E-Journal of Humanities, Arts and Social Sciences (EHASS)
ISSN – Online 2720-7722 | Print 2821-8949
Volume 4 Issue 1 – January 2023 pp 50-58
Available online at: https://noyam.org/journals/ehass/
DOI : https://doi.org/10.38159/ehass.2023414

Exploring the Quality of Multiple-Choice Question Type of Test Items in Information and Communication Technology Using Item Analysis

David Arhin[1], Ruth Annan-Brew [1], Ruth Owusuaah[1] & Winifred Opoku Bonsu[1]
1 Department of Education and Psychology, Faculty of Educational Foundations, University of Cape Coast, Ghana.

ABSTRACT
The study explored the quality of multiple-choice question types of test items in Information and Communication Technology (ICT) using item analysis. The study was an exploratory type where secondary data was collected and analysed. The secondary data was made up of ten (10) Multiple Choice Questions (MCQ) items on an ICT test instrument. Thirty learners' test instruments were sampled. The item analysis procedure included the item reliability coefficient, coefficient of variation, difficulty level, discrimination index, and distractor analysis. The study revealed that the test items had a low-reliability power and the class is homogenous. It was also found that ten per cent (10%) of the items were highly appropriate, fifty per cent (50%) needed revision, and forty per cent (40%) needed to be discarded. The study also found that some of the distractors were not appropriate and hence needed review while others need to be replaced. It was, therefore, recommended that item analysis should embrace race by teachers since it will foster effectiveness and efficiency carry out their duties in teaching and learning. Moreover, the Ministry of Education in Ghana must liaise with various heads of educational institutions to contract measurement and evaluation experts to organize workshops for their staff on how to do item analysis.

Keywords: *Reliability coefficient, Discrimination index, Difficulty level, Distractor analysis, Test items*

Correspondence
David Arhin
Email: david.arhin@stu.ucc.edu.gh

Publication History
Received 19th September, 2022
Accepted 5th January, 2023
Published online 26th January, 2023

INTRODUCTION
Amedahe and Asamoah-Gyimah explain item analysis as the method of assessing learners' answers to each question to determine the quality of the items.[1] Item analysis, they added, includes item reliability, difficulty level, and discrimination index of individual items, as well as key and distractors' analyses. According to Crocker and Algina, distractor analysis is a time-consuming task that necessitates a high level of professionalism and expertise, and as part of item analysis, the focus is always on item difficulty to the exclusion of item discrimination and distractor analysis due to their (item discrimination and distractor analysis) nature.[2] Distractor analysis, in particular, is a type of item

[1] Francis. K. Amedahe and Kenneth Asamoah-Gyimah, *Introduction to Measurement and Evaluation(7th ed),* (Cape Coast: Hampton Press, 2016).
[2] Linda Crocker and James Algina, *Introduction to Classical and Modern Test Theory.* (ERIC, 1986).

analysis that is comparable to item difficulty and item discrimination. The current study looked at item reliability, difficulty level, and discrimination index of individual items, as well as the analysis of distractors.

Throughout the school year, teachers at all levels of education development give a variety of official teacher-made examinations. As a result, tests are essential instruments in the educational process. When norm-referenced tests are produced for instructional purposes, strict attention to the principles of test creation, administration, and analysis as well as reporting is critical. The teacher is given a sensitive and prominent role in the testing and assessment process in Ghana. Anamuah-Mensah and Quaigrain mention that "It is vital for the classroom teacher to be fully educated with formalized testing methodologies to provide good facilitation of the mammoth work that impinges on his or her profession."[3] As a result, teachers must be well-versed in assessment methodologies to assure the validity and dependability of learners' test results and development.

However, in teacher-made classroom exams, the above-mentioned test item analysis procedures are rarely applied. This might be because in Ghanaian schools and classrooms, multiple-choice, matching, and short-response questions are widely utilized in objective examinations. Tests serve a vital function in providing feedback to teachers on their educational actions, hence the test's quality is crucial. After administering and scoring a test, a teacher must determine the quality of the test items and if they accurately represent the learners' performance in the course concerning the specific learning goals taught overcome time.[4] Objectives test items must be created using a Table of Specification (ToS) or Test Blueprint that incorporates cognitive domain behaviours (knowledge, comprehension, application, analysis, synthesis, and evaluation). As a result, it is critical to know if these behaviours have been properly assessed. Test item analysis would guarantee that adaptations are made depending on learner requirements, instructional pace, course topic covering, and the creation of a more effective and pleasant learning environment in the classroom. Teachers' expertise in assessment and evaluation according to Xu and Liu, is not a static process but rather a complex, dynamic and continuing activity.[5] Item analysis serves some objectives that include, according to Suruchi and Rana: identifying defective test items and 2. indicating areas where learners have or have not learned.[6]

There is also the essence of item analysis which is to look for defects and figure out how to fix them before finally administering the questions in subsequent tests.[7] As a result, item review comes naturally after item analysis. When an item cannot be reviewed or revised, it is usually discarded and replaced. The study, therefore, explored the quality of multiple-choice question types of test items in Information and Communication Technology (ICT) using item analysis through an exploratory type where secondary data was collected and analysed.

EMPIRICAL REVIEW

In similar studies, Toksoz and Ertunc sought to examine multiple-choice questions used to assess grammar, vocabulary, and reading comprehension and administered to learners in preparatory classes at a public university.[8] By employing the frequency of displaying the distribution of preparatory learners' replies, four hundred and fifty-three (453) learners' responses were studied in terms of item facility, item discrimination, and distractor efficiency. According to the findings, the majority of the

[3] J. Anamuah-Mensah and K. A. Quaigrain, "Teacher Competence in the Use of Essay Test," *The Oguaa Educator University of Cape Coast* 12 (1998): 31–42.

[4] Eric Anane and Kenneth Asamoah-Gyimah, *Assessment in Education* (Cape Coast: UCC Press, 2019).

[5] Yueting Xu and Yongcan Liu, "Teacher Assessment Knowledge and Practice: A Narrative Inquiry of a Chinese College EFL Teacher's Experience," *Tesol Quarterly* 43, no. 3 (2009): 492–513.

[6] S Suruchi and S S Rana, "Test Item Analysis and Relationship between Difficulty Level and Discrimination Index of Test Items in an Achievement Test in Biology," *PIJR* 3, no. 6 (2014): 56–58.

[7] Fahmi Ishaq El-Uri and Naser Malas, "Analysis of Use of a Single Best Answer Format in an Undergraduate Medical Examination," *Qatar Medical Journal* 2013, no. 1 (2013): 1; P. Gochyyev and D. Sabers, "Item Analysis," *J. Res. Methods. Https://Srmo. Sagepub. Com/View/Encyc-of-Researchdesign*, no. 199 (2012).

[8] Sibel Toksöz and Ayşe Ertunç, "Item Analysis of a Multiple-Choice Exam," *Advances in Language and Literary Studies* 8, no. 6 (2017): 141–46.

products have a modest degree of item facility. Furthermore, the findings reveal that twenty-eight per cent (28٪)of the items have a low item discrimination value. Finally, the frequency data was studied in terms of distractor effectiveness, and it was discovered that some of the exam's distractors are ineffective and should be altered. Also, Quaigrain and Arhin investigated the relationship between the item and test quality and the difficulty index (p-value) and discrimination index (D**52**) with distractor efficiency (DE).[9] The research involved 247 first-year learners at Cape Coast Polytechnic, Ghana pursuing a Diploma in Education. In the Educational Measurement course, a fifty-question multiple-choice test was given at the end of the semester. Using the Kuder–Richardson 20 coefficient, it was discovered that the test's internal consistency reliability was 0.77. (KR-20). The average score was 29.23 with a 6.36 standard deviation. The DI was 0.22 (SD 0.17), and the mean difficulty index (p) value was 58.46 % (SD 21.23 %). At p-values between 40 and 60 %, DI was found to be at its peak. The average DE was 55.04 % (SD 24.09%). To increase the quality of the assessment, items with moderate difficulty and strong discriminating power with functional distractors should be included in future examinations. Using DI, it was discovered that 30 (60%) of the test items fell into the "fairly good" or "acceptable" range of values.

Odukoya, Adekeye, Igbinoba, and Afolabi carried out a study with the main goal of determining the item difficulty and distractive indices of university courses.[10] A total of 112–1956 undergraduate learners took part in the research. The study was conducted after the incident, using secondary data as an instrument. The bulk of the questions (between 65 and 97 % of the 70 items fielded in each course) did not fulfil psychometric standards in terms of difficulty and distractive indices and hence had to be edited or eliminated. Given the importance of these courses, it was stressed that item analyses be used while creating these assessments. Finally, Asamoah and Ocansey presented a comprehensive study of thirty (30) examination questions in core mathematics in terms of item discrimination and distractor analysis.[11] The test assessed the achievement of senior high school learners in basic mathematics in some of the areas that they were taught. Following the discrimination and distractor analyses, it was discovered that twenty-seven (27) of the thirty (30) items discriminated favourably, two negatively, and one with no discrimination. The distractor analysis of all 30-item possibilities was also calculated and discussed. The studies revealed that there were solid analyses with a large sample size on average. However, no study has been carried out focusing on coursework in ICT.

METHODOLOGY

The study was an exploratory type where secondary data from Juansa Presbyterian JHS of Asante Akim North Municipality in Ghana, were collected and analysed. The secondary data were made up of ten (10) MCQ items on an ICT test instrument. Thirty learners' test instruments were sampled. Amedahe and Asamoah-Gyimah asserted that for the discrimination index, assessors are to compute twenty-seven percent (27٪) of the total number of learners if they are more than forty (40) and use the figure to sample the upper group and lower group as they discard the middle group.[12] However, the thirty (30) learners were split into two equal parts to represent the upper and the lower groups of fifteen (15) each after arranging their scores for the MCQ section in descending order.

RESULTS/FINDINGS

[9] Kennedy Quaigrain and Ato Kwamina Arhin, "Using Reliability and Item Analysis to Evaluate a Teacher-Developed Test in Educational Measurement and Evaluation," ed. Sammy King Fai Hui, *Cogent Education* 4, no. 1 (January 1, 2017): 1301013, https://doi.org/10.1080/2331186X.2017.1301013.

[10] Jonathan A. Odukoya et al., "Item Analysis of University-Wide Multiple Choice Objective Examinations: The Experience of a Nigerian Private University," *Quality & Quantity* 52, no. 3 (May 14, 2018): 983–97, https://doi.org/10.1007/s11135-017-0499-2.

[11] Daniel Asamoah and Moses K K Ocansey, "Item Discrimination and Distractor Analysis: A Technical Report on Thirty Multiple Choice Core Mathematics Achievement Test Items," *International Journal of Research and Scientific Innovation (IJRSI)* 6 (2019).

[12] Amedahe and Asamoah-Gyimah, *Introduction to Measurement and Evaluation* .

Table 1: ***Reliability Coefficient Computation***

LEARNER	SCORE (10)	GROUP	d = x-M	d^2
1	10	UPPER	3	9
2	9		2	4
3	9		2	4
4	9		2	4
5	8		1	1
6	8		1	1
7	8		1	1
8	8		1	1
9	8		1	1
10	8		1	1
11	8		1	1
12	8		1	1
13	8		1	1
14	7		0	0
15	7		0	0
16	7	LOWER	0	0
17	7		0	0
18	7		0	0
19	7		0	0
20	7		0	0
21	7		0	0
22	7		0	0
23	7		0	0
24	7		0	0
25	6		-1	1
26	6		-1	1
27	6		-1	1
28	6		-1	1
29	6		-1	1
30	4		-3	9
Total	220			44
	M = 7			s^2 = 4.4 / s = 2.1

$$\frac{n}{n-1} \times (1 - \frac{M \times (n - M)}{n \times s^2})$$

n – sample size, s^2 – variance of the test, M – mean of the test

$$\_30 \times (1 - 7x(30 - 7)) = 30 \times (1 - 161) = 30 \times (1 - 1.220) = 1.034 \times (-0.220) = -0.23$$

30-1 30 x 4.4 29 132 29

Table 1 shows the reliability coefficient r = – 0.23 which was computed using Richard Kuderson formulae. The r of the items was low and in a negative direction. Yuniarti, Setiawan, and Hariyanto presented that $0.0 \leq r < 0.40$ is a low-reliability coefficient.[13] The study determined the

[13] Nurhening Yuniarti, Ahmad Luthfi Setiawan, and Didik Hariyanto, "The Development and Comprehensive Evaluation of Control System Training Kit as a Modular-Based Learning Media," *TEM Journal* 9, no. 3 (2020): 1234.

variation of achievement among the learners. With that, the coefficient of variation (CV) was computed as $CV = \frac{s}{\mu} \times 100 = \frac{2.1}{7} \times 100 = 0.3 \times 100 = 30\%$. Based on the CV = 30%, the class was identified as homogeneous meaning that the learners have the same abilities in terms of academic achievement. Estey indicated that when CV ≤ 33%, then the class is homogenous while CV > 33% represents the heterogeneous class.[14]

Table 2: ***Analysis of Difficulty Level of Items***

Item	Number of learners who got the item right	P
1	**14**	**0.47**
2	24	0.80
3	**21**	**0.70**
4	**12**	**0.40**
5	24	0.80
6	26	0.87
7	**16**	**0.53**
8	29	0.97
9	29	0.97
10	23	0.77

Table 2 presents the results of the item difficult level of the ten (10) MCQs in an ICT test. The degree of difficulty of the MCQ items in proportion to the testees' cognitive capacity is indicated by the item difficulty level.[15] It is calculated by determining the percentage of test subjects who properly answered the question. When the index falls below 0.3 (ρ<.3), the item is considered too tough. When the index is greater than 0.7 (ρ>.7), the item is considered too easy. An item is easily adjudged easy when the index is above 0.7 (ρ>.7). From the table, it is evident that six items: 2, 5, 6, 8, 9, and 10 were easy. None of the 10 items was adjudged to be difficult with four items: 1, 3, 4, and 7 falling into the ideal range of difficult level. The next table presents the discrimination indices of the 10 items.

Table 3: ***Analysis of Discrimination Indices of Items***

Item	U	ρU	L	ρL	D
1	8	0.53	6	0.40	0.1
2	**14**	**0.93**	**10**	**0.67**	**0.3**
3	12	0.80	9	0.60	0.2
4	7	0.47	5	0.33	0.1
5	**14**	**0.93**	**10**	**0.67**	**0.3**
6	13	0.87	13	0.87	0.0
7	**11**	**0.73**	**5**	**0.33**	**0.4**
8	15	1.00	14	0.93	0.1
9	15	1.00	14	0.93	0.1
10	13	0.87	10	0.67	0.2

[14] Young K. Estey, *Educational Statistics* (Cape Coast: UCC Press, 2012).

[15] C. Boopathiraj and K. Chellamani, "Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education," *International Journal of Social Science & Interdisciplinary Research* 2, no. 2 (2013): 189–93.

According to Sabri, a discriminatory index displays an item's ability to distinguish between high and low-performing Testees.[16] Item discrimination determines whether those who performed well across board also performed well on a specific item. An item should be able to distinguish between groups with higher and lower scores. The extreme group technique which compares those who have done extremely well with those who have done very poorly, is one way to measure an item's potential to discriminate. According to Amedahe and Asamoah-Gyimah, a test item's discriminating power is its capacity to distinguish between learners who have performed well (the upper group) and those who have performed poorly (the lower group).[17]

To begin, select the Testees who scored in the top one-quarter of the class (upper quartile) and those who scored in the bottom one-quarter of the class (lower quartile). Calculate the percentage of people in the top and lower quartiles who properly answered a test item. Finally, to calculate the item's discrimination index, subtract the proportion of Testees who got the item right in the poorest-performing group from the proportion of Testees who got the item right in the top-performing group (D). Item discriminations of D = 0.3 or more are regarded as appropriate; D = 0 indicates that the item has no discriminatory power; D = 1.0 indicates that the item has perfect discriminatory power; nevertheless, D = - values indicate that the item has very poor discriminatory power.[18] From Table 3, it is evident that items 2, 5, and 7 discriminated appropriately while item 6 had no discriminatory power. It can also be noticed that the rest of items: 1, 3, 4, 8, 9, and 10 had weak discriminatory power.

Table 4: ***Distractor Analysis of Items***

Item	**Distribution of options among learners in the upper and lower group**								
1	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	4	6	10	8	6	14	3	3	6
2	**A***		Total	B		Total	C		Total
	U	L		U	L		U	L	
	14	10	24	0	4	4	1	1	2
3	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	1	1	2	12	9	21	2	5	7
4	A		Total	B		Total	**C***		Total
	U	L		U	L		U	L	
	2	3	5	5	7	12	8	5	13
5	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	1	3	4	14	11	25	0	1	1
6	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	3	1	4	12	13	25	0	1	1
7	A		Total	B		Total	**C***		Total
	U	L		U	L		U	L	
	2	5	7	11	5	16	2	5	7
8	**A***		Total	B		Total	C		Total

[16] Shafizan Sabri, "Item Analysis of Student Comprehensive Test for Research in Teaching Beginner String Ensemble Using Model Based Teaching among Music Students in Public Universities," *International Journal of Education and Research* 1, no. 12 (2013): 1–14.

[17] Amedahe and Asamoah-Gyimah, *Introduction to Measurement and Evaluation* .

[18] Anane and Asamoah-Gyimah, *Assessment in Education*.

	U	L		U	L		U	L	
	15	14	29	0	1	1	0	0	0
9	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	1	0	1	13	14	27	1	1	2
10	A		Total	**B***		Total	C		Total
	U	L		U	L		U	L	
	0	2	2	12	9	21	3	4	7

The distractor analysis determines the power of the distractor or foil (i.e., the incorrect options in an MCQ) in distracting the testees. The distractor analysis was done as a distribution of the learner's choice. According to Anane and Asamoah-Gyimah, a good distractor is best chosen more by testees who are not informed about the item or less by those who are informed about the item.[19] When a distractor distracts no testee or more of the testees from the upper group than from the lower group, it is concluded that such a distractor is poor and should be reviewed or discarded and replaced depending on the severity. The items were three-option types hence, there were two distractors in each item. In Table 4, distractors C in item 1, C in item 2, A in item 3, and C in item 9 attracted an equal number of testees from both groups hence needing to be reviewed. Again, distractor A in item 6, B in item 7, and a in item 9 attracted more testees from the upper group than the lower group while distractor C in the item did not attract any testees. It is, therefore, appropriate to discard and replace the four distractors. It is obvious from the analyses that sixty per cent (60%) of the twenty (20) distractors were appropriate, twenty per cent (20%)needs to be reviewed and 20 per cent needs to be discarded and replaced.

DISCUSSION

It was revealed that the test items have a low-reliability power. For the difficult level, six items: 2, 5, 6, 8, 9, and 10 were easy while four items: 1, 3, 4, and 7 were adjudged to be appropriate. It was also found that items 2, 5, and 7 discriminated appropriately while item 6 had no discriminatory power. It can also be noticed that items: 1, 3, 4, 8, 9, and 10 had weak discriminatory power. In summary, item 7 was the only item that lies in the ideal range of both the difficulty level and discrimination index. Items 2 and 5 could discriminate appropriately but were identified to be easy. Items 1, 3, and 4 were good in terms of difficulty level but had a weak discriminatory power. It is therefore appropriate to keep item 7, revise items 1, 2, 3, 4, and 5, and discard items 6, 8, 9, and 10. It can be revealed that ten per cent (10%) of the items were highly appropriate, 50 fifty per cent (50%) needs revision, and forty per cent (40%) needs to be discarded. The findings agree with Odukoya et.al., who found out that the bulk of the questions (between sixty-five and ninety-seven per cent (65-97%) of the seventy (70) items fielded in each course) did not fulfil psychometric standards in terms of difficulty and distractive indices, and hence had to be edited or eliminated.[20] However, the findings do not agree with the views of Asamoah and Ocansey as well as Quaigrain and Arhin who discovered that 30 (60%) of the test items fell into the "fairly good" or "acceptable" range of values.[21]

The study also found out that some of the distractors were not appropriate and hence need review while others need to be replaced. The finding confirms the views of Toksoz and Ertunc who discovered that some of the exam's distractors are ineffective and should be altered.[22]

[19] Anane and Asamoah-Gyimah.

[20] Odukoya et al., "Item Analysis of University-Wide Multiple Choice Objective Examinations: The Experience of a Nigerian Private University," 983-997.

[21] Asamoah and Ocansey, "Item Discrimination and Distractor Analysis: A Technical Report on Thirty Multiple Choice Core Mathematics Achievement Test Items," 1301013.

[22] Toksöz and Ertunç, "Item Analysis of a Multiple-Choice Exam."

RECOMMENDATIONS

From the findings and the discussion, it is recommended that item analysis should be embraced by teachers since it will foster effectiveness and efficiency as they carry out their duties in teaching and learning. Moreover, it will help the teacher who is a test item constructor to finetune their skills and competencies in test item construction. Item analysis is a procedure to aid test item constructors to write good and improved test items.[23] Finally, the Ministry of Education must liaise with various heads of educational institutions to contract measurement and evaluation experts to organize workshops for their staff on how to do item analysis.

CONCLUSION

The study explored the quality of multiple-choice question types of test items in Information and Communication Technology (ICT) using item analysis. In the study, the reliability of the test item was low and in a negative direction in a homogeneous class. Four items: 1, 3, 4, and 7 fall into the ideal range of difficulty level, and items 2, 5, and 7 are discriminated appropriately. It was, therefore, appropriate to keep item 7, revise items 1, 2, 3, 4, and 5, and discard items 6, 8, 9, and 10. Moreover, some distractors were identified as not to be plausible. From the findings and the discussion, it was concluded that the learners previewed the items. The items that were identified to be reviewed or discarded are a good call. Considering the thorough nature of the analyses, one can conclude that the study has indeed called on the test item constructor to look into how the item's construction and the preparation of testees were carried out.

CONFLICT OF INTEREST

All the authors unanimously accept equal responsibilities and no one claims a greater portion of the contribution.

BIBLIOGRAPHY

Amedahe, Francis. K., and Kenneth Asamoah-Gyimah. *Introduction to Measurement and Evaluation* . Cape Coast: Hampton Press, 2016.

Anamuah-Mensah, J, and K A Quaigrain. "Teacher Competence in the Use of Essay Test." *The Oguaa Educator University of Cape Coast* 12 (1998): 31–42.

Anane, Eric, and Kenneth Asamoah-Gyimah. *Assessment in Education*. Cape Coast: UCC Press, 2019.

Asamoah, Daniel, and Moses K K Ocansey. "Item Discrimination and Distractor Analysis: A Technical Report on Thirty Multiple Choice Core Mathematics Achievement Test Items." *International Journal of Research and Scientific Innovation (IJRSI)* 6 (2019).

Boopathiraj, C, and K Chellamani. "Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education." *International Journal of Social Science & Interdisciplinary Research* 2, no. 2 (2013): 189–93.

Crocker, Linda, and James Algina. *Introduction to Classical and Modern Test Theory.* ERIC, 1986.

El-Uri, Fahmi Ishaq, and Naser Malas. "Analysis of Use of a Single Best Answer Format in an Undergraduate Medical Examination." *Qatar Medical Journal* 2013, no. 1 (2013): 1.

Estey, Young K. *Educational Statistics*. Cape Coast: UCC Press, 2012.

Gochyyev, P, and D Sabers. "Item Analysis." *J. Res. Methods. Https://Srmo. Sagepub. Com/View/Encyc-of-Researchdesign*, no. 199 (2012).

Odukoya, Jonathan A., Olajide Adekeye, Angie O. Igbinoba, and A. Afolabi. "Item Analysis of University-Wide Multiple Choice Objective Examinations: The Experience of a Nigerian Private University." *Quality & Quantity* 52, no. 3 (May 14, 2018): 983–97. https://doi.org/10.1007/s11135-017-0499-2.

Quaigrain, Kennedy, and Ato Kwamina Arhin. "Using Reliability and Item Analysis to Evaluate a

[23] Gochyyev and Sabers, "Item Analysis."

Teacher-Developed Test in Educational Measurement and Evaluation." Edited by Sammy King Fai Hui. *Cogent Education* 4, no. 1 (January 1, 2017): 1301013. https://doi.org/10.1080/2331186X.2017.1301013.

Sabri, Shafizan. "Item Analysis of Student Comprehensive Test for Research in Teaching Beginner String Ensemble Using Model Based Teaching among Music Students in Public Universities." *International Journal of Education and Research* 1, no. 12 (2013): 1–14.

Suruchi, S, and S S Rana. "Test Item Analysis and Relationship between Difficulty Level and Discrimination Index of Test Items in an Achievement Test in Biology." *PIJR* 3, no. 6 (2014): 56–58.

Toksöz, Sibel, and Ayşe Ertunç. "Item Analysis of a Multiple-Choice Exam." *Advances in Language and Literary Studies* 8, no. 6 (2017): 141–46.

Xu, Yueting, and Yongcan Liu. "Teacher Assessment Knowledge and Practice: A Narrative Inquiry of a Chinese College EFL Teacher's Experience." *Tesol Quarterly* 43, no. 3 (2009): 492–513.

Yuniarti, Nurhening, Ahmad Luthfi Setiawan, and Didik Hariyanto. "The Development and Comprehensive Evaluation of Control System Training Kit as a Modular-Based Learning Media." *TEM Journal* 9, no. 3 (2020): 1234.

ABOUT AUTHORS

David Arhin is currently an MPhil student in Measurement and Evaluation awaiting Viva Voce at the Department of Education and Psychology, Faculty of Educational Foundations, University of Cape Coast, Ghana. His research focuses on Assessment, Mathematics, ICT, Teacher Education, Basic Education, Data Analysis, Monitoring, and Evaluation.

Ruth Annan-Brew is currently a Lecturer at the Department of Education and Psychology, Faculty of Educational Foundations, University of Cape Coast, Ghana. She holds a PhD in Measurement and Evaluation. Her research focuses on Assessment, Evaluation, Gender, Teacher Education and Psychology.

Ruth Owusuah is currently an MPhil student in Measurement and Evaluation at the Department of Education and Psychology, Faculty of Educational Foundations, University of Cape Coast, Ghana. Her research focuses on Assessment, Mathematics, ICT, Teacher Education and Supervision.

Winfred Bonsu Owusu is currently an MPhil student in Measurement and Evaluation awaiting Viva Voce at the Department of Education and Psychology, Faculty of Educational Foundations, University of Cape Coast, Ghana. His research focuses on Assessment, Mathematics, ICT, Teacher Education and Supervision.