

**PSYCHOMETRIC PROPERTIES OF SOCIAL STUDIES COGNITIVE
ACHIEVEMENT TEST FOR JUNIOR SECONDARY SCHOOLS IN EKITI
STATE.**

DARAMOLA, C. OLADAYO (PhD)

Department of Social Science Education, College of Education
Bamidele Olumilua University of Education, Science and Technology, Ikere Ekiti
daramola.clement@bouesti.edu.ng dayoasoju@gmail.com

Abstract

This study was designed to develop and validate Social Studies Cognitive Achievement Test for junior secondary three (JS3) students in Ekiti State. This study was guided by five research questions and used the instrumentation research design. The population for the study was made up of 99 840 JS3 Social Studies students in public secondary schools in Ekiti State. The sample comprised of 650 JS3 students. The instrument developed and validated was Social Studies Cognitive Achievement Test for JS3. The reliability coefficient of the test was 0.76 obtained using Kuder Richardson Formula 20 (KR-20). One hundred items were developed and validated, and after validation and trial testing it was reduced to seventy (70) items which were analyzed to determine the item difficulty, item discrimination and effectiveness of distracters. It was found that 58 items had appropriate item difficulty index, 70 items had positive discrimination index, and 70 items distracted effectively. Mean and standard deviation were used to analyse research question five, and it was found that the items of the developed Social Studies Achievement Test did not deviate from the core curriculum; the developed Social Studies Achievement Test for junior secondary schools has high psychometric properties in terms of difficulty and discrimination indices, the instrument as high reliability coefficient. Based on these result, it was recommended, among others, that the developed Social Studies Achievement Test should be used by Social Studies teachers in Ekiti State and beyond; the items of the Social Studies Achievement Test developed should serve as a template for the development of other achievement tests in Social Studies, and indeed other subjects.

Keywords: development, validation, reliability, social studies, test, achievement.

INTRODUCTION

In Nigeria today, the essence of education cannot be over emphasised, because education helps in contributing greatly to the nation's overall wellbeing and development. On this note, different subjects have been outlined in the curriculum of the educational system, which are to be studied in the various levels of education in Nigeria, so as to attain the required goals of education, as a life wire for effective development of the country, Nigeria. At the secondary level of education in Nigeria, Social Studies is studied as one of the main areas of development. Social Studies is best suited to respond to 'what makes us humans', and this subject covers many disciplines that deal with different areas of life. Social Studies in education helps students become informed citizens. Knowledge of the socio-political and economic conditions of their country and of the world, and knowledge of the rights and obligations of individuals as individuals in society help students become responsible citizens. Through Social Studies, they gain knowledge of human behaviour and cultural values that separate individuals from each other. They create higher levels of people, be it social, economic or lifestyle differences Social Studies aim at creating educated individuals who can be responsible citizens of their nation. Teaching children Social Studies increases the possibility of their becoming more aware, more responsible as human beings. The good performance of the students in the society depends on how well this subject is being taught in schools and how well the students have achieved the objectives of Social Studies. Again to ascertain if the students have achieved the desired goals of learning Social Studies, there is the need to determine the extent to which learning objectives are achieved through tests and other educational evaluation methods.

Tests assist the teacher to know the extent of the learners' mastery of the objectives in the subject or programme. Asuru (2015) defined test as a set of tasks, questions, situations, intended to elicit particular types of behaviour; it is an evaluation instrument used to measure skills, knowledge, intelligence, ability, aptitude, attitudes, interests, attention span, motivation, competence, and the like, hence it could take any form. Asuru further defined achievement test as an ability test that is designed to appraise what the individual has learned to do as a result of planned previous experience or training often provided in school. It is specifically used to find out the extent to which a testee has achieved, gained or mastered certain information or skill after he has been exposed to some training. Through achievement test, the teacher is able to know the students' progress in the subject and whether or not the stated objectives have been realized.

In any educational endeavour, there must be criteria for measuring success or achievement. In addition, the criteria must be uniform and standardized. In Nigerian education, it is advocated that materials used in measuring achievements must be such that least error is encountered in the course of the measurement. Instruments used for achievement measurement must possess the needed psychometric properties (appropriate item difficulty, discrimination and distracter indices) if they are to serve the purpose they are constructed to serve. The locally constructed teacher-made tests are usually the most common tools for both assessment and promotion of students into new classes. Most teacher-made tests have been found to be carelessly constructed or not constructed based on test blueprints or tables of specification, they also lack psychometric properties of validity, reliability, appropriate difficulty, discrimination and effectiveness of distracters (Ugwu, 2012, Chime, 2012, Adams, 1981, Inomiesa, 1988 and Nkpono, 2001) To a large extent this claim is true, going by what is observed from the test items of some of the teachers of Social Studies. The test items are not well spread among the topics while some appear to repeat items. Some of the question papers are also hand written while some

questions do not reflect the subject content of Social Studies. This is a clear case of invalidity. Ubolom, Uzoeshi, Amini & Vipene (2011) stated that the importance of test in the life of testees makes it imperative for tests of higher standard to be used in collecting information which reflect their true potentialities. The qualities of a good test therefore could be broadly classified into two main areas, namely; intrinsic and extrinsic qualities. The intrinsic qualities constitute the validity, reliability and usability, while extrinsic qualities embrace such characteristics as administration and scoring of tests which test development incorporate.

Test development is the set of activities involved in constructing and evaluating a test of some psychological function. The steps include specifying the construct of interest, deciding the test's function (diagnosis, description of skill level, prediction of recovery), choosing a method (performance, behavioural observation, self-report), designing item content, evaluating the reliability and validity of the test, and modifying the test to maximize its utility. Development of a test deals with planning of the test through the item writing to the trial testing stage. Validity of a test implies that a test measures what it sets out to measure and nothing else. A test that is valid is one that is truthful, accurate and relevant in measuring what it intends to measure. A test which is valid measures the content of what it sets to measure and no other thing (Ubolom et al 2011). The reliability of the test refers to the consistency of the results obtained by the same person when tested with the same instrument at different times or with different sets of equivalent tests. Test reliability, therefore, seeks to establish how possible it is to reproduce the same or similar scores when the individual is measured again with the same or equivalent test (Obilor, 2018). The more nearly the scores are reproduced the higher the correlation coefficient and the more reliable the test. Usability or employability of a test is simply based on some common sense and practical considerations on whether a test can be used or not. These considerations include availability of equivalent form of the test, simplicity of instructions, ease of administration, ease of scoring, ease of interpretation and application of the test scores and test economy.

Most teacher-made achievement tests do not ensure these intrinsic and extrinsic qualities of validity, reliability, usability, administration and scoring. This study brings to focus the needed steps in test construction which are test planning, item writing, trial testing, item analysis, item selection, estimation of validity, computation of reliability values, and timing and printing of the final form of the test. Test planning includes all the preparatory processes in test construction, which is the first stage in the construction of achievement test. The processes are: stating and defining the objectives, outlining the content covered during instruction, and developing a test blueprints or table of specifications. Stating objectives behaviourally plays important roles in the teaching-learning encounter. Generally, they are the expected behavioural changes of the learner as a result of the teaching-learning encounter. They guide both the teaching and the assessment processes. The essence of stating the objectives is to determine the extent of the instructional objectives to be achieved. The objectives should be stated in specific behaviours that the students are expected to exhibit at the end of the lesson. They are usually stated using action verbs which clearly indicate specific and direct observable behaviours.

Outlining the content of instruction encompasses the various teaching content areas, units, topics, and sub-topics that constitute the course or subject as specified in course outline or syllabus (Asuru, 2015). To ensure that a test adequately samples the topics and sub-topics covered in the content, an outline of the content to be covered has to be made. This involves the breaking of the content into smaller units. For instance, a test on part one of educational measurement and evaluation for first degree students could be broken into the following

units: basic concepts, types of tests, qualities of a good test, test construction, test administration and scoring, continuous assessment and soft skills.

The Table of Specification is a two dimensional chart showing list of instructional objectives, content areas and types of items in its dimensions (Obilor, 2019). It also specifies the proportion of questions allotted to each of the behavioural objectives and topics of the content. Preparing the table of specification includes four main steps which are as follows: (a) Determine which instructional objectives to include. (b) Determine which content areas to include. (c) Determine the item types to include. (d) Prepare the 2- way chart (Table of Specification).

The second step of test construction is item writing which is the couching of the items in the required language and format. It constitutes a very important aspect of test construction. Item writing requires both practical experience and professional judgement. It is worthy of note that 50% more items than needed should be written. After the item writing, the pool of items should be reviewed by another expert in the subject area of specialization to spot any ambiguities, grammatical faults. Then the pool of items should be produced for trial testing.

The trial testing or trial-run of the pool of items is the third step of the test construction. This is akin to the test run-of a newly built industrial plant to determine if the installation conforms to standard before actual production begins. Just as in the industrial settings, trial testing is a very important stage in test construction because it is a stage of quality assurance/quality control and thus, determines whether each item and as well the entire test will be fit-for-purpose. This is determined both judgementally and statistically during item analysis.

The items should be administered to a fairly representative sample of testees similar to those for whom the test is intended in terms of the content and objectives of their programme. The essence of the trial testing is to generate empirical data about the individual adequacy of each item and thus provide information for item analysis. According to Wiggins (1998) and Riaz (2008), item analysis is about how difficult an item is and how well it can discriminate between the good and the poor students. In other words, item analysis provides a numerical assessment of item difficulty, item discrimination and effectiveness of distracters. It provides objective, external and empirical evidence for the quality of the items. The objective of item analysis is to identify problematic or poor items which might either confuse the respondents or do not have a clear correct response or a distracter might well be competing with the keyed answer. Item analysis comprises item difficulty, item discrimination, and effectiveness of distracter.

The fourth step of test construction is item analysis. Item analysis is the process of “testing the item” to ascertain specifically whether the item is functioning properly in measuring what the entire test is measuring (Obilor, 2019). Item analysis begins after the test has been administered and scored. It involves detailed and systematic examination of the testees’ responses to each item to determine the difficulty level and discriminating power of the item. The procedure for item analysis involves arranging the scores in ascending order from the highest to the lowest scores. The scores of the highest 27% and lowest 27% testees are selected and use. For instance, if the test was trial tested on 450 students, the scripts will be arranged in descending order of scores. One hundred and one highest scores (27% of 450) will be selected to make up the highest scorers and 121 lowest scores (27%, of 450) to constitute lowest scorers. The rest scripts will be discarded. These highest 121 and lowest 121 scorers will be used for item analysis.

Item difficulty is simply the percentage of students who answer an item correctly. In this case, it is also equal to the item mean. The item difficulty index ranges from 0 to 100; the higher the value, the easier the item (question). Item discrimination refers to the ability of an item to

differentiate among students on the basis of how well they know the material being tested. Ideally, more testees in the upper group should get each item correct than those in the lower group, while effectiveness of distracters is one that attracts students with misconceptions or errors in thinking and reasoning, generally those with lower overall ability. It is expected that all the distracters in each item should operate effectively. This means that every distracter must be chosen by at least one person each from the upper and lower groups.

Item selection is the fifth step of test construction in which the number of items needed is selected. Items should first be ranked in descending order based on the magnitude of their discriminating power. Those with zero and negative values should be rejected outright. In the case of item difficulty, there is no specific value to be selected, but it depends on the use of the test. In most achievement tests, values of between 40% and 60% (.4 and .6) levels of difficulty are selected to cater for a wide talents or achievement range. It is on the basis of rejecting some items that the need to write more items than needed is premised.

Estimating the validity and reliability values of the test is the sixth step. In this step the researcher is to use any of the following methods: Validity: content, face, construct and criterion; Reliability: Test-Retest, Split-Half, Cronbach Alpha, Kuder Richardson, and Parallel Form. Timing and printing of the final form of the test is the last stage in test construction and should be done considering the age, sex, disposition, location and educational level of testees. Purpose of the test must be considered too when timing the test.

Statement of the Problem

In the school system, the curriculum demands that different assessments be carried out in the course of the instruction to guide effective teaching, learning and to assess the level of mastery of the students. Among these is the routine weekly, end-of-term and/or end-of-session test. Unfortunately, most teachers do not know how to develop valid and reliable tests. The researcher has been observed and other researchers too (Chime, 2012; Ugwu, 2012; Obilor, 2019) that the teacher-made tests lack basic psychometric properties and as a result they are not very appropriate for the assessment of students. This may be due to several reasons including teachers' lack of the requisite skills of test development. It is to check this anomaly that this study sought to establish the requisite test construction template that will equip teachers and other stakeholders to be able to produce and use, quality test items that possess the desired psychometric properties.

Purpose of the Study

The purpose of this study was to develop and validate Social Studies Achievement Test (SSAT) for junior secondary schools in Ekiti State. Specifically, the study intended to:

1. Develop Social Studies Achievement Test (SSAT) for junior secondary three (3) students in Ekiti State. Validate the developed test in terms of content of the Social Studies Curriculum.
2. Trial-test the developed test by giving it out to an equivalent group.
3. Perform item analysis to determine the item difficulty, discrimination and the effectiveness of distracters.
4. Determine the reliability coefficient

Research Questions

The study was guided by the following research questions:

1. What Social Studies Achievement Test has been developed by the researcher?
2. To what extent are the developed items valid in terms of content of the Social Studies

Curriculum?

3. What are the item difficulty, discrimination indices and effectiveness of distracters of the items?
4. What is the reliability coefficient of the Social Studies Achievement Test (SSAT)?

METHODOLOGY

The study adopted the instrumentation research design. Instrumentation research design was used because the study involved the development and validation of Social Studies Achievement Test (SSAT) for evaluating the cognitive learning outcomes of junior secondary students in Social Studies. The population of the study was 99840 junior secondary school distributed as follows: Ekiti East – 55647; Ekiti South-East – 21305; and Ekiti West – 22888 (Ekiti State Universal Basic Education Board, 2018). The sample size comprised all the 650 Junior Secondary 3 (JS3) students offering Social Studies in Ekiti State made up of 280 students (Ekiti East), 178 students (Ekiti South-East), and 192 students (Ekiti West).

The instrument used for data collection was the researcher-developed Social Studies Achievement Test (SSAT) with a reliability coefficient of 0.76 computed using Kuder-Richardson Formula 20 (K-R20). The test was constructed based on the JS3 Social Studies Curriculum. The topics covered were: contents of Social Studies, family life education, roles of extended family members in child development, human trafficking, harmful traditional practices in Nigeria, promoting peaceful living in our society, social conflict, conflict management and resolution strategies, controlling cultism in our society, preventing drug trafficking, common crimes and associated punishment, and Crimes and National Security.

RESULTS

Research Question 1: What Social Studies Achievement Test has been developed by the researcher?

Research Question 2: To what extent are the developed items valid in terms of content of the Social Studies Curriculum?

The draft SSAT was submitted to 47 experts in Social Studies, Measurement and Evaluation, and the teachers of Social Studies from all the sampled junior secondary schools in Ekiti State, for detailed editing, careful and critical review of the test items. This was done to avoid the inclusion of irrelevant items. To further ensure content validity of the SSAT, a table of specification (test blueprint) was developed and used as a guide in constructing the items. The said table of specification was also made available to the Social Studies experts for the validation exercise. Table 1 displays the Table of Specification for the SSAT for JS3 in Ekiti State.

Table 1: Table of Specification for Social Studies Achievement Test for JS3

| Contents | Behavioural Objectives | | | | | Eval. 0% Total |
|----------|------------------------|-------------------|-------|-------|--|----------------|
| | Know.4 4% | Comp. | Appl. | Anal. | | |
| | | Synth. 20% 10% | 16% | 10% | | |

| | | | | | | | |
|------------------------------------------------------------|-----------|-----------|-----------|-----------|-----------|----------|------------|
| Contents of Social Studies (6%) | 4 | 1 | NA | 1 | NA | NA | 6 |
| Roles of extended family Members in child development (6%) | 4 | 2 | NA | NA | NA | NA | 6 |
| Human trafficking (10%) | 4 | 2 | 2 | 2 | N.A | N.A | 10 |
| Preventing human trafficking (8%) | 3 | 2 | 2 | N.A | 1 | N.A | 8 |
| Harmful traditional practices (8%) | 4 | 1 | 1 | 1 | 1 | N.A | 8 |
| Promoting peaceful living in our society(8%) | 3 | 2 | 2 | 1 | NA | NA | 8 |
| Social conflict (10%) | 4 | 2 | 2 | 1 | 1 | NA | 10 |
| Managing and resolving (10%) | 4 | 2 | 1 | NA | 3 | NA | 10 |
| Controlling cultism in our society (10%) | 4 | 1 | 2 | 1 | 2 | NA | 10 |
| Preventing drug trafficking (10%) | 4 | 2 | 1 | 1 | 2 | NA | 10 |
| Common crime and Associated punishment (6%) | 3 | 1 | 2 | NA | NA | NA | 6 |
| Crime and National Security (8%) | 3 | 2 | 1 | 2 | NA | NA | 8 |
| Total | 44 | 20 | 16 | 10 | 10 | 0 | 100 |

Note: NA = Not Applicable, while Know, Comp, Appl, Anal, Synth and Eval represent respectively Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation.

Research Questions 3: What are the item difficulty, discriminating indices, and effectiveness of distracters of the items?

Table 2: Item Analysis Showing the Difficulty, Discriminating Indices and Distracters

| S/No | Items | Keys | Upper 27% | Lower 27% | Difficulty Index | Discriminating Index | Distracters |
|------|-------|------|-----------|-----------|------------------|----------------------|-------------|
| 1 | 1 | D | 160 | 82 | 0.69 | 0.44 | A, B and C |
| 2 | 2 | C | 158 | 49 | 0.59 | 0.62 | A, B and D |
| 3 | 3 | C | 104 | 57 | 0.46 | 0.27 | A, B and D |
| 4 | 4 | D | 147 | 59 | 0.59 | 0.50 | A, B and C |
| 5 | 5 | C | 163 | 54 | 0.62 | 0.62 | A, B and D |
| 6 | 6 | A | 158 | 62 | 0.63 | 0.55 | B, C and D |
| 7 | 7 | B | 148 | 54 | 0.58 | 0.54 | A, C and D |
| 8 | 8 | B | 85 | 40 | 0.36 | 0.26 | A, C and D |
| 9 | 9 | A | 138 | 60 | 0.57 | 0.44 | B, C and D |
| 10 | 10 | A | 145 | 72 | 0.62 | 0.42 | B, C and D |
| 11 | 11 | A | 153 | 56 | 0.58 | 0.55 | B, C and D |
| 12 | 12 | B | 150 | 55 | 0.59 | 0.54 | A, C and D |
| 13 | 13 | B | 83 | 47 | 0.37 | 0.21 | A, C and D |
| 14 | 14 | B | 138 | 51 | 0.54 | 0.50 | A, C and D |
| 15 | 15 | A | 141 | 53 | 0.55 | 0.50 | B, C and D |
| 16 | 16 | A | 143 | 56 | 0.57 | 0.50 | B, C and D |
| 17 | 17 | C | 157 | 66 | 0.64 | 0.52 | A, B and D |
| 18 | 18 | C | 140 | 61 | 0.57 | 0.45 | A, B and D |
| 19 | 19 | D | 139 | 48 | 0.53 | 0.52 | A, B and C |
| 20 | 20 | A | 128 | 49 | 0.51 | 0.45 | B, C and D |
| 21 | 21 | C | 113 | 48 | 0.46 | 0.37 | A, B and D |
| 22 | 22 | D | 118 | 47 | 0.47 | 0.41 | A, B and C |
| 23 | 23 | A | 123 | 51 | 0.50 | 0.41 | B, C and D |
| 24 | 24 | D | 142 | 70 | 0.61 | 0.41 | A, B and C |
| 25 | 25 | D | 152 | 57 | 0.60 | 0.54 | A, B and C |
| 26 | 26 | A | 114 | 62 | 0.50 | 0.30 | B, C and D |
| 27 | 27 | D | 80 | 35 | 0.33 | 0.26 | A, B and C |
| 28 | 28 | D | 128 | 51 | 0.51 | 0.44 | A, B and C |
| 29 | 29 | A | 133 | 60 | 0.55 | 0.42 | B, C and D |
| 30 | 30 | A | 138 | 74 | 0.61 | 0.37 | B, C and D |
| 31 | 31 | A | 140 | 63 | 0.58 | 0.44 | B, C and D |
| 32 | 32 | C | 155 | 58 | 0.61 | 0.55 | A, B and D |
| 33 | 33 | A | 141 | 68 | 0.60 | 0.42 | B, C and D |
| 34 | 34 | A | 114 | 67 | 0.52 | 0.42 | B, C and D |
| 35 | 35 | D | 133 | 52 | 0.53 | 0.46 | A, C and C |
| 36 | 36 | B | 156 | 52 | 0.59 | 0.59 | A, C and D |
| 37 | 37 | D | 150 | 51 | 0.57 | 0.56 | A, B and C |
| 38 | 38 | C | 154 | 53 | 0.59 | 0.57 | A, B and D |
| 39 | 39 | D | 151 | 60 | 0.60 | 0.52 | A, B and C |
| 40 | 40 | D | 90 | 48 | 0.39 | 0.24 | A, B and c |
| 41 | 41 | A | 153 | 56 | 0.58 | 0.55 | B, C and D |

| | | | | | | | |
|----|----|---|-----|----|------|------|------------|
| 42 | 42 | A | 163 | 54 | 0.62 | 0.62 | B, C and D |
| 43 | 43 | D | 158 | 49 | 0.59 | 0.62 | A, B and C |
| 44 | 44 | A | 124 | 58 | 0.52 | 0.37 | B, C and D |
| 45 | 45 | B | 150 | 5 | 0.59 | 0.53 | A, C and D |
| 46 | 46 | C | 155 | 58 | 0.61 | 0.55 | A, B and D |
| 47 | 47 | B | 136 | 56 | 0.55 | 0.48 | A, C and D |
| 48 | 48 | D | 152 | 60 | 0.61 | 0.53 | A, B and C |
| 49 | 49 | B | 154 | 58 | 0.61 | 0.55 | A, C and D |
| 50 | 50 | D | 150 | 55 | 0.59 | 0.31 | A, B and C |
| 51 | 51 | C | 120 | 50 | 0.48 | 0.40 | A, B and D |
| 52 | 52 | A | 120 | 70 | 0.54 | 0.28 | B, C and D |
| 53 | 53 | D | 80 | 35 | 0.33 | 0.26 | A, B and C |
| 54 | 54 | A | 85 | 40 | 0.36 | 0.26 | B, C and D |
| 55 | 55 | A | 100 | 52 | 0.43 | 0.27 | B, C and D |
| 56 | 56 | C | 136 | 56 | 0.55 | 0.48 | A, B and D |
| 57 | 57 | B | 152 | 60 | 0.61 | 0.53 | A, C and D |
| 58 | 58 | D | 97 | 60 | 0.49 | 0.21 | A, B and C |
| 59 | 59 | C | 83 | 47 | 0.37 | 0.21 | A, B and D |
| 60 | 60 | C | 104 | 57 | 0.46 | 0.27 | A, B and D |
| 61 | 61 | A | 147 | 59 | 0.59 | 0.50 | B, C and D |
| 62 | 62 | C | 91 | 48 | 0.39 | 0.24 | A, B and D |
| 63 | 63 | D | 128 | 49 | 0.51 | 0.45 | A, B and C |
| 64 | 64 | A | 113 | 48 | 0.46 | 0.37 | B, C and D |
| 65 | 65 | C | 98 | 40 | 0.39 | 0.33 | A, B and D |
| 66 | 66 | A | 123 | 51 | 0.50 | 0.41 | B, C and D |
| 67 | 67 | A | 142 | 70 | 0.36 | 0.41 | B, C and D |
| 68 | 68 | B | 88 | 40 | 0.36 | 0.27 | A, C and D |
| 69 | 69 | D | 114 | 62 | 0.50 | 0.30 | A, B and C |
| 70 | 70 | A | 85 | 41 | 0.36 | 0.25 | B, C and D |

Table 2 shows the item difficulty and discrimination indices of each of the items of the instrument. In terms of item difficulty, 58 items out of 70 items have item difficulty indices ranging from 0.43 to 0.61 which satisfied the acceptable range, while 12 items fell within 0.33 to 0.37 which indicated that the items are not adequate for inclusion. In terms of item discrimination the 70 items have item discrimination ranges from 0.21 to 0.62 which also satisfied the acceptable range. All the 70 items distracted effectively.

Research Question 4: What is the reliability coefficient of the Social Studies Achievement Test (SSAT)?

The reliability coefficient of the SSAT is 0.76 computed using Kuder Richardson Formula 20 (KR-20). This shows that the Social Studies Achievement Test developed is highly reliable.

Summary of Findings

1. A Social Studies Achievement Test containing 58 items was developed.
2. The developed test met the criteria of content validity.
3. Fifty eight (58) items of the Social Studies Achievement Test have appropriate item difficulty indices, 70 items have positive discrimination indices, and all the items distracted effectively.
4. The developed Social Studies Achievement Test has a high reliability coefficient of

0.76 using Kuder Richardson formula.

Discussions

The results of the study show that 58 items out of 70 items have item difficulty indices ranging from 0.43 to 0.61 which satisfied the acceptable range, while 11 items fell within 0.33 to 0.37 which indicated that the items are not ideal. In terms of item discrimination, 70 items have item discrimination indices ranging from 0.21 to 0.62 which also satisfied the acceptable range. This implies that most of the items are ideal and are acceptable because they have appropriate difficulty indices and positive discrimination indices. These results are similar to the findings by Inomiesa (1998) who constructed and validated an achievement test on upper primary science. The 102 items have facility indices which ranged from 0.30 to 0.70 and discrimination indices of 0.20 to 0.60. Also the results are similar to those of Adams (1981) who constructed and validated an achievement test in Integrated Science for Nigerian Secondary JS1 students. The following psychometric measures were established for the test: an average discriminative index of 46.33, a difficulty indices range of 18.5 to 82.5. This work is similar to that of Chime (2012) and Adams (1981) who also agreed that in their various achievement test items, the distractors distract effectively.

The developed Social Studies Achievement Test for junior secondary schools in Ekiti State has a very high reliability coefficient of 0.76 using Kuder-Richardson formula. In a similar manner, the studies by Effiong (2006), Nkpone (2001) and Inosmiesa (1988) had reliability coefficients of 0.99, 0.89 and 0.87 respectively using Kuder Richardson formula. These values of reliability coefficient were considered high reliability, thus the present study is equally considered to have a high level of reliability coefficient, thereby making the SSAT very valuable.

CONCLUSION

From the results obtained, the following conclusions were drawn:

1. The Social Studies Achievement Test (SSAT) was developed by the researcher.
2. The validated items were congruent with the Social Studies curriculum.
3. The developed and validated Social Studies Achievement Test exhibited a good measure of difficulty and discrimination indices, and the items distracted effectively.
4. The Test has a high reliability coefficient of 0.76 computed using KR-20 method.

RECOMMENDATIONS

Based on the findings of the study, the following recommendations were made:

1. The developed Social Studies Achievement Test should be used by Social Studies teachers.
2. The items of SSAT developed should serve as a template to develop other achievement tests in Social Studies (and other subjects) for junior secondary schools.
3. Regular sensitization workshop, seminars and conferences should be organized for teachers in order for them to be acquainted with techniques needed for construction of valid assessment instruments.
4. Educational inspectors should improve upon their duties by embarking on consistent, planned and objective inspection.

REFERENCES

Adams, I. K. (1998). *Construction and validation of an achievement test in Integrated Science for the Nigeria secondary school class one*. Unpublished M.ed. Thesis. Ibadan: University of Ibadan.

Asuru, V. A. (2015). *Measurement and evaluation in education and psychology*. Port Harcourt: Minson Nigeria Limited.

Chime, U. M. (2012). *Development and validation of economics achievement test for senior secondary school students*. Unpublished M.Ed Thesis, Faculty of Education, University of Nigeria, Nsukka.

Effiong, E. J. (2006). *Development and Validation of alternative to practical tests for measurement skills in electronic devices and circuits on technical colleges*. Unpublished Ph.D Thesis, University of Nigeria, Nsukka.

Inomiesa, E. A. (1988). *The development, validation and use of standardized instrument for the continuous assessment of pupil achievement in upper primary science*. Unpublished Ph.D Thesis, Department of Education, University of Nigeria Nsukka

Nkpone, H. L. (2001). *Development and standardization of Physics achievement test for senior secondary school students*. Unpublished M.Ed Thesis, Faculty of Education, University of Nigeria, Nsukka.

Obilor, E. I. (2018). *Fundamentals research methods and statistics in education and social sciences*. Port Harcourt: Sabcos Printers and Publishers.

Obilor, E. I. (2019). *Essentials of measurement and evaluation*. Port Harcourt: Sabcos Printers and Publishers.

Onunkwo, G. I. N. (2002). *Fundamentals of educational measurement and evaluation*.

Owerri: Cape Publishers International Ltd.

Riaz, M. N.(2008).*Test Construction: Development and Standardization of Psychological Tests in Pakistan*. Islamabad: HEC

Ubolom, W. J., Uzoeshi, K. C., Amini, C. M. & Vipene, J. B. (2011). *Fundamentals of measurement and evaluation*. Port Harcourt: Celwil Publisher.

Ugwu, S. N. (2012). *Development and validation of criterion referenced achievement test in Biology*. Unpublished M.Ed Thesis, Faculty of Education, University of Nigeria, Nsukka.

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.