

Research Article DOI: <https://doi.org/10.47434/JEREDA>. eISSN: 2735-9107


## ANALYSIS OF ITEM WRITING FLAWS (IWFs) EVIDENT IN OBJECTIVE FORMATS EXAMINATION QUESTIONS IN FEDERAL COLLEGE OF EDUCATION (TECHNICAL) ASABA, NIGERIA

<sup>1</sup>Peter Olorunmowaju Ajayi, <sup>2</sup>Vivian Ozofu Idusogie

\*<sup>1</sup>Department of Educational Psychology, Federal College of Education (Technical), Nigeria

<sup>2</sup>Department of Educational Evaluation and Counselling Psychology, University of Benin, Nigeria; <sup>2</sup>[ozyvivid@yahoo.com](mailto:ozyvivid@yahoo.com)

\*<sup>1</sup>Corresponding Author: [ajayipeter76@gmail.com](mailto:ajayipeter76@gmail.com)

 <https://orcid.org/0000-0003-3595-9898>

Received: 18<sup>th</sup> August, 2020; Revised: 23<sup>rd</sup> September, 2020; Accepted: 30<sup>th</sup> October, 2020

### ABSTRACT

**Purpose:** The purpose of this study was to ascertain IWFs contained in objective formats examination questions for assessment of students in the Nigeria Certificate in Education (NCE) programme in 2016/2017, 2017/2018 and first semester 2018/2019 academic sessions.

**Methodology:** Descriptive cross-sectional design was adopted which enabled the researchers to estimate prevalence (number of cases) of IWFs associated with objective questions constructed for end-of-semester examinations. The researchers retrieved 57 objective question papers administered in end-of-semester examinations centrally conducted in the College within the period of three academic sessions (2016/2017, 2017/2018 and first semester 2018/2019). 19 common violations of item writing principles were selected from literature and used in assessing the quality of 57 objective questions. The study classified test items in each of the 57 objective question papers into standard and flawed categories such that if an item is flawed, the exact type of flaw(s) (including options) was recorded.

**Results:** The results showed that short answer, multiple choice and alternate response questions are the types of objective test formats constructed by lecturers; there was a high rate of violation of standard item writing rules in most objective semester examination questions constructed by lecturers; and the nature of the four most frequently violated IWFs were related to irrelevant difficulty which tended to make questions or tasks more difficult than intended.

**Recommendations/Classroom Implications:** College management should organize seminar or workshop at regular interval to train or update lecturers' knowledge on test construction skills and tips.

**Keywords:** Alternate response items, Assessment, Item writing flaws, Item Writing Guidelines, Multiple-choice items, Objective test items, Short answer items



### Cite paper as:

Ajayi, P. O., & Idusogie V. O. (2020). Analysis of item writing flaws (IWFs) evident in objective formats examination questions in federal college of education (technical) Asaba, Nigeria. *Journal of Educational Research in Developing Areas*, 1 (3), 181-191. <https://doi.org/10.47434/JEREDA.1.3.2020.181>.



Copyright © 2020 The author(s) of this article retain(s) the copyright.

## **PUBLIC INTEREST STATEMENT**

Teachers occupy the center stage in the choice and construction of tests or examinations used for assessment of and for learning. The calls for teachers to construct quality objective test items that will yield accurate measure of students' achievement have been stressed in literature. However, teacher-constructed achievement examinations are largely criticized for item construction flaws. This study will help lecturers in Colleges of Education in particular and teachers in general to identify common item writing flaws committed in the construction of objective questions and fosters standard item writing skills that will yield accurate and effective measure of learning outcomes.

## **INTRODUCTION**

Teachers' choice of suitable objective test formats of assessment technique is essential but construction of quality test items that will make certain accurate and effective measure of learning outcomes is of paramount importance. In this light, development of quality objective test items makes possible effective and accurate measure of students' learning outcome. This underscores the call on teachers to construct quality (standard) assessment instruments capable of yielding accurate information about testees' ability and it has remained the emphasis of experts and stakeholders in assessment. However, construction of quality objective questions is very demanding; nevertheless, use of flawed objective questions in assessment of learning tends to contaminate achievement measure of students' learning outcome.

Examination is an obligatory assessment of learning conducted at the end of semester in Colleges of Education (COEs) in Nigeria. In 2012, the National Commission for Colleges of Education (NCCE) in its curriculum implementation framework for Nigeria Certificate in Education (NCE) reviewed the methods of assessment to ensure harmony and parity in assessment procedure among Colleges of Education by prescribing essay and objective test techniques of assessment. The NCCE guidelines for examinations however, recommended a minimum of 25 objective question items to be set on a 1 credit courses, 50 for 2 credit courses and 75 for 3 credit courses. With this guideline in place, many lecturers in Colleges of Education have embraced the construction of objective test item formats for

assessment of students in semester examinations.

The goal of assessment as enumerated in the National Policy on Education (NPE), inter alia, is to improve the credibility of examinations conducted in Nigeria and enhance global competitiveness of the products of the Nigerian educational system (Federal Ministry of Education, 2014). It is therefore incumbent on lecturers to construct and use quality objective examination items that are capable of yielding credible scores that can serve a number of purposes but not limited to grading, selection and certification. However, there seems to be a rising consensus among experts that teacher-made-tests (TMTs) are beset with flaws leading to ineffective and inaccurate achievement measure of students. According to Tarrant, Knierim, Hayes and Ware (2006) in Nedeau-Cayo, Laughlin, Rus and Hall (2013) and Rush, Rankin and White (2016) observed that teacher-developed examinations across many disciplines are excessively rife with item writing flaws (IWFs) while Drasgow, Luecht and Bennett (2006) in Royal, Hedgpeth and Posner (2018) averred that writing quality test items has remained a long standing problem confronting teachers. However, Mehrens and Lehmann in Ali and Ruit (2015) attributed the poor quality of examinations constructed by teachers to lack of adequate training in test construction. The assertions above suggest that teachers contravene standard item writing guidelines which casts aspersion on the quality and trustworthiness of objective examination questions constructed by lecturers in Colleges of Education.

There seems to be little empirical basis for the construction of test items as such, existing rules on how to construct effective test are guiding principles (Nedeau-Cayo, Laughlin, Rus & Hall, 2013). However, Haladyna, Downing and Rodriguez in Pais, Silva, Guimarães, Povo, Coelho, Silva-Pereira, Lourinho, Ferreira and Severo (2016) prescribed a 31 items-writing guidelines (IWGs) for effective construction of objective examination questions which are classified into: *Content Concerns, Formatting Concerns, Style Concerns, Stem Writing, and Writing the Choices*. Item writing flaws (IWFs) arises when one or more generally accepted standard item writing rules are violated by test item writers. Pham, Besanko and Devitt (2018) averred that construction of objective test questions are susceptible to errors that may occur through item writing flaws (IWFs), which are defined as violations of standard IWGs. In spite of established IWGs for constructing quality and effective objective test, item writing flaws are observed to be prevalent in faculty-prepared examinations which could render examination questions easier or more difficult than intended (Rush, Rankin & White, 2016).

Frequently committed IWFs are broadly classified into two as issues related to testwiseness; and issues related to irrelevant difficulty (Khan, Danish, Awan, & Anwar, 2013). Frequently committed IWFs related to testwiseness give rise to artificial ease and tend to make items easier than intended. The IWFs in this category are: the use of: grammatical clues (using a word that gives hint on the correct response); logical clues (arrangement of correct options using predictable patterns); absolute terms (always, never, only, and all) usually render a statement false; extra detail in correct option (longest option is the correct option); implausible distracters or decoys (when one or more decoys are obviously incorrect); and convergence strategy (Rush, Rankin & White, 2016). In addition to IWFs related to testwiseness (flaws that make items easier for

students to answer correctly based on their test taking strategies or skills) are: implausible distracters or decoys used to create item uniformity (when one or more decoys are obviously incorrect); mutually exclusive distracters (when two out of four options are known to be wrong answers) which render items easier than anticipated, thereby favouring test-wise students based on their test-taking strategies.

On the other hand, IWFs related to irrelevant difficulty lead to artificial difficulty which make test more difficult than expected. This category of frequently committed IWFs are the use of; "all except" or "none except" in the stem, stem negation (not, except, not true, true except, incorrect), none of the above (NOTA) and all of the above (AOTA) or combination of NOTA and AOTA within response options (complex or K-type items-combination of alternatives ("A. I & II; B. I & III; combination of AOTA and NOTA in decoys; or use of "two of the above"), heterogeneous options (dissimilar number of options or repeated elements of correct answer included within other options), and numeric data not stated correctly. Similarly, Rush, et al. listed other IWFs related to irrelevant difficulty which are the use of; awkward stem structure (complete the sentence, fill in the blank-placing a response at the end of a sentence), irrelevant or misleading information in the stem; response options are a series of true-false statements, vague or generalizing terms (sometimes, frequently, often, occasionally, typically and potentially), and unfocused stem (broad and open-ended questions that do not pose specific problem, distracters are unrelated or distantly related to a single learning objective, did not ask a direct question or required the examinee to read all answer options before being able to answer). They tend to render questions unnecessarily complex and prevent hardworking students from demonstrating mastery of the material. Generally, IWFs lead to 10-15% misclassification of tested students as failed rather than passed (Downing in

Pais, Silva, Guimarães, Povo, Coelho, Silva-Pereira, Lourinho, Ferreira & Severo, 2016) while Pais et al. aver that violation of IWGs relating to content concerns, stem writing and writing the choices have negative impact on the psychometric properties of test.

One quality of a good test is its capability to comprehensively cover almost all the aspects of the instructional content taught. Construction of table of content specifications is a strategy for attaining comprehensive test. Test construction requires utilization of skills that can enable test designers to construct test with accuracy, suitable use of language, impartial communication, items validation and good grading criteria (Silker, 2003; Ovat & Ofem, 2017). However, when item writers fail to use test blueprint, the tendency of writing lopsided items that do not comprehensively cover the instructional content on one hand and Bloom's cognitive domain on the other leads to questions focusing on lower cognitive levels. It is on the basis of the foregoing that this study analyzed item writing flaws (IWFs) evident in objective formats of semester examination in Federal College of Education (Technical) Asaba, Delta State

#### **STATEMENT OF THE PROBLEM**

Examination questions constructed by lecturers for the assessment of students at the end of the semester undergo mandatory internal or external moderation process to ensure development of quality items that will guarantee effective and accurate measure of learning outcomes. Despite this quality assurance procedure, one of the investigators observed a persistent general complains by students in examination halls on issues not limited to ambiguous questions and unclear instructions. This seems to be evidence of item writing flaws which tend to contaminate the quality and effectiveness of tests constructed, by polluting students' scores thereby affecting the accuracy and the valid interpretation of examination results. However, the existence or otherwise of item writing

flaws (IWFs) evident in objective examination questions constructed by lecturers in Federal College of Education (Technical) Asaba, Delta State, Nigeria has remained empirically uncertain and undetermined. Hence, this unsatisfactory state of affairs created the gap which the present study sought to fill.

#### **PURPOSE OF THE STUDY**

In specific terms, the study sought to:

1. determine the types of objective question formats constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State.
2. examine the prevalent rate of item writing flaws in objective questions constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State.
3. identify the nature of item writing flaws associated with objective questions constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State.

#### **RESEARCH QUESTIONS**

1. What are the types of objective formats constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?
2. What is the prevalent rate of item writing flaws in objective questions constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?
3. What is the nature of item writing flaws associated with objective questions constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?

#### **METHODOLOGY**

##### **Research Design**

Descriptive cross-sectional design was adopted in the conduct of this study. According to Kesmodel (2018),

descriptive cross-sectional design is most relevant when there is the need to provide estimates of the prevalence (number of cases) of certain phenomena, attitudes, knowledge and behaviour. In the light of the present study, prevalence of IWFs was estimated using the end-of-semester objective examination questions to establish a possible link between the quality of test constructed and high rate of pass or failure of students in examinations.

### **Population and Sample**

The population of the study comprised 57 retrieved objective examination questions constructed by lecturers and administered to NCE I (n = 32), NCE II (n = 17) and NCE III (n = 8) questions papers over the period of three academic sessions (2016/2017, 2017/2018 and first semester 2018/2019) in centrally conducted examinations in Federal College of Education (Technical), Asaba. Purposive sampling technique was adopted which enabled the researchers to use the entire population as sample due to its manageable size.

### **Instrument for Data Collection**

Secondary data comprising 57 objective examination question papers administered to students in different departments over the period of three academic sessions (2016/2017, 2017/2018 and first semester 2018/2019) was collected for the study using a designed checklist (pro forma). The checklist was validated by two experts in Measurement and Evaluation, Federal College of Education (Technical), Asaba. The suggestions of the experts improved the veracity of the instrument. Similarly, the validity of the retrieved objective question papers for the study

## **RESULTS**

**Research Questions 1:** What are the types of objective formats constructed by

were determined through internal and external moderation exercise for year one and second, and third year respectively and were adjudged suitable for administration.

### **Procedure of Data Analysis**

Three expert judges comprising one subject expert and two in the field of measurement and evaluation in Federal College of Education (Technical) Asaba were engaged to review each test items and identify any of the empirically suggested 19 frequently occurring IWFs contained in each question papers. The quality of the items in each of 57 objective examination questions was examined and classified into standard or flawed, if flawed the exact type of item flaw or flaws contained within the question (including options) was recorded. Controversial test items and disagreements concerning multiple flaws within an item among expert judges were resolved through a consensus process.

### **Method of Data Analysis**

The retrieved 57 objective question papers used for analysis cut across all the seven Schools in the College viz., (School of Education, 26; School of Science Education, 6; School of Vocational Education, 1; School of Technical Education, 1; School of Early Childhood and Primary Education, 8; School of Adult and Non-Formal and Special Education, 11; School of Business Education, 3; and 1 VTE cutting across Schools of Business, Vocational and Technical). However, some of the examinations written overlap for some students. Frequency counts and percentage mean and standard deviation statistics were used for analysis of data.

lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?

**Table 1: Types of objective question formats constructed by lecturers for semester examinations**

Objective Question Formats	Frequency ( <i>f</i> )	Percentage (%)	Mean ( $\bar{X}$ )	STD
Short Answer Questions (SAQs)	34	59.6	11.40	13.96
Multiple Choice Questions (MCQs)	16	28.1		
Alternate Response Questions (ARQs)/SAQs	3	5.3		
MCQs/SAQs	2	3.5		
MCQs/ARQs/SAQs	2	3.5		
<b>Total</b>	<b>57</b>	<b>100</b>		

Table 1 reveals objective question formats constructed for semester examinations in Federal College of Education (Technical) Asaba for which a mean and standard deviation scores ( $\bar{X} = 11.40$ ;  $SD = 13.96$ ) were obtained indicating that the short answer objective questions (SAQs) and multiple choice questions (MCQs) points of data are larger than the mean score and a high standard deviation score ( $\bar{X} = 11.40 < SD = 13.960$ ) suggests a wide variability in the objective question formats constructed by lecturers.

Similarly, 34 out of 57 objective question formats constructed by lecturers for end of semester examinations are short answer objective questions (SAQs) representing 59.6%, 16 of 57 objective

question formats constructed by lecturers are multiple choice questions (MCQs) representing 28.1%, 3 of the 57 objective question formats are combination of alternate response questions (ARQs) and SAQs representing 5.3%. 2 of the 57 objective questions constructed are MCQs/SAQs and MCQs/ARQs/SAQs represent 3.5% respectively. Analysis of data suggests dominance of short answer questions followed by multiple choice questions and alternate response questions constructed alone or combination of two or more objective question formats are constructed for end-of-semester examinations but matching item questions are avoided.

**Research Questions 2:** What is the prevalent rate of item writing flaws in objective questions constructed by

lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?

**Table 2: Frequency of item writing flaws contained in each objective examination questions evaluated**

No of Objective Question Papers Evaluated ( $\bar{X}$ )	No of IWFs identified ( <i>f</i> )	Percentage (%)	Mean ( $\bar{X}$ )	STD
2	0	0.0	2.50	1.37
21	1	6.7		
25	2	13.3		
3	3	20.0		
5	4	26.7		
1	5	33.3		
<b>Total</b>	<b>57</b>	<b>100</b>		

Table 2 reveals the occurrence of IWFs in each of the objective examination question papers evaluated with a mean and standard deviation scores ( $\bar{X} = 2.50$ ;  $SD = 1.37$ ). This indicates that any point of data ( $f = 0$ )

implies no flaws, any point of data less than the obtained mean score ( $f < \bar{X} = 2.50$ ) indicate low rate of flaws and all points of data ( $f \geq 2.50$ ) shows high rate of flaws. The standard deviation ( $SD = 1.37$ ) shows a generally close spread of

IWFs across majority of the evaluated objective question papers.

In the same way, out of the 57 objective question papers evaluated, 2 (0.0%) contained no IWFs respectively, 21 objective question papers contained 1(6.7%) IWF each, 25 objective question papers contained 2(13.3%) IWFs each, 3 objective question papers contained 3(20.0%) IWFs each, 5 question papers contained 4 (26.7%) IWFs each while 1 question paper contained 5(33.3%) IWFs. Cumulatively, of the 57 objective examinations question papers evaluated, 55(96.5%) are flawed as they contained at least a minimum of one and maximum of five (1-5) IWFs whereas 2 (3.5%) of the objective question papers are standard due to none violation of IWGs.

Analysis of data therefore, suggests a minimal rate of (1 – 2) IWFs in 46 (20.0%) of the 55 flawed objective examination question papers but a high rate of (3 – 5) IWFs in 9 (80.0%) flawed objective question papers. Generally, 55 (96.5%) of the 57 end of semester objective examination question papers evaluated were fraught with one or more IWFs.

**Research Question 3:** What is the nature of item writing flaws associated with objective questions constructed by lecturers for end of semester examinations in Federal College of Education (Technical) Asaba, Delta State?

**Table 3: Nature of item writing flaws committed in objective examination question papers constructed by lecturers**

Types of IWFs	f	%	Mean (X)	STD
1. Grammatical clues (Using a word that gives hint on the correct response)	1	.95	1.86	1.95
2. Logical clues (Arrangement of correct options using predictable patterns)	2	1.9		
3. Use of absolute terms (Always, never, only, and all)	0	0		
4. Extra detail in correct option (Longest option is the correct option)	0	0		
5. Implausible distracters or decoys (when one or more decoys are obviously incorrect)	4	3.6		
6. Mutually exclusive distracters (when two out of four option are known to be wrong answers)	5	4.8		
7. Use of all of the above (AOTA) in the options	1	.95		
<b>IWFs related to testwiseness committed</b>	<b>13</b>	<b>12.38</b>		
8. Use of negative stem (Not true, true except, incorrect, not, except)	3	2.9	7.67	10.76
9. Use of none of the above (NOTA)	2	1.9		
10. Response options are a series of true or false statements (Which of the following statements is correct or incorrect?)	0	0		
11. Complex stems (K-Type) (selection of combination of options- A. I & II B. I & III, et cetera) or combination of AOTA and NOTA in decoys	2	1.9		
12. Use of vague or generalizing terms (Sometimes, frequently, often, occasionally, typically)	0	0		
13. Unfocused stem (broad and open-ended questions that do not pose specific problem, unrelated response options or distantly related to a single learning objective)	8	7.6		
14. More than one correct answer or no answer	14	13.3		
15. Poor formatting (wrong vocabulary, incorrect punctuation, horizontal order of options instead of vertical and mix-up of items of different test formats)	27	25.7		
16. Heterogeneous options/Dissimilar number of options (use of different response options in MCQs)	1	.95		
17. Irrelevant or misleading information in the stem	3	2.9		
18. Numeric data not stated correctly	1	.95		
19. Awkward stem structure (complete the sentence or fill-in-the-blank by placing the response at the end of a sentence)	31	29.5		
<b>IWFs related to irrelevant difficulty committed</b>	<b>92</b>	<b>87.62</b>		
<b>Total</b>	<b>105</b>	<b>100</b>		

The results presented in Table 3 reveals the nature of IWFs associated with the 57 evaluated objective examination question papers used for end of semester examination. Item 1 – 7 — constitute testwiseness related nature of IWFs committed which makes it possible for students to easily answer questions correctly than envisaged were violated 13 times except for item 3 and 4 with a mean and standard deviation score ( $X = 1.86$ ;  $SD = 1.95$ ) indicative of a slightly close spread of testwiseness related nature of IWFs across item 1, 2, 5, 6 and 7. On the other hand, item 8 – 19 comprise irrelevant difficulty related nature of IWFs involved which make questions more difficult for students to answer correctly than intended were compromised 92 times except for item 10 and 12 with a mean and standard deviation score ( $X = 7.67$ ;  $SD = 10.76$ ) signifying a high widespread of irrelevant difficulty related nature of IWFs in nearly all items.

Testwiseness related nature or category of IWFs committed by lecturers is 13 (12.38%). The respective IWFs in this category are "grammatical clues," and "use of AOTA" are 1 (.95%) respectively. "Use of absolute terms" and "extra detail in the correct option" have 0(0.0%) respectively. "Logical clues" accounts for 2 (1.9%) whereas "implausible distracters or decoys" 4(3.6%) and "mutually exclusive distracters" 5 (4.8%).

Irrelevant difficulty related category or nature of IWFs committed in the evaluated objective question papers constitutes 92(87.62%). Irrelevant difficulty related IWFs of high occurrence are "awkward stem structure" 31 (29.5%), "poor formatting" 27 (25.7%), "more than one correct answer or no answer" 14 (13.3%) and "unfocused stem" 8 (7.6%). Others are "use of negative stem" and "use of irrelevant and misleading information in the stem" account for 3 (2.9%) respectively. Use of "none of the above (NOTA)," and "complex stem (K-Type)" account for 2 (1.9%) respectively. "Heterogeneous options" and "numeric data not stated correctly," represent 1 (.95%)

respectively whereas "response options are a series of true or false statements" and "use of vague or generalizing terms" account for 0(0%) respectively. Analysis of data suggests that testwiseness related and irrelevant difficulty related natures of IWFs are committed but irrelevant difficulty related IWFs are more evident and pervasive.

## DISCUSSIONS

The study revealed that short answer questions followed by multiple choice questions and alternate response questions alone or combination of two or more objective question formats are constructed for end-of-semester examinations. This pattern of finding is attributable to the perceived missing link in the NCCE prescription of objective method of assessment with emphasis on number of items (questions) to be constructed on the basis of course credit only without consideration for suitability or otherwise of some of the existing objective test formats for students at NCE level. Lecturers seem to take advantage of the gap in policy to construct objective test format that seems convenient but not necessarily suitable. Hence, lecturers' preference for short answer format of objective examination questions rather than multiple choice questions is inconsistent with Udoh (2016) and Omorogiuwa (2010) who reported that multiple-choice test format (MCQs) is the most extensively used objective tests in educational testing. The choice of short answer method of assessment tends to influence students' choice of learning approach (Biggs, 2003; Reid, Duvall and Evans, 2007 in Tariq, Tariq, Maqsood, Jawed, & Baig, 2017). It is also inconsistent Pais, et al. (2016) who reported the most common methods of assessment are multiple choice questions (MCQs), extended matching questions (EMQs), short essay questions (SEQ), among others. This is not in agreement with Lee (2012) who avers that multiple choice test items are suitable for measuring learning at knowledge, comprehension, application, analysis, synthesis and evaluation levels of



cognitive domains. The use of alternate response questions seems unsuitable for students in tertiary institutions as it has been adjudged to be good for young children (Okoye, 2015).

The second finding showed a minimal rate of (1 – 2) IWFs in 46 (20.0%) of the 55 flawed objective examination question papers but a high rate of (3 – 5) IWFs in 9 (80.0%) flawed objective question papers. However, 55 (96.5%) of the 57 end of semester objective examination question papers evaluated were fraught with one or more IWFs. This finding is in conformity with, Tariq, Tariq, Maqsood, Jawed and Baig (2017) who reported that tertiary educational institutions have lecturers with diverse background with few trained and well experienced in test construction while others are new in teaching leading to the likelihood of teacher-made-objective tests that are not of the desired quality hence consisting of quite a lot of IWFs. Similarly, this finding is consistent with Rush, et al. (2016) who reported that item writing flaws are prevalent in faculty-prepared examinations which could render examination questions easier or more difficult than intended. This is also in line with Tarrant, Knierim, Hayes and Ware (2006) in Nedeau-Cayo, Laughlin, Rus and Hall (2013) and Rush, Rankin and White (2016) who reported that teacher-developed examinations across many disciplines are excessively rife with item writing flaws (IWFs) while Drasgow, Luecht and Bennett (2006) in Royal, Hedgpeth and Posner (2018) reported that writing quality test items has remained a long standing problem confronting teachers. In line with reported findings of previous related studies, this study shows that teacher-made-tests largely breach standard item writing guidelines thereby eroding the quality and trustworthiness of objective examination questions. Mehrens and Lehmann in Ali and Ruit (2015) attributed the poor quality of examinations constructed by teachers to lack of adequate training in test construction leading to inaccurate assessment of students. In credence to the assertion above, Tariq, Tariq, Maqsood, Jawed, and

Baig (2017) affirmed that designing of a high-quality MCQs requires skill made possible by training and practice without which teachers are more liable to commit more IWFs in the construction of assessment instruments.

The study found that testwiseness related and irrelevant difficulty related natures of IWFs are associated with objective examination questions constructed by lecturers but irrelevant difficulty related IWFs are more evident and pervasive. The testwiseness related natures of IWFs leads to misclassification of tested students as passed while irrelevant difficulty related natures of IWFs which is more evident and pervasive leads to misclassification of students as failed rather than passed. This finding is in line with IWFs lead to 10-15% misclassification of tested students as failed rather than passed (Downing in Pais et al., 2016) while Pais et al. aver that violation of IWGs relating to content concerns, stem writing and writing the choices have negative impact on the psychometric properties of test. This is related to Baig, Ali, Ali and Huda (2014) reported that poor quality questions tend to promote the assessment of superficial learning approach. Also, this finding is in conformity with Downing (2005) in Tariq, Tariq, Maqsood, Jawed, and Baig (2017) who reported that IWFs occur when test writers stray from accepted item writing guidelines distorting students' performance such that it becomes easier or difficult for them to respond to items correctly.

## **CONCLUSION**

The study concludes that lecturers' choice of convenient objective test formats such as alternate response questions (ARQs) is unsuitable for testing students at NCE level and promotes construction of low quality items. However, multiple choice questions (MCQs) only or in combination with short answer questions (SAQs) and matching item questions (MIQs) are adjudged suitable for assessment of NCE level students depending on the topics and courses. Again, lecturers' low level skill in

standard objective test item writing contributes to established flawed objective examination question papers constituting threats to quality and effectiveness of assessment instruments. Finally, the study also concluded that testwiseness related nature of IWFs and the pervasive irrelevant difficulty nature of IWFs render assessment outcomes inaccurate thereby making pass rate a positive false and fail rate a negative false. This occurs through misclassification of many tested students as "passed" rather than "failed" and several tested students as "failed" instead of "passed" respectively.

### Recommendations

Based on the findings and conclusion of this study, it is recommended that:

1. Adoption of one best answer (OBA) multiple choice questions (MCQs) with at least 3-response options as a standard and suitable objective method of assessment by NCCE and Management of Colleges of Education; or a 75% MCQs, 15% Short Answer and 10% Matching items as a standard and suitable objective method of assessment.
2. Use of experts and experienced lecturers in regular training of lecturers to acquire requisite skills in construction of quality objective examination questions to abate flaws that benefit testwise students and at the same time, preclude flaws related to irrelevant difficulty.
3. Specialists and experienced lecturers in test construction and should be engaged as moderators of objective examination questions by Departments and/or Schools for easy identification of any possible IWFs that may jeopardize the quality of such assessment instruments.

**Conflict of Interest:** The authors hereby declare no conflicts of interest.

**Acknowledgements:** This work was not funded by any organization or group. However, the authors express their unalloyed gratitude to all authors cited in this work and Dr. Ayo-Vaughan Adewunmi for his useful review, comments and suggestions.

**Disclaimer Statement:** The authors hereby declare that this study was independently carried out by the authors and expression of ideas and opinions therein are purely that of the authors and do not represent the official position of FCE (T) Asaba.

### Authorship and Level of Contribution:

**POA** coined the title, raised the research questions, designed the study, collected and compiled the examination question papers and interpreted data.

**VOI** contributed in review of extant literature, data analysis and editing the manuscript. All authors contributed to manuscript revision before publication.

### REFERENCES

- Ali, S. H., & Ruit, K. G. (2015). The impact of item flaws, testing at low cognitive level, and low distractor functioning on multiple-choice question quality. *Perspective of Medical Education*, 4, 244 - 251.  
<https://doi.org/10.1007/s40037-015-0212-x>.
- Baig, M., Ali, S. K., Ali, S., & Huda, N. (2014). Evaluation of multiple choice and short essay question items in basic medical sciences. *Pakistani Journal of Medical Science*, 30, 3-6.
- Federal Ministry of Education (2012). *National Commission for Colleges of Education curriculum implementation framework for Nigeria certificate in education*. Nasara Press Ltd.
- Federal Ministry of Education (2014). *National policy on education (6<sup>th</sup> ed.)*. NERDC Printing Press.
- Kesmodel, U. S. (2018). Cross-sectional studies: What are they good for? *Acta Obstet Gynecol Scand*, 97, 388-393.

- <https://doi.org/10.1111/aogs.13331>.
- Khan, H. F., Danish, K. F., Awan, A. S., & Anwar, M. (2013). Identification of technical item flaws leads to improvement of the quality of single best multiple-choice questions. *Pakistani Journal of Medical Science*, 29(3), 715-718. <http://dx.doi.org/10.12669/pjms.293.2993>.
- Lee, O. K. (2012). *Test planning: Constructing language tests*. <https://www.smartt2.ums.edu.my>.
- Nedeau-Cayo, R., Laughlin, D., Rus, L. & Hall, J. (2013). Assessment of item-writing flaws in multiple-choice questions. *Journal for Nurses in Professional Development*, 9(2), 52-57. <https://doi.org/10.1097/NND.0b013e318286c2f1>.
- Okoye, R. O. (2015). *Educational and psychological measurement and evaluation* (2<sup>nd</sup> edition). Erudition Publishers.
- Omorogiuwa, O. K. (2010). *An introduction to educational measurement and evaluation*. Perfect Touch Print.
- Ovat, S., & Ofem, U. J. (2017). Teachers' variables and application of test blue prints in learners' assessment in secondary schools in Cross River State. *International Journal of Scientific Research in Education*, 10 (1), 112-118.
- Pais, J., Silva, A., Guimarães, B., Povo, A., Coelho, E., Silva-Pereira, F., Lourinho, I., Ferreira, M. A. & Severo, M. (2016). Do item-writing flaws reduce examinations psychometric quality? *BioMed Central Research Note*, 9 (399), 1-7. <https://doi.org/10.1186/s13104-016-2202-4>.
- Pham, H., Besanko, J., & Devitt, P. (2018). *Examining the impact of specific types of item writing flaws on student performance and psychometric properties of the multiple choice question*. <https://doi.org/10.15694/mep.2018.0000225.1>.
- Royal, K. D., Hedgpeth, M. W. & Posner, L. P. (2018). A simple method for discerning item construction flaws in health professions examinations. *Health Profession Education*, <https://doi.org/10.1016/j.hpe.2018.03.001>.
- Rush, B. R., Rankin, D. C & White, B. J. (2016). The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Medical Education*, 16, 250. <https://doi.org/10.1186/s12909-016-0773-3>.
- Silker, R. T. (2003). *Teachers and tests*. Basil Blackwell.
- Simon, M. K., & Goes, J. (2013). *Dissertation and scholarly research: Recipes for success*. Seattle, WA: Dissertation success LLC. [www.dissertationrecipes.com](http://www.dissertationrecipes.com)
- Tariq, S., Tariq, S., Maqsood, S., Jawed, S. & Baig, M. (2017). Evaluation of cognitive levels and item writing flaws in medical pharmacology internal assessment examinations. *Pakistan Journal of Medical Science*. 33 (4), 866-870. <https://doi.org/10.12669/pjms.334.12887>.
- Udoh, E. G. (2016). Students' perception of multiple choice item format and mathematics achievement test in junior secondary schools in Uyo Educational Zone of Akwa Ibom State, Nigeria. *Academia Journal of Educational Research*, 4 (7), 111-116. <https://doi.org/10.15413/ajer.2016.0117>.