Development and Content Validation of Multiple-Choice Basic Science Test (MCBST) for Junior Secondary Schools in IMO State

Onyejiekwe, Moses Chidube; Prof. I. M. Opara & Dr. A. Amadioha

Department of Educational Psychology, Guidance and Counselling, Faculty of Education, University of Port Harcourt, Rivers State, Nigeria

DOI: https://doi.org/10.68050/JAMS.2026.326

Received: 01 Sept 2026; Accepted: 02 Sept 2026; Published: 06 August 2026

ABSTRACT

This study focused on the development and content validation of a Multiple-Choice Basic Science Test (MCBST) for Junior Secondary School One (JSS I) students in Imo State, Nigeria. The study adopted an instrumentation research design. The population comprised 1,508 JSS I Basic Science students in 291 public secondary schools in Imo State, while 500 students were used for trial testing. An initial pool of 150 multiple-choice items was developed from the Basic Science curriculum using a table of specifications. The items were subjected to face and content review by three validators: two specialists in Measurement and Evaluation and one Basic Science teacher with a B.Sc.Ed in Integrated Science. Based on the validators' comments and recommendations, 50 items were discarded and the remaining 100 items were retained for trial testing. Following trial testing, item analysis based on difficulty and discrimination indices resulted in the retention of 60 items for the final MCBST. The final 60-item test was organized across six content areas and the intended cognitive levels of Bloom's taxonomy. Evidence of content validity was obtained from curriculum alignment, the table of specifications, and expert judgment. The internal consistency reliability of the 100-item trial version was estimated using the Kuder-Richardson Formula 20 (KR-20), yielding a coefficient of 0.84. The findings provided evidence that the developed MCBST has appropriate content representation and satisfactory internal consistency for assessing Basic Science achievement among JSS I students in Imo State. The study recommended systematic use of test blueprints and expert review in classroom test development.

Keywords: Test development, content validity, Basic Science, multiple-choice test, instrumentation research

INTRODUCTION

Assessment remains a fundamental component of the educational process, providing critical information about students' learning outcomes, instructional effectiveness, and the alignment between curriculum objectives and classroom practices. It involves the systematic collection, analysis, and interpretation of information for educational decision-making. Effective assessment enables teachers to monitor students' progress, identify learning difficulties, evaluate instructional effectiveness, and make informed decisions concerning teaching and learning. As noted by Onunkwo (2002), assessment involves interpreting students' performance in relation to their overall functioning and learning outcomes. Similarly, Kpolovie (2010) emphasized that assessment extends beyond data collection to include meaningful interpretation of educational measurements.

Among the various assessment tools, multiple-choice tests (MCTs) are widely used because of their objectivity, ease of scoring, and capacity to cover a broad content area within a relatively short period. These characteristics make them particularly useful for classroom and large-scale educational assessment. In science education, where conceptual understanding, application of knowledge, problem-solving, and interpretation of scientific information are important, appropriately constructed multiple-choice tests can provide an efficient means of assessing students' achievement. According to Nitko and Brookhart (2014), well-constructed multiple-choice items can measure different levels of cognitive functioning when they are appropriately aligned with instructional objectives.

Teacher-made tests play an important role in classroom assessment. Such tests are designed to measure students' achievement in relation to specific instructional objectives and may serve diagnostic, formative, and summative purposes. However, the quality of teacher-made tests depends largely on teachers' competence in test construction and their adherence to appropriate assessment procedures. Empirical studies have shown that inadequate assessment knowledge may affect the quality of test items developed by teachers. For instance, Magno (2003) reported that teachers' competencies influence the quality of test items they develop. Similarly, Kinyua and Okunya (2014) reported limitations in teachers' skills in constructing valid and reliable classroom tests.

One major challenge associated with teacher-made tests is inadequate adherence to systematic test-development procedures. Some tests are constructed without a table of specifications, resulting in inadequate or disproportionate representation of content areas and cognitive levels. Poorly worded, ambiguous, or irrelevant items may also result in inaccurate measurement of students' achievement. Nworgu (2015) emphasized the importance of systematic planning in test construction, including the alignment of test items with instructional objectives and the use of a table of specifications.

Validity is an important consideration in educational measurement because it concerns the evidence supporting the interpretations and intended uses of test scores. The *Standards for Educational and Psychological Testing* state that validity concerns the degree to which evidence and theory support the interpretations of test scores for their proposed uses (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). In achievement testing, evidence of content validity is particularly important because the items should adequately represent the content domain and instructional objectives the test is intended to assess.

Content validity concerns the extent to which the items of an assessment instrument adequately represent the relevant content domain and intended instructional objectives. Kpolovie (2010) explained that content validity involves systematic examination of test items to determine whether they adequately represent the domain of behaviour or knowledge to be measured. Asuru (2015) similarly described content validity as the systematic examination of test items to determine the extent to which they cover a representative sample of the subject matter and stated instructional objectives. Obagah and Inko-Tariah (2014) noted that when test items constitute a representative sample of the course content, the test has evidence of content validity.

The determination of content validity is commonly supported by expert judgment and a table of specifications. Experts examine test items to determine whether they adequately represent the relevant subject matter, instructional objectives, and cognitive demands. A table of specifications provides a systematic framework for distributing items across content areas and cognitive levels. Thus, the use of a table of specifications helps to reduce the possibility of over-representation or under-representation of particular areas of the curriculum.

In the context of science education, the need for appropriate assessment instruments is particularly important. Basic Science serves as a foundational subject at the Junior Secondary School level in Nigeria. It introduces students to fundamental scientific concepts and provides a basis for further learning in the sciences. Nwanekezi and Arokoyu (2014) noted the importance of Basic Science in developing scientific literacy and critical thinking among learners. Osborne and Dillon (2008) similarly emphasized the broader importance of science education in preparing learners to participate effectively in a scientifically and technologically oriented society.

Despite the importance of Basic Science, concerns remain about the quality of achievement tests used in schools. Some teacher-made tests may not adequately reflect curriculum content or the cognitive objectives of instruction because of weaknesses in test construction and validation. Agu et al. (2013) and Nworgu (2015) observed that weaknesses in test construction can affect the quality of educational assessment and consequently influence instructional decisions.

The development of a good achievement test involves a number of systematic stages. These include defining the purpose of the test, identifying the content domain, preparing a table of specifications, writing test items, subjecting the items to expert review, conducting trial testing, performing item analysis, and selecting appropriate items for the final instrument. The table of specifications serves as a blueprint for ensuring adequate representation of content areas and cognitive objectives. Oladunjoye and Ogunsola (2014) emphasized the usefulness of a test blueprint in aligning test items with instructional objectives.

Item analysis is another important stage in test development. It involves examining individual items using indices such as difficulty and discrimination. The difficulty index provides information about the proportion of students who answer an item correctly, while the discrimination index provides information about the extent to which an item differentiates between relatively higher- and lower-performing students. Udo and Eze (2021) noted that item analysis contributes to the improvement of the quality of test items and the effectiveness of assessment instruments.

Reliability is also an important characteristic of an assessment instrument. It concerns the consistency of scores produced by an instrument under specified conditions. For dichotomously scored achievement-test items, internal consistency may be estimated using the Kuder–Richardson Formula 20 (KR-20). Kaplan and Saccuzzo (2005) emphasized the importance of reliability in educational measurement because consistent measurement is necessary for meaningful interpretation of test scores.

The development of achievement tests may be informed by Classical Test Theory (CTT) and other psychometric approaches. Classical Test Theory views an observed score as comprising a true-score component and an error component. Within this framework, item difficulty, item discrimination, and reliability can be examined to improve the quality of an assessment instrument.

Given the importance of Basic Science and the need for appropriately developed assessment instruments, there is a need for tests that adequately represent the intended curriculum content and cognitive objectives. A carefully developed Multiple-Choice Basic Science Test can provide teachers and researchers with a systematic means of assessing students' achievement in selected Basic Science content areas.

Therefore, this study focused on the development and establishment of evidence of content validity and internal consistency reliability of a Multiple-Choice Basic Science Test (MCBST) for Junior Secondary School One students in Imo State, Nigeria.

Statement of the Problem

Despite the importance of Basic Science in the Nigerian education system, concerns remain about the quality of assessment instruments used to measure students' achievement at the Junior Secondary School level. Some teacher-made multiple-choice tests may be constructed without adequate adherence to systematic test-development procedures, including curriculum analysis, use of a table of specifications, expert review, and empirical item analysis.

When assessment instruments do not adequately represent the content and cognitive objectives of instruction, the resulting scores may not provide appropriate evidence of students' achievement. Inadequate item construction may also result in ambiguous items, inappropriate difficulty levels, poor discrimination, and disproportionate representation of content areas.

The need to address these concerns is particularly relevant in Basic Science, where students are expected not only to recall scientific facts but also to understand concepts, apply knowledge, analyze information, and make appropriate judgments. An assessment instrument that emphasizes only factual recall may therefore provide inadequate representation of students' learning outcomes.

The absence of a systematically developed and content-validated Basic Science achievement test for the target population therefore constitutes an assessment concern. This study was designed to address this concern through the development of a Multiple-Choice Basic Science Test for JSS I students in Imo State.

Aim and Objectives of the Study

The major aim of the study was to develop a Multiple-Choice Basic Science Test (MCBST) and establish evidence of its content validity for Junior Secondary School One (JSS I) students in Imo State, Nigeria.

Specifically, the study sought to:

1. Develop a Multiple-Choice Basic Science Test (MCBST) for JSS I students.

2. Establish evidence of content validity of the developed MCBST using a table of specifications and expert judgment.

Research Questions

The following research questions guided the study:

1. How many items were retained for the final Multiple-Choice Basic Science Test (MCBST)?

2. What evidence supports the content validity of the developed Multiple-Choice Basic Science Test (MCBST)?

METHODOLOGY

Research Design

The study adopted an instrumentation research design. Instrumentation research is concerned with the systematic development and validation of measurement instruments for use in educational and related settings. According to Nworgu (2015), instrumentation research focuses on the development and validation of instruments for measuring specified characteristics. Similarly, Abonyi (2011) described instrumentation research as a research approach through which instruments required for educational practice are developed and, where necessary, validated. The design was considered appropriate because the major purpose of the study was to develop a Multiple-Choice Basic Science Test and obtain evidence regarding its content validity and internal consistency reliability for JSS I students in Imo State.

Area of the Study

The study was conducted in Imo State, Nigeria. Imo State is located in the South-Eastern region of Nigeria and comprises three senatorial zones, namely Owerri, Orlu, and Okigwe. Basic Science is taught as a core subject at the Junior Secondary School level in the state. The study focused on public secondary schools in Imo State.

Population of the Study

The population of the study comprised 1,508 Junior Secondary School One (JSS I) Basic Science students drawn from 291 public secondary schools in Imo State, Nigeria, based on records obtained from the Planning, Research and Statistics Department of the Imo State Universal Basic Education Board (IMSUBEB, 2025). JSS I students were selected because they represent the first year of the junior secondary school level and constituted the target population for the developed Basic Science achievement test.

Sample and Sampling Procedure

A sample of 500 JSS I students was used for the trial testing of the MCBST. The students were selected from public secondary schools in the study area using the sampling procedure adopted by the researcher. The trial-testing sample was used to obtain empirical information concerning the functioning of the 100-item trial version of the MCBST. The responses obtained from the students were subsequently subjected to item analysis and reliability analysis.

Three validators participated in the preliminary face and content validation of the instrument. The validators comprised Prof. I. M. Opara and Dr. A. Amadioha, both specialists in Measurement and Evaluation in the

Department of Educational Psychology, Guidance and Counselling, University of Port Harcourt, and one practicing Basic Science teacher with a B.Sc.Ed in Integrated Science.

Instrument for Data Collection

The instrument used for the study was the Multiple-Choice Basic Science Test (MCBST) developed by the researcher. The test was designed to measure cognitive achievement in selected Basic Science content areas among JSS I students. The initial version of the MCBST consisted of 150 multiple-choice items, with four response options labelled A–D. Each item had one correct answer and three distractors. The development of the instrument followed a systematic sequence involving curriculum analysis, preparation of a table of specifications, item writing, expert review, item revision, trial testing, item analysis, and final item selection.

Development of the Test Blueprint

A table of specifications was prepared to guide the distribution of items across the six selected content areas and the specified cognitive levels of the final 60-item MCBST. The six content areas were Sanitation (Family Health), Nutrition (Family Health), Drug and Substance Abuse I (Family Health), Drug and Substance Abuse II (Family Health), Environmental Pollution, and Living and Non-Living Things. Each content area contributed 10 items. The cognitive distribution comprised 6 Knowledge items (10%), 6 Comprehension items (10%), 18 Application items (30%), 12 Analysis items (20%), 12 Synthesis items (20%), and 6 Evaluation items (10%).

Content Area	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total
Sanitation (Family Health)	1	1	3	2	2	1	10
Nutrition (Family Health)	1	1	3	2	2	1	10
Drug and Substance Abuse I (Family Health)	1	1	3	2	2	1	10
Drug and Substance Abuse II (Family Health)	1	1	3	2	2	1	10
Environmental Pollution	1	1	3	2	2	1	10
Living and Non-Living Things	1	1	3	2	2	1	10
Total	6	6	18	12	12	6	60
Percentage	10%	10%	30%	20%	20%	10%	100%

Table 1 showed that each content area contributed 10 items to the final instrument and that all six specified cognitive levels were represented. The distribution was designed to provide equal content-area representation while giving greater weight to Application, Analysis, and Synthesis.

Item Writing

Guided by the content domain and test blueprint, an initial pool of 150 multiple-choice items was constructed from the selected Basic Science content areas. Each item consisted of a stem and four response options labelled A–D, with one correct answer and three distractors. During item writing, attention was given to clarity of language, relevance to the Basic Science curriculum, suitability for JSS I students, alignment with the intended cognitive objectives, and appropriateness of the distractors.

Preliminary Face and Content Validation

The initial 150-item MCBST was subjected to preliminary face and content validation by three validators. The validators comprised Prof. I. M. Opara and Dr. A. Amadioha, both specialists in Measurement and Evaluation, and one Basic Science teacher with a B.Sc.Ed in Integrated Science.

The validators examined the items with respect to their relevance to the Basic Science curriculum, clarity of expression, suitability for JSS I students, alignment with instructional objectives, adequacy of content coverage, and appropriateness of the response options. Their comments, corrections, and recommendations were used to revise the items. Items considered appropriate were retained, items requiring clarification were modified, and items considered unsuitable or insufficiently aligned with the intended content and objectives were discarded. Following the expert review, 50 items were eliminated from the initial 150-item pool, resulting in a 100-item trial version.

The test-development process was therefore: 150 initial items → 100 items after expert review → 60 items after trial testing and item analysis.

Trial Testing

The 100-item trial version of the MCBST was administered to 500 JSS I students in public secondary schools in Imo State. The test was administered under similar examination conditions using the same instructions and response format. Responses were scored dichotomously, with one mark awarded for a correct response and zero for an incorrect response. The responses were subsequently used for item analysis and estimation of internal consistency reliability.

Item Analysis

The responses obtained from the 500 students were subjected to item analysis. The analysis focused on the difficulty index and discrimination index of the items. The difficulty index was used to determine the proportion of students who answered each item correctly, while the discrimination index was used to determine the extent to which each item differentiated between relatively higher- and lower-performing students. Items that met the psychometric criteria used in the study were retained, while items that did not meet the criteria were eliminated. At the end of the item-analysis process, 60 items were retained to constitute the final version of the MCBST. The numerical cut-off values used for item retention are not reported here because they were not part of the reported study record.

Evidence of Content Validity

Evidence of content validity was obtained from three complementary sources: alignment of the test content with the selected Basic Science curriculum, the table of specifications showing the intended representation of content areas and cognitive levels, and expert judgment. The three validators examined the items for relevance, clarity, suitability for JSS I students, alignment with instructional objectives, adequacy of content coverage, and appropriateness of response options. No numerical content-validity index is reported because no such index was computed in the study.

Reliability of the Instrument

The internal consistency reliability of the MCBST was estimated using the Kuder–Richardson Formula 20 (KR-20) because the items were scored dichotomously. The KR-20 analysis was based on responses obtained from the 500 students who participated in the trial testing and yielded a coefficient of 0.84. This coefficient indicates satisfactory internal consistency within the trial-testing sample. The KR-20 coefficient provides evidence of internal consistency reliability; it is not, by itself, evidence of content validity.

Method of Data Analysis

Data obtained during the development and validation of the MCBST were analyzed using expert judgment, item analysis, and reliability analysis. Expert judgment was used during the preliminary face and content validation to examine relevance, clarity, appropriateness, and curriculum alignment. Trial-test responses were analyzed using difficulty and discrimination indices, and items that did not meet the psychometric criteria used in the study were eliminated. The KR-20 coefficient was used to estimate internal consistency reliability.

RESULTS

Research Question One

How many items were retained for the final Multiple-Choice Basic Science Test (MCBST)?

Research Question One was answered by examining the stages of development and refinement of the MCBST. The initial item pool comprised 150 multiple-choice items. Following preliminary face and content validation by three validators, 50 items were eliminated and 100 items were retained for trial testing. The 100-item trial version was administered to 500 JSS I students. The responses were subjected to item analysis involving difficulty and discrimination indices, after which 60 items were retained for the final MCBST.

Stage of Development	Number of Items
Initial item pool	150
Items retained after expert review	100
Items retained after trial testing and item analysis	60
Final MCBST	**60**

The result shows that the final MCBST comprised 60 items.

FINAL VERSION OF THE MULTIPLE-CHOICE BASIC SCIENCE TEST (MCBST)

Class: JSS I Sex: Male / Female Time Allowed: 1 hour 30 minutes Number of Items: 60

Instruction: Answer all questions. For each question, choose the option that best answers the question from the alternatives lettered A–D.

SECTION A: SANITATION (FAMILY HEALTH)

1. Sanitation refers mainly to the maintenance of

A. cleanliness and proper waste disposal

B. expensive household equipment

C. large quantities of food

D. recreational facilities

2. Why is handwashing with soap and clean water important after using the toilet?

A. It makes the hands smell pleasant.

B. It removes harmful microorganisms that may cause disease.

C. It increases the amount of water in the body.

D. It makes the toilet easier to use.

3. After using the toilet at school, Chika wants to prevent the spread of germs. What should she do first?

A. Wipe her hands on her clothes.

B. Wash her hands thoroughly with soap and clean water.

C. Shake hands with her classmates.

D. Touch the classroom door immediately.

4. A family notices that flies frequently gather around an uncovered refuse container near their house. Which action would best solve the problem?

A. Add more refuse to the container.

B. Keep the refuse container covered and dispose of waste properly.

C. Pour water over the refuse every morning.

D. Move the refuse closer to the kitchen.

5. A community has no proper toilet facilities, and many residents defecate in nearby bushes. Which measure would most directly improve sanitation in the community?

A. Construct and use hygienic toilet facilities.

B. Increase the number of roadside shops.

C. Encourage people to walk farther from their homes.

D. Burn all household furniture.

6. Two households live in the same community. Household A covers its food, disposes of waste properly, and washes hands regularly. Household B leaves food uncovered and dumps waste near its house. Which conclusion is most reasonable?

A. Household B is more likely to experience sanitation-related diseases.

B. Household A is more likely to experience sanitation-related diseases.

C. Both households have exactly the same sanitation practices.

D. Sanitation practices have no relationship with disease prevention.

7. A school records frequent cases of diarrhoea among pupils. Investigation shows that pupils rarely wash their hands after using the toilet and that the school water source is contaminated. Which factors are most directly related to the problem?

A. Poor hand hygiene and unsafe water

B. Excessive exercise and inadequate sleep

C. Lack of sports facilities and classroom space

D. Poor handwriting and overcrowded textbooks

8. A school wants to reduce the spread of sanitation-related diseases. Which plan would be most effective?

A. Provide toilets, safe water, handwashing facilities, and proper waste disposal.

B. Provide only more classrooms.

C. Ask pupils to avoid drinking water during school hours.

D. Increase the number of school uniforms.

9. Which combination of activities would best promote good sanitation in a community?

A. Open defecation, littering, and burning plastics indoors

B. Proper waste disposal, use of hygienic toilets, and regular handwashing

C. Dumping refuse in streams, uncovered food, and stagnant water

D. Leaving sewage untreated and washing hands without soap

10. A community leader proposes dumping household waste into a nearby stream because "the flowing water will carry it away." How should this proposal be evaluated?

A. It is appropriate because flowing water removes all harmful substances.

B. It is appropriate because streams cannot be polluted by household waste.

C. It is inappropriate because the waste can pollute the water and spread disease.

D. It is inappropriate only when the stream contains fish.

SECTION B: NUTRITION (FAMILY HEALTH)

11. Which class of food is mainly responsible for growth and repair of body tissues?

A. Proteins

B. Vitamins

C. Carbohydrates

D. Water

12. A balanced diet is best described as a diet that

A. contains only foods that provide energy

B. contains the required nutrients in appropriate proportions

C. contains mainly fruits and vegetables

D. contains foods that are eaten only once a day

13. A pupil has rice and wants to make the meal more balanced. Which combination would provide additional protein and protective nutrients?

A. Rice, fish and vegetables

B. Rice, sugar and sweets

C. Rice, soft drink and biscuits

D. Rice, salt and cooking oil

14. A child frequently eats foods rich in carbohydrates but rarely eats protein-rich foods. Which food should the child's parents add to the diet?

A. Fish

B. Sugar

C. Cassava starch

D. Soft drink

15. A student becomes weak after spending several hours playing in the sun without drinking water. Which action would best help the student?

A. Give the student clean drinking water.

B. Give the student more salt without water.

C. Prevent the student from eating for the rest of the day.

D. Give the student only cooking oil.

16. Three pupils eat different lunches. Pupil A eats rice and soft drink. Pupil B eats rice, beans, vegetables and fruit. Pupil C eats sweets and a soft drink. Which pupil has the most balanced meal?

A. Pupil A

B. Pupil B

C. Pupil C

D. All three pupils

17. A child has swollen feet, a swollen abdomen, poor growth and weakness. The child eats very little protein. Which nutritional condition is most consistent with these observations?

A. Kwashiorkor

B. Scurvy

C. Night blindness

D. Rickets

18. A school wants to design a healthy lunch programme for pupils. Which menu would provide the best combination of nutrients?

A. Rice, beans, vegetables and fruit

B. Sweets, biscuits and soft drink

C. Fried snacks and carbonated drink

D. Sugar, salt and water

19. Which plan would best help a family maintain good nutrition throughout the week?

A. Eat only one type of food every day.

B. Combine different food groups and include fruits and vegetables regularly.

C. Avoid all protein-containing foods.

D. Replace meals with sugary drinks.

20. A pupil says, "Eating a large quantity of one type of food every day is better than eating different foods." Which response is most appropriate?

A. The statement is correct because quantity is all that matters.

B. The statement is correct because all foods contain the same nutrients.

C. The statement is incorrect because the body requires different nutrients from different food groups.

D. The statement is incorrect only for adults.

SECTION C: DRUG AND SUBSTANCE ABUSE I (FAMILY HEALTH)

21. Drug abuse is best described as

A. the inappropriate or harmful use of drugs or substances

B. the use of all medicines prescribed by doctors

C. the preparation of food with herbs

D. the drinking of clean water

22. Why can taking a medicine without proper medical advice be dangerous?

A. The medicine may be inappropriate or taken in an unsafe dose.

B. Medicines always improve health when taken without advice.

C. Medicines cannot affect the body negatively.

D. All medicines have exactly the same effects.

23. A friend offers Musa a tablet and tells him that it will help him stay awake during examination preparation. Musa does not know what the tablet contains. What should he do?

A. Take two tablets to obtain a stronger effect.

B. Refuse the tablet and seek advice from a responsible adult or health professional.

C. Give the tablet to another pupil.

D. Mix the tablet with an energy drink.

24. A student is pressured by friends to smoke a substance during a party. Which response demonstrates good decision-making?

A. Accept it to avoid embarrassment.

B. Refuse the substance and leave the situation if necessary.

C. Try it once because it cannot cause harm.

D. Take it secretly after the party.

25. A person has been prescribed medicine by a doctor. The person decides to double the prescribed dose because the illness has not improved quickly. What should the person do instead?

A. Double the dose again.

B. Stop all treatment permanently.

C. Consult the doctor or another qualified health professional.

D. Combine the medicine with another unknown drug.

26. A teenager has recently started associating with a group that encourages drug use. He begins missing classes, losing interest in normal activities, and frequently asking for money. Which combination of observations may indicate a developing substance-abuse problem?

A. Regular attendance, good concentration and healthy eating

B. Withdrawal from normal activities, truancy and unusual demands for money

C. Participation in sports, good grades and regular sleep

D. Reading books, attending classes and helping at home

27. A student claims that "peer pressure is harmless because everyone makes their own decisions." Which evidence best challenges the claim?

A. Friends can influence a person's choices, especially when the person wants to fit into a group.

B. Friends can never influence behaviour.

C. Peer pressure affects only teachers.

D. People are never influenced by their social environment.

28. A school wants to prevent drug abuse among students. Which programme would provide the most comprehensive approach?

A. Punish students without providing information or support.

B. Provide drug education, counselling, peer-support activities and responsible-adult guidance.

C. Ignore drug-related behaviour unless students are arrested.

D. Discuss drug abuse only with students who have already been affected.

29. Which combination of personal strategies would best help a teenager resist pressure to abuse drugs?

A. Assertiveness, good decision-making and choosing positive friends

B. Isolation, secrecy and accepting every suggestion from friends

C. Experimentation, secrecy and avoiding responsible adults

D. Taking unknown medicines and copying peers

30. A student argues that taking an unprescribed drug is safe because "a friend took it and nothing happened." How should this argument be evaluated?

A. It is valid because everyone's body responds to drugs in exactly the same way.

B. It is valid because one person's experience proves that a drug is safe.

C. It is weak because a drug can affect different people differently and may be harmful when misused.

D. It is valid if the drug is inexpensive.

SECTION D: DRUG AND SUBSTANCE ABUSE II (FAMILY HEALTH)

31. Which of the following is an example of a substance that can be abused?

A. Heroin

B. Clean water

C. Fresh fruit

D. Oxygen in normal air

32. Drug dependence means that a person

A. may develop a strong physical or psychological need for a drug

B. never experiences any effect from a drug

C. uses food as medicine

D. takes every medicine correctly

33. A student notices that an older sibling has developed a strong craving for a drug and becomes uncomfortable when unable to obtain it. What should the student do?

A. Encourage the sibling to take a larger dose.

B. Keep the problem secret from everyone.

C. Encourage the sibling to seek help from a qualified health professional and a trusted adult.

D. Obtain the drug for the sibling.

34. A teenager frequently uses alcohol and then operates a motorcycle. Which immediate action would reduce the risk of injury?

A. Encourage the teenager to ride faster.

B. Prevent the teenager from riding and arrange safe transportation.

C. Give the teenager another alcoholic drink.

D. Allow the teenager to ride if the road is empty.

35. A person has been using a substance heavily and wants to stop. Which action is most appropriate?

A. Stop seeking help and manage the problem secretly.

B. Seek professional assistance and follow an appropriate treatment plan.

C. Replace the substance with another unknown drug.

D. Increase the amount used before stopping.

36. Four students describe their behaviours: I. One student takes medicine according to a doctor's prescription. II. One student takes a friend's prescription medicine. III. One student deliberately takes more medicine than prescribed. IV. One student uses an illegal drug for pleasure. Which group represents inappropriate drug use?

A. I only

B. II, III and IV

C. I and II only

D. I, II, III and IV

37. A student who frequently abuses drugs begins to show poor concentration, declining academic performance, changes in behaviour and loss of interest in normal activities. What does this pattern suggest?

A. The student is experiencing only improved academic motivation.

B. The drug use may be negatively affecting the student's physical, mental and social functioning.

C. The student has developed better study habits.

D. The changes are evidence that drug abuse improves school performance.

38. A community wants to reduce substance abuse among young people. Which combined strategy is most likely to be effective?

A. Drug education, counselling, family support, recreational opportunities and access to appropriate health services

B. Keeping young people uninformed about drugs

C. Punishment alone without education or rehabilitation

D. Encouraging young people to experiment with substances under supervision

39. Which plan would best help a student who is trying to avoid drug abuse?

A. Develop positive friendships, participate in healthy activities, seek guidance when necessary, and avoid risky situations.

B. Spend more time with people who encourage drug use.

C. Keep drug-related problems secret from responsible adults.

D. Experiment with substances to learn about their effects.

40. A person claims that drug abuse is a personal choice and therefore cannot harm other people. Which evaluation is most accurate?

A. The claim is completely correct because drug abuse affects only the user.

B. The claim is incorrect because drug abuse can affect families, relationships, school performance, safety and the wider community.

C. The claim is correct whenever the drug is inexpensive.

D. The claim is correct when the user is an adult.

SECTION E: ENVIRONMENTAL POLLUTION

41. Which type of pollution occurs when harmful substances contaminate the atmosphere?

A. Air pollution

B. Soil pollution

C. Water pollution

D. Noise pollution

42. Why can untreated sewage discharged into a river be harmful?

A. It can introduce disease-causing organisms and pollutants into the water.

B. It always makes the water safer for drinking.

C. It increases the amount of clean drinking water.

D. It prevents all aquatic organisms from growing.

43. A factory releases thick smoke into the atmosphere every day. Which form of pollution is most directly involved?

A. Air pollution

B. Soil pollution

C. Water pollution

D. Noise pollution

44. Farmers in an area use excessive quantities of chemical fertilizers and pesticides. Which practice would help reduce the resulting environmental problems?

A. Apply chemicals in excessive quantities.

B. Follow recommended application rates and use environmentally safer practices where appropriate.

C. Pour unused chemicals into nearby streams.

D. Store chemicals in open containers near water sources.

45. A community has a large amount of plastic waste that is regularly dumped into a drainage channel. During rainfall, the drains become blocked and flooding occurs. What should the community do?

A. Dump more plastic into the drains.

B. Reduce plastic waste, dispose of it properly, and keep drainage channels clear.

C. Cover the drains with more plastic.

D. Wait for the rain to carry the waste away.

46. A river downstream from a settlement contains dead fish, foul-smelling water and large quantities of untreated sewage. Which explanation best accounts for these observations?

A. The sewage has introduced pollutants that have degraded water quality.

B. The river has become cleaner because of the sewage.

C. The dead fish prove that the water contains more oxygen.

D. The smell indicates that the river is free from microorganisms.

47. An industrial area records high levels of smoke and carbon monoxide from vehicle and factory emissions. Residents report breathing difficulties. Which relationship best explains the situation?

A. Increased air pollutants can reduce air quality and affect respiratory health.

B. Increased air pollutants always improve respiratory health.

C. Carbon monoxide increases the oxygen-carrying capacity of blood.

D. Smoke from industries has no effect on human health.

48. A local government wants to reduce environmental pollution. Which combination of measures would be most effective?

A. Proper waste management, treatment of sewage, emission control and environmental education

B. Increased dumping of waste in rivers

C. Open burning of all household waste

D. Allowing factories to release untreated waste into the environment

49. Which plan would best reduce soil, water and air pollution in a farming community?

A. Use chemicals responsibly, prevent waste from entering water bodies, and avoid unnecessary burning.

B. Increase indiscriminate use of pesticides and burn plastics in open spaces.

C. Dump farm chemicals into streams after use.

D. Burn vegetation and dispose of waste in nearby rivers.

50. A factory owner argues that releasing untreated industrial waste into a river is acceptable because "the river is large enough to dilute it." Which evaluation is most appropriate?

A. The argument is acceptable because dilution removes every pollutant.

B. The argument is unacceptable because pollutants can remain harmful and damage aquatic life and human health.

C. The argument is acceptable whenever the river is deep.

D. The argument is acceptable if the waste is released at night.

SECTION F: LIVING AND NON-LIVING THINGS

51. Which of the following is a characteristic of living things?

A. Reproduction

B. Rusting

C. Melting

D. Breaking

52. Why is a tree classified as a living thing?

A. It has a definite shape and colour.

B. It performs life processes such as growth, respiration and reproduction.

C. It remains in one place throughout its life.

D. It is made of hard material.

53. A potted plant placed in a dark room for several weeks becomes weak and pale. Which action would best help the plant?

A. Move it to a suitable place where it can receive adequate light.

B. Keep it in complete darkness.

C. Remove all its leaves.

D. Stop watering it permanently.

54. A student observes that a seedling grows taller and develops new leaves after being planted in suitable soil and given water and light. Which characteristic of living things is demonstrated?

A. Growth

B. Melting

C. Rusting

D. Freezing

55. An iron nail is left in a damp environment and develops a reddish-brown coating. Which statement best explains the observation?

A. The nail has become a living organism.

B. The iron has undergone a chemical change known as rusting.

C. The nail has started reproducing.

D. The nail has begun to respire.

56. A student compares a cat, a tree, a stone and a plastic bottle. The cat and tree grow and reproduce, while the stone and plastic bottle do not. What conclusion can be drawn?

A. The cat and tree exhibit characteristics of living things.

B. The stone and plastic bottle are living because they occupy space.

C. All four objects are living because they have mass.

D. The plastic bottle is living because it changes shape when heated.

57. A learner says that a moving car is a living thing because it moves from one place to another. Which evidence best disproves the learner's claim?

A. A car moves but does not grow, reproduce or carry out biological life processes.

B. A car has wheels.

C. A car can move faster than many animals.

D. A car occupies space.

58. A teacher asks pupils to design an investigation that demonstrates that plants are living things. Which plan would provide the best evidence?

A. Observe a plant over time and record changes such as growth and development of new leaves.

B. Place a stone beside the plant and compare their colours.

C. Measure only the weight of the flowerpot.

D. Observe the plant for one minute and record whether it changes position.

59. A science club wants to demonstrate differences between living and non-living things. Which set of observations would provide the strongest evidence?

A. Growth, respiration, reproduction and response to stimuli in living organisms compared with their absence in non-living objects.

B. Colour, shape and size alone.

C. Weight and ability to occupy space only.

D. Ability to be touched and seen only.

60. A pupil concludes that a candle flame is a living thing because it grows, moves and appears to reproduce when one flame lights another. How should this conclusion be evaluated?

A. It is correct because movement and apparent growth are sufficient to prove life.

B. It is correct because anything that reproduces is living.

C. It is incorrect because a flame does not carry out biological life processes such as cellular reproduction and respiration.

D. It is incorrect only because a flame cannot move from place to place.

Research Question Two

What evidence supports the content validity of the developed Multiple-Choice Basic Science Test (MCBST)?

Research Question Two was answered using the table of specifications and expert judgment. Evidence of content validity was established through curriculum-based test construction, the blueprint, and expert review. Three validators participated in the review: Prof. I. M. Opara and Dr. A. Amadioha, both specialists in Measurement and Evaluation, and one Basic Science teacher with a B.Sc.Ed. in Integrated Science. The validators examined the items for relevance to the Basic Science curriculum, clarity of expression, suitability for JSS I students, alignment with instructional objectives, adequacy of content coverage, and appropriateness of the response options. Following the validators' recommendations, 50 items were eliminated from the initial 150-item pool and 100 items were retained for trial testing. The 100-item trial version was subsequently subjected to empirical item analysis, resulting in the retention of 60 items.

The final 60-item MCBST was distributed equally across the six selected content areas and across the specified cognitive levels.

Content Area	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total
Sanitation (Family Health)	1	1	3	2	2	1	10
Nutrition (Family Health)	1	1	3	2	2	1	10
Drug and Substance Abuse I (Family Health)	1	1	3	2	2	1	10
Drug and Substance Abuse II (Family Health)	1	1	3	2	2	1	10
Environmental Pollution	1	1	3	2	2	1	10
Living and Non-	1	1	3	2	2	1	10

Living Things							
Total	6	6	18	12	12	6	60

Table 3 shows that each content area contributed 10 items, representing 16.67% of the final instrument. At the cognitive level, Knowledge and Comprehension each accounted for 6 items (10%), Application accounted for 18 items (30%), Analysis accounted for 12 items (20%), Synthesis accounted for 12 items (20%), and Evaluation accounted for 6 items (10%). The distribution demonstrates representation of the selected content areas and different levels of cognitive demand.

Reliability Result

The internal consistency reliability of the MCBST was estimated using the Kuder–Richardson Formula 20. The analysis of the responses obtained from the 500 students who participated in the trial testing yielded a KR-20 coefficient of 0.84. The result indicates satisfactory internal consistency among the dichotomously scored items within the trial-testing sample.

DISCUSSION OF RESULTS

Development and Selection of the MCBST

The finding for Research Question One showed that the MCBST was developed through a progressive process involving an initial pool of 150 items, expert review, trial testing, item analysis, and final item selection. The initial item pool was larger than the required final test length, thereby providing an opportunity for items to be reviewed and refined before final selection.

The first reduction occurred during expert review. The three validators examined the items for curriculum relevance, clarity, suitability for the target population, alignment with instructional objectives, and appropriateness of the response options. Following the expert review, 50 items were eliminated and 100 items were retained for empirical trial testing.

The second reduction occurred after trial testing. The 100-item version was administered to 500 JSS I students, and the responses were analyzed using item difficulty and discrimination indices. Sixty items were subsequently retained for the final MCBST.

The progression from 150 to 100 and finally to 60 items illustrates the usefulness of combining expert review and empirical item analysis in the development of an achievement test. Expert review provides qualitative information concerning content relevance and item appropriateness, whereas empirical item analysis provides information concerning the functioning of individual items when administered to students.

This finding is consistent with Haladyna (2004), who emphasized the importance of generating an adequate pool of items and refining the pool through systematic review and empirical analysis. Crocker and Algina (2008) and Nitko and Brookhart (2014) similarly emphasized the importance of item refinement and analysis in the development of educational tests.

The final 60-item MCBST was equally distributed across the six selected content areas. Each content area contributed 10 items. The instrument also included six cognitive levels, with greater representation given to Application, Analysis, and Synthesis. The inclusion of these levels means that the final test was not restricted to factual recall; it included items requiring students to use knowledge in situations, examine relationships among information, formulate solutions or plans, and make judgments based on given information.

Evidence of Content Validity

The finding for Research Question Two showed that the MCBST had satisfactory evidence of content validity based on curriculum alignment, the table of specifications, and expert judgment. The table of specifications provided a systematic framework for representing the selected content areas and cognitive levels. Each of the six content areas contributed 10 items, while the cognitive distribution comprised Knowledge (10%), Comprehension (10%), Application (30%), Analysis (20%), Synthesis (20%), and Evaluation (10%).

The participation of three validators provided both measurement and subject-matter perspectives. Two validators were specialists in Measurement and Evaluation, while the third was a practicing Basic Science teacher with a B.Sc.Ed in Integrated Science. Their review focused on content relevance, clarity, suitability for JSS I students, alignment with instructional objectives, content coverage, and response-option appropriateness.

The elimination of 50 items following expert review indicates that the validation process contributed to the refinement of the initial item pool. Items considered inappropriate or insufficiently aligned with the intended content and objectives were removed, while other items were revised based on the validators' recommendations.

The finding is consistent with the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014), which emphasizes the importance of evidence supporting interpretations and intended uses of test scores. In achievement testing, alignment between test content and the intended content domain is an important source of validity evidence.

The finding is also consistent with Nworgu (2015), who emphasized systematic planning and the use of a table of specifications in educational test construction. Asuru (2015) similarly emphasized the importance of examining test items to determine whether they adequately represent the subject matter and instructional objectives.

Thus, the present study provides evidence that the MCBST was developed through a structured content-representation process involving curriculum analysis, blueprinting, expert judgment, item revision, trial testing, and empirical item analysis.

Internal Consistency Reliability of the MCBST

The trial-testing responses produced a KR-20 coefficient of 0.84, indicating satisfactory internal consistency among the dichotomously scored items within the trial-testing sample. This reliability evidence concerns the consistency of scores produced by the items under the conditions of the trial testing. It should be distinguished from content-validity evidence: the KR-20 coefficient does not establish content validity. In the present study, evidence of content validity was obtained from curriculum alignment, the table of specifications, and expert judgment.

CONCLUSION

Based on the findings of the study, it was concluded that a 60-item Multiple-Choice Basic Science Test (MCBST) was developed for JSS I students in Imo State through a systematic process of item construction, expert review, trial testing, item analysis, and reliability estimation.

The development process began with an initial pool of 150 items. Following expert review by three validators comprising two specialists in Measurement and Evaluation and one Basic Science teacher with a B.Sc.Ed in Integrated Science, 100 items were retained for trial testing. The 100-item trial version was administered to 500 JSS I students, and the responses were subjected to item analysis. Based on the item-analysis results, 60 items were retained for the final instrument.

The final MCBST provided equal representation of the six selected content areas: Sanitation, Nutrition, Drug and Substance Abuse I, Drug and Substance Abuse II, Environmental Pollution, and Living and Non-Living Things. Each content area contributed 10 items. The cognitive distribution comprised Knowledge (10%), Comprehension (10%), Application (30%), Analysis (20%), Synthesis (20%), and Evaluation (10%).

The study provided evidence of content validity through curriculum alignment, the use of a table of specifications, and expert judgment. In addition, the trial-testing responses yielded a KR-20 coefficient of 0.84, indicating satisfactory internal consistency. It is therefore concluded that the developed MCBST provided satisfactory evidence of content validity and internal consistency reliability within the context for which it was developed. This conclusion should not be interpreted as evidence that all other forms of validity or full standardization have been established.

RECOMMENDATIONS

1. Basic Science teachers may use the developed MCBST as an assessment instrument for measuring students' achievement in the selected Basic Science content areas covered by the test.

2. Teachers should be encouraged to use a table of specifications when constructing classroom-based achievement tests to ensure adequate representation of content areas and cognitive objectives.

3. Schools and educational authorities should provide opportunities for teachers to develop their competencies in test construction, item writing, validation, and item analysis.

4. Further studies should be conducted to obtain additional validity evidence for the MCBST, including evidence based on relationships with other variables and the internal structure of the instrument.

5. Further psychometric studies may apply more advanced approaches, including Item Response Theory, to the MCBST to provide additional information about item functioning and test characteristics.

REFERENCES

1. Abonyi, O. S. (2011). Instrument in behavioural research: A practical approach. Fulladu Publishing Company.
2. Agu, N. N., Onyekuba, C., & Anyichie, C. A. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill inventory. Educational Research and Reviews, 8(8), 431– 439.
3. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.
4. Asuru, V. A., & Owhor, G. (2022). Psychometric analysis of basic science objective tests in junior secondary schools. Nigerian Journal of Educational Research and Evaluation, 21(2), 44–59.
5. Asuru, V. A. (2015). Measurement and evaluation in education and psychology. Pearl Publishers International.
6. Crocker, L., & Algina, J. (2008). Introduction to classical and modern test theory. Cengage Learning.
7. Elliott, S. N., Kratochwill, T. R., Cook, J. L., & Travers, J. F. (2000). Educational psychology: Effective teaching, effective learning. McGraw-Hill.
8. Haladyna, T. M. (2004). Developing and validating multiple-choice test items (3rd ed.). Lawrence Erlbaum Associates.
9. Kaplan, R. M., & Saccuzzo, D. P. (2005). Psychological testing: Principles, applications, and issues (4th ed.). Thomson Wadsworth.
10. Kinyua, M. M., & Okunya, L. O. (2014). Assessment of teachers' competencies in constructing classroom tests in secondary schools in Kenya. International Journal of Education and Research, 2(6), 493–502.
11. Kpolovie, P. J. (2010). Advanced research methods. Springfield Publishers.
12. Magno, C. (2003). Teachers' competencies in the development of test construction skills. Philippine Journal of Counseling Psychology, 5(1), 47–61.
13. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry. American Psychologist, 50(9), 741–749.
14. Nitko, A. J., & Brookhart, S. M. (2014). Educational assessment of students (7th ed.). Pearson.
15. Nwanekezi, A. U., & Arokoyu, A. A. (2014). Teaching integrated science constructively. Soteria Publishing House.

16. Nworgu, B. G. (2015). Educational research: Basic issues and methodology. University Trust Publishers.
17. Nworgu, B. G., & Ofojebe, W. N. (2017). Validation of multiple-choice test items in basic science: A content analysis approach. Journal of Science Education and Research, 5(3), 38–49.
18. Obagah, M. O. N., & Inko-Tariah, D. C. (2014). Educational measurement and evaluation. Rodi Printing and Publishing Company.
19. Oladunjoye, O. O., & Ogunsola, A. (2014). Development and validation of an achievement test in chemistry for senior secondary schools in Nigeria. African Educational Research, 2(3), 123–130.
20. Olatoye, R. A., & Aderogba, A. M. (2020). Assessment of higher-order thinking in basic science:A study of teacher-constructed tests. International Journal of Educational Assessment, 15(1), 22–34.
21. Onunkwo, G. I. N. (2002). Fundamentals of educational measurement and evaluation. Cape Publishers International.
22. Osborne, J. F., & Dillon, J. (2008). Science education in Europe: Critical reflections.
23. Popham, W. J. (2013). Classroom assessment: What teachers need to know (6th ed.). Pearson.
24. Reckase, M. D. (2009). Multidimensional item response theory. Springer.
25. Sidhu, K. S. (2008). New approaches to measurement and evaluation. Sterling Publishers.
26. Thorndike, R. M., & Thorndike-Christ, T. S. (2010). Measurement and evaluation in psychology and education (8th ed.). Pearson.
27. Udo, I. N., & Eze, U. N. (2021). An empirical analysis of multiple-choice science items using point-biserial correlation. African Journal of Educational Research and Development, 14(1), 59–72.