what a ‘Hypertext’ is.

19. I understand the copyright issues regarding on-line documents.

20. I can use graphics software such as ‘Photo Shop’ and/or ‘Paint Shop.’
21. **‘Paint Shop.’**
I can explain what an ‘Internet Firewall’ is.

22. I can download electronic documents, pictures, and/or movie/sound files from the Web onto my computer.

23. I can create movie files for ‘Real Player’ or ‘Windows Media Player.’

24. I can edit audio/video files.

25. I can create FLASH/Dynamic HTML files.

26. Having my own homepage is quite useful.

27. Making my homepage is difficult.

28. I am interested in new web design methods.

29. I am not confident with editing multimedia.

30. I can use my homepage to express myself to others.

31. I use multimedia (‘Power Point’ and computer projector) for my class presentations.
32. Using the Internet search engines (‘YAHOO,’ ‘GOOGLE,’ ‘ALTAVISTA,’ etc.) is convenient for me.

33. I tend to use the Internet when creating or planning something.

34. I subscribe to one or more on-line newspaper(s) and/or magazine(s) written in English.

35. I can make appropriate and legal citations of on-line documents.

36. I cannot explain the differences among various Internet search engines.

37. I use the Internet to read the news.

38. I can find books or documents using on-line libraries.

39. The Internet is my primary source of information on other cultures/cultures other than my own.

40. The Internet is helpful for my acquisition of second/foreign languages.

41. I do not use the Internet for my homework/projects.
42. I am following the current events using the Internet.

43. I use electronic/on-line dictionaries as part of my acquisition of other languages.

44. The Internet is very helpful for my studies.

45. The Internet is the most effective way to gather information.

PART B: d-VOCI Attitudes towards d-VOCI

The following questions are designed to measure your attitudes towards d-VOCI. Your honest answers will help for us to develop more effective and efficient English Oral Proficiency Assessment.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strongly disagree</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Slightly disagree</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I don’t know</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Slightly agree</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Strongly agree</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>Not applicable</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

46. I want to take the d-VOCI test again.

47. It was interesting to take the d-VOCI test.

48. It was fun to take the d-VOCI test.
I believe that this d-VOCI test fairly represents my English Proficiency better than other tests such as writing, reading, and Listening tests.

49. The interview questions were difficult to answer.

50. I couldn’t understand the interview questions.

52. The multimedia in the software helped me to answer the Interview questions.

53. I believe that the d-VOCI will help me improve my English oral proficiency.

54. The procedure of the test (handling computer and proceeding Test, etc) was difficult.

55. I’ve had taken an English Speaking Test before (other than d-VOCI). a) Yes   b) No

(10 9 8 7 6 5 4 3 2 1) (Only for those who said “YES” to question 55)

Please describe your previous English Oral Proficiency Test in terms of where, by whom, when.

56. Name:
57. Gender: (male) (female)
58. Years of English learning: (years)
59. d-VOCI: 

60. On a scale of one through ten, my proficiency in English would rank (   ).

(1: novice ....5: intermediate ....10: superior)

Thank you very much!!
Appendix E: Script for the English d-VOCI

1. Could you introduce yourself? Tell me something about yourself and your family.

2. What subject would you like to major in KMA? And why do you want to study that subject?

3. What you would like to do after you graduate from KMA?

4. What is your hobby? How would you convince your friend to join you in your hobby?

5. Now, look at the picture on your screen. What do you see in this picture? How would you describe it?

6. Dialog
   Wonseok: Hi, Cheol-Soo! When did you get back from China?
   Cheol-Soo: Hi, Won-Seok! I just got back in town last week.
   Wonseok: Really?

   Question:
   Imagine that your friend Cheol-Soo just came back from a trip to China. Ask him at least three questions about the trip.

7. Dialog
   Man 1: Excuse me!
   Man 2: Yes?
Man 1: Would you tell me how to get to Kyung-Bok Palace?
Man 2: Sure, hmm. Let’s see...

Question:
Help the stranger to find Kyung-Bok Palace by giving him a direction referring the map you have.

8. KMA places a heavy emphasis on English. What do you think is the best way to learn English?

9. Dialogs
Won-Seok: What’s wrong with you Cheol-Soo?
Cheol-Soo: Hi, Wonseok! Just another stressful day... I don’t know why my English test score is so bad...

Won-Seok: Hmm...

Question:
What are some things that can cause stress in your daily life? What do you do to lower your stress?

10. Tell about a current event you have read or heard from the media. What was interesting about the event?
Appendix F: Main Screen of d-VOCI

Welcome to Digital-VOCl.

Press the ENTER key or click the mouse to continue.
Appendix G: d-VOCI log-in screen
Appendix H: d-VOCI interview screen
Appendix I: d-VOCI video screen
Appendix J: Summary of Reliability Estimates for Electronic Literacy scales

Electronic Communication

**RELIABILITY ANALYSIS - SCALE (ALPHA)**

<table>
<thead>
<tr>
<th>Statistics for Scale</th>
<th>Mean</th>
<th>Variance</th>
<th>Std Dev</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>53.9514</td>
<td>141.8368</td>
<td>11.9095</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Means</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.5968</td>
<td>3.0347</td>
<td>3.9583</td>
<td>.9236</td>
<td>1.3043</td>
<td>.1165</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Variances</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.3441</td>
<td>1.1526</td>
<td>1.7373</td>
<td>.5847</td>
<td>1.5073</td>
<td>.0237</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inter-item Correlations</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.4309</td>
<td>.2274</td>
<td>.7022</td>
<td>.4749</td>
<td>3.0887</td>
<td>.0110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item-total Statistics</th>
<th>Scale</th>
<th>Corrected Scale</th>
<th>Corrected Item if Item Deleted</th>
<th>Corrected Item if Item Deleted</th>
<th>Corrected Total Item-Total Correlation</th>
<th>Corrected Multiple Item-Multiple Correlation</th>
<th>Corrected Alpha if Item Deleted</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>Mean</td>
<td>Variance</td>
<td>Item-</td>
<td>Item-</td>
<td>Correlation</td>
<td>Correlation</td>
<td>Alpha</td>
<td>Alpha</td>
</tr>
<tr>
<td>Scale</td>
<td>Mean</td>
<td>Variance</td>
<td>Item-</td>
<td>Item-</td>
<td>Correlation</td>
<td>Correlation</td>
<td>Alpha</td>
<td>Alpha</td>
</tr>
<tr>
<td>Scale</td>
<td>Mean</td>
<td>Variance</td>
<td>Item-</td>
<td>Item-</td>
<td>Correlation</td>
<td>Correlation</td>
<td>Alpha</td>
<td>Alpha</td>
</tr>
<tr>
<td>Scale</td>
<td>Mean</td>
<td>Variance</td>
<td>Item-</td>
<td>Item-</td>
<td>Correlation</td>
<td>Correlation</td>
<td>Alpha</td>
<td>Alpha</td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Sq.</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between People</td>
<td>1352.1773</td>
<td>143</td>
<td>9.4558</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within People</td>
<td>1765.6000</td>
<td>2016</td>
<td>.8758</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Measures</td>
<td>234.7704</td>
<td>14</td>
<td>16.7693</td>
<td>21.9307</td>
<td>.0000</td>
</tr>
<tr>
<td>Residual</td>
<td>1530.8296</td>
<td>2002</td>
<td>.7647</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3117.7773</td>
<td>2159</td>
<td>1.4441</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Grand Mean 3.5968

Reliability Coefficients 15 items

**Alpha** = .9191  Standardized item alpha = .9191
Cyberspace Construction

### Reliability Analysis - Scale (Alpha)

<table>
<thead>
<tr>
<th>Statistics for</th>
<th>Mean</th>
<th>Variance</th>
<th>Std Dev</th>
<th>N of Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>47.5208</td>
<td>153.6219</td>
<td>12.3944</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Means</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item Variances</td>
<td>Mean</td>
<td>Minimum</td>
<td>Maximum</td>
<td>Range</td>
<td>Max/Min</td>
<td>Variance</td>
</tr>
<tr>
<td>Inter-item Correlations</td>
<td>Mean</td>
<td>Minimum</td>
<td>Maximum</td>
<td>Range</td>
<td>Max/Min</td>
<td>Variance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item-total Statistics</th>
<th>Scale</th>
<th>Scale</th>
<th>Corrected</th>
<th>Item-</th>
<th>Squared</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>if Item Deleted</td>
<td>if Item Deleted</td>
<td>Total</td>
<td>Multiple</td>
<td>if Item Deleted</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Q16 | 44.8611 | 134.2183 | .6844 | .7135 | .9226 |
| Q17 | 44.8403 | 134.4289 | .7235 | .7497 | .9217 |
| Q18 | 44.4306 | 130.3728 | .7708 | .6587 | .9199 |
| Q19 | 44.3819 | 132.8391 | .6891 | .5214 | .9224 |
| Q20 | 44.5972 | 129.9066 | .8050 | .6788 | .9189 |
| Q21 | 43.5764 | 135.1969 | .6163 | .5990 | .9246 |
| Q22 | 43.6319 | 132.6398 | .6987 | .6853 | .9221 |
| Q23 | 44.7431 | 136.5839 | .6464 | .5237 | .9238 |
| Q24 | 44.6667 | 134.1678 | .6669 | .6133 | .9231 |
| Q25 | 43.9306 | 135.0441 | .5929 | .5244 | .9254 |
| Q26 | 44.4861 | 140.0418 | .4856 | .3298 | .9280 |
| Q27 | 44.3333 | 134.4615 | .6477 | .5257 | .9236 |
| Q29 | 44.4306 | 138.9002 | .5098 | .3683 | .9274 |
| Q30 | 43.7431 | 136.3181 | .5401 | .5173 | .9270 |

### Analysis of Variance

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Sq.</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between People</td>
<td>1464.5292</td>
<td>143</td>
<td>10.2415</td>
<td></td>
<td>.9154</td>
</tr>
<tr>
<td>Within People</td>
<td>1845.4667</td>
<td>2016</td>
<td>.9154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Measures</td>
<td>373.2528</td>
<td>14</td>
<td>26.6509</td>
<td>36.2550</td>
<td>.0000</td>
</tr>
<tr>
<td>Residual</td>
<td>1472.2139</td>
<td>2002</td>
<td>.7354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3309.9958</td>
<td>2159</td>
<td>1.5331</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Mean</td>
<td>3.1681</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Reliability Coefficients

**Alpha = .9282**  
**Standardized item alpha = .9282**
### Academic Problem-Solving Research

**RELIABILITY ANALYSIS - SCALE (ALPHA)**

<table>
<thead>
<tr>
<th>Statistics for</th>
<th>Mean</th>
<th>Variance</th>
<th>Std Dev</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>60.3472</td>
<td>108.6618</td>
<td>10.4241</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Means</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.0231</td>
<td>3.3542</td>
<td>4.3819</td>
<td>1.0278</td>
<td>1.3064</td>
<td>.0947</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item Variances</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1067</td>
<td>.8737</td>
<td>1.3492</td>
<td>.4755</td>
<td>1.5443</td>
<td>.0255</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inter-item Correlations</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
<th>Max/Min</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.4008</td>
<td>.0741</td>
<td>.7220</td>
<td>.6478</td>
<td>9.7372</td>
<td>.0140</td>
</tr>
</tbody>
</table>

**Item-total Statistics**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Variance</th>
<th>Item- Deleted</th>
<th>Item- Deleted</th>
<th>Total Correlation</th>
<th>Multiple Correlation</th>
<th>if Item Deleted Correlation</th>
<th>Alpha Deleted</th>
<th>if Item Correlation</th>
<th>Alpha if Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Corrected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q31</td>
<td>56.5972</td>
<td>97.9625</td>
<td>.4819</td>
<td>.3243</td>
<td></td>
<td></td>
<td></td>
<td>.9057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q32</td>
<td>55.9931</td>
<td>95.6713</td>
<td>.6627</td>
<td>.5186</td>
<td></td>
<td></td>
<td></td>
<td>.8999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q33</td>
<td>56.2708</td>
<td>94.9122</td>
<td>.6170</td>
<td>.4974</td>
<td></td>
<td></td>
<td></td>
<td>.9011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q34</td>
<td>56.9931</td>
<td>98.6223</td>
<td>.3767</td>
<td>.3399</td>
<td></td>
<td></td>
<td></td>
<td>.9104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q35</td>
<td>56.4306</td>
<td>94.8063</td>
<td>.5904</td>
<td>.4109</td>
<td></td>
<td></td>
<td></td>
<td>.9021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q36</td>
<td>56.7014</td>
<td>98.6585</td>
<td>.4737</td>
<td>.3486</td>
<td></td>
<td></td>
<td></td>
<td>.9059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q37</td>
<td>56.1181</td>
<td>94.0349</td>
<td>.7128</td>
<td>.5878</td>
<td></td>
<td></td>
<td></td>
<td>.8980</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q38</td>
<td>56.3194</td>
<td>95.0581</td>
<td>.5925</td>
<td>.4827</td>
<td></td>
<td></td>
<td></td>
<td>.9020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q39</td>
<td>56.1667</td>
<td>91.9301</td>
<td>.7352</td>
<td>.6205</td>
<td></td>
<td></td>
<td></td>
<td>.8966</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q40</td>
<td>56.7361</td>
<td>94.9648</td>
<td>.5609</td>
<td>.4061</td>
<td></td>
<td></td>
<td></td>
<td>.9033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q41</td>
<td>56.1736</td>
<td>95.4871</td>
<td>.5281</td>
<td>.4665</td>
<td></td>
<td></td>
<td></td>
<td>.9046</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q42</td>
<td>56.2847</td>
<td>93.5477</td>
<td>.6685</td>
<td>.5552</td>
<td></td>
<td></td>
<td></td>
<td>.8992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q43</td>
<td>56.0972</td>
<td>92.4660</td>
<td>.7231</td>
<td>.6753</td>
<td></td>
<td></td>
<td></td>
<td>.8972</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q44</td>
<td>56.0139</td>
<td>95.0907</td>
<td>.6692</td>
<td>.6526</td>
<td></td>
<td></td>
<td></td>
<td>.8995</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q45</td>
<td>55.9653</td>
<td>95.9918</td>
<td>.6409</td>
<td>.4744</td>
<td></td>
<td></td>
<td></td>
<td>.9006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Analysis of Variance**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Sq.</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between People</td>
<td>1035.9093</td>
<td>143</td>
<td>7.2441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within People</td>
<td>1528.9333</td>
<td>2016</td>
<td>.7584</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Measures</td>
<td>190.9259</td>
<td>14</td>
<td>13.6376</td>
<td>20.4053</td>
<td>.0000</td>
</tr>
<tr>
<td>Residual</td>
<td>1338.0074</td>
<td>2002</td>
<td>.6683</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2564.8426</td>
<td>2159</td>
<td>1.1880</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Grand Mean** 4.0231

**Reliability Coefficients** 15 items

<table>
<thead>
<tr>
<th>Alpha</th>
<th>Standardized item alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>.9077</td>
<td>.9094</td>
</tr>
</tbody>
</table>

151