PSYCHOMETRIC PROPERTIES OF INFORMATION AND COMMUNICATION TECHNOLOGY MULTIPLE CHOICE EXAMINATION OF SULE LAMIDO UNIVERSITY, KAFIN-HAUSA, JIGAWA STATE NIGERIA

MUHAMMAD GADDAFI SHUAIBU
`Sule Lamido University Kafin Hausa, Jigawa State, Nigeria.
muhammadgaddafishuaibu12@gmail.com

Abstract

This study investigated the psychometric properties of 2018 GSP 122 ICT multiple choice examination of Sule Lamido University Kafin Hausa Jigawa State Using Item Response Theory. Three objectives were raised in relation to the study, with three research questions. Ex-post factor research design was employed as research design; six hundred and nineteen (619) student's responses were the total population of the study. Enumeration sampling technique was used as sampling technique and the researcher used all the population of the study as sample of the study. The instrument used for the study was the responses of students in GSP 122 ICT multiple choice examination of SLUK. The analyses were done using Item Response Theory PRO soft ware. The findings revealed that most of the items were difficult, the items have moderate discrimination indices and the most of the items have very low guessing indices. It was recommended that constructors in the School of General and Entrepreneurship Studies SLUK should be well grounded on item analysis to ensure the production of high quality items and further studies should be carried out to generate more empirical data to have a clearer picture of the quality of the items.

Keywords: Item Response Theory, Item Difficulty, Item Discrimination, Guessing Parameter.

Introduction

The measurement of individuals' traits, or mental properties such as abilities and attitudes, has been a long-lasting quest that dates back to 1882 with Galton's pioneering work developing rating scales and questionnaires, and Thorndike's contributions to psychometric theory and its application to educational measurement. This quest continues today (Sijtsma, 2011). Measurement in an educational setting serves several purposes, namely, planning, monitoring, and evaluation instruction.

Today for an individual to function effectively in the furtherance of national development and sustainability he/she must possess the basic knowledge of Information and communication technology (ICT). According to Nwana (2008) Information and communication technologies (I.C.Ts) are electronic technologies used for information storage and retrieval. Ngwu (2014) defined I.C.T as electronic base technology that is used to retrieve stormed process and package information as well as provide access to knowledge. The development of microcomputer optical disc, the establishments of telecommunication network television, internets have assisted in broadening people's knowledge and facilitating effective communication.

In SLUK, all the efforts invested in teaching and learning of the various concepts of the ICT would have been wasted if there is nothing like the process of testing and measurements of various attributes of learners. The students' failure recorded in 2018 GSP 122 (ICT) examination in SLUK need to be addressed if the noble objectives of the course are to be achieved. A number of questions need to be asked as to the possible causes of the failure in these examinations; were the items too difficult? It should

be noted that very low indices implies very difficult item, while very high indices implies very easy item. How difficult was each item to the students? If the difficulty indices to a student is <0.50 is refers to an easy items and the difficulty indices of > 0.50 is refers to difficult items (Baker, 2001). Did the test items adequately discriminate between high and low achievers? The discrimination indices of an item indicate how well an item discriminates between the strong and weak students. Guessing is the third parameter of IRT and it is called c parameter. Guessing means giving an answer or making a judgment about something without being sure of all the facts (Gao & Stokes, 2008). In this study an Item Difficulty parameter (b): Ranges from -3 to +3 was used in finding out difficulty in dices. And Item discrimination parameter (a) was based on 0.01 as item with no discrimination power, 0.01 – 0.34 as very low, 0.35 – 0.64 low, 0.65-1.34 moderate, 01.35- 1.69 high, > 1.70 very high, and +∞ as perfect discrimination indices. While guessing parameter (c): Very low (0.00 & below), low (0.01-0.25) and high (0.26 & above) (Baker, 2001) and (DeMars 2010).

Item analysis is about how difficult an item is and how well it can discriminate between the good and the poor students. In other words, item analysis provides a numerical assessment of item difficulty and item discrimination. Item analysis is the assessment of the essential qualities of the test items which helps in building reliability and validity into the test from the start (Singh, 2008). Item analysis can be both qualitative and quantitative. Qualitative item analysis focuses on issues related to the content of the test e.g. Test analysis. Quantitative analysis involves measurement of item difficulty and item discrimination (Ebel & Fresbie, 2011). The outcome of item analysis helps the teacher to improve on item selection for the test by eliminating unreliable items, substituting for poor items, or by recasting poorly stated questions for better effect. The essential qualities normally considered in item analysis are item difficulty/easiness, item discrimination and distracter analysis (Anikweze, 2012). Borsboom, (2006) opined that item analysis provides three kinds of important information about the quality of test items.

Item difficulty is a measure of whether an item was too easy or too difficult.

Item discrimination is a measure of whether an item discriminated between candidates who knew the test well and candidates who did not.

Distractor Index measures the effectiveness of alternatives, that is, to determine whether distractors (incorrect but plausible options) tend to be chosen by the less able examinees and not by the more able examinees.

Hence, only those items which are valid and match the purpose of the test are selected whereas others are either eliminated or revised through the process of item analysis (Singh, 2008). It involves qualitative and quantitative techniques. Qualitative- This approach depends on the judgment of the experts/ reviewers about various characteristics of the items. Suitability of the item content and arrangement in accordance to the purpose of the test and the target population it is designed for. Clarity in the items involved no language error; whether basic rules of item writing are followed or not (Urbina, 2014).

Quantitative- The psychometric properties of the items are assessed with the use of statistical procedures (Urbina, 2014). According to DeMars (2010) difficulty is defined in both CTT and IRT in terms of the likelihood of correct response, not in terms of the perceived difficulty or amount of effort required. In CCT, the difficulty index, P, is the proportion of examinees who answer the item correctly (sometimes P-value). In IRT, the difficulty index, b, is on the same metric as the proficiency distribution in a designated group has a mean of 0 and standard deviation of 1. The item difficulty identifies to which about 50% of the examinees (or a little more, depending on the model) are expected to answer the item correctly. It is denoted by 'p' and the subscript is used to denote the item number, for example p1. The difficulty level of an item ranges from 0 (no one got the correct response) to 1 (everyone got the correct response) and is different for each item. The higher difficulty level items with high trait levels and lower difficulty items target lower trait levels thus helping in differentiating between the responses (Penfield,

2013). Therefore, optimal levels for difficulty generally are midway between 100 percent respondents scoring correctly and those expected by chance alone (Kaplan & Saccuzzo, 2009).

The second item parameter is discrimination; the discrimination indices of an item indicate whether or not the item is measuring the same ability as the test measures. It shows how well an item discriminates between the strong and weak students. It is a measure of correlation between the item and total test score. Like the coefficient of correlation, the DI ranges between -1.00 and +1.00 (Sidhu, 2012).

The 3-parameter model of Item Response Theory is guessing, which means giving an answer or making a judgment about something without being sure of all the facts (Gao &Stokes, 2008). Guessing is a standard test-taking strategy presented to examinees taking a multiple choice assessment. This strategy provides an opportunity to have an item counted correct even when the examinee has insufficient knowledge of the subject matter. If test scores are based simply on the number of questions answered correctly, then a random guess increases the chance of a higher score (Mehrens & Lehmann, 2004).

Statement of the Problem

As a result of failure experienced by School of General and Entrepreneurship studies in it is 2018 ICT Examination in Sule Lamido University Kafin Hausa, the researcher convinced that it is of great importance to investigate the reasons behind this failure. The 2018 Examination Results Summary of the GSP 122 (ICT) indicated that only two third of the students that sat for the examination pass the course, while almost one third failed. Also out of this number only 6.05% scored A, 13.88% scored B, 24.74% scored C, 23.12% scored D and 33.2% of the students failed, these has shown that about one third of the students failed the course. When compared to the proceedings ICT Examination results summary, there was a marginal declined in performance of the students. Consequently, it's against this background that the study analyzed the difficulty level, discrimination indices and probability of guessing in all the three parameters of the multiple choice test items of the 2018 ICT Examination among students in the different Faculties of SLUK.

Objectives of the Study

The study was guided by the following objectives;

i. To find out the b parameter of each item of 2018 GSP 122 in SLUK, Jigawa State.
ii. To find out the a parameter of each item of 2018 GSP 122 in SLUK, Jigawa State.
iii. To find out the c parameter of each item of 2018 GSP 122 in SLUK, Jigawa State.

Research Questions

This study answered the following questions:

i. What is the b-parameter of each item of 2018 GSP 122 in SLUK, Jigawa State?
ii. What is the a-parameter of each item of 2018 GSP 122 in SLUK, Jigawa State?
iii. What is the c-parameter of each item of 2018 GSP 122 in SLUK, Jigawa State?

Methodology

The researcher employed ex post facto research design since the focus of this study was to analyzed a-parameter, b-parameter and c-parameter of GSP 122 ICT SLUK. The population of the study comprised two categories. The first category is the observation unit, which contains all the undergraduate students of SLUK who sat for 2018 GSP 122 ICT Second category of the population is the Unit of analysis which was the items in the GSP 122 ICT of 2018. The researcher used the entire population of the study as sample of the study. This means the researcher used one thousand one hundred and thirty seven (1137) students as sample of the study. The researcher used census/enumeration sampling technique.

The data collection instruments of this study are the examinations items of GSP 122 ICT of 2018. There are all Multiple Choice Examinations set and administered by the School of General and entrepreneurship studies, SLUK, Jigawa State. These examinations consisted of 60 multiple choice items

for each GSP course, it also has five (5) options (A, B, C, D and E) and the administration time for this examinations were two hours (2hrs). The analysis was computed through the means of IRT PRO 2.1 for windows.

Results

Answers to Research Questions

Research Question 1: What is the b-parameter of each item of 2018 GSP 122 in SLUK, Jigawa State?

Table 1: Item Parameter b Estimates of 2018 GSP122 Examination

Item No.	Estimate	Item No.	Estimate	Item No.	Estimate
1	0.9	21	13.16	41	-0.8
2	-3.11	22	0.62	42	-2.57
3	0.07	23	-0.5	43	2.75
4	2.77	24	-0.48	44	10.64
5	1.57	25	1.44	45	0.29
6	0.93	26	-0.92	46	0.49
7	3.96	27	-1.11	47	1.17
8	3.9	28	-0.93	48	1.12
9	2.77	29	-0.93	49	0.99
10	2.17	30	-0.9	50	-0.31
11	0.47	31	-1.65	51	0.07
12	-0.29	32	5.88	52	1.72
13	-0.17	33	1.38	53	2.22
14	0.22	34	-1.69	54	-0.05
15	0.1	35	-1.42	55	1.15
16	0.4	36	0.14	56	1.44
17	-0.1	37	0.79	57	0.8
18	-2.16	38	-1.24	58	0.1
19	1.57	39	-1.04	59	-0.16
20	0.22	40	-1.63	60	-0.09

The table 1 above presents the item parameter b estimate of 2018 GSP122 examination in SLUK, Jigawa State based on three-parameter logistic (3PL) model. The parameter b was estimated by running the person-by-item- matrix via IRTPRO package. From the table, it can be seen that 31(51.67%) items are within the range of +3 and have positive difficult estimate and are termed as difficult items. These items are: 1, 3, 4, 5, 6, 9, 10, 11, 14, 15, 16, 19, 20, 22, 25, 33, 36, 37, 43, 45, 46, 47, 48, 49, 51, 52, 53, 55, 56, 57 and 58. On the other hand, 24(40%) items are within the range of -3 and have negative difficulty estimates and are termed as easy items. These items are: 2, 12, 13, 17, 18, 23, 24, 26, 27, 28, 29, 30, 31, 34, 35, 38, 39, 40, 41, 42, 50, 54, 59 and 60. However, 5(8.33%) items exceeded the range of -3 to +3 and therefore lack item information functions (Items 7, 8, 21, 32 and 44)

Research Question 2: What is the a-parameter of each item of 2018 GSP 122 in SLUK, Jigawa State?

Table 2: Item Parameter a Estimates of 2018 GSP122 Examination

Item No.	Estimate	Remarks	Item No.	Estimate	Remarks	Item No.	Estimate	Remarks
1	0.6	Low	21	0.06	V. Low	41	0.78	Moderate
2	0.14	V. Low	22	1.13	Moderate	42	0.26	V. Low
3	1.21	Moderate	23	1.28	Moderate	43	0.21	V. Low
4	1.78	V. High	24	0.5	Low	44	0.2	V. Low
5	1.16	Moderate	25	1.77	V. High	45	0.96	Moderate
6	0.7	Moderate	26	1.29	Moderate	46	0.76	Moderate
7	0.22	V. Low	27	1.19	Moderate	47	1.24	Moderate
8	0.43	Low	28	1.08	Moderate	48	2.51	V. High
9	1.78	V. High	29	0.45	Low	49	0.62	Low
10	1.95	V. High	30	1.4	High	50	0.88	Moderate

11	1.8	V. High	31	0.78	Moderate	51	1.13	Moderate
12	1.6	High	32	0.16	V. Low	52	0.37	Low
13	1.37	High	33	0.39	Low	53	0.57	Low
14	1.11	Moderate	34	0.65	Moderate	54	0.86	Moderate
15	4.44	V. High	35	1.27	Moderate	55	0.36	Low
16	1.26	Moderate	36	1.54	High	56	1.65	High
17	1.11	Moderate	37	0.36	Low	57	0.92	Moderate
18	0.7	Moderate	38	0.9	Moderate	58	1.6	High
19	1.16	Moderate	39	0.95	Moderate	59	1.03	Moderate
20	0.68	Moderate	40	0.6	Low	60	1.47	High

From the table 2 above, it can be seen that none of the items fall under none (≤ 0) value. The items that fall under very low (.01-.34) values are 7(11.67%) (2, 7, 21, 32, 42, 43 and 44), the items that have low discriminating are 11(18.33%) (.35-.64) and these items are: 1, 8, 24, 29, 33, 37, 40, 49, 52, 53 and 55. The items that falls under moderate (.65-1.34) values are 28(46.67%) items (3, 5, 6, 14, 16, 17, 18, 19, 20, 22, 23, 26, 27, 28, 31, 34, 35, 38, 39, 41, 45, 46, 47, 50, 51, 54, 57 and 59). Furthermore, 7(11.67%) items fall under high (1.35-1.69) values and the items are 12, 13, 30, 36, 56, 58 and 60. Equally 7(11.67%) items are within the very high (≥ 1.7) values. Thus, these items are: 4, 9, 10, 11, 15, 25 and 48 respectively.

Research Question 3: What is the c-parameter estimate of each item of 2018 GSP122 Examination in SLUK, Jigawa State?

Table 3: Item Parameter c Estimates of 2018 GSP122 Examination

Item No.	Estimate	Remarks	Item No.	Estimate	Remarks	Item No.	Estimate	Remarks
1	-0.54	V. Low	21	-0.79	V. Low	41	0.62	High
2	0.45	High	22	-0.71	V. Low	42	0.67	High
3	-0.09	V. Low	23	0.64	High	43	-0.58	V. Low
4	-4.93	V. Low	24	0.24	Low	44	-2.1	V. Low
5	-1.81	V. Low	25	-2.54	V. Low	45	-0.28	V. Low
6	-0.65	V. Low	26	1.2	High	46	-0.38	V. Low
7	-0.89	V. Low	27	1.32	High	47	-1.45	V. Low
8	-1.66	V. Low	28	1	High	48	-2.81	V. Low
9	-4.93	V. Low	29	0.42	High	49	-0.61	V. Low
10	-4.23	V. Low	30	1.26	High	50	0.27	High
11	-0.85	V. Low	31	1.29	High	51	-0.08	V. Low
12	0.46	High	32	-0.92	V. Low	52	-0.64	V. Low
13	0.23	Low	33	-0.54	V. Low	53	-1.27	V. Low
14	-0.24	V. Low	34	1.11	High	54	0.04	Low
15	-0.45	V. Low	35	1.79	High	55	-0.41	V. Low
16	-0.5	V. Low	36	-0.21	V. Low	56	-2.37	V. Low
17	0.11	Low	37	-0.28	V. Low	57	-0.73	V. Low
18	1.51	High	38	1.11	High	58	-0.15	V. Low
19	-1.81	V. Low	39	0.99	High	59	0.16	Low
20	-0.15	V. Low	40	0.97	High	60	0.13	Low

The table 3 above indicated that 36(60%) items have very low guessing values that ranges between (-0.00 & below) and these items are: 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 16, 19, 20, 21, 22, 25, 32, 33, 36, 37, 43, 44, 45, 46, 47, 48, 49, 51, 52, 53, 55, 56, 57 and 58. By implication, the chance of getting an answer correctly by mere guessing is very low on those items. 6(10%) items (13, 17, 24, 54, 59 and 60) have low value that ranges from (0.01-0.25). 18(30%) fall under the range of 0.26 and above as the high category. Thus these items are: 2, 12, 18, 23, 26, 27, 28, 29, 30, 31, 34, 35, 38, 39, 40, 41, 42 and 50. By implication it means that the chance of getting an answer correctly by mere guessing is high regarding

those items.

Discussion

The study found that most of the items are within the range of +3 and have positive difficult. This finding is different from the finding of the work of Sayyah et. al (2012) and that of Olutola (2015) in which they both found out the discrimination indices of their items have low discrimination. In research question two, the study found that majority of the items have moderate discrimination and this supports the work of Bashir (2018) who found out that the items in 2014 WAEC English Examination have moderately discriminates between high and low achievers. This finding is contrary with the work of Mustapha (2018) who found that only 19 items discriminated well between high and low achievers on 2014 Jigawa State Mathematics Qualifying Examination in Hadejia Education Zone while 21 items did not.

In research question three, the c parameter estimates found that majority of the have very low guessing indices. This corroborated with the findings of Rabi,u (2011) in which the study found that items had an average guessing value of 0.06, a low c-index. The low average index indicates that items of this group has low guessing ability and since the purpose of the test is a certification therefore the items should have been a bid higher guessing probability so as to measure the achievement of the examinee appropriately.

Conclusion

From the findings of this study, it was concluded that a large number of items in the 2018 GSP 122 ICT multiple choice examination of SLUK were difficult items. Even though the items were moderate in discriminating between high and low achievers and most of the items gave room for guessing correctly.

Recommendations

Based on the finding from this study, it is also recommended that the university together with the School of General and Entrepreneurship Studies and the lecturers that teach the courses to publish standard textbooks, which will guide the lecturers and at the same time provide the students with standard reading materials.

The finding from the study also recommended that the language and method of teaching the both GSP 122 and GSP 121 courses should be appropriate to the level and background of the students; and the instructional materials use should be more culturally relevant, for effective learning to take place.

References

Akindele, B.P. (2003). The development of an item bank for *selection tests into Nigerian Universities:* An exploratory study (Unpublished doctoral thesis) University of Ibadan, Ibadan, Nigeria.

Anikweze, C. M. (2012). *Measurement and Evaluation for Teacher Education* (3rded.).: Constellation and Evaluation

Baker, F. B. (2001) *Basics of Item Response Theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation.

Bashir, A. (2018). Analysis of the Psychometric Properties of 2014 WAEC English Examination in Minjibir Education Zone, Kano State. An Unpublished M.Ed. Dissertation, Bayero University Kano, Nigeria.

Borsboom, D., (2006). Attack of the psychometrians. *Psychometrika* 71(3) 425-440.

DeMars, C. (2010). *Item Response Theory, Understanding Statistic Measurements.* New york: Oxford Press

Ebel, R. L. & Freshbie, D. A (2011*) Essentials of Educational Measurement* (5th ed.). Prentice Hall of India Private Limited.

Gao, J. & Stokes, S.L. (2008). Bayessian IRT guessing models for partial guessing behaviours *Psychometrika* , 73(2).

Mustapha, S. H. (2018). Analysis of Items On 2014 Jigawa State Mathematics Qualifying Examination in Hadejah Education Zone. An Unpublished M.Ed. Dissertation, Bayero University Kano State.

Ngwu, O.G (2014), Assessment of availability and utilization of information and communication technology resources in teaching in federal college of education Ela Anufu Enugu Nigerian. ICELW 2014.

Nwana. S.E, (2012) Challenges in the application of e-learning by secondary school teacher in Anambra state, Nigeria, *African journal of teacher education* (AJOTE) , 2(1), 14.22.

Obinne, A.D.E. (2008). Psychometric properties of senior certificate biology examinations conducted by West African Examinations council: Application of item response theory. Unpublished doctoral dissertation, University of Nigeria, Nsukka.

Penfield, R. D. (2013). Item Analysis. In Geisinger et al. (Eds.), *APA Handbook of Testing and Assessment in Psychology* (Vol. 1, pp. 121-138). American Psychological Association.

Rabi,u A. (2021). Assessment ofqualityof general studies programme of gsp 1201 (use of english) examination of bayero university kano using item response theory. M.Ed. Dissertation, Bayero University Kano State.

Sidhu, S. K. (2012), *New Approaches to Measurement and Evaluation*.: SPV Limited.

Singh, A. K. (2008). *Tests, Measurements and Research Methods in Behavioural Sciences*(Rev. ed.).: Bharati Bhawan.

Urbina, S. (2014). *Essentials of Psychological Testing* (2nd ed.). Hoboken, NJ: John Wiley & Sons.