Utilizing Most Complex type of Test in English as Second Language Course in a Public Junior High School in the Philippines

Eliza G. Barasi

College Of Business And Financial Science, University Of Makati, Philippines.
*Corresponding Author: * Eliza G. Barasi*

Abstract: This research paper investigated the most complex type of test taken by the ESL grade 7 student-participants through test analysis. A total of 30 achievement test results were the instruments used in this study. The data were analyzed and interpreted employing the Item Difficulty Index (I.D.) and the Discrimination Index (D.I.). The findings showed that the student-participants hardly accomplish or answer the tests on True or False, Vocabulary, Fill in the Blanks, and Multiple Choice. Among these types of tests, the most complex is the Multiple Choice.

Keywords: most complex type of test, Item Difficulty Index, Discrimination Index

I. INTRODUCTION

Tests are essential components of language teaching and learning. This claim was supported by [1] when they underscored the significance of administering the various types of tests. They said, "Testing becomes an integral part of teaching because it provides significant information or inputs about the growth and achievement of learner's difficulties, styles of learning, and anxiety levels." It means that tests mirror the students' achievement of what they can learn and what they hardly learn. Thus, learning styles and anxiety levels of the students should likewise be looked into. Through this, teachers can gauge and monitor the learning abilities as well as the learning deficiencies of the students [1, 2,3].

Similarly, teachers can strategize and use their time effectively and efficiently in filling the gap between their students' learning achievements and deficiencies. It is indeed true that effective teaching strategies and effective testing are inseparable [4,5, 12]. Moreover, a "curriculum is what constitutes a total teaching-learning program composed of overall aims, syllabuses, materials, methods, and testing. It provides a framework of knowledge and capabilities, selected to be appropriate to a particular level." This means that possessing effective teaching skills as well as using good materials and assessment tools is imperative. This claim was supported by [1] when they mentioned that the purpose of the test is not only limited to evaluate the progress and achievement of learners but also to evaluate the effectiveness of the teaching materials and methods used.

The definition of the test was unlocked by [6] as a "systematic procedure for observing persons and describing them with either a numerical scale or a category system. Thus, the test may give either qualitative or quantitative information." Another definition of the test was provided by [7] as a "method to determine a student's ability to complete certain tasks or demonstrate mastery of a skill or knowledge of content." He also provided some types of tests/tasks, like multiple-choice tests, or a weekly spelling test. Said tests/tasks can be used interchangeably with assessment or even evaluation. More often than not, teachers use tests or tasks to assess the cognitive levels of the students.

Also, educational institutions all over the world use assessment and evaluation tools in the form of examinations. University of Waterloo [8] provided some types of tests among which are the following: multiple-choice, true/false, matching type, short response, and essay.

II. BACKGROUND OF THE STUDY

Sax posited as to [9] that formulating "fair and systematic evaluations of others' performance can be a challenging task." He also highlighted that "judgments cannot be made solely based on intuition, haphazard guessing, or custom."On the other hand, the teacher who will formulate the test should understand the underlying principles on the development of the test. This includes the process of restricting, defining and selecting the content that should be covered by a test. The process may use the table of specifications (TS),

"which enable the test-developer, see the content to be measured, the level at which it should be measured and how it should be measured." Besides, discussing the guidelines concerning the construction of test items is imperative. The process of choosing effective objective and essay type items or other types of tests is important. Likewise, the test items or questions can be revised carefully before administering the test [10, 11]. It was also emphasized that there are two key concepts in testing: first, is the use of TS. According ly, TS is "essentially a blueprint for a test with the purpose to ensure that all intended outcomes, and only the intended outcomes, are measured and that the test includes the appropriate number of items for each measured item.", while the second concept is the use of Test validation (TV). The TV is the process of administering and revising over and over again until acceptable levels of validity and reliability are achieved.

After administering the test, the teacher may immediately proceed with the Item Analysis (IA). IA was defined by the Educational Development Center [13] as a "statistical technique that helps instructors identify the effectiveness of their test items.

In developing quality assessment and specifically effective multiple-choice test items, IA plays an important role in contributing to the fairness of the test along with identifying content areas that may be problematic for students." In other words, employing IA is essential in validating the learning contents and competencies in which the students may find difficult or problematic.

This research study tried to shed light on the efficacy of learning through most complex type of test taken by the ESL student-participants at the Fort Bonifacio High School (FBHS), Makati City. The objectives of this study were to answer these questions: as the most complex type of test taken by the ESL Grade 7 students at FBHS, and identify the least complex type of test taken by the ESL Grade 7 students at FBHS. The researcher used four (4) types of test items in this study, namely: True or False, Vocabulary, Fill in the Blanks, and Multiple Choice then he administered the same to 30 ESL grade 7 students. The researcher determined the most complex and the least complex types of tests based on the results of the Item Analysis through Item Difficulty Index (I.D.) and Discrimination Index (D.I.).

III. METHODOLOGY

3.1 Participants

Thirty (30) grade 7 ESL student-participants enrolled at the Fort Bonifacio High School were selected to participate in this research investigation. Table 1 shows the demographic distribution of the student-participants according to gender. As can be seen, the majority of the student-participants are female. On the other hand, their male counterparts got a frequency of 14 or almost forty-seven percent.

Table-I: Demographic Distribution of the Student-Participants of the Study According to Gender

Gender	Frequency (f)	Percentage (%)
Male	14	46.66
Female	16	53.33
TOTAL	30	100

3.2 Research Design

The descriptive method of research was used by the researcher because it is believed to be useful and reliable due to its significant features like it is a "fact-finding study with adequate and accurate interpretation of the findings". It likewise emphasizes "what exist such as current conditions, practices, situations, or any phenomena [14,15, 16, 17].

3.3 Instrument

The instrument used in this study was the 40-item achievement test formulated by the researcher to elicit the most complex and the least complex types of tests taken by the ESL grade 7 students based on the results of the test item analysis. The researcher used the results of thirty (30) achievement tests taken by the student-participants administered by the researcher in one of the classrooms at FBHS. The achievement test is composed of four (4) types of tests, namely: True or False, Vocabulary (Similar or Opposite), Fill in the Blanks, and Multiple Choice. The researcher believed that the instruments he used in this study could answer the questions in the objectives of the study.

3.4 Data Collection

The researcher sought permission from the principal of FBHS to conduct a study in one of the Secondary Public Schools in Makati City. The researcher made the class aware that the results of the tests will be used as the primary sources of this study. The student-participants of this study were given 35-45 minutes to accomplish their respective tests.

3.5 Data Analysis

After the data gathering, the researcher completed the following procedures; first, the test papers were

checked and arranged the results of the tests from the highest to the lowest; secondly, identify the highest 27% and the lowest 27% of the examinees (total of 16 examinees); thirdly, computed the Item Difficulty Index (ID) using this formula:

I.D. = $\dfrac{\text{No. of students marking the items correct}}{\text{Total no. of examinees}}$ x 100   (1)

The desirable I.D. is 20% to 80%.
The Discrimination Index (D.I.) was computed using the formula:

D.I. = $\dfrac{RH - RL}{NR}$ x 100   (2)

where,
RH = no. of students in the high scoring group with correct items
RL = no. of students in the low scoring group with correct items
NR = total no. of examinees
The Ideal D.I. = 15% to 70%.

IV. RESULTS AND DISCUSSION

This section presents the results and discussion of the data gathered and utilized in the study based on the research objectives. Table 2 provides the results of the analysis on Item Difficulty Index.

Table- II: Summary of the Results of the Analysis on Item Difficulty Index

Item No.	Computation	Percentage	Remarks
1	16/16 X 100	100%	Too Easy
2	6/16 X 100	37.50%	Ok
3	13 /16 X 100	81.25%	Too Easy
4	5/16 X 100	31.25%	Ok
5	5 /16 X 100	31.25%	Ok
6	10/16 X 100	62.50%	Ok
7	6/16 X 100	37.50%	Ok
8	8/16 X 100	50%	Ok
9	5 /16 X 100	31.25%	Ok
10	2/16 X 100	12.50%	Too Difficult
11	9/16 X 100	56.25%	Ok
12	11/16 X 100	68.75%	Ok
13	12/16 X 100	75%	Ok
14	11/16 X 100	68.75%	Ok
15	6 /16 X 100	37.50%	Ok
16	7/16 X 100	43.75%	Ok
17	5 /16 X 100	31.25%	Ok
18	8/16 X 100	50%	Ok
19	12 /16 X 100	75%	Ok
20	10/16 X 100	62.50%	Ok
21	12 /16 X 100	75%	Ok
22	11/16 X 100	68.75%	Ok
23	8 /16 X 100	50%	Ok
24	16/16 X 100	100%	Too Easy
25	7/16 X 100	43.75%	Ok
26	9/16 X 100	56.25%	Ok
27	4 /16 X 100	25%	Ok
28	12/16 X 100	75%	Ok
29	5 /16 X 100	31.25%	Ok
30	10/16 X 100	62.50%	Item is on average difficulty
31	8 /16 X 100	50%	Item is on average difficulty
32	2/16 X 100	12.50%	Item is very difficult
33	4 /16 X 100	25%	Item is on average difficulty
34	5/16 X 100	31.25%	Item is on average difficulty
35	6/16 X 100	37.50%	Item is on average difficulty
36	8/16 X 100	50%	Item is on average difficulty
37	5/16 X 100	31.25%	Item is on average difficulty
38	8/16 X 100	50%	Item is on average difficulty

| 39 | 4/16 X 100 | 25% | Item is on average difficulty |
| 40 | 1/16 X 100 | 6.25% | Item is very difficult |

As can be seen, almost all of the items are marked "Ok" except for item nos. 1, 3, 10, 24, 32, and 40. Only item nos. 1, 3, and 24 were marked "Too Easy" while item nos. 10, 32 and 40 were marked "Too Difficult". The "Too Easy" and "Too Difficult" test items amounted to 15 percent of the total number of the items.

All the students belonging to both the highest 27% and the lowest 27% answered the item nos. 1 (True or False) and 24 (Fill in the Blanks) correctly while the item nos. 10 (Fill in the Blanks), 32 (Multiple Choice) and 40 (Multiple Choice) found difficult by almost all of the student-participants. Like in the item nos. 10 and 32, only two (2) students got each item correctly while in item no. 40, only one (1) student answered it correctly.

It could be noted, however, that no student belonging to the lowest 27% correctly answered the items marked with "Too Difficult". To obtain the "most significant discrimination between groups" is to get the highest 27% and the lowest 27% of the total number of examinees.

Based on the above findings, it can be inferred that the most complex type of test taken by the ESL grade 7 students is the Multiple Choice while the least complex type of test is the True or False.

Table-III: Summary of the Results of the Analysis on Discrimination Index

Item No.	Computation	Percentage	Remarks
1	8-8/16 X 100	0%	-
2	4-2/16 X 100	12.50%	-
3	8-5/16 X 100	18.75%	Acceptable
4	3-2/16 X 100	6.25%	-
5	3-2/16 X 100	6.25%	-
6	5-5/16 X 100	0%	-
7	5-1/16 X 100	25.00%	Acceptable
8	7-1/16 X 100	37.50%	Acceptable
9	5-0 /16 X 100	31.25%	Acceptable
10	2-0/16 X 100	12.50%	-
11	4-5/16 X 100	6.25%	-
12	7-4/16 X 100	18.75%	Acceptable
13	8-4/16 X 100	25%	Acceptable
14	8-3/16 X 100	31.25%	Acceptable
15	5-1 /16 X 100	25%	Acceptable
16	2-5/16 X 100	18.75%	Acceptable
17	4-1/16 X 100	18.75%	Acceptable
18	6-2/16 X 100	25%	Acceptable
19	7-5 /16 X100	12.50%	-
20	7-3/16 X 100	25%	Acceptable
21	6-6/16 X 100	0%	-
22	8-3/16 X 100	31.25%	Acceptable
23	3-5 /16 X 100	12.50%	-
24	8-8/16 X 100	0%	-
25	3-4/16 X 100	6.25%	-
26	5-4/16 X 100	6.25%	-
27	1-3 /16 X 100	12.50%	-
28	5-7/16 X 100	12.50%	-
29	3-2 /16 X 100	6.25%	-
30	6-4/16 X 100	12.50%	-
31	7-1/16 X 100	37.50%	Acceptable
32	2-0/16 X 100	12.50%	-
33	2-2 /16 X 100	0%	-
34	5-0/16 X 100	31.25%	Acceptable
35	6-0/16 X 100	37.50%	Acceptable
36	6-2/16 X 100	25%	Acceptable
37	3-2/16 X 100	6.25%	-
38	5-3/16 X 100	12.50%	-
39	3-1/16 X 100	12.50%	-
40	1-0/16 X 100	6.25%	-

Table III displays the results of the analysis on the item Discrimination Index. As can be seen, 17 items were marked "Acceptable". This finding would mean that 23 items need a little modification to make them " Acceptable". These items are the following: 1-2, 4-6, 10-11, 19, 21, 23-30, 32-33, 37-40. Osman (n.d.) emphasized one of the two key concepts in testing which is the use of Test validation (TV).

The TV is the process of administering and revising over and over again until acceptable levels of validity and reliability are achieved. Furthermore, Garett (1958) said that "the nearer the percentage is to fifty, the more ideal is the difficulty and the more valid is the test."

V. CONCLUSION AND RECOMMENDATION

Based on the significant findings of this study, the following conclusions were drawn, firstly, the most complex type of test taken by the ESL grade 7 students at FBHS is the Multiple Choice. Secondly, the least complex type of test taken by the ESL grade 7 students at FBHS is the True or False. Lastly, 17 out of 40 test items were marked "Acceptable".

It is significant then that the teachers who will be tasked to formulate the test should understand the underlying principles on the development of the test or the so-called test construction. This includes the process of restricting, defining and selecting the content that should be covered by a test. The process may use the table of specifications (TS) "which enable the test-developer, see the content to be measured, the level at which it should be measured and how it should be measured." In other words, having superficial knowledge, at least, on the guidelines or parameters concerning the construction of test items is imperative.

Also, right after administering the test, teachers may immediately proceed with the Item Analysis (IA) to check the validity and reliability of the test items.

IA plays a pivotal role in contributing to the fairness of the test along with identifying content areas that may be problematic for students. Employing IA then is essential in validating the learning contents and competencies in which the students may find difficult or problematic.

In light of the significant findings in this study, the researcher provided the following recommendations: Firstly, the revision of the "Too Easy" and "Too Difficult" test items is imperative. Secondly, test items which were not marked "Acceptable" may be revised or modified.

Lastly, for the future researchers, employing Triangulation is highly recommended to illicit the most complex and the least complex types of tests using the following instruments, namely: the constructed/formulated tests, structured interviews with the students to elicit their perceptions on the most complex and the least complex types of tests, and another structured interviews with the ESL teachers to elicit their perceptions on the most complex and the least complex types of tests.

Besides, since the researcher of this study utilized a limited number of participants, the future researchers may involve 100-150 participants, or it varies depending on the total number of population to be surveyed, and they may also include 5-6 or more types of tests in their study.

REFERENCES

[1]. Desheng, C. & Varghese, A. (2013). Testing and evaluation of language skills. IOSR Journal of Research & Method in Education (Vol 1, Issue 2). Department of English and Foreign Languages, Bharathiar University, Coimbatore. Retrieved on 06 December 2015 from http://www.iosrjournals.org/iosr-jrme/papers/Vol-1%20Issue-2/F0123133.pdf?id=1662

[2]. Dasig Jr, D. (2014, February). User Experience of Embedded System Students on Arduino and Field Programmable Gate Array (FPGA). In Proc. of the Second Intl. Conf. on Advances in Applied Science and Environmental Engineering-ASEE (pp. 124-128).

[3]. Austria, D., Milani, M., Dasig Jr, D. D., & Valderama, A. M. C. (2015, May). Exploratory Study on Learner-Driven Blended Learning Environment. In 2nd International Conference in Education, Psychology, and Social Science (ICEPSS) International Research Enthusiast Society Inc.(IRES Inc.) May (pp. 22-23).

[4]. Dasig Jr, D. D., Gatpandan, M. P., Valderama, A. M. C., Taduyo, M. A. B., Traballo, R. C., Gatpandan, P., & Dasig, D. D. (2015). Management of Knowledge Containers and Learning Strategies in the Context of Case-Based Reasoning. Proceedings Journal of Education, Psychology and Social Science Research.

[5]. Gatpandan, P. H., Rosas, M.F. & Dasig, Jr., Microsoft office specialist and microsoft technology associate certification: An integrated curriculum for technical skills validation.

[6]. Nitko, A. J. (n.d.). Test and types of test. Retrieved on 05 December 2015 frpm http://www.hec.gov.pk/InsideHEC/Divisions/LearningInnovation/Documents/Learning%20Portal/NCES/NCES%20Presentations/TEST%20AND%20TYPES%20OF%20TEST.pdf

[7]. Overton, T. (n.d.). Assessing learners with special needs. Retrieved on 05 December 2015 from http://www.slideshare.net/norazmi3003/testing-assessment-measurement-and-evaluation-definition

[8].   University of Waterloo (2015). Center for teaching excellence. ON Canada: University Avenue West Waterloo. Retrieved on 17 December 2015 from https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/developing-assignments/exams/questions-types-characteristics-suggestions

[9].   Matlock-Hetzel, S. (1997). Basic concepts in item and test Analysis. Retrieved on 09   December 2015 from http://ericae.net/ft/tamu/Espy.htm

[10].   Dasig Jr, D., & Pascua, S. (2016, March). Effects of Digital Learning Objects in Teaching Realtime System. In Society for Information Technology & Teacher Education International Conference (pp. 1488-1498). Association for the Advancement of Computing in Education (AACE).

[11].   Osman, R. M. (n.d.). Education evaluation and testing. Retrieved on 05 December 2015 from http://oer.avu.org/bitstream/handle/123456789/78/Educational%20Evaluation%20and%20Testing.pdf?sequence=1

[12].   Dasig Jr, E. D. D., Reyes, E. L. R., & Dasig, D. D. Perceptions and Lived Experiences of a Project-based Learning in Computer Engineering Education.

[13].   Educational Development Center (n.d.). Scantron guide. Retrieved on 08 December 2015 from https://carleton.ca/edc/wp-content/uploads/Item-Analysis.pdf

[14].   Valderama, A. M. C., Mangaba, J. B., Traballo, R. C., Gatpandan, P. H., Taduyo, M. A. B., Dasig, D. D., ... & Vinluan, A. A. Efficacy of Learning Scaffolds and Learner-User Experience (UX) in the Zone of Proximal Development. In 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-8). IEEE.

[15].   Dasig Jr, D. D. (2014). A study on the sectors of economy serviced by pre-industry system developers among companies in Metro Manila: A tool for business reengineering. arXiv preprint arXiv:1409.7277.

[16].   Dasig Jr, D. (2017). A frontier in organizational and business process innovation in service management through lean six sigma Kaizen project implementation. Journal of Administrative and Business Studies, 3(6), 263-283.

[17].   Rondon, M. V. (2013). Impact of the CIES learning center among BEEd graduates. Makati City, Philippines: University of Makati.