**RESEARCH PAPER**

# Development of a Valid and Reliable Summative Test in Plane Trigonometry

Joseph Balinag Tandas*

BSEd Chairperson, College of Education, Ifugao State University-Potia Campus, Philippines

*Corresponding author: tandas_joseph@yahoo.com

Received: 27 August 2020; Accepted: 22 January 2021; Published: 29 January 2021

**To cite this article (APA):** Tandas, J. B. (2021). Development of a Valid and Reliable Summative Test in Plane Trigonometry. *Journal of Science and Mathematics Letters, 9*(1), 46-59. https://doi.org/10.37134/jsml.vol9.1.5.2021

**To link to this article:** https://doi.org/10.37134/jsml.vol9.1.5.2021

## Abstract

The study was carried out to develop a valid and reliable test in Plane Trigonometry for college students. Four research questions were drawn to guide the establishment of validity and reliability for the Plane Trigonometry summative test. It is a multiple choice objective test of four options with 50 items. All students enrolled in plane trigonometry were used as participants to determine the reliability of the test. Three (3) experts/judges were used for the content validity of the test. After validation to the three experts/judges, the result showed that the Plane Trigonometry test has a very high validity. The test item validity was determined through difficulty and discrimination indices. The test has a reliability coefficient of 0.909 established through the use of Cronbach's Alpha. The test is valid and reliable for assessing students' content knowledge in Plane Trigonometry. Thus, the constructed test can be used for institutional entrance examination in the field of mathematics education.

**Keywords**: Multiple Choice Objective Test; Validity; Reliability; Plane Trigonometry

## INTRODUCTION

Trigonometry is an inseparable part of mathematics in higher education. It takes some subjects of arithmetic and geometry as any source. In other words, it is a product of algebraic techniques, geometrical realities and trigonometric relationships. Mathematics, particularly trigonometry is one of the school subjects that most students hate and struggle with. Trigonometry is an area of mathematics that students believe to be particularly difficult and abstract compared with the other subjects of mathematics (Gur, 2009). Thus, there is a need for an assessment to evaluate the content knowledge of the students.

Test is an assessment that measures student's learning. If a test is good, it should reflect student's proficiency level. On the other hand, an ill-conceived test may reveal what the students do not know and what they have not been taught (Hasan, 2014). A good test has some characteristic qualities: validity and reliability. It is important to note that designing test instrument involves specifications, test construction, try-out, analysis and revision. Most teachers hurriedly copy questions from any past question paper to compose their summative tests. As a result, teachers do not establish validity and reliability for such tests (Priyambodo, 2016).

The unreliable summative test administered during examinations, is often used by the teacher to assess the students. The use of poorly designed summative test is a major problem as it affects students' interest and achievement in a certain subject. It has already been pointed out that poorly designed tests could make the students loose interest in a particular subject (Osadebe, 2001; Ohuche & Akeju, 1988 in Osadebe, 2014). Similarly, Onunkwo's (1998) observation that most examiners find it easier to construct test items in the lower cognitive levels (knowledge and comprehension) than the higher cognitive levels (application, analysis, synthesis and evaluation). However, the construction of test items is an art that only few people seem to master (Nunnally, 1981; Osadebe, 2001 in Osadebe, 2014). Therefore, there is the need for experts to construct enough valid and reliable tests for use in schools.

The construction of valid and reliable test demand a special knowledge. There are literature on the test construction. When a teacher constructs a test, it is said to be a teacher made test that is poorly prepared. Then when an expert constructs a valid and reliable test, it is called a standardized test. Be it as it may, a teacher can construct a test if well guided. The problem of teachers for constructing poor test is a major issue in education that requires special attention. However, teachers should consult an expert before using a test (Osadebe, 2012). Thus, non-standardized and standardized tests could be produced by experts. The construction and standardization of test are special areas for those who have the background (Nunnally, 1981; Osadebe, 2001 in Osadebe, 2014).

A well-respected leader in test design, Popham (2014) states that in general, constructing classroom texts with care will be sufficiently reliable for the decisions a teacher will base on the test result. A teacher, on the other hand, needs to be at least knowledgeable about the fundamental meaning of reliability. However, Popham (2014) suggested that any classroom tests must always undergo reliability testing.

The researcher decides to construct a valid and reliable summative test in plane trigonometry for teachers as an area of need. The test will only be given out when needed so as to avoid misuse. It should be used in assessing students' achievement after teaching and learning plane trigonometry particularly when the content areas have been covered.

The irregular education students of Ifugao State University-Potia Campus are currently enrolled in Math 12-Plane Trigonometry. The school gave a chance to those students following the old curriculum to finish the course. Thus, the students respondents are the combination of irregular first year to irregular fourth year students.

Therefore, the main purpose of the study was to validate summative test in plane trigonometry for teachers and others to use particularly when the content areas of plane trigonometry curriculum have been covered and use to prepare students for external examinations such as qualifying examination or used for pre-test in the college.

Specifically, the study focused on how to establish item analysis, validity and reliability for the summative test in plane trigonometry.

## Framework of the Study

Gabuyo (2014) said that one of the most important functions of a teacher is to assess the performance of the students. This is very complicated task because you will consider many activities such as the timing of the assessment process, the format of the assessment tools and the duration of the assessment procedures. He stressed that after designing the assessment tools, package the test, administer the test to the students, check the test papers, score and then record them. Return the test

papers and then give feedback to the students regarding the result of the test. After constructing the test items and putting them together, then the next step is to administer the test to the students. The administration procedures greatly affect the performance of the students in the test. After the examination, the next activity that the teacher needs to do is to score the test papers, record the result of the examination, return the test papers and last to discuss the test items in the class so that you can analyze and improve the test items for future use. Thus, he said that after administering and scoring the test, the teacher should also analyze the quality of each item in the test. Through this you can identify the item that is good, item that needs improvement or items to be removed from the test. But when do we consider that the test is good? How do we evaluate the quality of each item in the test? Why is it necessary to evaluate each item in the test?

With this, item analysis helped teachers determine the quality of a test. Item analysis is the process of examining the student's response to individual item in the test. One of the purposes of item analysis is to improve the quality of the assessment tools. Through this we can identify the item that is to be retained, revised or rejected and also the content of the lesson that is mastered or not.

Tabulation is done to determine the level of difficulty or item difficulty, and discriminating power of the test items or item discrimination.

The above two indices help in item selection for the final draft of the test. Another step which leads the calculation of item difficulty and item discrimination of a test is item selection based upon the judgment of competent persons as to the suitability of the item for the purposes of the test (Aggarwal, 1986 in Boopathiraj & Chellaman, 2013). There are several methods of item analysis described in various texts exclusively based on construction of tests.

Item difficulty may be defined as the proportion of the examinees that marked the item correctly. Item difficulty is the percentage of students that correctly answered the item, also referred to as the p-value. The range is from 0% to 100%, the higher the value, the easier the item. P values above 0.90 are very easy items and might be a concept not worth testing. P-values below 0.20 indicate difficult items and should be reviewed for possible confusing language or the contents needs re-instruction. Optimum difficulty level is 0.50 for maximum discrimination between high and low achievers. For example an item answered correctly by 70% examinees has a difficulty index of 0.70. If 90% of a standard group pass an item, it is easy; if only 10% pass, the item is hard or too difficult. Generally, items of moderate difficulty are to be preferred to those which are much easier or much harder. The higher the value of the index of difficulty, the easier the item is. Hence, more students got the correct answer and more students mastered the content measured by that item.

Item discrimination or the discriminating power of a test item refers to the degree to which success or failure on an item indicates possession of the ability being measured. It determines the extent to which the given item discriminates among examinees in the function or ability measured by the item. This value ranges between 0.0 and 1.00. Higher the value, more discrimination of the item is. A highly discriminating item indicates that the students who had high tests scores got the item correct whereas students who had low test scores got the item incorrect. It also refers to the number of students in the upper group who got an item correctly minus the number of students in the lower group who got an item correctly. Divide the difference by either the number of the students in the upper group or number of students in the lower group or get the higher number if they are not equal. Discrimination index is the basis of measuring the validity of an item. This index can be interpreted as an indication of the extent to which over-all knowledge of the content area or mastery of the skills is related to the response on an item.

There are three kinds of discrimination index: positive discrimination, negative discrimination, and zero discrimination. Positive discrimination happens when more students in the

upper group got the item correctly than those in the lower group. Negative discrimination occurs when more students in the lower group got the item correctly than the students in the upper group. Zero discrimination happens when a number of students in the upper group and lower group who answer the test correctly are equal, hence the test item cannot distinguish the students who performed in the overall test and the students whose performance are very poor.

The following are the steps in solving difficulty index and discrimination index: (1) Arrange the scores from highest to lowest; (2) Separate the scores into upper group and lower group. There are different methods to do this: (a) if a class consists of 30 students who takes an exam, arrange their scores from highest to lowest, then divide them into two groups. The highest score belong to the upper group. The lowest score belong to the lower group. (b) Other literature suggested to use 27%,30%, or 33% of the students for the upper group and lower group. However, in the Licensure Examination for Teachers (LET) the test developers always used 27% of the students who participated in the examination for the upper and lower groups; (3) Count the number of those who chose the alternatives in the upper and lower group for each item and record the information; (4) Compute the value of the difficulty index and the discrimination index; (5) Make an analysis for each item.

Reliability refers to the consistency with which it yields the same rank for individuals who take the test more than once (Kubiszyn & Borich, 2007, in Gabuyo 2014). That is, how consistent test results or other assessment results from one measurement to another. We can say that a test is reliable when it can be used to predict practically the same scores when test administered twice to the same group of students and with a reliability index of .60 or above. The reliability of a test can be determined by means of Pearson product correlation coefficient, Spearman-Brown formula and Kuder-Richardson formulas.

Test constructors believed that every assessment tool should possess good qualities. Most literatures consider the most common technical concepts in assessment are the validity and reliability. For any type of assessment whether traditional or authentic it should be carefully developed so that it may serve whatever purpose it may have (Boopathiraj & Chellaman, 2013).

The main purpose of the study was to validate summative test in plane trigonometry for teachers and others to use particularly when the content areas of plane trigonometry curriculum have been covered. Specifically, the study focused on how to establish item analysis, validity and reliability for the summative test in plane trigonometry.

To understand better the conceptual framework, the research paradigm was presented below.

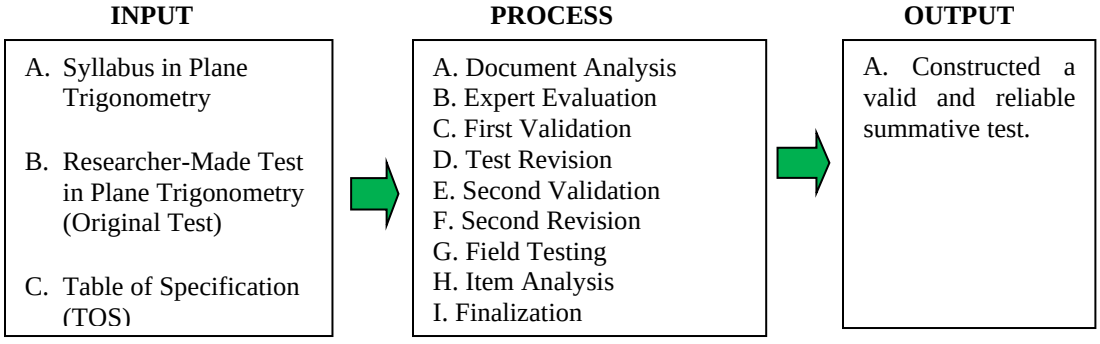


**Figure 1.** Research paradigm.

The researcher used the input, process, and output for the conduct of the study. The researcher includes syllabi in Plane Trigonometry, researcher-made test in plane trigonometry, and table of

specification as input of the study. For the process, firstly, document analysis was done to look into the contents of the syllabus and the original test with the table of specification. Second, the selected three experts evaluated the original test. Third, first validation was done to withdraw those items not included. Fourth, conducted test revisions. Fifth, second validation was done to the test revised. Sixth, revision was done from to the second validation. Seventh, conducted field testing among respondents. Eighth, Item analysis was done to check difficulty and discrimination and computed reliability coefficient. Ninth, Finalization of the test. Constructed a valid and reliable summative test was the output of the study.

**Statement of the Problem**

The purpose of this study is to develop a valid and reliable test in Trigonometry. Specifically, it seeks to answer the following questions:
1. What summative test in plane trigonometry will be developed?
2. What is the level of validity of the summative test in plane trigonometry?
3. What is the reliability of the summative test in plane trigonometry?
4. What is the difficulty index of the test in plane trigonometry?
5. What is the discrimination index of the test in plane trigonometry?

## METHODOLOGY

This study utilized a developmental research design to develop a summative test in plane trigonometry subject at the Ifugao State University- Potia Campus. This study took place at the Ifugao State University (IFSU) – Potia Campus. The irregular education students were currently enrolled in Math 12 (Plane Trigonometry) - This subject is designed to complete the general education courses in the college of education. There are only thirty irregular education students enrolled in the course- 7 males and 23 females, all of them were taken as the participants of this study. There are also three experts/judges used for the content validation of the test in plane trigonometry. They were the teachers who permanently teaches mathematics for a minimum of three (3) years.

The study is instrumentation in nature. It is instrumentation because it is a study aimed at the construction, validation and production of valid and reliable summative test for teachers and others to use in assessing students' content knowledge in Plane Trigonometry. The plane trigonometry test is a multiple choice objective test of 50 items. The test items were drawn in a blue print or table of specifications based on the content of plane trigonometry curriculum and the educational objectives of Bloom taxonomy of Remembering, Understanding, Applying, Analyzing, Evaluating and Creating. There are four options (A, B, C, D) for each item: made up of one correct answer (key) and three wrong answers (distracters). The distracters are plausible and each was randomly distributed.

The researcher conducted the first field testing for the detection and elimination, and revision of the original test with 70 item to come up with 50 item test in plane trigonometry. After the elimination and revision of the test, the researcher floated the second field testing with the same students.

Tabulation was done to determine the level of difficulty or item difficulty and discriminating power of the test items or item discrimination. These two indices help in item selection for the final

draft of the test. Scores of the students were entered in Microsoft Excel sheet and it was arranged in descending order. Since there are only 30 examinees, divide the respondents with 50% upper group and 50% lower group of students for the item analysis. Item analysis was conducted on the test items, where the index of difficulty and discrimination was computed, based on the result of the test of the students.

The weighted mean was used to determine the extent of validity of the plane trigonometry test. **Cronbach's alpha** was used to measure the internal consistency and reliability of the test in plane trigonometry. If the coefficient is reliable, then it is valid.

ADDIE model was an instructional design that were also suited in the construction of the summative test. Analyze, design, development, implementation, and evaluation was the five phases of ADDIE model. First, analyze the learners in which recognize and investigate what was included in the syllabus and in the original test as well. It is also here that the items matches the level of skill and intelligence that each student/participant shows. Second, design where it assess the tools that was used if items in the test corresponds to the level of cognition in the table of specification. Third, development where designers make use of the data collected from the two previous stages, and use this information to revise items. Fourth, implementation where it reflects the continuous modification and field testing of the test to make sure maximum efficiency and positive results are obtained. Fifth, Evaluate if the goals have been met, and to establish what will be required moving forward in order to further the efficiency and success rate of the test. Do item analyze and finalization.

The following were used for the level of difficulty, level of discrimination and extent of validity of test. Table 1 was used to check the difficulty level of an item. This is ranges from range 0.00-0.20 to range 0.81-1.00 or Very Difficult to Very Easy respectively (Gabuyo, 2014). Moreover, Table 2 was used to check the level of discrimination of an item. It ranges from 0.19 and below to 0.40 and above or Poor item to Very good item respectively (Ebel & Frisbie (1986); Hetzel (1997) in Gabuyo (2014)). Table 3 was used for the extent of validity of the test item. Range 4.20-5.00 was interpreted very high or no flaws observed and nothing more to be desired to make it better. Range 3.40-4.19 was interpreted as high or very little flaws are observed and minor rewording of few items needed. Range 2.60 -3.39 was interpreted as moderate or some flaws are observed and the overall usefulness is diminished only slightly. Range 1.80 – 2.59 was interpreted as fair or several flaws are observed and overall usefulness is diminished greatly. Range 0.00-1.79 was interpreted as poor or major revision is needed to make it useful. Table 4 was used to interpret the level of reliability of the test. Ranges from below 0.50 to above 0.9 or questionable reliability to excellent reliability (***Scorepak: Item Analysis)***.

**Table 1.** Level of difficulty of an item.

| Index Range | Difficulty Level |
|---|---|
| 0.00-0.20 | Very Difficult |
| 0.21-0.40 | Difficult |
| 0.41-0.60 | Average/Moderately Difficult |
| 0.61-0.80 | Easy |
| 0.81-1.00 | Very Easy |

**Table 2.** Level of discrimination.

| Index Range | Discrimination Level |
|---|---|

| | |
|---|---|
| 0.19 and below | Poor item, should be eliminated or need to be revised |
| 0.20-0.29 | Marginal item, needs some revision |
| 0.30-0.39 | Reasonably good item but possibly for improvement |
| 0.40 and above | Very good item |

**Table 3.** Extent of validity.

| Scale/Range | Qualitative Interpretation |
|---|---|
| 4.20 -5.00 – Very High | No flaws observed; nothing more to be desired to make it better. |
| 3.40-4.19 – High | Very little flaws are observed; minor rewording of few items needed. |
| 2.60-3.39 – Moderate | Some flaws are observed; the overall usefulness is diminished only slightly. |
| 1.80-2.59 – Fair | Several flaws are observed; overall usefulness is diminished greatly. |
| 0.00-1.79 – Poor | Major revision is needed to make it useful. |

**Table 4.** Level of reliability coefficient.

| Reliability Coefficient | Interpretation |
|---|---|
| Above 0.90 | Excellent reliability |
| 0.81-0.90 | Very good for a classroom test |
| 0.71-0.80 | Good for classroom test. There are probably few items needs to be improved |
| 0.61-0.70 | Somewhat low. The test needs to be supplemented by other measures (more test) to determine grades |
| 0.51-0.60 | Suggests need for revision of test, unless it is quite short (ten or fewer items). Needs to be supplemented by other measures (more test) for grading |
| 0.50 and below | Questionable reliability. This test should not contribute heavily to the course grade, and it needs revision. |

## RESULTS AND DISCUSSION

The presentation, analysis and interpretation of the data was presented below which aims to develop a valid and reliable test in Plane Trigonometry. Specifically, the study focused on how to establish item analysis, validity and reliability for the summative test in plane trigonometry.

### Table of Specification of the summative test in Plane Trigonometry

Table 5 shows the table of specifications used as a basis in the development of the summative test. The TOS is used to identify the achievement domains being measured and to ensure that a fair and representative sample of questions appears on the test. Teachers cannot measure every topic or objective and cannot ask every question they might wish to ask. A TOS allows the teacher to construct a test that focuses on the key areas and weights those different areas based on their importance. It provides the teacher with evidence that a test has content validity and that it covers what should be covered. TOS are typically designed based on the list of course objectives, the topics covered in class, the amount of time spent on those topics, textbook chapter topics, and the emphasis and space provided in the text. In some cases, a great weight will be assigned to an extremely important concept, even if relatively little class time was spent on the topic. It benefits students in

two ways: (1) it improves the validity of teacher-made tests; and (2) it can improve student learning as well.

In this study, the researcher included six learning competencies in the summative test as reflected in table the table below. Introductory concepts has twelve items: eight of which are under *remembering*, four is *understanding*. Trigonometric functions of an angle has seven items: two *remembering,* two *understanding,* and three *analysing*. Right triangles has sixteen items: three of which are *remembering,* four *applying,* and nine are *evaluating*. Oblique triangles has fifteen items: nine *applying*; *five analyzing,* and *one evaluating*. The TOS shows that the summative test has a total of 50 items: most of the test questions fall under *remembering and applying* with 13 items each, 6 items are under *understanding*, 8 items are under *analyzing*, and 10 items are under *evaluating*. This was in accordance with the research study of Alade and Omoruyi (2014) that teachers should endeavors to construct a well test blue print that will help improve the validity of teacher evaluation based on given assessment, teachers must ensure that the test constructed measure an adequate sampling of the class at all level of domains.

## Level of Validity of the Test in Plane Trigonometry

Table 6 presents the level of validation of the three experts/ judges in plane trigonometry test. It further shows that the overall weighted mean is 4.60 or very high. The following indicator in the instrument such as "Each of the items is comprehensive; it covered all areas that are important to the study" and "The items are formulated in accordance to the competencies and objectives of the lessons" obtained a weighted mean of 4.33 interpreted as high. This means that very little flaws are observed; minor rewording of few items needed. The indicator "The items do not overlap with each other; no duplication of items was done" has a weighted mean of 5.00 interpreted as very high and the other indicators are rated with a weighted mean of 4.67 interpreted as very high. This means that no flaws observed; nothing more to be desired to make it better. This is in consonance with the research study of Alfrits Roul Sinadia and Surya Jatmika (2020) that all items were considered valid based on validity analysis.

**Table 5.** Table of specification in plane trigonometry.

| Topic/Content | Types of Test | S K I L L S | | | | | | Total | % |
|---|---|---|---|---|---|---|---|---|---|
| | | Remembering | Understanding | Applying | Analyzing | Evaluating | Creating | | |
| **Chapter I: Introductory concepts**<br>1. kinds of angles<br>2. radian measure<br>3. relationship between degree and radian measure | Multiple Choice | **8**<br>8,10,11,17, 18,36,41,42 | **4**<br>1,2,3,5 | | | | | **12** | **24.00%** |
| **Chapter II: Trigonometric functions of an angle** | | **2**<br>4,37 | **2**<br>6,16 | | **3**<br>7, 12,20 | | | **7** | **14.00%** |
| **Chapter III. Right triangles** | | **3**<br>38,39,40 | | **4**<br>9,14,15,44 | | **9**<br>13,19, 43, 45-50 | | **16** | **32.00%** |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Chapter IV: Oblique triangles** | | | | **9** 22,23,24, 26,27,28, 30,31,32 | **5** 25,29,33, 34,35 | **1** 21 | | **15** | **30.00%** |
| TOTAL | | **13** | **6** | **13** | **8** | **10** | | **50** | |
| Percentage | | **26%** | **12%** | **26%** | **16%** | **20%** | | | **100.00%** |

**Table 6.** Level of validity of the test in trigonometry.

| Indicator | Weighted Mean | Qualitative Interpretation |
|---|---|---|
| 1. The directions of the test are clear. | 4.67 | Very High |
| 2. Each of the items is clearly stated. | 4.67 | Very High |
| 3. The items are interrelated. | 4.67 | Very High |
| 4. Options per item are plausible enough. | 4.67 | Very High |
| 5. Each of the items is comprehensive; it covered all areas that are important to the study. | 4.33 | High |
| 6. Each item is focused on the particular thought or idea. | 4.67 | Very High |
| 7. The items are objective; the responses to be elicited are neither biased nor reactive. | 4.67 | Very High |
| 8. The items are formulated in accordance to the competencies and objectives of the lessons. | 4.33 | High |
| 9. The items are systematically arranged according to the degree of difficulty, from easy to difficult. | 4.33 | High |
| 10. The items do not overlap with each other; no duplication of items was done. | 5.00 | Very High |
| **Overall** | **4.60** | **Very High** |

## Reliability Coefficient of the Test in Plane Trigonometry

The Table 7 shows the reliability coefficient of the instrument using Cronbach's Alpha. Thus, a reliability coefficient of .909 was obtained. This helped to establish the internal consistency of the test. The result reveals that the test has a high reliability and should be used for assessing students' content knowledge in plane trigonometry. This is strengthened by the research study of Alfrits Roul Sinadia and Surya Jatmika (2020) that the consistency index of the test already met the minimum reliability coefficient required.

**Table 7.** Reliability coefficient of the test.

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .909 | .909 | 50 |

## Difficulty Index and Level of the Test in Trigonometry

Table 8 lists the difficulty index of the test. The value of difficulty index of the 50 items test was ranges from 0.43 to 0.77. There 28 items belonged to average or moderately difficult. These items are 2,4,5,6,7,8,9,10,11,12,14,15,16,17,20,21,22,23,25,28,30,36,38,44,46,48,49,50 and there are 22 items belonged to easy. This means that there are more students got the correct answer and more students mastered the content. This is supported by the research findings of Boopathiraj and Chellamani (2013) that most of the items were falling in acceptable range of difficulty.

**Table 8.** Difficulty index and level of the test.

| ITEM | HIGHEST SCORE | LOWEST SCORE | DIFFICULTY INDEX | DIFFICULTY LEVEL |
|---|---|---|---|---|
| 1 | 12 | 7 | 0.63 | EASY |
| 2 | 9 | 4 | 0.43 | MODERATE |
| 3 | 13 | 7 | 0.67 | EASY |
| 4 | 11 | 6 | 0.57 | MODERATE |
| 5 | 10 | 4 | 0.47 | MODERATE |
| 6 | 11 | 6 | 0.57 | MODERATE |
| 7 | 10 | 5 | 0.50 | MODERATE |
| 8 | 11 | 4 | 0.50 | MODERATE |
| 9 | 9 | 4 | 0.43 | MODERATE |
| 10 | 11 | 7 | 0.60 | MODERATE |
| 11 | 11 | 6 | 0.57 | MODERATE |
| 12 | 9 | 5 | 0.47 | MODERATE |
| 13 | 12 | 7 | 0.63 | EASY |
| 14 | 11 | 6 | 0.57 | MODERATE |
| 15 | 11 | 6 | 0.57 | MODERATE |
| 16 | 13 | 5 | 0.60 | MODERATE |
| 17 | 10 | 5 | 0.50 | MODERATE |
| 18 | 12 | 7 | 0.63 | EASY |
| 19 | 12 | 7 | 0.63 | EASY |
| 20 | 10 | 6 | 0.53 | MODERATE |
| 21 | 11 | 5 | 0.53 | MODERATE |
| 22 | 11 | 7 | 0.60 | MODERATE |
| 23 | 10 | 6 | 0.53 | MODERATE |
| 24 | 14 | 8 | 0.73 | EASY |
| 25 | 11 | 6 | 0.57 | MODERATE |
| 26 | 12 | 7 | 0.63 | EASY |
| 27 | 13 | 7 | 0.67 | EASY |
| 28 | 12 | 6 | 0.60 | MODERATE |
| 29 | 14 | 9 | 0.77 | EASY |
| 30 | 11 | 7 | 0.60 | MODERATE |
| 31 | 12 | 7 | 0.63 | EASY |
| 32 | 12 | 7 | 0.63 | EASY |

| | | | | |
|---|---|---|---|---|
| 33 | 12 | 7 | 0.63 | EASY |
| 34 | 13 | 8 | 0.70 | EASY |
| 35 | 12 | 7 | 0.63 | EASY |
| 36 | 13 | 5 | 0.60 | MODERATE |
| 37 | 13 | 8 | 0.70 | EASY |
| 38 | 10 | 5 | 0.50 | MODERATE |
| 39 | 13 | 6 | 0.63 | EASY |
| 40 | 13 | 8 | 0.70 | EASY |
| 41 | 14 | 9 | 0.77 | EASY |
| 42 | 14 | 7 | 0.70 | EASY |
| 43 | 13 | 8 | 0.70 | EASY |
| 44 | 10 | 5 | 0.50 | MODERATE |
| 45 | 13 | 8 | 0.70 | EASY |
| 46 | 10 | 6 | 0.53 | MODERATE |
| 47 | 12 | 8 | 0.67 | EASY |
| 48 | 12 | 4 | 0.53 | MODERATE |
| 49 | 11 | 6 | 0.57 | MODERATE |
| 50 | 12 | 4 | 0.53 | MODERATE |

**Discrimination Index and Level of the Test in Plane Trigonometry**

Table 9 presents the discrimination index and level of the test in plane trigonometry. The value of discrimination index of the test ranges from 0.27 to 0.53. There are 13 items, 29 items, and 8 items belonged to very good item, good item, and marginal item respectively. The 13 items are "very good items". Result shows that items 3,5,21,24,27 has a value of discrimination index of 0.40, items 8,39,43 obtain a value of discrimination index 0.47, and item 16, 36,48, and 50 obtained a value of discrimination index 0.53. The 8 items which are marginal items with discrimination index of 0.27 are 10,12,20,22,23,30,46,47. This means that the 8 items needs some revision. The 29 items namely 1,2,4,6,7,9,11,13,14,15,17,18, 19,25,26,29,31,32,33,34 are good items with a value of discrimination index 0.33. This means that the 29 items are reasonably good items but possibly for improvement. This is supported by the research findings of Boopathiraj and Chellamani (2013) that most of the items were falling in acceptable range of discrimination.

**Table 9.** Discrimination index and level of the test.

| ITEM | HIGHEST SCORE | LOWEST SCORE | DISCRIMINATION INDEX | DISCRIMINATION LEVEL |
|---|---|---|---|---|
| 1 | 12 | 7 | 0.33 | GOOD ITEM |
| 2 | 9 | 4 | 0.33 | GOOD ITEM |
| 3 | 13 | 7 | 0.40 | VERY GOOD ITEM |
| 4 | 11 | 6 | 0.33 | GOOD ITEM |
| 5 | 10 | 4 | 0.40 | VERY GOOD ITEM |
| 6 | 11 | 6 | 0.33 | GOOD ITEM |
| 7 | 10 | 5 | 0.33 | GOOD ITEM |

| | | | | |
|---|---|---|---|---|
| 8 | 11 | 4 | 0.47 | VERY GOOD ITEM |
| 9 | 9 | 4 | 0.33 | GOOD ITEM |
| 10 | 11 | 7 | 0.27 | MARGINAL ITEM |
| 11 | 11 | 6 | 0.33 | GOOD ITEM |
| 12 | 9 | 5 | 0.27 | MARGINAL ITEM |
| 13 | 12 | 7 | 0.33 | GOOD ITEM |
| 14 | 11 | 6 | 0.33 | GOOD ITEM |
| 15 | 11 | 6 | 0.33 | GOOD ITEM |
| 16 | 13 | 5 | 0.53 | VERY GOOD ITEM |
| 17 | 10 | 5 | 0.33 | GOOD ITEM |
| 18 | 12 | 7 | 0.33 | GOOD ITEM |
| 19 | 12 | 7 | 0.33 | GOOD ITEM |
| 20 | 10 | 6 | 0.27 | MARGINAL ITEM |
| 21 | 11 | 5 | 0.40 | VERY GOOD ITEM |
| 22 | 11 | 7 | 0.27 | MARGINAL ITEM |
| 23 | 10 | 6 | 0.27 | MARGINAL ITEM |
| 24 | 14 | 8 | 0.40 | VERY GOOD ITEM |
| 25 | 11 | 6 | 0.33 | GOOD ITEM |
| 26 | 12 | 7 | 0.33 | GOOD ITEM |
| 27 | 13 | 7 | 0.40 | VERY GOOD ITEM |
| 28 | 12 | 6 | 0.40 | VERY GOOD ITEM |
| 29 | 14 | 9 | 0.33 | GOOD ITEM |
| 30 | 11 | 7 | 0.27 | MARGINAL ITEM |
| 31 | 12 | 7 | 0.33 | GOOD ITEM |
| 32 | 12 | 7 | 0.33 | GOOD ITEM |
| 33 | 12 | 7 | 0.33 | GOOD ITEM |
| 34 | 13 | 8 | 0.33 | GOOD ITEM |
| 35 | 12 | 7 | 0.33 | GOOD ITEM |
| 36 | 13 | 5 | 0.53 | VERY GOOD ITEM |
| 37 | 13 | 8 | 0.33 | GOOD ITEM |
| 38 | 10 | 5 | 0.33 | GOOD ITEM |
| 39 | 13 | 6 | 0.47 | VERY GOOD ITEM |
| 40 | 13 | 8 | 0.33 | GOOD ITEM |
| 41 | 14 | 9 | 0.33 | GOOD ITEM |
| 42 | 14 | 7 | 0.47 | VERY GOOD ITEM |
| 43 | 13 | 8 | 0.33 | GOOD ITEM |
| 44 | 10 | 5 | 0.33 | GOOD ITEM |
| 45 | 13 | 8 | 0.33 | GOOD ITEM |
| 46 | 10 | 6 | 0.27 | MARGINAL ITEM |
| 47 | 12 | 8 | 0.27 | MARGINAL ITEM |
| 48 | 12 | 4 | 0.53 | VERY GOOD ITEM |
| 49 | 11 | 6 | 0.33 | GOOD ITEM |

| | | | | |
|---|---|---|---|---|
| 50 | 12 | 4 | 0.53 | VERY GOOD ITEM |

## CONCLUSIONS

1. The TOS is very important in constructing a competency-based summative fair tests that give accurate information about students learning. It helps identify the achievement domains being measured.
2. The Plane Trigonometry test instrument can be used to evaluate the effectiveness of instruction and assess students' content knowledge. Moreover, the three experts/judges believed that very little flaws are observed and minor rewording of few items needed.
3. The overall scores of the students are reliable The Plane Trigonometry summative test instrument used in the study has an excellent reliability. Thus, the students belonged to a heterogeneous class.
4. The test items are considered as moderate and easy. Thus, the majority of the respondents in the upper class mastered the lesson.
5. The test items are *good items* and *very good items*. Thus, some items are reasonably good items but need revision.

## RECOMMENDATIONS

1. Construction of summative assessment must always use the table of specification for evidence that the test content has validity and it covers what should be covered.
2. Teachers may use summative test to assess students' achievement when they have covered some of the content areas of the course plane trigonometry.
3. The test should be used to prepare students for internal and external examinations.
4. There is a need to observe content knowledge of the lower class respondents. It is recommended that tutorial or remediation classes may be organized for better mathematical achievement.
5. Reword and revise some items for the students to better understand the questions. The principles of test construction must be followed.
6. Teachers of the different schools are encourage to conduct validity and reliability of their test.

## REFERENCES

Alade, O. M & Omoruyi, I. V. (2014). Table of specification and its relevance in educational development assessment. *European Journal of Educational and Development Psychology. 2*(1), 1-17.

Alfrits Roul Sinadia & Surya Jatmika(2020). Development of the academic achievement test for undergraduate students. *Advances in Social Science, Education and Humanities Research*, *438*, 191-195.

Boopathiraj, C. & Chellamani, K. (2013). Analysis of test items on difficulty level and discrimination index in the test for research in education. *International Journal of Social Science & Interdisciplinary Research,2*(2), 189-193.

Gabuyo, Y. (2014). *Assessment of learning 1 textbook and revie*w. Rex Book Store Inc.

Gur, H. (2009) Trigonometry learning. *New Horizons in Education*, *57*(1), 67-80. Retrieved from https://eric.ed.gov/?q=trigonometry&pg=2&id=EJ860819

Hasan, R. *Qualities of a good test. english literature essay*. Retrieved from http://allrfree.blogspot.com/2010/03/qualitiesof-good-test.html

Nworgu, B. G. (2003). *Educational measurement and evaluation theory and practice*. Nsukka: University Trust Publishers.

Onunkwo, G.I.N. (1998). Cognitive hierarchies of the junior secondary certificate examination questions in integrated science. *Review of Education, Xv*(1), 200-208.

Osadebe, P. U. (2001). *Construction and validation of test.* A seminar paper presented at the University of Port Harcourt.

Osadebe, P.U., & Kpolovie, P.J. (2008). Validation of educational research and evaluation in Nigeria. *Journal of Vocational Science And Educational Developmen*t. *8*(1), 103-108.

Osadebe, P.U (2012) *Procedures for construction, validation and standardization of test*. A Seminar Paper Presented at Delta State University, Abraka.

Osadebe, P.U. (2013). Evaluation techniques, DELSU. *Journal of Educational Research and Development*, *12*(1), 56-63.

Osadebe, P.U. (2014). Standardization of test for assessment and comparing of students' measurement. *International Education Studies*, 7(5) 94-103.

Popham, W.J. (2014). *Classroom assessment: What teachers need to know* (7th ed.). Boston, MA: Pearson.

Priyambodo, Efran (2016). Validity and realibility of chemistry systemic multiple choices questions (CSMCQs). *International Journal of Evaluation and Research in Education (IJERE), 5*(4), 306~309.