

## **KOREAN UNIVERSITY STUDENTS' RANKS OF DIFFICULTY AND FREQUENCY OF ATTITUDE ON TEST-TAKING**

**Dr. Ian Done D. Ramos\***

\*The Graduate School of the Department of English Language and Literature, The University of Suwon, 17-gil Wauan, Bongdam-eup, Hwaseong City, Gyeonggi-do, South Korea 445-743. Tel: +82) 31-220-2503/2287.

---

**ABSTRACT:** *This study determined Korean university students' ranks of difficulty as well as rank and frequency of attitude on test-taking. Specifically, difficulties on 1) test instructions, 2) test contents, 3) test types, and 4) frequency of attitude towards testing were determined by employing the exploratory-quantitative-qualitative-interpretative type of research design. Through students' survey questionnaire and focused group discussion as well as teachers' interview, the findings are concluded that: 1) having been taught basic grammar points and vocabulary with the use of the learning materials, the students could still not communicate confidently and spontaneously in conversation practice that made them not able to display desirable competence in some written tests and activities; 2) students' attitudes made an impact in learning the language as most of them did not see the importance of it for future life, because they did not exert effort to learn by studying religiously and participating in all activities; 3) the students were not able to develop their linguistic and communicative potentials because only an hour and fifty minutes were spent for each class in a week, covering the four macro skills with grammar and vocabulary enrichments; and, 4) for the teachers, tests were difficult to construct for validating language proficiencies because extremely poor and extremely good students were made to attend in one class altogether. It is implied that the students were not emotionally- and intellectually-prepared. To establish remedy, the teachers should provide 1) elaborated lecture-discussion to internalize knowledge, 2) appropriate teaching techniques to assimilate and accommodate language inputs, 3) sufficient examples of a test with various test types to demonstrate well their language skills, and 4) constant dialogue with them to establish rapport in a language testing environment and language classroom learning in general. The recommendations made by Lado (1961) and Jianrattanapong (2011) may also work in this study.*

**KEYWORDS:** Language Testing, Personality Factor, Short History of English Language Testing, Types of Testing Based On Purpose and Interpretation

---

### **INTRODUCTION**

In South Korea, it is implied that the best teachers are the ones who can make a good test, so universities accept students and companies hire job applicants based on examination results. The goal of testing is always to arrive at intelligent decisions as expounded by Carroll (1961):

"sometimes, these decisions affect only the future design of use of the tests themselves, in which case we are dealing with solely experimental uses of tests. Sometimes, the decisions have to do with the retention or alteration of courses of training as when one decides that poor test results are due to ineffective training. Most often, the decisions have to do with the management of the educational careers of individuals." (p. 314)

Nam's (2008) study indicated that "many Korean students do not study English with the ultimate goal of achieving communicative proficiency but rather obtain a high grade to enter a prestigious university." This is also supported by Kroeker's (2009) study in a South Korean University indicating that "the primary role of English conversation classes is not to build communicative competence." As a consequence, Korean university students tend to show consciousness, anxiety, and hesitations intensively on test-taking and test results, based on general observation. Both studious and non-studious students cannot deny these feelings because written test results are the common accepted measurement of their total language performance, not necessarily the actual conversation or oral skills.

To pass the examination or test, English teachers teach with "the excessive emphasis on grammar and reading" (Zsoldos, 2010). Along with this, there are local English teachers "who do not have the adequate command in English" were also tasked to teach the class (Peters, 2010). It could then be implied that local teachers (or even foreign teachers perhaps) lack the important foundations of teaching. Thus, students also lack motivation and retention span to absorb language inputs and develop linguistic and communicative skills, and these may also affect their attitudes towards testing.

To investigate students' real issues on language testing at the university level, this study is guided with the following questions: 1) what difficulties do first and second year students and English teachers encounter in the English language classes in terms of: 1.1) test instructions, 1.2) contents of test, 1.3) vocabulary, and 1.4) types of tests?; 2) what attitudes related to these difficulties are shown by first and second year students in English classes?; 3) how can these difficulties be addressed in relation to their attitude that is observed in English classes of the first and second year levels?; and 4) corollary to the above questions, what inputs for test-taking challenges can be proposed to address the students' dilemma.

This study aims to determine specific, detailed findings because grades of English language courses taken at the university are primarily the basis for job employment in Korea aside from obtaining good result in TOEIC or other standard exams. It is also important to note that their test-taking strategies and emotional stability should reach a desirable level of maturity so as to survive on test-taking.

## **LITERATURE REVIEW**

### **Short History of English Language Testing**

Testing has been a part of human's experience since the day Adam and Eve in the Genesis chapter were tested by a serpent. Testing is one way to measure amount of trust, existing performance, attitude or behavior, and general ability to sustain in a difficult situation.

As for modern language testing, its history began to grow when soldiers were required to take foreign language skills due to World War I, and this is the time the US army started to construct and develop the Army Alpha tests to be used for people's linguistic ability measurement (Kluitmann, 2008, p. 5). He reported that with "Vilareal's Test of Aural Comprehension in 1947 and Lado's Measurement in English as a Foreign Language in 1949" (Kunnan 1999, p. 707), the language testing became an important topic for classroom practice as basis for scientific investigation.

Lado and other scholars have continued the investigation of language testing. Kluitmann (2008) reported that: 1) in 1961, Lado used 'discrete point testing' based on his structuralist approach, assuming that the four macro skills, such as listening, reading, writing, and speaking, are "independent from one another and can therefore be assessed separately" (p. 5); 2) in the late 1970s, John Oller proposed pragmatic approach to testing based on the concept that "language proficiency was indivisible and could consequently not be tested in part" (pp. 5-6); and 3) during the 1980s, the communicative testing approach (first designed by Canale and Swain) came in targeting 'actual testing' which means that "we should aim at providing candidates with authentic stimuli and tasks as well as testing them while they are engaged in some sort of communication, be it active or passive" (p. 6).

The development of English language testing has further continued up to the present. Types of testing were formulated and now are being improved by various scholars for different purposes.

### **Language Testing**

There is no question that language testing is considered as a valid, reliable, and usable tool to measure one's language performance if it is done objectively and efficiently. According Rattu (2011), language testing "can be used to prepare 1) learning and to evaluate 2) students; it can also be used as a reinforcement 3) for attaining the course objectives and even sometimes as the communicative evaluation 4) of the teaching and learning process." Allison (1999) explains that language testing "involves the assessment of some or all aspects of language ability of individuals in some context (not necessarily that of a class) and some set of purposes (not necessarily common to all parties)" (p.5).

To assess students' language ability, teachers or evaluators have various reasons why tests are conducted. Students' performance is measured based on the following types

of test: proficiency, diagnostic, placement, achievement, and language aptitude (see Dewi & Nastiti, 2012; Butt, 2007; McNamara, 2000).

*Proficiency Test* (e.g. TOEFL) is given when the purpose is to measure students' universal knowledge of a language based on their collective training or exposure. The test contains standardized multiple-choices item on different language areas. *Diagnostic Test* (e.g. teacher-made test) is taken by students when the purpose is to determine their specific strengths and weaknesses of a language use and knowledge-based. Often, this test is designed in response to students' real needs by developing a language program, syllabus, and materials. *Placement Test* (e.g. ESL placement test or a university entrance exam) refers to a measurement intended to determine students' appropriate placement into a particular section of a class at the university or any institutions. A student's result from the test determines content, materials, and activities to be used to improve his/her proficiency. *Achievement Test* (e.g. major exams at the university) is used to determine whether objectives of a particular course have been met at the end of instruction period. It is summative in nature in the sense that students' amount of learning and skills performance are measured with what has been taught since the beginning of instruction. And, *Language Aptitude Test* (e.g. the University of Oxford Aptitude Test) is given to predict students' success if they are to be trained in a foreign language. The challenge is how easily and quickly a learner acquire or learn a new language in a given language training program duration.

These types of tests based on purpose go along with two kinds of tests based on score interpretation. These tests are norm-referenced and criterion-referenced (see the Glossary of Education Reform, 2013; James, McInnis, & Devlin, 2002; Stiggins, 1994). *Norm-Referenced Test* is utilized when scores to be interpreted are based on the content that discriminates low language performers from high language performers. In other words, student's score rank obtained from his/her collective experience or training and knowledge from formal learning setting is compared with others, and that determines the exact interpretation in norm-referenced. Whereas, *Criterion-Referenced Test* is utilized when scores to be interpreted are based on the knowledge and skills a student has obtained from particular objectives of a language curriculum or program. In other words, students are expected to have mastered a language and now their scores are being interpreted as how much they know.

The Language Policy Division (2002) in Strasbourg notes that any tests should consider the proper balance among the following: "1) level of difficulty (in terms of the mean difficulty of the test tasks/items and the range of difficulty covered); 2) content (in terms of the topics or subject matter); 3) coverage (in terms of the representativeness of tasks and testing focus); 4) gradedness (in terms of whether the test becomes progressively more difficult); and, 5) item types or test tasks (in terms of the differing cognitive demands they make on test-takers)" (p. 23)

With the considerations mentioned above, Korean students are aware of the test types that have made them study to pass exams related to employment and academic requirements. Kim Eun-gyong (2008) reported the Samsung Economic Research Institute's (SERI) findings on students' enthusiasm on test-taking:

"Koreans spend about 15 trillion won (\$15.8 billion) on English learning per year. Koreans also topped the applicant list of the Test of English as a Foreign Language (TOEFL) between 2004 and 2005 as about 102,340 out of the 554,942 applicants were Koreans. They also paid 700 billion won toward English examination fees. Enthusiasm for English study has also seen a large number of children, teenagers and even adults going to English-speaking countries like the United States, Australia and Britain to study. Last year some 250,000 under 29 years old went abroad for studying." (p. 1)

For university entrance exam or other classroom exams, schools have also made their own standardized exam or teacher-made exam, considering these types of tests based on purpose and interpretation. However, these are not only used to determine student's progress of ability, but also other forms of assessment such as observations, checklists, every day interaction, and many more to effectively identify students' strengths and weaknesses in the classroom language learning.

### **Empirical Studies on Testing and Attitude towards Testing**

It is important to note that difficulties on language testing on various areas are not only derived from test items or test types themselves, but also their attitude towards testing.

Hoshino's (2013) study made use of multiple-choice vocabulary tests in sentence level, considering the degree of difficulty, from simplest to most difficult. He measured up the effect of different types of distractors committed by 372 students from seven universities in Japan. The results were: "tests with syntagmatically related distractors were generally the most difficult, and tests with unrelated distractors, the easiest; the paradigmatically related distractors remained in the middle" (p. 1). He emphasized that "this difference disappeared when the test takers could not use contextual information, which indicates that test takers strongly rely on contextual information in taking multiple choice vocabulary tests in context" (p.1).

Iseni's (2011) study made use of taking notes and reviews of errors during tests, interviews, dictations, writing paragraphs, and different essays -- which had been observed from 50 Albanian university students for four years. It was suggested that "the written work errors should be corrected more carefully, because if they are left without correction, these errors can become fossilized" (p. 1). He emphasized that it is "important to attempt to find the most appropriate strategy to correct errors and mistakes and the best way to assess our students' writings" as well as "their work regarding the attitude that should be held towards errors and mistakes of our students" (p.1).

Li, Zhong, and Suen's (2012) study investigated 150 undergraduate students at a university in Beijing in terms of their perceptions on the CET (College Entrance Test) impact directed to their English-learning orientation and affective states. They revealed the findings that: "the majority of the respondents indicated that the CET had

a greater impact on what they studied than on how they studied" (p.1). In addition, "about half of the students reported a higher level of self-efficacy in regard to their overall English ability and some specific English skills as a result of taking or preparing for the CET; however, many students also reported experiencing increased pressure and anxiety in relation to learning English" (p.1).

Joy's (2013) study investigated 133 students' level of anxiety related to pre-, during- and post-test stages in India. The finding is that "test anxiety is a major factor that can impair test performance and lead to underachievement" (p.6). She revealed that the level of anxiety among these students is too huge that resulted to fainting students' performance level that also affects their way of living. Thus, it was recommended that "educating the students on the coping strategies to tackle test anxiety would at least reduce unwanted pressure which is looming large in the contemporary educational setup" (p.6).

Thus, it is implied that students' difficulties on tests are also affected by their attitude, and this attitude towards testing is correlated with personality factor. Krashen (2002) hypothesizes that "the self-confident or secure person will be more able to encourage intake and will also have a lower filter." This is supported by H. D. Brown's (1977, p. 352; cited in Krashen, 2002) view stating: "Presumably, the person with high self-esteem is able to reach out beyond himself more freely, to be less inhibited, and because of his ego strength, to make the necessary mistakes involved in language learning with less threat to his ego."

## **METHODOLOGY**

The research design of this current study is exploratory-quantitative-interpretative. This means that the design is non-experimental, the data collected is primarily quantitative, and the analysis is highly interpretative. The number of respondents was determined by purposive-cluster random sampling with the Slovin's formula. This formula was also used to determine the number of students in the FGD as well as the teachers who were interviewed. Frequency count was determined by the percentage formula. Additionally, it is qualitative approach, since triangulation (considering multiple methods of data collection and multiple sources of information) was considered. Multiple methods of data collection included survey questionnaire, interview, and focused group discussion (FGD). Multiple sources of information were 70 first year students and 58 second year students; 5 professional English teachers and 5 content professors who were made to teach English; and 7 FGD groups (7 to 10 members per group).

## **RESULTS**

Tables 1 to 4 discuss students' difficulties on testing gathered from one of the universities in Jeollabuk-do province, South Korea.

Specifically, Table 4 presents the difficulties as for the test instructions.

**Table 1: Rank of Difficulty on Test Instructions**

	Easy	%	Moderate	%	Difficult	%	Difficulty rank
1. Answering fast in all types of test	16	13	55	42	57	<b>45</b>	<b>1</b>
2. Answering items in each test type of both major exams (midterms and finals) and quizzes	25	20	79	61	24	<b>19</b>	<b>2</b>
3. Following the instructions of both major exams (midterms and finals) and quizzes	19	15	85	66	24	<b>19</b>	<b>2</b>

In the three skills shown above table, *answering fast in all types of test* ranks first (45%); *answering items in each test type of both major exams (midterms and finals) and quizzes* and *following the instructions of both major exams (midterms and finals) and quizzes* second (19%).

Table 2 presents the difficulties as for the contents or test items of quizzes, tests or major exams.

**Table 2: Rank of Difficulty on Test Contents**

	Easy	%	Moderate	%	Difficult	%	Difficulty rank
1. Doing well in practical conversations	9	7	54	42	65	<b>51</b>	<b>1</b>
2. Doing well in test items not found in the textbook and workbook	13	10	67	52	48	<b>38</b>	<b>2</b>
3. Doing well in test items taken from the textbook and workbook	29	23	84	65	15	<b>12</b>	<b>3</b>

The table above shows that among the three skills, *doing well in practical conversations* is at the first rank (51%); followed by *doing well in test items not found in the textbook and workbook* second (38%); and, *doing well in test items taken from the textbook and workbook* third (12%).

Table 3 presents the difficulties as for the types of test in quizzes or major exams.

**Table 3: Rank of Difficulty on Test Types**

	Easy	%	Moderate	%	Difficult	%	Difficulty rank
1. Doing well in dialog	12	9	64	50	52	<b>41</b>	<b>1</b>
2. Doing well in correcting errors	16	13	62	48	50	<b>39</b>	<b>2</b>
3. Doing well in word order	13	10	68	53	47	<b>37</b>	<b>3</b>
4. Doing well in question-response	20	16	70	54	39	<b>30</b>	<b>4</b>
5. Doing well in vocabulary	9	7	83	65	36	<b>28</b>	<b>5</b>
6. Doing well in affirmative and negative statements	27	21	70	55	31	<b>24</b>	<b>6</b>
7. Doing well in reading comprehension	23	18	76	59	29	<b>23</b>	<b>7</b>

In the above table, among the seven skills, *doing well in dialog* is the most difficult; followed by *doing well in correcting errors*; *doing well in word order*; *doing well in question-response*; *doing well in vocabulary*; *doing well in affirmative and negative statements*; and, *doing well in reading comprehension* -- which rank first (41%), second (39%), third (37%), fourth (30%), fifth (28%), sixth (24%), and seventh (23%), respectively.



Table 4 presents the attitudes shown during tests or exams.

**Table 4: Rank and Frequency of Attitude on Tests**

	always	%	usually	%	sometimes	%	seldom	%	never	%	Rank
1. Not finishing the exams or tests when I think they are difficult	1	1	6	5	24	19	33	25	64	<b>50</b>	<b>1</b>
2. Easily complaining when I get low score in the tests, even when I didn't study hard	0	0	6	5	22	17	37	29	63	<b>49</b>	<b>2</b>
3. Studying harder when my score in the tests are good	47	<b>36</b>	43	34	28	22	8	6	2	2	<b>3</b>

It is shown in the above table that among the three situations, *not finishing the exams or tests when I think they are difficult* is the most frequently done (first rank with 50% indicating *never*); followed by *easily complaining when I get low score in the tests, even I didn't study hard* (second rank with 49% indicating *never*); lastly, *studying harder when my score in the tests are good* (third rank with 36% indicating *always*).

## DISCUSSION

Analysis and interpretation of the data below were drawn from the students' FGDs and teachers' interviews.

The data shown in Table 1 (*Rank of Difficulty on Test Instructions*) explain that the poor students could not answer fast in all types of tests, even they were oriented with them as found in the workbook. They were in deficit of understanding the test items, even test construction that implies the inclusion of test instruction was simplified or modified. As for both major exams (midterms and finals) and quizzes, some students could survive though in answering items in each test type because an example was provided before they began to answer, and they were given time for such task. However, as for following the instructions on both major exams (midterms and finals) and quizzes, few students failed on where to write their answers; they wrote them on the questionnaire instead on the answer sheets. Lado (1961) posits that "the decision in the choice of skill versus element will depend to a considerable extent on the

purpose for which the test is made and the available time for testing, plus the conditions under which the test must be given" (p. 27). Element refers to the specific test items from which students react during the test.

The data shown in Table 2 (*Rank of Difficulty on Test Contents*) explain that students could hardly do well in practical conversations due to awkwardness and withdrawal, especially when a teacher tried to converse with them. Such difficulty justifies that practical conversation tests are the most difficult to perform because students were only introduced, oriented, and trained with the structure or patterned dialogues and substitution drills found in the learning materials. What is worse, they could not do well in test items not found in the textbook and workbook, they even hesitated to answer them, because they were oriented with book-based that made them not apply the amount of learning into similar test items or practical situations. In fact, they easily gave up when tests were found beyond their ability. However, most of them could do well in test items whether taken from the textbook and workbook due to their personal reviews. They found substitution drill less difficult because they had improved study habits and were aware of their test-taking strategies. In fact, they felt happy of their accomplishment when they were able to successfully substitute missing information with their own ideas. According to Lado (1961), "a test that has good language content and appears to use essentially linguistic situations will not be valid if it introduces a very heavy intelligence factor, or a heavy memory factor, or some other element that cancels out the language content and the valid situation" (p. 30).

The data shown in Table 3 (*Rank of Difficulty on Test Types*) clearly explain that dialogue was the most difficult for most students, even when words for substitution were presented as choices. Some students could not do well in correcting errors due to limited recognition or analysis of syntactic structure, while others could survive though in word order and question-response. Still others could do well in vocabulary due to translation and context clues that were provided, and in affirmative and negative statements due to patterns that were presented. Reading comprehension was the least difficult due to explicit choices that were found in the texts. In this aspect, the Language Policy Division (2002) argues that "the difficulty of an item cannot be assumed to be a simple result of the linguistic relationship of the text and the answer. Both stimulus and response have their own linguistic features, and the task that bridges them may involve some cognitive complexity in addition to the demands of the language" (p. 24).

Moreover, such difficulties may be drawn from students' poor study habits and less interests in classroom challenges. External factors like point basis and attendance may affect their academic performance positively, but not as much as developing communicative competence. However, other students like from the Nursing Department did well in the exams because they were basically good in accommodating the test types with their respective test items. The Language Policy Division (2002) comments that "the most important criterion for measuring the value of an item type is its appropriateness for use in testing language in a particular situation and for a specified purpose. The item type which provides the most direct

means of measuring the desired learning outcome tends to be the best item type to choose" (p. 29).

The data shown in Table 4 (*Rank and Frequency of Attitude on Tests*) explain that some students never decided of not finishing the exams or tests when they thought these were difficult. In fact, they used up the time given. Moreover, they never easily complained when they got low score in the tests, even they did not study hard. However, at some point in a semester, they blamed the tests as difficult; while others assertively complained, as higher grade is essential in job seeking. Usually, they just studied harder when their scores were good. In fact, they (not all though) sought more for oral practice to develop competence. It is then implied that "when allowed to learn in their favorite way, unpressured by learning environment or other factors, students often use strategies that directly reflect their preferred learning" (Oxford, 2012, p. 127)"

In other words, such frequencies of attitudes clarify that despite students' weaknesses in the entire English classes, they still attended English classes, since A+ is an investment to motivate themselves for attendance and participation in the graded activities, as clearly emphasized in the FGDs and teacher interview. They tried their best by all means, while few did not care much. Some of them confessed their feelings and emotions towards tests and activities when positive feedback was given to them.

### **IMPLICATION TO RESEARCH AND PRACTICE**

Often, most teachers teach English classes following what has already been established in their lesson plans or syllabi, and this is the basis of their tests. To check and balance the teaching inputs with students' exam results, it is important to note that teachers when teaching English classes should first consider students' emotional and intellectual preparedness. This could be done by providing 1) elaborated lecture-discussion to internalize knowledge, 2) appropriate teaching techniques to assimilate and accommodate language inputs, 3) sufficient examples of a test with various test types to demonstrate well their language skills, and 4) constant dialogue with them to establish rapport in a language testing environment and language classroom learning in general. In this way, students would not feel that testing is always difficult. Ramos' (2014) study published in the *International Journal of English Language Education-USA* emphasized that "since Koreans are becoming more aware of other styles of education, their actual experiences, and their directions towards globalization" (p. 172), they would realize the importance of study habits and language test drills on their own. This is one input of making them emotionally- and intellectually-prepared.

### **CONCLUSION AND RECOMMENDATION**

As emphasized in the FGDs and teacher interviews, causes of the findings discussed earlier are concluded as in the following: 1) having been taught basic grammar points and vocabulary with the use of the learning materials, the students could still not communicate confidently and spontaneously in conversation practice that made them not able to display desirable competence in some written tests and activities; 2)

students' attitudes made an impact in learning the language as most of them did not see the importance of it for future life, because they did not exert effort to learn by studying hard and participating in all activities; 3) the students were not able to develop their linguistic and communicative potentials because only an hour and fifty minutes were spent for each class in a week, covering the four macro skills with grammar and vocabulary enrichment; and, 4) for the teachers, tests were difficult to construct for validating language proficiencies because extremely poor and extremely good students were made to attend in one class altogether.

Lado (1961) recommends that "even when we need to give individual tests, we will attempt to devise techniques to do it in the shortest possible time to permit language teachers to use such a test" (p. 28). The use of positive washback can also be recommended. Jianrattanapong's (2011) study on the Thai University Entrance Examination results indicated that it is better "to apply direct tests in the writing part of the exams" (p. 59). She emphasized the possibility to overcome the shortcomings like "subjective and large-scale scoring, for instance, to have effective rating scales, to have well trained and certified raters, to use scoring technologies which have proven their reliability by standardized tests like TOEFL, IELTS" (p.59) as well as teacher-made tests on their own classroom objectives. This is a proper assessment to determine students' strengths and weaknesses in language learning, since the teacher respondents in this study were seemed to have no idea or actual experience in constructing good test items and proper scoring of students' performance.

## **FUTURE RESEARCH**

Since students' emotional and intellectual preparedness as well as teachers' skills on test item construction and student performance scoring were not well taken into consideration, this study will then provide TESOL practitioners, educators, curriculum developers, English majors, and other researchers an instrument to evaluate or investigate the language testing impact in an EFL environment. More future detailed studies related to this current study is to be prioritized seriously so as to arrive at an effective teaching model that directly targets students' test-taking strategies. In other words, teaching model can be designed with test strategy model to assist students in their test-taking challenges. These two models should match each other in curriculum objectives set by a support system. Support system may refer to teachers, a curriculum developer, and students aiming to obtain meaningful language learning and testing.

## **ACKNOWLEDGEMENT**

This paper is supported by the Department of English and Literature of the University of Suwon, Association of Filipino Educators in Korea (AFEK), Ramos' family, Kim Jin Hyuk, Uhm Sunmee, and Kim Joong Woo.

**REFERENCES**

- Allison, D. (1999). *Language testing and evaluation*. Singapore: Singapore University Press.
- Brown, H. D. (1977). *Cognitive and affective characteristics of good language learners*. Paper presented at Los Angeles Second Language Acquisition Research Forum, UCLA. February 1977.
- Butt, F. A. (2007). *Kinds of test*. Retrieved January 30, 2014, from <http://fazalalibutt.blogspot.kr/>
- Carroll, J B (1961). Fundamental considerations in testing for English language proficiency of foreign students. In *Testing Center for Applied Linguistics* (p. 314). Washington, DC. Reprinted in Allen, H B & Campbell R N. (1972). *Teaching English as a Second Language: A Book of Readings*. McGraw Hill, New York.
- Dewi, R. K. & Nastiti, N. S. (2012). Kind of test. Retrieved January 11, 2014, from <http://thejoyoflanguageassessment.wordpress.com/2012/12/19/kind-of-test/>
- Glossary of Education Reform. (2013). *Norm-referenced test*. Retrieved January 29, 2014, from <http://edglossary.org/norm-referenced-test/>
- Hoshino, Y. (2013). Relationship between types of distractor and difficulty of multiple-choice vocabulary tests in sentential context (p.1). *Language Testing in Asia*, 3(16). Heidelberg, Germany: SpringerOpen. Retrieved January 11, 2014, from <http://www.languagetestingasia.com/content/pdf/2229-0443-3-16.pdf>.  
<http://dx.doi.org/10.1186/2229-0443-3-16>
- Iseni, A. (2011). Assessment, testing and correcting students' errors and mistakes (p. 1). *Language Testing in Asia*, 1(3). Heidelberg, Germany: SpringerOpen. Retrieved January 11, 2014, from <http://www.languagetestingasia.com/content/pdf/2229-0443-1-3-60.pdf>
- James, R., McInnis, C., & Devlin, M. (2002). *A comparison of norm-referencing and criterion-referencing methods for determining student grades in higher education*. Retrieved January 29, 2014, from <http://www.cshe.unimelb.edu.au/assessinglearning/docs/NormCritRef.pdf>
- Jianrattanapong, A. (2011). Positive washback from Thai university entrance examinations (p. 59). *Language Testing in Asia*, 1(1). Heidelberg, Germany: SpringerOpen. Retrieved January 11, 2014, from <http://www.languagetestingasia.com/content/pdf/2229-0443-1-1-50.pdf>
- Joy, J. L. (2013). The altitude of anxiety among second language learners (p. 6). *Language Testing in Asia*, 3(10). Heidelberg, Germany: SpringerOpen. Retrieved January 11, 2014, from <http://www.languagetestingasia.com/content/pdf/2229-0443-3-10.pdf>.  
<http://dx.doi.org/10.1186/2229-0443-3-10>
- Kim, E. (2008). History of the English education in Korea (p. 1). *The Korean Times*. Retrieved January 30, 2014, from [http://www.koreatimes.co.kr/www/news/special/2009/10/181\\_21843.html](http://www.koreatimes.co.kr/www/news/special/2009/10/181_21843.html).
- Kluitmann, S. (2008). *Testing English as a Foreign Language*. Retrieved January 23, 2014, from

[http://www.sprachenmarkt.de/fileadmin/sprachenmarkt/wissen\\_images/Wissenschaftliche\\_Arbeit\\_Sebastian\\_Kluitmann\\_TOEIC-KMK.pdf](http://www.sprachenmarkt.de/fileadmin/sprachenmarkt/wissen_images/Wissenschaftliche_Arbeit_Sebastian_Kluitmann_TOEIC-KMK.pdf)

- Krashen, S. (2002). *Second language acquisition and second language learning*. Retrieved August 10, 2011, from [http://sdkrashen.com/SL\\_Acquisition\\_and\\_Learning/index.html](http://sdkrashen.com/SL_Acquisition_and_Learning/index.html).
- Kroeker, R. (2009). *The reality of English conversation classes: A study in a South Korean university*. England, Birmingham: University of Birmingham. Retrieved January 4 2014, from <http://www.finchpark.com/courses/grad-dissert/articles/speaking/Thesis-Kroeker.pdf>
- Kunnan, A.J. (1999). *Language testing: Fundamentals*, (p. 707). In: Spolsky, Bernard & Asher, R.E. (Eds.), 707-711.
- Lado, R. (1965). *Language testing – The construction and use of foreign language tests* (4th ed.). London: Lowe and Brydone.
- Lado, R (1961). *Language testing: The construction and use of the foreign language tests* (pp. 27-30). McGraw-Hill Book Company.
- Language Policy Division. (2002). *Language examining and test development* (pp. 23-29). Retrieved January 29, 2014, from [http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Publications/Language\\_examining\\_EN.pdf](http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Publications/Language_examining_EN.pdf)
- Li, H., Zhong, Q., & Suen H. K. (2012). Students' perceptions of the impact of the college English test (p. 1). *Language Testing in Asia*, 2(3). Heidelberg, Germany: SpringerOpen. Retrieved January 11, 2014, from <http://www.languagetestingasia.com/content/pdf/2229-0443-2-3-77.pdf>. <http://dx.doi.org/10.1186/2229-0443-2-3-77>
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Nam, H. (2008). A case study of improving English writing skills from reading journals. *The Journal of the Research Institute of Korean Education*
- Oxford, R. L. (2012). Language learning strategies in a nutshell: Update and ESL suggestions. In Richards, C. J., & Renanya W. A. (Eds.), *Methodology in language teaching* (p. 127). New York: Cambridge University Press.
- Peters, W. (2010). *Improving English education in Korea*. Retrieved June 13, 2010, from <http://www.seoulstyle.com/art-english-education-korea.htm>
- Ramos, I. D. (2014). The English majors' expectations, experiences, and potentials: Inputs toward Korea's globalization (p. 172). *International Journal of English Language Education*, 2(1). Las Vegas, USA: Macrothink Institute. Retrieved January 29, 2014, from <http://www.macrothink.org/journal/index.php/ijele/article/view/4963/4036>. <http://dx.doi.org/10.5296/ijele.v2i1.4963>
- Rattu, B. R. (2011). *Language testing and evaluation*. Retrieved February 2, 2014, from <http://www.scribd.com/doc/50833648/Language-Testing-and-Evaluation-Test>
- Stiggins, R.J. (1994). *Student-centered classroom assessment*. New York: Merrill
- Zsoldos, L (2010). The state of English education in Korea. In *ALAK Newsletter*. Chung-Ang University, South Korea.

