Development and Factorial Validation of Basic Science and Technology Performance Test for Primary School Pupils

Matthew Cyril

Department of Science and Vocational Education

Faculty of Education and Extension Services

Usmanu Danfodiyo University, Sokoto, Sokoto State, Nigeria

E-mail: mmmcyril@gmail.com

matthew.cyril@udusok.edu.ng

GSM: 08032210388

Abstract - The paper was designed to develop and factorially validate a basic science and technology performance test (BSTPT) for primary school pupils. The study adopted instrumentation research design. Percentage weights were assigned to each content area understudy. Initial pools of seventy (70) multiple-choice items were selected based on already constructed test blueprint. Kuder-Richardson (KR20) formula was used to determine the internal consistency reliability estimate of the instrument. Simple random sampling technique was used to draw three hundred and eighty- four (384) male and female pupils from three public and three private schools within Sokoto Metropolis, upon which BSTPT was trial tested for factor analysis. After the trial testing, thirty - six (36) items survived. Prior to factor analysis using principal components analysis (PCA), the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of 0.3 and above. The Kaiser- Meyer-Olkin value was 0.82, exceeding the recommended value of 0.6 and Bartlett's Test of Sphericity reached statistical significance of 0.000 supporting the factorability of the correlation matrix. These were further trial tested on a sample of 150 pupils and scored out for item analysis. Item analysis results revealed that thirty-two (32) items of the BSTPT falls within the acceptable difficulty index of $0.3 < K < 0.8$ and discrimination index of $D \geq 0.4$ An internal consistency reliability coefficient of 0.87 was established. Based on these findings, BSTPT instrument developed is reliable and dependable for use. It was recommended as follows: Percentage weights should be assigned to content area before constructing table of specification for content validity. Prior to performing factor analysis, suitability of data should be assessed to ascertain inspection correlation matrix and Bartlett's test of sphericity. Only items found within the acceptable difficulty index of $0.3 < K < 0.8$ and discrimination index of $D \geq 0.4$ should be included in the final form of the instrument. Instrument must exhibit a considerable internal consistency reliability coefficient for intended use.

Index Terms - Percentage weight, test blueprint, factor analysis, item analysis, basic science and technology.

Introduction

Basic Science and Technology (BST) is one of the core subjects in the new 9-year Basic Education Curriculum (BEC) for Nigerian Primary and Junior Secondary Schools Education Program. One of the basic objectives of BST as enshrined in the curriculum for primary and junior secondary school is knowledge and skills acquisition. On the average, education programs are goal oriented and assessments are used to measure and evaluate the intended knowledge content and expected skills to be acquired. No doubt then, that studies by Kirton, Hallam,

Peffers, Robertson and Stobart, [1] indicated that assessment in science has a profound and positive impact on teaching and learning. Assessment according to Morrison [2] is an integral part of instruction, as it determines whether or not the goals of education are being met.

The researcher sees assessment as a process by which learning can be improved using a systematically developed and professionally validated reliable instruments. It should be systematic because development of assessment instrument is not to be haphazardly done. It must follow the acceptable and recognized stages of instrument development, which include: analysis of content, review of instructional objectives, assignment of percentage weights, construction of table of specification, items generation, face validity, factor analysis, item analysis and determination of internal consistency reliability coefficient. In the view of Nworgu [3] a test is a structured situation comprising a set of items (questions or statements) with preferred responses, given to an individual(s) to determine the amount of the relevant trait or attribute they possess. To Williams [4], test indicators are signs or activities that show whether the students' performance is satisfactory or not. Therefore, test at all levels of education especially primary school level is an appropriate procedure for assessing pupils' performance.

In this study however, the paper and pencil approach using a multiple-choice objective test which [5] described as an appropriate measurement tool was adopted to assess primary school pupils' performance in basic science and technology within Sokoto metropolis. Primary education has been described as the basic and foremost right of every child [2]. In contrast to public schools which are fully funded, conducted and maintained by government, Collins [6] described private schools as schools under the financial and managerial control of a private body, individuals or charitable trust, accepting mostly fee-paying pupils.

The problem

The attainment of the teaching objectives of primary education, to some extent depends on the effective assessment of learning outcome, because assessment of learning outcomes provide feedback on the extent to which learning objectives are met. These can only be achieved with properly developed assessment instruments that are valid, reliable and established within the acceptable facility indexes. A close observation of most teacher made tests at Basic Science and Technology levels are not properly developed, as many of such instrument's lacks sound psychometric properties. This problem has resulted in having pupils who completed primary school without having acquired the basic knowledge required for further studies, particularly in science. Consequent upon this, pupils would be forced to drop the study of science and technology because of the difficulty they may face in continuing with science, occasioned by the poor method of assessments. In view of these possible problems, it is imperative to develop and factorially validate a test instrument that will determine students' performance in science and technology. The inevitable consequences of these prompted the study on the development and factoral validation of basic science and technology performance test for primary school pupils.

Objectives of the study

The following are the objectives of the study.

1. determine the mean and percentage weight of content areas of Basic Science and Technology,
2. establish Test Blue-Print of the instrument,
3. determine the factorability of correlation matrix,
4. determine the difficulty index of the items,
5. determine the discrimination index of the items,
6. determine the internal consistency reliability coefficient of the instrument,

Research Questions

1. What percentage weights are assigned to content areas of Basic Science and Technology?
2. What is the Test Blue-Print of the instrument?
3. What is the factorability of correlation matrix?
4. What is the difficulty index of the items?
5. What is the discrimination index of the items?
6. What is the internal consistency reliability coefficient of the instrument?

Significance of the Study

The theoretical significance of this study is the fact that item difficulty and item discrimination which are two statistics in Classical Test Theory (CTT) are central to this study. The theory provides a general framework of logic and mathematical models that underlie standard practices in test construction and use, which are seldom followed in the development of teacher made tests at primary school levels. The study will be of practical significance to the following groups: Primary school pupils, teachers of basic science and technology, curriculum developers, examination bodies, researchers, etc. For instance, the study will significantly assist teachers assess pupil's performance in basic science and technology. It will also provide teachers with feedback in order to improve the effectiveness of their own teaching.

Test Theory

The two currently and most popular statistical frameworks identified for addressing measurement problems such as test development are Classical Test Theory (CTT) and Item Response Theory (IRT) Hambleton and Jones [7]. The theory of interest in this paper is the Classical Test Theory (CTT). It is a theory about test scores that introduces three concepts: test score (often called the observed score), true score and error score. Research studies [8,9] have shown that most works in CTT are focused on models at the test scores level. That is, the models have linked test scores to true scores. However, item statistics that represent item difficulty and item discrimination power have been adopted, and their connection to test statistics, such as test-score, mean, standard deviation and reliability are well-known and are used in test development process to produce tests with the desired statistical properties. These item statistics (and their variations) have proven valuable in test development process. They are well known and have a long track record.

Method

The research design employed in this study is instrumentation research design. An instrumentation research design is a study in which a researcher develops and validates a test instrument as its major focus based on a certain curriculum [10].

Development and Content validity of BSTPT

Content analysis was first carried out to provide a summary of the intentions of the curriculum expressed in content terms [11]. Content analysis enables a researcher to be acquainted at the very beginning with the content area to be covered by the scope of the study. The content areas of basic science and technology provided on the new 9-year basic education curriculum include: basic science, basic technology, information communication technology (ICT) and physical and health education (PHE). In this paper however, the content areas of interests are: observing, classifying, measuring, communicating, inferring, predicting and number relationship, which are treated under the following topics in the basic science and technology curriculum: Measurement, Importance of Colours, Reflection of Light, the Water Cycle, Acid and Bases, Materials, Heat and Temperature, Force, Meaning and Forms of Energy, Common Domestic and Wild Animals, Uses of Soap, etc.

The second step in the development of a test is the review of instructional objectives. Instructional objectives are those behavioural changes, which a teacher expects to notice in his students after they have been exposed to a particular topic [12]. One of the major instructional objectives to be achieved by this study as reflected in the national core curriculum for basic science and technology is acquisition of basic knowledge in science and technology.

Judges are appointed to assign weight for an intended 70 items performance test in basic science and technology. These include: the researcher, three teachers of basic science and technology and three lecturers of science education who independently suggested weights for the content areas on table 1.

Table 1: Mean and percentage weights assigned to content areas of basic science and technology

S/N	Content Area	Judges' Weight A	B	C	D	E	F	G	S Total	S^2 Rating	Mean	% weight	No. of Items
1.	Observation	4	5	2	6	3	7	6	33	1089	4.67	20	14
2.	Classification	4	4	4	5	5	5	6	33	1089	4.67	20	14
3.	Measuring	3	4	4	5	3	6	4	28	784	4.0	16	12
4.	Communication	2	3	2	4	2	3	3	19	361	2.67	12	8
5.	Prediction	1	1	3	2	2	3	4	16	256	2.33	10	7
6.	Inferring	1	2	2	3	3	2	3	16	256	2.33	10	7
7.	Number Relationship	3	2	3	2	2	3	4	19	361	2.67	12	8
		Total								4196	23.34	100	70

Table 1 indicates that 20% of the 70 items gives 14 numbers of items. This implies that 14 items are to be developed from Observation in that order, as would be reflected on the test blueprint (Table of specification). Anene and Ndubuisi [12] described test blueprint as a means of ensuring content validity. The use of table of specification for constructing an instrument according to Ali (2006) will ensure a fair and comprehensive coverage of all the contents and learning outcomes to be included in the instrument. Akujo and George [13] defined content validity of a test as its ability to measure the subject matter content in relation to the instructional objectives. It relates the content and the instructional objectives to be achieved. This is similar to what [14] described as the master plan and should be readily available to guide the test constructors in allocating number of items per content area.

Table 2: Test blueprint

Content/Objectives		**Knowledge** 60%	**Comprehension** 25%	**Application** 15%	**No. of** Items
Observation	20%	8	4	2	14
Classification	20%	8	4	2	14
Measurement	16%	7	3	2	12
Communication	12%	5	2	1	8
Prediction	10%	4	2	1	7
Inferring	10%	4	2	1	7
Number Relationship	12%	5	2	1	8
Total	100%	41	19	10	70

Table 2 present the content validity of the intended instrument in which 70 items are to be generated. Items generation or writing is done based on the format of test adopted by the researcher. In this paper, the multiple-choice format was adopted in generating the seventy 70 items. After writing the items, they are sent for face validation. Anyaokoha [15] stated that a test is said to have face validity if it looks like going to measure what is made to measure. Face validity involves a cursory glance of a test. A test that looks like it measures what is expected to measure has face validity [16]. This was achieved by subjecting the instrument to the scrutiny of relevant experts in measurement and evaluation as well as subject expects in the area of the study. After the face validity of the instrument, items were reviewed to identify those that survive the scrutiny. Those that survived are trial-tested by administering the test on an equivalent sample of the group for which the test is developed [17]. This was done in other to establish construct validity of the instrument through factor analysis.

Factor analysis of BSTPT

The 70 originally generated items of the multiple-choice basic science and technology performance test (BSTPT) were trial tested on 384 male and female samples drawn from public and private primary schools within Sokoto metropolis. Data from the dichotomously scored items were subjected to principal components analysis (PCA) using SPSS Version 20. Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of 0.3 and above. The Kaiser-Meyer-Olkin value was 0.82, (table 3) exceeding the recommended value of 0.6 (Kaiser 1970, 1974) and Bartlett's Test of Sphericity [18] tests the null hypothesis that the correlation matrix is an identity matrix. That is, a matrix in which all of the diagonal elements are 1 and all off diagonal elements are 0. Bartlett's test (table 3) reached statistical significance of 0.000 supporting the factorability of the correlation matrix. Principal components analysis revealed the presence of 10 components (table 4) with eigenvalues exceeding 1, explaining 25.4%, 7.5%, 4.9%, 4.5%, 4.1%, 3.7%, 3.5%, 3.3%, 3.0% and 2.8% of the variance respectively.

Table 3: **KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.820
Bartlett's Test of Sphericity	Approx. Chi-Square	2033.018
	df	630
	Sig.	.000

Table 4: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.126	25.349	25.349	9.126	25.349	25.349
2	2.701	7.503	32.853	2.701	7.503	32.853
3	1.762	4.893	37.746	1.762	4.893	37.746

4	1.632	4.534	42.280	1.632	4.534	42.280
5	1.479	4.109	46.389	1.479	4.109	46.389
6	1.340	3.722	50.112	1.340	3.722	50.112
7	1.263	3.507	53.619	1.263	3.507	53.619
8	1.196	3.323	56.942	1.196	3.323	56.942
9	1.088	3.023	59.965	1.088	3.023	59.965
10	1.021	2.836	62.801	1.021	2.836	62.801

Extraction Method: Principal Component Analysis.

Scree Plot

An inspection of the scree plot in figure (1.0) revealed a clear break after the second component. Using Catell [19] scree test, the researcher decided to retain two components for further investigation.

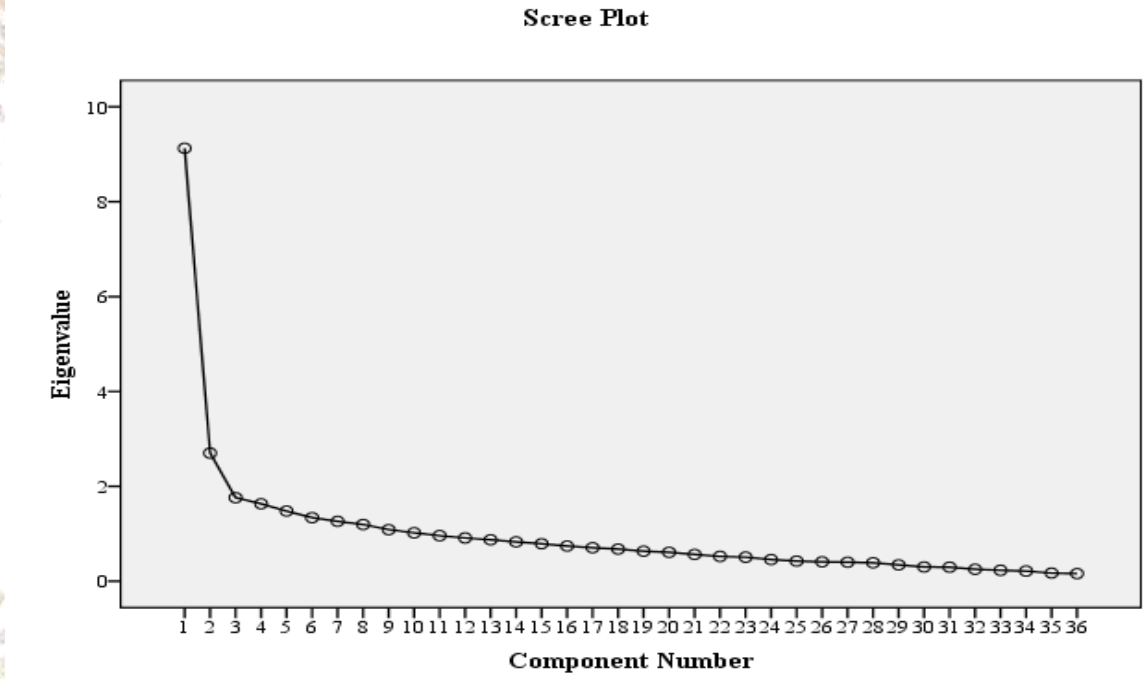

Figure 1.0: Scree Plot

This was further supported by the results of Parallel Analysis (table 5), which showed only two components with eigenvalues exceeding the corresponding criterion values for a randomly generated data matrix of the same size (36 variables × 384 respondents).

Table 5: Parallel Analysis

Component Number	Actual eigenvalue from PCA	Criterion value from parallel analysis	Decision
1	9.126	2.0677	Accept
2	2.701	1.9372	Accept

The two-component solution explained a total of 32.9% of the variance, with Component 1 contributing 25.4% and Component 2 contributing 7.5%. Table 5: Parallel analysis result further support rotation for 2 components as shown on figure 1.0

It is assumed that the underlying processes are uncorrelated; hence Varimax rotation was employed to maximize high correlations between factors and variables, and minimize low ones by maximizing the variance of the loadings within factors, across variables [20]. To aid in the interpretation of these two components, Varimax rotation was performed on the data structure in other to obtain a simple structure. Simple structure is a condition in which variables load at near 1 (in absolute value) or at near 0 on an eigenvector. Variables that load near 1 are clearly important in the interpretation of the factor, and variables that load near 0 are clearly unimportant. Simple structure thus simplifies the task of interpreting the factors. The rotated solution revealed the presence of simple structure as having made the criteria. Results of the factor analysis reduced the number of items on the instrument to 36. These were also trial tested on yet another sample of 150 pupils, and the data obtained were used for item analysis.

Item analysis refers to a statistical technique that helps researchers or test developers identify the effectiveness of the items in their instruments [21]. It is also defined as the specific methods used in education to evaluate test items, typically for the purpose of test construction and revision [22]. Very difficult items were reviewed or eliminated and those with negative discrimination indexes were deleted from the list of items on the instrument. Difficulty index is a measure of the level of how easy or tough each item in the instrument is. It is a measure of the rate of the number of pupils who got the item correct to the number of pupils who took the test [14]. Discrimination index on the other hand is an indicator that shows how sharply a test item differentiates between the more or less able students [14]. In other words, it is a measure on how the items are able to discriminate between the good students and the not so good students.

Out of the 36 items, only 32 items survived item analysis with difficulty index that ranged between 0.45 and 0.77, which is very much within the acceptable moderate difficulty index of $0.3 < K < 0.8$ recommended by Sang [23] and discrimination index that ranged between 0.37 and 0.52, which is very much within the acceptable moderate discrimination index of $D > 0.4$ also as recommended by Sang [23].

Reliability

Kuder-Richardson KR20 formula was used to determine the internal consistency reliability of the developed 32 items BSTPT for primary school pupils. The internal consistency reliability coefficient of BSTPT as calculated was found to be 0.868, which is reliable enough for intended use.

Results

1. Table 1 provided the following percentage weights: 20% of the 70 items gave 14 number of items, 16% gave 12 items, 12% gave 8 items and 10% gave 7 items in that order.
2. Test Blue-Print is provided on table 2
3. Bartlett's test of sphericity reached statistical significance of 0.00 supporting the factorability of correlation matrix.
4. Difficulty index: Item difficulty is the percentage of pupils that correctly answered the item. It is also referred to as the p-value; and ranges from 0 to1. High p-values mean the item is easy and low p-values mean the item is difficult [24]. P-values above 0.90 are very easy items and items with p-values < 0.3 are too difficult and as such, were reviewed either by modifying the distracters, re-writing the item content entirely or both. This was done because such items may not assess the objective it was designed to assess or the question could have been worded ambiguously [25]. Optimum difficulty level is 0.5 for maximum discrimination between high and low achievers. Item answered correctly by 70% examiners has a difficulty index of 0.70 If only 10% pass an item, it implies that the item it is hard or has low difficulty index. Generally, items of moderate difficulty index are to be preferred to those which are much easier or much harder [26].

Items within difficulty index of 0.4 to 0.8 were included in the final form of the test instrument, which had thirty-two (32) items in all. This is in line with the findings of, Shaibu and Mari [27], Thompson and Levitov [28], and Singh, Sharma and Brijesh [29]. It also corresponds with the work of Nworgu [30]. Based on these criteria, the researcher had items with difficulty index as follows: three (3) items representing 9.38% had difficulty index of 0.4, eleven (11) items representing 34.38% had difficulty index of 0.5, fourteen (14) items representing 43.75% had difficulty index of 0.6, and four (4) items representing 12.5% had difficulty index of 0.7. These selections are in line with test development techniques of the CTT adopted in this study.

5. **Discrimination Index**

Items with negative discrimination were deleted. This is in line with Matlock-Hetzel [31], Boopathiraj and Chellamani [26]. Items within discrimination index of 0.3 to 0.5 were included in the final form of the test instrument, which had twenty-three (32) items in all. This is in line with the findings of Shaibu and Mari [27], Ebel and Frisbie [32], Singh, Sharma and Brijesh [29] and Nworgu [30]. Based on these criteria, the researcher had items with discrimination index as follows: three (3) items representing 9.4% had discrimination index of 0.3, nineteen (19) items representing 59.4% had discrimination index of 0.4, and ten (10) items representing 31.2% had discrimination index of 0.5 Items within these discrimination indices were selected in line with the

provisions of test development techniques of the CTT adopted in this study. This implies that the instrument has a sound facility indexes as required by the theory.

6. Reliability of the developed Instrument

The valid thirty-two (32) items of BSTPT exhibited a high KR20 reliability coefficient of 0.87, well above 0.77, which was found suitable in a study by Shaibu and Mari [27]. This reliability coefficient is close to 0.88 obtained by Temiz, Taser and Tan [33]. This is also considered appropriate because to obtain a high reliability is an indication that the instrument has high internal consistency hence BSTPT is reliable and dependable.

Conclusion

On the basis of the major findings of this study, the following conclusions have been made.

1. Items with high difficulty index of $K > 0.9$ and very low difficulty index of $K < 0.3$ were modified, whereas items within the difficulty index of $0.3 < K < 0.8$ were accepted without modifications.
2. Items with positive medium discrimination index of $0.20 < D < 0.40$ were modified and items with positive low discrimination index of $0 < D < 0.20$ were rewritten. Items with positively high discrimination index of $D \geq 0.40$ were accepted without revision. However, those items with negative discrimination were out rightly eliminated.
3. Validation and items analysis of the BSTPT revealed that only 32 items survived. All the 32 items were within the acceptable difficulty index of $0.3 < K < 0.8$ and acceptable discrimination index of $0.4 = D > 0.4$ Hence, they were retained for the developed and factorially validated Basic Science and Technology Performance Test (BSTPT) for Primary School Pupils.
4. Content validity revealed percentage weights of 20%, 16%, 12%, 10% and a corresponding number of items per content area as follows: 14, 12, 8 and 7 respectively.
5. The internal consistency reliability analysis of BSTPT using Kuder-Richardson KR20 indicates that the instrument has a high internal consistency coefficient of 0.87 and is therefore reliable and dependable for use at Primary School level.

Recommendations

1. Percentage weights should be assigned to content areas before constructing table of specification for content validity.
2. Validity improves the quality of a test; hence test developers should endeavour to establish both face and content validity of their test instrument, and avoid the practice of self-validation. Expert validates in the field of study should be contacted to face validate the instrument, while test blueprint should be employed to establish content validity.
3. Difficulty index and discrimination index have great significance in test development. Test developers should be very careful while selecting or generating items of instrument. Only items found within the

acceptable difficulty index of 0.3 < K < 0.8 and discrimination index of D= > 0.4 should be included in the final form of the instrument.

4. Prior to performing factor analysis, suitability of data should be assessed to ascertain inspection correlation matrix and Bartlett's test of sphericity to conform with acceptable values and statistical significance respectively.
5. Instrument must exhibit a considerable internal consistency reliability coefficient for intended use.

References

[1] Kirton, A., Hallam, S., Peffers, J., Robertson, P. & Stobart, G. (2007). Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools. *British educational research journal,* 33(4), 605-627

[2] Morrison, G. R. (2010). *Fundamentals of early childhood education.* Pearson education Inc. US

[3] Nworgu, B. G. (2015). *Educational research.* Basic issues and methodology. 3rd edition University Trust Publishers Nsukka, Enugu State, Nigeria

[4] Williams, P. S. (2009). Skill acquisition: A tool foryouth development and empowerment. *Nigerian association of teachers of technology (NATT) 22nd annual national conference* P. 184 – 188

[5] Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short – Answer questions in a marketing context. *Journal of marketing education.* 25(1), 31-36

[6] Collins, W. (2012). *English dictionary.* Harpercollins publishers. Digital edition. Retrieved from www.dictionary.reference.com/browse

[7] Hambleton, R. K. & Jones, R. W. (1993). *Comparison of classical test theory and item response Theory and their applications to test development.* http://www.educationalmeasurement/issues&practice

[8] Millman, J. & Greene, J. (1989). *The specification and development of tests of achievement and Ability.* 3rd edition, New York: Macmillan.

[9] Henrysson, S. (1971). *Gathering, analyzing and using data on test items.* 2nd edition Washington D C: American council on education.

[10] Ali, A. (2006). *Conducting research in education and the social sciences.* Tashiwa Networks Ltd.

[11] Izard, J. (1997). Trial testing and item analysis in test construction. Paris: International Institute for Educational Planning

[12] Anene, G. U. & Ndubuisi, O. G. (2015). *Educational measurement and evaluation: Theory and practice.* 2nd edition. University Trust Publishers, Nsukka Enugu State, Nigeria

[13] Akujo, C. C. & George, C. C. (2010). Assessment of technological work skills required by Marketing education graduates in the world of work. *Nigerian vocational association Journal, 15(1), 206 – 213*

[14] Izukanne, M. I. (2008). *Educational psychology in measurement and evaluation.* University of Nigeria press Ltd.

[15] Anyakoha, E. U. (2009). *Development research skills, concepts and conceptual framework.* Nsukka: Great AP express publishers ltd

[16] Olayiwola, A. O. (2010). *Procedures in educational research.* Kongo (Nig.) Ltd. NN6 Lokoja Road/Lagos street Kaduna.

[17] Ezeh, D. N. (2015). *Educational measurement and evaluation: Theory and practice.* 2nd edition. University Trust Publishers, Nsukka Enugu State, Nigeria

[18] Shrestha, N. (2021). Factor Analysis as a Tool for Survey Analysis. *American journal of applied Mathematics and Statistics Vol. 9 No.1, 4-11* http://pubs.sciepub.com/ajams/9/1/2

[19] Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1, 245-276.*

[20] Tabacchnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. 5th Edition, Pearson Education inc. USA

[21] Crocker, L. & Algina, J. (2008). *Introduction to classical and modern test theory.* Cengage Learning USA.

[22] Eaves, S. & Erford, B. (2009). *Item analysis.* Retrieved from www.education.com/reference/ Article/item-analysis.

[23] Sang, M. S. (1994). Penilaian, Pemulihan Dan PengayaanDalamPendidikan". *SubangJaya:* Kumpulan BudimanSdn. Bhd.

[24] Schwarz, J. (2011). *Research methodology: Tools, applied data analysis (with SPSS).* Lecture 02: measurement scales and item analysis. Lucerne university of applied sciences and arts Retrieved from http://www.schwarzpartners.ch/applieddatanalysis

[25] Hong, T., Purzer, S. & Cardella, M. E. (2011). A psychometric re-evaluation of the design, Engineering and technology (det) survey. *Journal of engineering education.* 100(4), 800-818

[26] Boopathiraji, C. & Chellamani, K. (2013). Analysis of test items on difficulty level and Discrimination index in the test for research in education. *International journal of social Science and interdisciplinary research.* 2(2), 189-193

[27] Shaibu, A. A. M. & Mari, J. S. (2002). Development and standardization of an instrument for Assessing students' acquisition of science process skills in integrated science at the Junior secondary level. *Zaria Journal of Studies in Education, 3(1), 121-132*

[28] Thompson, B. & Levitov, J. E. (1985). Using microcomputers to score and evaluate test items. *Collegiate Microcomputer, 3, 163-168*

[29] Singh, Y. K., Sharma, T. K. & Brijesh, U. (2012). *Educational technology: Techniques of tests and evaluation.* New Delhi: APH Publishing corporation

[30] Nworgu, B. G. (2015). *Educational measurement and evaluation.* Theory and practice. 2nd edition. University Trust Publishers Nsukka, Enugu State, Nigeria

[31] Matlock-Hetzel, S. (1997). *Basic concepts in item and test analysis.* Retrieved from http://ericae.net/ft/tamu/Espy.htm

[32] Ebel, R. L. & Frisbie, D. A. (1986). *Essentials of educational measurement.* Englewood cliffs, NJ: Pentice - Hall

[33] Temiz, B. K., Tasar, M. F. & Tan, M. (2006). Development and validation of a multiple format Test of science process skills. *International education journal.* 7(7), 1007-1027

BASIC SCIENCE AND TECHNOLOGY PERFORMANCE TEST (BSTPT) FOR PRIMARY SCHOOL PUPILS

School type *(Please tick where appropriate):* Private [ ] Public [ ]

Sex *(Please tick where applicable):* Male [ ] Female [ ]

Instruction: ***Tick any option you consider to be the correct answer, please.***

1. Tongue is the organ in the mouth that we use for _____
- a. Smelling things
- b. Observing objects
- c. Tasting sugar
- d. Chewing

2. Colours are classified into two groups:
- a. Black and white colours
- b. Primary and Secondary colours
- c. Red and Blue Colours
- d. Yellow and Green Colours

3. The part of a Fish that helps it to swim in water is called _____
- a. Eyes
- b. Fins
- c. Head
- d. Mouth

4. Which of the following is NOT a domestic animal?
- a. Goat
- b. Chicken
- c. Dog
- d. Lion

5. How many minutes make 1 hour?
- a. 6o minutes
- b. 60 seconds
- c. 30 minutes
- d. 120 minutes

6. The instrument used in measuring temperature is called _____
- a. Barometer
- b. Rain guage
- c. Wind vane
- d. Thermometer

7. All of the following colours are primary colours except
- a. Blue
- b. Red
- c. Magenta
- d. Green

8. For any object to float on water and fly in air, it has to
- a. Be light in weight
- b. Small as stone
- c. Be big
- d. Have weight

9. The following are examples of objects that fly in Air except_____
- a. Balloon
- b. Kite
- c. Bird
- d. Dog

10. Birds fly with the help of their_____
- a. Wing

b. Beak
c. Tail
d. Legs

11. All the following objects produces sound **Except**
- a. Ringing bell
- b. Measuring tape
- c. Blowing flute
- d. Clapping of hands

12. Clinical thermometer is used to measure
- a. Human body temperature
- b. Room temperature
- c. Boiling water
- d. Rainfall

13. Which of the following organs of the body is used for tasting?
a. Lips
b. Teeth
c. Tongue
d. Saliva

14. Soil is classified into _____ types.
a. 2
b. 3
c. 4
d. 6

15. The difference between domestic animals and wild animals is _____
a. They both live at home
b. Domestic animals live in the forest while wild animals live at home
c. They both live in the forest
d. Domestic animals live at home while wild animals live in the forest.

16. Plants which grow on other plants are called

a. Parasites
b. Epiphytes
c. Symbiosis
d. Endocytes

17. What is the next number in the series: 3, 9, 27, _____?
a. 81
b. 90
c. 100
d. 101

18. Liquid substances that can dissolve solid objects are called _____
a. Solutes
b. Solvents
c. Liquids
d. Gas

19. There are 10 oranges in a box. Bello has 5 boxes and 3 more oranges. How many oranges have Bello in all?
a. 30
b. 50
c. 53
d. 18

20. An unripe fruit has _____taste?
a. Bitter
b. Sour
c. Salt
d. Sweet

21. Which of these is not a farm animal?
a. Cattle
b. Sheep
c. Donkey
d. Cat

22. The Rainbow is made up of _____ number of colours?
a. 6
b. 3
c. 7
d. 5

23. Living things are things that can reproduce young ones of their kind while non-living things _____
a. Can move from one place to another.
b. Cannot reproduce young ones of their kinds.
c. Can respire
d. Can excrete

24. A point in the body where two or more bones meet is called
a. Junction
b. b. Joint
c. Junctures
d. Tendon

25. The wearing away of the earth surface by wind or rain water is called
a. Pollution
b. Erosion
c. Evaporation
d. Sedimentation

26. Arrange these numbers: 273 269 305 410 285 from the biggest to the smallest:
a. 269 305 410 285 273
b. 305 410 285 273 269
c. 410 305 285 273 269
d. 410 285 273 269 305

27. The mixture of solute and solvent forms a _____
a. Solution
b. Condensation
c. Evaporation
d. Evolution

28. The disease that occurs as a result of protein deficiency is
a. Malaria
b. Cholera
c. Kwashiorkor
d. Cancer

29. 2 + 4 - 3 = _____
a. 3
b. 4
c. 6
d. 2

30. Rita has 6 sweets. She gave half of them to her brother. How many sweets does she have left?
a. 2
b. 3
c. 4
d. 5

31. The process of making water free from germs is called _____
a. Purification
b. Boilification
c. Saponification
d. Esterification

32. The raw material used in making cement is called
a. Basalt
b. Granite
c. Goldstone
d. Limestone

BSTPT KEY

Every correct answer caries a mark

1 C	**9 D**	**17 A**	**25 B**
2 B	**10 A**	**18 B**	**26 C**
3 B	**11 B**	**19 C**	**27 A**
4 D	**12 A**	**20 B**	**28 C**
5 A	**13 C**	**21 D**	**29 A**
6 D	**14 B**	**22 C**	**30 B**
7 C	**15 D**	**23 B**	**31 A**
8 A	**16 B**	**24 B**	**32 D**